37

Adaptive control of converters dominated distribution network based on digital twins

Hui Cai

Ilmenauer Beiträge zur elektrischen Energiesystem-, Geräte- und Anlagentechnik

TECHNISCHE UNIVERSITÄT
ILMENAU

Hui Cai

Adaptive control of converters dominated distribution network based on digital twins

Ilmenauer Beiträge zur elektrischen Energiesystem-, Geräte- und Anlagentechnik (IBEGA)

Herausgegeben von
Univ.-Prof. Dr.-Ing. Dirk Westermann
(Fachgebiet Elektrische Energieversorgung) und
Univ.-Prof. Dr.-Ing. Frank Berger
(Fachgebiet Elektrische Geräte und Anlagen)
an der Technischen Universität Ilmenau.

Band 37

Hui Cai

Adaptive control of converters dominated distribution network based on digital twins

Universitätsverlag Ilmenau
2024

Impressum

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über http://dnb.d-nb.de abrufbar.

Diese Arbeit hat der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität Ilmenau als Dissertation vorgelegen.

Tag der Einreichung:	10. Februar 2023
1. Gutachter:	Prof. Dr.-Ing. Dirk Westermann (Technische Universität Ilmenau)
2. Gutachter:	Univ.-Prof. Dr. sc. Andreas Ulbig (RWTH Aachen)
3. Gutachter:	Prof. Dr.-Ing. Jens Haubrock (Hochschule Bielefeld)
Tag der Verteidigung:	23. September 2023

Technische Universität Ilmenau/Universitätsbibliothek
Universitätsverlag Ilmenau
Postfach 10 05 65
98684 Ilmenau
https://www.tu-ilmenau.de/universitaetsverlag

ISSN 2194-2838
DOI 10.22032/dbt.59235
URN urn:nbn:de:gbv:ilm1-2023000383

Acknowledgment

At this moment of completing my dissertation, I would like to express my heartfelt gratitude to those who have supported and accompanied me.

First and foremost, I would like to thank my parents who gave me the opportunity and support to further my education in Germany.

Next, I would like to thank my doctoral father Mr. Prof. Dirk Westermann for the opportunity to work in his research group. He also gave me support and professional comments throughout all the challenging moments of completing this work.

Furthermore, I would like to express my gratitude to Dr. Teng Jiang for giving me strength and encouragement to reach the end of this journey.

Moreover, I am genuinely grateful for my colleagues Xinya and the rest of my colleagues as well as my friends, whose support and accompaniment have been very important to me.

Lastly, I would like to thank myself for the perseverance in completing this research work.

Ilmenau, October 2022

Hui Cai

Kurzfassung

Mit der zunehmenden Einspeisung und Verteilung von erneuerbaren Energien im Stromnetz wächst die Anzahl der Stromrichter (SR). Dies kann zu einer Destabilisierung des Systems aufgrund mehrerer Faktoren führen, von denen einer auf die Reglerparameter der Stromrichter zurückzuführen ist. Es erfordert eine Parametrierung und Optimierung der genannten Parameter. Die vorliegende Arbeit widmet sich der zentralen Forschungsfrage, wie die Stabilität eines Stromrichter-dominierten Verteilnetzes bei variierenden Betriebspunkten durch das dynamisch adaptive Regelverfahren verbessert werden kann. Für die Parametrierung der Regler wird ein Reinforcement Learning (RL)-Agent vorgeschlagen, der adaptiv geeignete Reglerparameter in Abhängigkeit von den Systemzuständen erzeugt, um die Stabilität zu gewährleisten und zu verbessern. In dieser Arbeit wird zunächst ein digitaler Zwilling (DZ) des untersuchten Netzes mit Hilfe von Parameterschätzungen konstruiert. Das DZ-Modell wird verwendet, um Trainingsdaten für einen auf einem künstlichen neuronalen Netz (KNN) basierenden Zustandsschätzer zu erzeugen, der den stabilen Zustand des Systems präzise und effizient bestimmen soll. Darüber hinaus wird ein Kleinsignal-Stabilitätsindikator (SI) unter Verwendung des Dämpfungsgrades des dominanten Eigenwerts für das DZ-Modell entwickelt. Da der SI dem RL-Agenten die Systemstabilitätsreserve während des RL-Trainings anzeigt, ist der Agent letztendlich in der Lage, optimale Reglerparameter für die Stromrichters auszugeben. Mittels numerischer Fallstudien wird die Funktionsfähigkeit des vorgeschlangenen Ansatzes verifiziert, dass die Netzstabilität durch die vorgeschlagene adaptive Kontrollmethode verbessert werden kann.

Abstract

With the increasing feed-in and distribution of renewable energies in the power system, the number of converters is growing. This can lead to the destabilization of the system due to several factors, one of which is the controller parameter of the converter. This requires parameterization and optimization of the mentioned parameters. The presented work is devoted to the main research question of how the stability of a converters dominated distribution grid is improved for varying operating points by the dynamic adaptive control method. To carry out the parameterization of the controllers, the Reinforcement Learning (RL) agent for adaptively creating proper controller parameters according to the system states to improve and ensure stability is proposed. In this work, a digital twin (DT) of the studied network is first constructed with the help of parameter estimation. The DT model is utilized to generate training data for an artificial-neural-network-based state estimator, which is dedicated to arcuately and efficiently determining the system's stable state. Moreover, a small signal stability indicator (SI) using the damping ratio of the dominant eigenvalue for the DT model is developed. With the SI indicating to the RL agent the system stability margin during the RL training, the agent can ultimately output optimal controller parameters for the converters. Numerical case studies are used to verify the viability of the proposed approach and that network stability can be improved by the proposed adaptive control method.

Table of contents

1 Introduction

Fossil fuels, such as coal, oil, and gas, are the largest contributors to global climate change, accounting for over 75% of global greenhouse gas emissions and nearly 90% of all carbon dioxide emissions [1]. To avoid the harsh effects of climate change, it is necessary to end our dependence on fossil fuels and invest in clean, sustainable, and reliable alternative energy sources. Germany is implementing the “Energiewende”, a national energy transition strategy through replacing the use of fossil fuels with renewable energies (RE) [2–5].

Electricity demand is expected to grow rapidly with the accelerated electrification of all energy-using sectors. By 2030, the share of electricity in Germany's final energy demand will exceed 26% [6]. RE will play a major role and is expected to account for 80% of total electricity generation in 2030 [7].

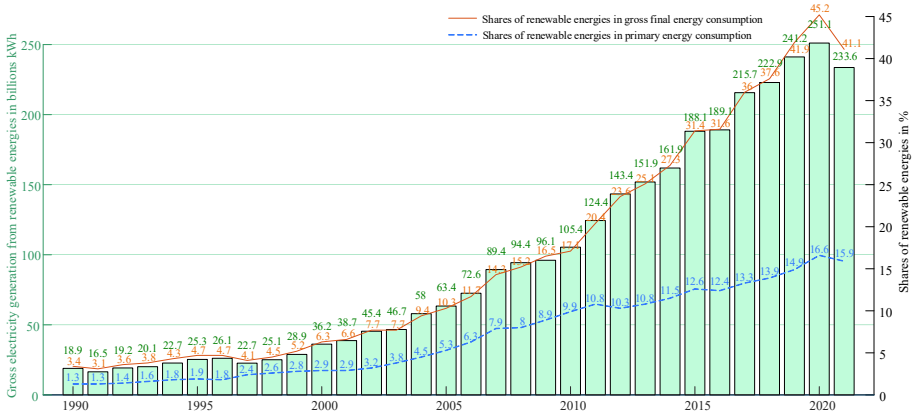


Fig. 1.1: Renewable energy development timelines 1990-2021 in Germany [8]

Fig. 1.1 shows that the share of RE has been increasing significantly over the past two decades. It is a fact that more than 40% of the electricity consumed in Germany last year came from RE plants. However, regarding total energy consumption, the share of RE is significantly lower, at less than 16%. In 2020 Germany's share of RE in electricity consumption reached a historical high of 45.2%, with its share dropping for the first time in 20 years in 2021 due to the slowdown in the expansion of RE as a result of the COVID-19 epidemic. Moreover, because of weaker wind conditions, RE generation fell by 4.1% last year to 233.6 billion kWh. To achieve the energy transition goals, RE expansion and integration into the power system will accelerate over the next eight years.

The development of RE can lead to stability problems in the electricity system [9]. The Renewable Energies Act (Erneuerbare Energien Gesetz EEG) stipulates that electricity from renewable sources should be fed into the grid before conventional electricity [10]. In addition, conventional power plants must reduce generation or be shut down during the oversupply [11]. The shutdown of large conventional coal-fired power plants and large three-phase generators weakens power system stability [12]. The mechanical inertia of rotating masses, such as turbines and generators,

can prevent the speed changes caused by failures in power systems (e.g., failure of large power plants or transmission lines) [13]. The speed of the turbine or generator determines the system frequency [14]. Therefore, the system can use the inertia of the rotating masses for frequency support. The generator of a wind power plant can simulate the inertia of the turbine of a conventional power plant. However, compared to the inertia of a conventional large-scale power plant, the inertia of a RE plant is smaller. Respective stability problems are most likely to occur in low-inertia grids [15].

The integration of the RE systems to the electricity grid is carried out via converters. The task of a converter is to set and maintain the voltage and current at its input and output to desired values to achieve the predetermined quantities, e.g., active and reactive power. Here, the control loop is introduced. Normally, the controlled system represents the part or the process to be controlled. The controlled variables are influenced within the controlled system. For a converter, its controlled system corresponds to the power supply system (AC-side controlled system) and the DC link (DC-side controlled system). Accordingly, the actuating variable y is the output voltage, or the output current of the converter and the actuator thus designates the converter itself, cf. [16–19].

Due to the control structure and the control parameters, the converter greatly influences the system's stability. [20]. System stability in the small signal range decreases with an increasing number of converters [21]. Currently, the standard optimization method for control parameters in the cascade control of the UR is static, i.e., it is not suitable for dynamic adaptation. Therefore, an adaptive and dynamic optimization method for the control parameters of the converter is developed to improve the characteristics of the excitation system.

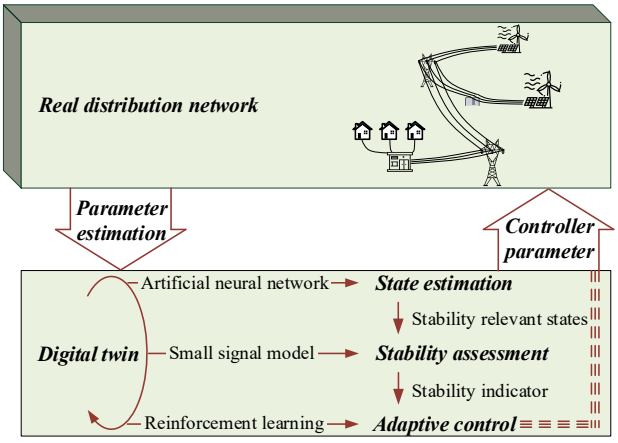


Fig. 1.2: Proposed concept: An automatic control of converters dominated distribution network

In this work, the parameterization of the controllers in converters for adapting the network state is carried out to optimize network behaviour and stability. For this purpose, the following research questions will be addressed:

> *Q1: How to use the digital twin of the distribution network to increase its observability, which includes a limited number of measurement points?*

Since the operating state of the system impacts the adaptive parameterization process, and the state of required system variables needed for the control concept are normally not directly observable, the system has only a limited number of measurement points. Therefore, a digital twin (DT) model that can be used to improve the observability of the converters dominated distribution network is needed. Therefore, the following two sub-questions arise:

> *Q1.1: How to generate a DT for a distribution network from a limited number of measurements?*
> *Q1.2: How to utilize the DT to estimate the state of the system, especially the state that is required for system stability evaluation?*

Once the first question has been settled, the second, which is the focus of this research, is also addressed.

> *Q2: How to apply the converters in the distribution network to improve network stability?*

Before answering this question, a stability indicator (SI) is required to indicate the stability of the network. Moreover, a method is proposed so that the system can be automatically detected for its stability level, which plays an important role in the following adaptive control of the converter. Again, two sub-questions are presented below.

> *Q2.1: Which stability indicator is suitable for detecting network stability and how can their detection be automated?*
> *Q2.2: Which methods will be applied in the adaptive control of the converters?*

To answer the above questions, this work is divided into six chapters. After an introduction, Chapter 2 presents state of the art, which emphasize the methods covered in the proposed concept, as shown in Fig. 1.2, and the related intellectual background as well as their current applications in power systems A basic understanding of the theoretical foundations of machine learning (ML), including the different types of models for ML and their three applications, is presented in Section 2.1. Stability issues due to converters are addressed in Section 2.2. At the same time, the converter, as well as its control structure and stability criteria, are introduced in the same section.

Chapter 3 first presents the creation of the DT by modelling the individual network components and applying the parameter estimation. Subsequently, section 3.2 is dedicated to realizing the state estimation based on the training of the artificial neural network (ANN). Section 3.3 describes the automatic stability detection method for the reference distribution network. This work uses an eigenvalue-based SI for stability detection, which is generated based on the state matrix of the small-signal model of the system. In addition, the method for determining the dominant oscillation frequency and the utilization of ANN as an indicator for fast stability detection are also presented in section 3.3. With a fast-generated SI, it can then be implemented into the DT model and use reinforcement learning (RL), a method of ML, to train the adaptive control of the model with the help of optimizing the converter controller parameters. This will be described in section 3.4. Then follows a summary of chapter 3.

Chapter 4 covers the design of numerical case studies for the DT approach, the ANN state estimator, automatic stability detection and network stabilization measures using the adaptive control approach. The simulation results and discussion corresponding to each section in chapter 4 are placed in chapter 5.

Chapter 6 gives a summary of this work and guidelines for potential further research. The parameters used and some of the results are detailed in the appendix.

2 State of the art

In recent years ML has been increasingly applied in power systems [22–25]. Chapter 2.1 introduces the basic theoretical foundations of ML, including models and algorithmics for three ML types: supervised learning, unsupervised learning, and reinforcement learning. In addition, the theoretical knowledge of three ML applications in this work is presented. In chapter 2.2, the four stability issues caused by converter in the network, namely the rotor angle stability issues, the frequency stability issue, the voltage stability issue and the converter-driven stability issue are presented in the first section 2.2.1 The structure of the converter and its control loop structure and functions are covered in section 2.2.2. The last section 2.2.3 describes the Lyapunov stability criterion and the dynamic behavior of a system reflected by its eigenvalues.

2.1 Machine Learning and its application

ML consists of different types of models that use different algorithmic techniques. Based on the type of data and desired outcome, ML can be divided into three categories of learning models: supervised learning, unsupervised learning, and reinforcement learning. The main applications of the three learning models are shown in the following Fig. 2.1.

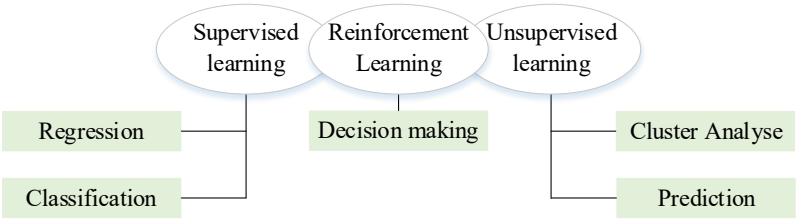


Fig. 2.1: ML classification and application

Supervised learning means the given data to train algorithms are labeled with the desired value. The machine learns by finding correlations between the inputs and outputs to classify data or to predict outcomes accurately. It is widely used for fault detection and state estimation in power systems. In the unsupervised learning models, unlabelled and unstructured are the features of the models´ data. The machine examines the input data and recognizes patterns and correlations using all relevant, accessible data. A common application of unsupervised learning is forecasting and pattern recognition of load. Rather than taking a specific dataset, the RL model is given a set of allowable actions, rewards, and potential states. The model learns through interacting with an operational environment and collects the system's responses in different situations. In RL models, a user-defined "reward" is numerical and programmed into the algorithm as an object that the system wants to optimize. RL has gained considerable attention in applications related to voltage and frequency control in power systems [26].

Artificial neural network

ANN is the model structure used for ML in this work. Compared to natural neurons, artificial neurons are simplified to a large extent. The basic model is Rosenblatt's perceptron [27]. Here x is the input vector and w_j is the weight vector representing the weights from each input to output

y_j. Thus, the activation is understood as the product of x and w_j. The signal is passed into the activation function after weighted summation. Here a nonlinear activation function is used. Through it, the ambiguity is reduced, and the correlation of different neurons is eliminated. The output of one neuron has a certain transfer function:

$$y_j = g(\sum_i w_{ij} x_i + b_{ij}) = g(\mathbf{w}_\mathbf{j}^\mathbf{T} x + \mathbf{b}) \tag{2.1}$$

Only "local" outputs can be obtained through the upper-mentioned networks from the inputs x and the weights w_j. In classification problems, the output must produce a probability between 0 und 1.

2.1.1 Digital twin of an electrical network

In the context of this work, DT will be understood as a real-time digital representation of physical components based on measurement data and analytical knowledge. It enables power suppliers to transform operations with implementable insights to improve commercial decisions. In other words, the power grid operator can improve operations, reduce unplanned outages and manage fluctuations in market conditions, fuel costs, and weather conditions [28, 29].

A DT can be used in a power system, for example, to estimate the current state in a digital power plant and to design the appropriate controller in converters, whose amount recently increased significantly in the power system due to the energy transition. By training with numerous data collected from the running power system, the DT can model and track the system's state. With the DT, plant operators can optimize immediate and transient control of the plant in terms of efficiency or performance. Furthermore, the operators can make informed decisions on components considering their lifespan, distribute loads over time, and perform the proper maintenance tasks at the ideal time. The DT makes it possible to evaluate different power system scenarios and increase efficiency in making decisions [29, 30]. DT can also demonstrate the actual and possible future system state, making it a solution for further improving the monitoring and control of the power systems [31].

A standard procedure for creating a DT based on parameter estimation is presented in Fig. 2.2 [32]. Within the scope of this work, a DT of a distribution network is built. In the first step, a network model is created digitally according to the known network topology. Then, a connection between the real power network and the virtual DT shall be established. Under the same setup X, a DT shall behave like its real network. For this purpose, parameter estimation is used to build the DT, whose goal is to find the best set of model parameters p that achieves the best fit of the behavior of the real power network Y and the DT Y′. With the known network topology and the estimated model parameters p, the DT of the network can be created and updated continuously according to the cycle of the measurement data.

Parameter estimation is often based on grey-box models. Grey-box models assume that the system's structure is known, but the partial system parameters are unknown; thus, building theoretical models from a set of differential equations with the optimized parameters could accurately describe the system's behavior [33–35].

In the following, an optimization technique is presented whereby the model parameters $\underline{p}$ can be estimated. The optimization goal is to minimize the error between the simulated $\underline{Y}_{\mathrm{s}}$ and measured outputs $\underline{Y}_{\mathrm{m}}$.

$$\underline{e}_k(\underline{p}) = \underline{Y}_{sk} - \underline{Y}_{\mathrm{m}k}, k = 1,2,...,N \tag{2.2}$$

where $\underline{Y}_{\mathrm{s}} = f(\underline{p})$.

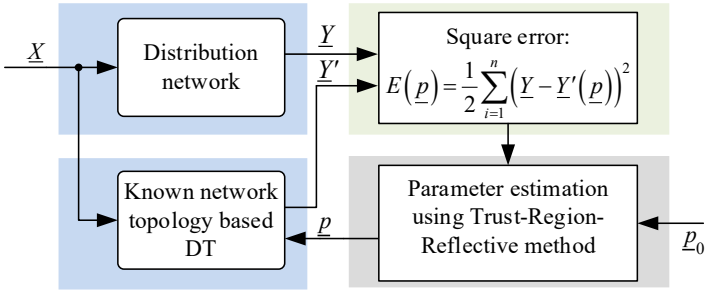


Fig. 2.2: Procedure of DT creation based on parameter estimation

There are different methods for estimating the parameters of a dynamic model. Here the widely used method of nonlinear-least squares is presented, characterized by a good convergence performance. This method determines the best functional match for the estimated data by minimizing the sum of the squares of errors $E(\underline{p})$.

$$E(\underline{p}) = \frac{1}{2}\sum_{k=1}^{N}\left[\underline{Y}_{sk} - \underline{Y}_{\mathrm{m}k}\right]^2 = \frac{1}{2}\sum_{k=1}^{N}\left[\underline{e}_k(\underline{p})\right]^2, k = 1,2,...,N \tag{2.3}$$

This minimization problem is solved using the Trust-Region-Reflective-Algorithm (TRR), which, compared to the methods of Gauss-Newton and Levenberg-Marquardt, offers acceptable accuracy with less effort [36, 37].

The TRR algorithm approximates the square error $E(\underline{p})$ in the vicinity of $\boldsymbol{N}$ of $\underline{p}_i$ with a quadratic function $L(\underline{h})$ by its Taylor series expansion at $\underline{p}_i$. The next iterative point $\underline{p}_{i+1}$ can be found within the trust region $\boldsymbol{N}$. The iteration steps $\underline{h}_i = \underline{p}_{i+1} - \underline{p}_i$ are determined by an approximate solution with the equation (2.4). This is referred to as a sub-problem of the TRR method [38]. The following equation (2.4)

$$E\left(\underline{p}_{i+1}\right) = E\left(\underline{p}_i + \underline{h}_i\right) \approx L\left(\underline{h}_i\right) \equiv E\left(\underline{p}_i\right) + \left[\underline{g}\left(\underline{p}_i\right)\right]^T \underline{h}_i + \frac{1}{2}\underline{h}_i{}^T \underline{H}\left(\underline{p}_i\right)\underline{h}_i \tag{2.4}$$

restricted to $\left\|\underline{h}_i\right\| \le \Delta_i$, where $\underline{g}$ is the gradient of $E(\underline{p})$

$$\underline{g} = \nabla E\left(\underline{p}\right) \tag{2.5}$$

and $\underline{H}$ is the Hessian matrix of $E(\underline{p})$.

$$\underline{H} = \nabla^2 E\left(\underline{p}\right) \tag{2.6}$$

Both are calculated at $\underline{p}_i$. The degree of reduction is defined to enable the confidence interval's size Δ_i to be adjusted.

$$\rho_i = \frac{E\left(\underline{p}_i\right) - E\left(\underline{p}_i + \underline{h}_i\right)}{L(0) - L\left(\underline{h}_i\right)} = \frac{\text{Actual reduction}}{\text{Predicted reduction}} \tag{2.7}$$

If this degree of reduction is ρ_i, the following applies to the different areas:

- $\rho_i > \frac{3}{4}$: The step size $\underline{h}_i$ is accepted and the size of the confidence region Δ_i is expanded to $2\Delta_i$ in the next iteration.
- $\frac{1}{4} \le \rho_i \le \frac{3}{4}$: The step size $\underline{h}_i$ is accepted and the size of the confidence region Δ_i remains unchanged in the next iteration.
- $\rho_i < \frac{1}{4}$: The step size $\underline{h}_i$ is rejected, i.e. the values of the searched parameters $\underline{p}_i$ remain the same, and the size of the confidence region is reduced to $\frac{1}{4}\Delta_i$ in the next iteration.

A solution to Eq.(2.4) can be found by determining the Lagrange multiplier $\lambda > 0$ in such a way that $\underline{H} + \lambda \cdot \underline{I}$ is positive definite. Here λ is defined in such a way that it increases with decreasing confidence area size Δ_i, namely $\lambda \to +\infty$ when $\Delta_i \to 0^+$. If $\underline{h}^*$ is a solution for Eq. (2.4), then $\underline{h}^*$ is a solution to an equation of the form

$$\left(\underline{H} + \lambda \cdot \underline{I}\right) \cdot \underline{h}^* = -\underline{g} \tag{2.8}$$

with $\lambda \cdot (\Delta_i - \left\|\underline{h}^*\right\|) = 0$. Then the step size can be calculated with the following equation.

$$\underline{h}^* = -\left(\underline{H} + \lambda \cdot \underline{I}\right)^{-1} \underline{g} \tag{2.9}$$

If y varies between 0 and $+\infty$, the corresponding search direction of the iteration step $\underline{h}_i$ changes between the Newton direction [39] and a multiple of the negative gradient, according to Fig. 2.1. The cause is presented below with the equations (2.10) and (2.11). If $\underline{H} = \nabla^2 E\left(\underline{p}_i\right)$ is positive definite and the confidence region size Δ_i is large enough, i.e. λ is approximately zero, is the solution to the sub-problem of the trust region.

$$\underline{h}^* = -\left(\nabla^2 E\left(\underline{p}\right)\right)^{-1} \nabla E\left(\underline{p}\right) \tag{2.10}$$

The above equation can be derived from Eq. (2.8) with the insertion of $\lambda = 0$, and Eq. (2.5) and (2.6) can be derived. That is, $\underline{h}^*$ is in the Newton direction. When Δ_i falls to zero, λ increases

to infinity, which means that the sub-term $\lambda \cdot \underline{I}$ increasingly dominates the term $\underline{H} + \lambda \cdot \underline{I}$. Therefore, the solution equation of Eq. (2.8) by an approximation $\underline{H} + \lambda \cdot \underline{I} \rightarrow \lambda \cdot \underline{I}$ with insertion of Eq. (2.5) can be derived.

$$\underline{h}^* \rightarrow -\frac{1}{\lambda} \nabla E\left(\underline{p}\right) \tag{2.11}$$

This means that when $\Delta_i \rightarrow 0^+$ the iteration step $\underline{h}_i$ is in the steepest descent direction [37, 40].

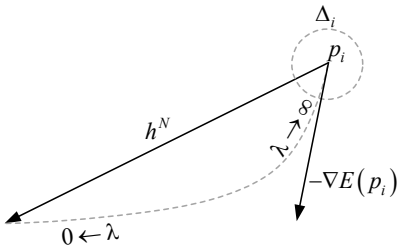


Fig. 2.3: Search direction of the iteration step

After determining the search direction and the width of the step $\underline{h}_i$, the estimated parameters $\underline{p}_i$ are calculated by

$$\underline{p}_{i+1} = \begin{cases} \underline{p}_i + \underline{h}_i & \text{when } \rho_i \geq \frac{1}{4} \\ \underline{p}_i & \text{when } \rho_i < \frac{1}{4} \end{cases} \tag{2.12}$$

The iteration of the TRR algorithm ends when the gradient $\underline{g} = 0$. As a result, the algorithm outputs the estimated model parameters $\underline{p}$. With these parameters, the DT model of the network can be set up, and further investigations can be carried out on the DT.

2.1.2 Data-driven state estimation

Assessing the power network's steady state allows system operators to carry out power system planning, operation, control, and stability analysis. Power flow calculation is a fundamental tool for analyzing the power systems in steady-state operation. It was originally performed for the transmission network, but with the penetration of distributed energy sources, the target has been extended to the distribution network [41]. The power flow problem is a time-consuming task that consists of a set of nonlinear algebraic equations. The conventional method to solve the power flow problem is the Newton-Raphson approach. Due to its superior computational speed over the conventional power flow method, the ANN algorithm is implemented to reduce the time in calculating power system steady state. An ANN model is then trained and utilized as a steady-state estimator in this work. The estimated state will be applied to predict the system stability when the operating points change, which requires a fast and exact steady-state estimation in the distribution network.

In recent years, the NNs have been proposed in many literatures for solving power flow problems. For example, a radial basis function NN using direct system measurements as input is proposed in [42], which is successfully applied to a 39-bus New England test system. The simulation results illustrate that the trained NN can be used to estimate the power flow output with high accuracy almost instantaneously under different operating conditions. [43] provides a counter-propagation NN to solve the power flow problem under different loading/contingency conditions, which is demonstrated by the computation of bus voltage magnitudes and voltage angles in the IEEE 14-bus system. Different ANN architectures, such as multilayer perceptrons, decision trees, and so on, for predicting line loading and voltage magnitude are compared in [44]. These models are tested in three realistic power systems, and the required training time, prediction times, and prediction errors are evaluated for each ANN model. A physics-guided MLP-NN to solve the power flow problem is proposed in [45], which forces both power flow solving and power flow modelling tasks. This MLP-NN can perform power flow analysis in situations where accurate system parameters and control logic are difficult to obtain.

In this work, the main goal of applying NN in state estimation is to achieve a real-time estimation of the system's steady state and even prediction of these states. The estimated states value will be utilized in calculating the system SI, which will be introduced in Chapter 2.2.3. An NN-steady-state estimator stores the system's structure and parameters into the network's weights and biases after training. Therefore, the bus voltages in the system can be computed straightforwardly, and the fast computation speed is then guaranteed.

This work proposes a feedforward NN using a scaled conjugate gradient (SCG) training algorithm for steady-state estimation. SCG was demonstrated in [46] to converge faster than the other standard learning algorithm and to handle large-scale problems efficiently.

Scaled Conjugate Gradient Method

The SCG method is a learning algorithm for ANN based on a class of optimization techniques known in numerical analysis as conjugate gradient methods. Compared to other conjugated gradient methods, SCG avoids the line search per learning iteration by using a Levenberg-Marquardt approach [47] to scale the learning rate [46].

With SCG, ANN is first trained by minimizing the global error E, which is defined as follows:

$$E\left(\underline{w}\right)=\frac{1}{L}\sum_{l=1}^{L}E_l \tag{2.13}$$

where, L = Total number of training patterns,

$\underline{w}$ = Vector of weights and bias,

E_l = Error for the training pattern l.

The errors E_l in a pattern l are calculated from the mean square error (1/2 is used because it makes the later derivation easier):

$$E_l\left(\underline{w}\right)=\frac{1}{2}\sum_{i=1}^{n}\left(t_i-o_i\left(\underline{w}\right)\right)^2 \tag{2.14}$$

where, n = Total number of the output neuron,

t_i = Target output of the i-th output neuron for sample *l*,

o_i = Actual output of the i-th output neuron for sample *l* as a function of the weights.

This training algorithm reduces global error by adjusting weights and biases [48].

The iteration of $\underline{w}$ is defined as:

$$\underline{w}_{k+1} = \underline{w}_k + \sigma_k \cdot \underline{p}_k \tag{2.15}$$

where, $\underline{w}_k$ = Vector of current weights and bias,

σ_k = current learning rate,

$\underline{p}_k$ = current conjugated weight vector.

$E_{qw}\left(\underline{y}\right)$ is defined as a quadratic approximation to E in the neighborhood of a point $\underline{w}$.

$$E_{qw}\left(\underline{y}\right) = E\left(\underline{w} + \underline{y}\right) \approx E\left(\underline{w}\right) + E'\left(\underline{w}\right)^{\mathrm{T}} \underline{y} + \frac{1}{2}\underline{y}^T E''\left(\underline{w}\right)\underline{y} \tag{2.16}$$

In order to determine the minimum for $E_{qw}\left(\underline{y}\right)$, the critical points for $E_{qw}\left(\underline{y}\right)$ must be found, i.e. the points $\underline{y^*}$ at which

$$\underline{y^*} E'_{qw}\left(\underline{y^*}\right) = E''\left(\underline{w}\right)\underline{y^*} + E'\left(\underline{w}\right) = 0 \tag{2.17}$$

The step from a starting point $\underline{y}_1$ to a critical point $\underline{y^*}$ can be expressed as a linear combination of $\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_N$.

$$\underline{y^*} - \underline{y}_1 = \sum_{i=1}^{N} a_i \cdot \underline{p}_i, a_i \in \Re \tag{2.18}$$

Theorem 1: Let $\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_N$ be a conjugate system and $\underline{y}_1$ a point in weight space. Let the points $\underline{y}_2, \ldots, \underline{y}_N$ be recursive and are defined by:

$$\underline{y}_{k+1} = \underline{y}_k + a_k \cdot \underline{p}_k \tag{2.19}$$

where, $a_k = \mu_k / \delta_k$,

$\mu_k = -\underline{p}_k^{\mathrm{T}} E'_{qw}\left(\underline{y}_k\right)$,

$\delta_k = \underline{p}_k^T E''\left(\underline{w}_k\right)\underline{p}_k$.

Theorem 2: Let $\underline{y}_1$ be a point in the weight space and $\underline{p}_l$ and $\underline{r}_l$ equal to the steepest descent vector $-E'_{qw}\left(\underline{y}_l\right)$. $\underline{p}_{k+1}$ is defined recursively by:

$$\underline{p}_{k+1} = \underline{r}_{k+1} + \underline{\beta}_k \cdot \underline{p}_k \tag{2.20}$$

where, $\underline{r}_{k+1} = E'_{qw}\left(\underline{y}_{k+1}\right)$,

$$\underline{\beta}_k = \left(\left|\underline{r}_{k+1}\right|^2 - \underline{r}_{k+1}{}^{\mathrm{T}}\,\underline{r}_k\right) / \underline{p}_k^{\mathrm{T}}\,\underline{r}_k .$$

The most important step in minimizing the global error E in the SCG is to estimate the learning rate α_k. The idea is to estimate the term $\underline{s}_k = E''\left(\underline{w}_k\right)\underline{p}_k$ with an unsymmetrical approximation[49].

$$\underline{s}_k = E''\left(\underline{w}_k\right)\underline{p}_k \approx \frac{E'\left(\underline{w}_k + \sigma_k \cdot \underline{p}_k\right) - E'\left(\underline{w}_k\right)}{\sigma_k} + \lambda_k \cdot \underline{p}_k \,, 0 < \sigma_k \ll 1 \tag{2.21}$$

By adjusting a scalar λ_k in each iteration with a view to the sign of $\delta_k = \underline{p}_k^{\mathrm{T}}\underline{s}_k$, which directly indicates whether $E''\left(\underline{w}_k\right)$ is positive definite. If $\delta_k \le 0$ then $E''\left(\underline{w}_k\right)$ is not positive definite and λ_k is increased and $\underline{s}_k$ is estimated again. If the new value of $\underline{s}_k$ in $\hat{\underline{s}}_k$ renamed and the λ_k is increased to $\hat{\lambda}_k$, then

$$\hat{\underline{s}}_k = \underline{s}_k + \left(\hat{\lambda}_k - \lambda_k\right)\underline{p}_k \tag{2.22}$$

Here, the new positive $\hat{\delta}_k$ can be calculated:

$$\hat{\delta}_k = \underline{p}_k^{\mathrm{T}}\hat{\underline{s}}_k = \underline{p}_k^{\mathrm{T}}\left(\underline{s}_k + \left(\hat{\lambda}_k - \lambda_k\right)\underline{p}_k\right) = \delta_k + \left(\hat{\lambda}_k - \lambda_k\right)\left|\underline{p}_k\right|^2 > 0$$

$$\Rightarrow \hat{\lambda}_k > \lambda_k - \frac{\delta_k}{\left|\underline{p}_k\right|^2} \tag{2.23}$$

Equation (2.23) is used to obtain positive $\hat{\delta}_k$. The scalar λ_k should be increased by more than $-\frac{\delta_k}{\left|\underline{p}_k\right|^2}$. Hence, $E''\left(\underline{w}_k\right)$ it will be positive definite. For an optimal solution, a suitable choice of the size of $\hat{\lambda}_k$ is:

$$\hat{\lambda}_k = 2\left(\lambda_k - \frac{\delta_k}{\left|\underline{p}_k\right|^2}\right) \tag{2.24}$$

This leads to the following equation.

$$\hat{\delta}_k = \delta_k + \left(\hat{\lambda}_k - \lambda_k\right)\left|\underline{p}_k\right|^2 = \delta_k + \left(2\lambda_k - 2\frac{\delta_k}{\left|\underline{p}_k\right|^2} - \lambda_k\right)\left|\underline{p}_k\right|^2 = -\delta_k + \lambda_k\left|\underline{p}_k\right|^2 > 0 \tag{2.25}$$

The learning rate is indicated by

$$\alpha_k = \frac{\mu_k}{\delta_k} = \frac{\mu_k}{-\underline{p}_k^T\,\hat{\underline{s}}_k + \lambda_k\left|\underline{p}_k\right|^2} \tag{2.26}$$

Equation (2.26) shows that the values of λ_k directly scale the learning rate a_k.

From equation (2.16) it is known that the values of the Hessian matrix $E''(\underline{w}_k)$ influence the quadratic approximation $E_{qw}(\underline{y})$ of $E(\underline{w}+\underline{y})$. In addition, equation (2.21) shows that λ_k scales the values of $E''(\underline{w}_k)$ i.e. λ_k is referred to as an influencing factor for the degree of approximation of $E_{qw}(\underline{y})$. Accordingly, it is necessary to raise and lower a mechanism of λ_k to get a good approximation $E_{qw}(\underline{y})$ even though $E''(\underline{w}_k)$ is positive definite. To assess the approximation, the dimension Δ_k is defined as:

$$\Delta_k = \frac{E(\underline{w}_k) - E(\underline{w}_k + \alpha_k \cdot \underline{p}_k)}{E(\underline{w}_k) - E_{qw}(\alpha_k \cdot \underline{p}_k)} = \frac{2\delta_k \left[E(\underline{w}_k) - E(\underline{w}_k + \alpha_k \cdot \underline{p}_k) \right]}{\mu_k{}^2} \tag{2.27}$$

Here, Δ_k is a measure of how well $E_{qw}(\alpha_k \cdot \underline{p}_k)$ approximates $E(\underline{w}_k + \alpha_k \cdot \underline{p}_k)$. In the sense that the closer Δ_k is to 1, the better the approximation. The raising and lowering of λ_k is according to the equation:

$$\begin{aligned} &\text{For } \Delta_k > 0.75 : \lambda_k = \frac{1}{4}\lambda_k \\ &\text{For } \Delta_k < 0.25 : \lambda_k = \lambda_k + \frac{\delta_k \left[1 - \Delta_k\right]}{\left|\underline{p}_k\right|^2} \end{aligned} \tag{2.28}$$

With $\underline{w}_{k+1} = \underline{w}_k + \alpha_k \cdot \underline{p}_k$ and $\underline{p}_{k+1} = \underline{r}_{k+1} + \beta_k \cdot \underline{p}_k$ also the learning rate a_k, which is scaled by the λ_k, desired vector of weights and bias $\underline{w}_{k+1}$ can be achieved. The training of ANN with SCG ends when $\underline{r}_k = -E'(\underline{w}_k) = 0$.

2.1.3 Control optimization

PID (proportional-integral-derivative) controllers have been in the field of automatic control for a long time and are fundamental in the industrial automation system. Over 95% of industrial controllers are of the PID type [50, 51]. In the power system area, PID controllers are applied in automatic generation control of multi-area multi-source interconnected power systems [52], in load frequency control for a single area power system for different types of turbines [53], as well as in energy storage system for effective microgrid operations [54] and so on. PID control has the following advantages to offer as an extensive feedback control technique. On the one hand, it has a simple algorithm structure and therefore is feasible, easy to implement, and conveniently adjusted. On the other hand, it responds well to unmeasured disturbances, ensuring the stability and reliability of its controlled system. To achieve a desired control effect on the controlled object, the PID controller parameters tuning is required in the control system [55]. Gaining appropriate controller parameters is still a problem in many practical industrial applications because of high order, time delay and nonlinearity of the controlled objects [56]. The mentioned characteristics are caused mainly by the large-scale use of power electronic devices. The

manual calculation of the parameters is laborious and time-consuming. Using traditional PID parameter tuning is difficult to achieve the optimal PID controller parameters amid continuous industrial development.

In recent years, many heuristics and innovative methods for obtaining the optimal PID controller parameters, such as Quasi-Affine Transformation Evolutionary algorithms, fuzzy logic, biogeography optimization algorithm, genetic algorithms, and so on, have been proposed as intelligent controls in successfully solving nonlinear complex optimization problems [57–60]. As an intelligent self-learning strategy, RL attracts growing attention in the power system control field [61]. RL has been developed in the field of ML, which is devoted to solving control or sequential decision-making. Saénz-Aguirre et al. presented an advanced yaw control strategy for wind turbines based on RL [62]. An innovative, intelligent control strategy based on RL for buck DC-DC converter with constant power loads is constructed in [63]. [64] contributes to applying RL in Multi-Carrier Energy Systems to provide flexibility to meet the residential thermal and electrical demand. RL also plays a great role in the field of PID control optimum. In [65], RL is utilized for training an adaptive PID controller of an inverted pendulum whose parameters change according to the changing state. The closed-loop stability is ensured during policy search by applying RL-based tuning methods in [66]. The presented novel stability-preserving framework is validated in a second order plus dead-time system.

Moreover, RL has been applied in the control of power electrical systems with the participation of converters. In [67], an online tuned PI controller as an actor in the current control of a converter is presented. The constructed controller is compared with the classical PI in different conditions. It is concluded that the proposed controller reduces the convergence time, and tracking error. A policy-free adaptive RL learning algorithm for cooperative secondary voltage control in an islanded microgrid is proposed in [68]. The result shows that the performance of each distributed generator in simultaneous tracking of the AC voltage reference is improved. However, there is little research on RL applications in PID control optimum in the multi-converters system, considering its stability and robustness. This work is aimed at filling this gap. Although there are other online PID controller tuning methods, for example particle swarm optimization algorithm [69], fruit fly optimization algorithm [70], simultaneous perturbation stochastic approximation [71] and genetic algorithm [72] and so on [73], RL is proven to be the only learning algorithm for online controller tuning under the condition of highly nonlinear and noisy systems [67]. In the following, the mathematical foundations and concepts of RL will be first introduced. Furthermore, the algorithm to be utilized is presented.

Elements of RL

RL is implemented to derive a policy, which defines how the agent behaves to maximize the satisfaction of a criterion. Meanwhile, a long-term sum of rewards gained by trial and error is obtained through interactions with a given environment. Fig. 2.4 shows the basic framework of RL, which includes two main elements: agent and environment. The agent consists of two parts: a policy and an RL-algorithm. The interaction between agent and environment is through the state, reward, and action.

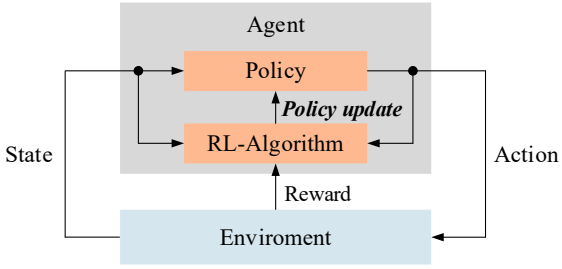


Fig. 2.4: RL basic framework

- Agent

The agent is a decision-maker that acts on the environment under a given state through a policy, which is implemented by a parameterized function, e.g. a deep neural network. Based on actions, observations and rewards, the RL-algorithm continuously updates the policy parameters. The agent's goal is to maximize the reward received from the environment and the cumulative reward calculated from it.

- Environment

The object that interacts with the agent is called the "Environment". It receives a control signal from the agent, creates a new state and returns the state and a reward to the agent after each simulation step. The step reward is utilized to evaluate how well the environment responds to appropriate control signals.

Basic process and algorithms of RL

Markov Decision Process (MDP) is the basic process of RL. To construct the MDP, the following elements need to be clarified:

$A \in \mathcal{A}$: Finite action set, representing the set of actions available to the agent. An action executed at the time *t* is marked as $a_t \in \mathcal{A}$.

$S \in \mathcal{S}$: Finite state set, standing for the set of possible states of the environment. A state at the time *t* is marked as $s_t \in \mathcal{S}$.

P : Transition probabilities $P(s'|s,a), s, s' \in S, a \in A$ are the possibility of the environment reaching a new state s' when the agent takes action a under a state s. In a deterministic case, a certain state s' is reached, therefore $P = 1$. In stochastic cases, the state for the next time step is undetermined, therefore, P is a probability distribution.

$R \in \mathcal{R}$: Reward, depending on S and A, which is a mathematical function to describe how good the current state is or how good an action is. The reward value is transferred from the environment to the agent at each time step. The role of the reward is in the renewal of the policy.

$\gamma \in [0,1]$: Discount factor, which determines the concern degree of RL agents for rewards in the distant future compared to that in the immediate future.

S_t is defined to depend on the state and action of the last time step according to a probability distribution, i.e. $p(S_{t+1} \mid S_t, A_t, ..., S_0, A_0) = p(S_{t+1} \mid S_t, A_t)$, which is called a Markov property. Because each time step commands a corresponding reward, a return function is needed to evaluate the performance of a trajectory:

$$G_t = R_{t+1} + \gamma R_{t+1} + ... = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{2.29}$$

To describe the reward expectation under a certain state and certain chosen action, the so-called action-reward function is as follows.

$$q(s,a) = E\left[G_t \mid S_t = s, A_t = a\right] \tag{2.30}$$

To maximize the reward, the transition between different states should not depend on the transition probabilities but follow an improved statute called a policy. This process is called MDP. MDP aims to find an optimal policy π to maximize the reward function.

In deterministic policy, a particular action is performed under each state. But in the stochastic case, the policy π is a probability distribution, i.e., all possible actions under respective states are emitted by the agent with a probability.

$$\pi(a \mid s) = P[A_t = a \mid S_t = s] \tag{2.31}$$

The reward function and action-reward function under MDP are as follows.

$$\upsilon_\pi(s) = E_\pi\left[G_t \mid S_t = s\right] \tag{2.32}$$

$$q_\pi(s,a) = E_\pi\left[G_t \mid S_t = s, A_t = a\right] \tag{2.33}$$

The corresponding Bellman equations are:

$$\upsilon_\pi(s) = E_\pi\left[R_{t+1} + \gamma \upsilon_\pi(S_{t+1}) \mid S_t = s\right] \tag{2.34}$$

$$q_\pi(s,a) = E_\pi\left[R_{t+1} + \gamma q_\pi(S_{t+1}, A_{t+1}) \mid S_t = s, A_t = a\right] \tag{2.35}$$

Since the state and the action occur in pairs, the relationship between the evaluation function and the action evaluation function must be examined.

In the RL, it is necessary to find an optimal agent with a policy π_* that allows the greatest reward function to be achieved, which means:

$$\upsilon_*(s) = \max_\pi \upsilon_\pi(s) \tag{2.36}$$

Because the above equations are non-linear, approaches such as value iteration, policy iteration, Q-learning, etc. can be used to find an optimal solution to the RL.

Monte Carlo method and temporal difference method

When exploring an environment, the reward is considered and adjusted with the estimate for the state assessment. The "Monte Carlo method" fits this way exactly. The principle of this method

is to solve complex probability problems with the help of sufficient repetition of the random experiment, and its weighted reward function is then calculated. After an episode, a renewal of the reward function $\upsilon_\pi(s_t)$ is carried out. To effectively estimate the value of a state's reward function, a large number of episodes are performed in that state, and then the mean value of the reward function is determined. The same state can appear several times in one or more test episodes. At the end of an episode, the algorithm starts the calculated reward accumulation. The update works by calculating the average reward accumulation received in that state.

The principle of the Monte Carlo update equation is that a given arithmetic mean υ is accumulated with a value *G*.

$$v := \frac{v \cdot (N-1) + G}{N} \tag{2.37}$$

The above equation is converted to addition form:

$$v := v + \frac{G - v}{N} \tag{2.38}$$

The variable in (2.38) is replaced with the variable in MDP, i.e., $\eta = \frac{1}{N(s)}$, $v = \upsilon_\pi(s_t)$ and $G = R_t$, then the equation for the relevant Monte Carlo update rule is obtained.

$$\upsilon_\pi(s_t) := \upsilon_\pi(s_t) + \eta[R_t - \upsilon_\pi(s_t)] \tag{2.39}$$

Here, R_t means the respective mean value of reward when the policy is kept in the state s from time t onwards. Characteristically, it is the principle of the Monte Carlo method that the renewals are carried out first after the completed episode, and the episodic learning process is its prerequisite.

In contrast, the temporal difference (TD) procedure works differently. In the case of exploration, it occurs that the reward determined at the time $t+1$ of the observation differs from the estimate provided. Therefore, the recorded reward should be renewed by assuring the estimate of the reward of a subsequent state. This algorithm is called the "TD method".

In a stochastic environment, the newly generated state of an action is uncertain. Exclusively, the probability of accumulated reward is considered when an action a_t is performed in the state s_t. The achieved state s_{t+1} is also uncertain in this case. One idea for a solution is to apply the learning rate η with $0 < \eta < 1$, which is also called "step size". The principle is that different rewards of the subsequent state are considered according to their probability distributions $\mathcal{P}_{ss'}^a$ from the Markov model. In the learning process, the learning rate η can always be renewed to converge the reward of a state. The convergence of the learning process can occur if the learning rate η is decreased as the number of episodes increases.

Based on the upper specification, Q-learning can be derived, which is an essential success in developing the RL. Characteristically, the renewal of the Q-function is composed of the reward received so far and the reward calculated from the scaled estimate of the subsequent state:

$$\hat{Q}(s_t, a_t) \leftarrow \hat{Q}(s_t, a_t) + \eta[r + \gamma \max_{a_{t+1}} \hat{Q}(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \tag{2.40}$$

Because Q-learning refers to an unknown environment, i.e., the transition possibility $\mathcal{P}_{ss'}^{a}$ is not given, this procedure is classified in the so-called "model-free methods".

Policy-Gradient and Actor-Critic-Method

As mentioned above, if TD algorithms calculate the reward of each state, the action can be executed according to the associated Q-value. But for some continuous-time scenarios, it is impossible to represent the Q-value of every possible state. Therefore, another idea is to try to find the optimal policy.

The algorithms of "policy gradient" can be understood as a method of searching for the optimal policy. And the policy is considered as a parametric control or a function $\pi(s, a, \theta)$ that can output an optimal action. Instead of finding the optimal action in a given situation, an optimal control strategy is searched in this case.

Here, the SoftMax distribution is offered to perform continuous changes and optimization to the policy. This produces a probability distribution over the various available actions according to their reward estimation.

$$\pi(s, a, \theta) = \frac{e^{h(s,a,\theta)}}{\sum_{b \in A}^{A} e^{h(s,b,\theta)}} \tag{2.41}$$

In some cases, Q-value cannot be formed for every possible state, so an action preference $\pi(s, a, \theta)$ is generated from the parameters, which does not denote a Q-value, but presents the preference in action choice.

To improve the control algorithm, the expected reward $J(\theta)$ should be introduced. On the one hand, it represents the relationship between control parameters θ and the evaluation of this control, it also describes the improvement or deterioration of the control when the parameters are changed. A higher value of $J(\theta)$ means the performance of the control.

The learning process, which is based on gradient ascent, aims to maximize the expected reward:

$$\theta_{t+1} = \theta_t + \eta \nabla J(\theta_t) \tag{2.42}$$

For episodic cases, the performance of the control can be identified with the value of the initial state under an initial control:

$$J(\theta) = V^{\pi_\theta}(s_0) \tag{2.43}$$

Then the gradient of expected reward can be presented:

$$\nabla J(\theta) = \nabla V^{\pi_\theta}(s_0) \tag{2.44}$$

The gradient depends on the policy and describes the property of the environment in which the agent operates. To estimate the gradient that depends only on control parameters, when the gradient is affected by the changes in the state distribution, the so-called "Policy Gradient Theorem" is used:

$$\begin{aligned}\nabla J(\theta) &\propto \sum_s \mu(s) \sum_a Q^\pi(s,a) \nabla \pi(a \mid s, \theta) \\ &= E_\pi \left[\sum_a Q^\pi(s_t, a) \nabla \pi(a \mid s_t, \theta) \right]\end{aligned} \tag{2.45}$$

The theorem represents the weighted sum over the expected reward obtained as it would appear under the policy π. The equation represents that the gradient is formed from the expected value over the sum of products from the Q-value and the policy gradient.

In this case, the selected action a_t at the time t, which is used to calculate the policy gradient, must be taken into account to form the sum.

$$\begin{aligned}\nabla J(\theta) &= E_\pi \left[\sum_a Q^\pi(s_t, a) \nabla \pi(a \mid s_t, \theta) \right] \\ &= E_\pi \left[\sum_a Q^\pi(s_t, a) \pi(a \mid s_t, \theta) \frac{\nabla \pi(a \mid s_t, \theta)}{\pi(a \mid s_t, \theta)} \right] \\ &= E_\pi \left[\sum_a Q^\pi(s_t, a_t) \frac{\nabla \pi(a \mid s_t, \theta)}{\pi(a \mid s_t, \theta)} \right]\end{aligned} \tag{2.46}$$

Observing the actions at the time t of a sample, which are selected according to the probability distribution, the weighted sum over all possible actions can be omitted.

For the pair (s_t, a_t) of a period *T*, the "measured" value consists of the cumulative subsequent rewards.

From equation (2.47), the information can be applied to the update rule of the traditional RL algorithm:

$$\begin{aligned}\theta_{t+1} &= \theta_t + G_t \frac{\nabla \pi(a_t \mid s_t, \theta)}{\pi(a_t \mid s_t, \theta)} \\ &= \theta_t + \eta G_t \Delta \ln \pi(a_t \mid s_t, \theta)\end{aligned} \tag{2.47}$$

The next problem is to calculate the gradient $\Delta \ln \pi(a_t \mid s_t, \theta)$. For small state-action space, the action preference $h(s,a,\theta)$ can be tabulated as $\theta_{s,a}$ assigned:

$$h(s,a,\theta_s) = \theta_{s,a}, \; where \; \theta_{s,a} \in \mathbb{R} \tag{2.48}$$

Logarithmize both sides of equation:

$$\ln \pi(s,a,\theta_s) = \theta_{s,a} - \sum_{b \in A}^{A} e^{\theta_{s,b}} \tag{2.49}$$

The Equation can be derived directly to calculate the gradient:

$$\frac{\partial \ln \pi(s,a,\theta_s)}{\partial \theta_{s,i}} = \begin{cases} 1 - \dfrac{e^{\theta_{s,a}}}{\sum_{b\in A}^{A} e^{\theta_{s,b}}} = 1 - \pi(s,a,\theta_s), \ when\ i = a \\ 0 - \dfrac{e^{\theta_{s,i}}}{\sum_{b\in A}^{A} e^{\theta_{s,b}}} = -\pi(s,a,\theta_s), \quad when\ i \neq a \end{cases} \tag{2.50}$$

This is a policy gradient algorithm under the Monte Carlo method, is called "REINFORCE". The learning step of REINFORCE consumes more time but runs more effectively. The step width must also be chosen appropriately.

Another way is an integration of the so-called baseline in the updated rule of policy gradient:

$$\theta_{t+1} = \theta_t + \eta_\theta [G_t - b(s_t)] \nabla \ln \pi(a_t \mid s_t, \boldsymbol{\theta}) \tag{2.51}$$

If the practically cashed reward has a large deviation from the expected baseline, the policy can be strongly adjusted with a baseline. If this baseline is set as 0, this update rule represents no difference to REINFORCE. As the baseline, the currently given estimate of the state value $\hat{V}(s_t)$ can be established:

$$\theta_{t+1} = \theta_t + \eta_\theta [G_t - \hat{V}(s_t)] \nabla \ln \pi(a_t \mid s_t, \boldsymbol{\theta}) \tag{2.52}$$

This means that if the value deviates less from the currently given estimation $\hat{V}(s_t)$ of the state value, a decreasing adaptation of the policy will take place; thus, the learning process will converge.

The definition of the expectation of the cumulative reward G_t is taken into account. The TD error can be used in online adaptation:

$$\theta_{t+1} = \theta_t + \eta_\theta [R_{t+1} + \gamma \hat{V}(s_{t+1}) - \hat{V}(s_t)] \nabla \ln \pi(a_t \mid s_t, \boldsymbol{\theta}) \tag{2.53}$$

This allows the policy gradient algorithm to be combined with the value optimization by TD methods by processing the TD online. The TD is used in two places. On the one hand, the policy is immediately improved by TD errors. The agent that behaves under this policy is called an "Actor". On the other hand, the estimation of the state value $\hat{V}(s_t)$ is updated. The component that outputs and updates state value is called "Critic". The algorithm with this basic principle is called the "Actor-Critic" algorithm [74].

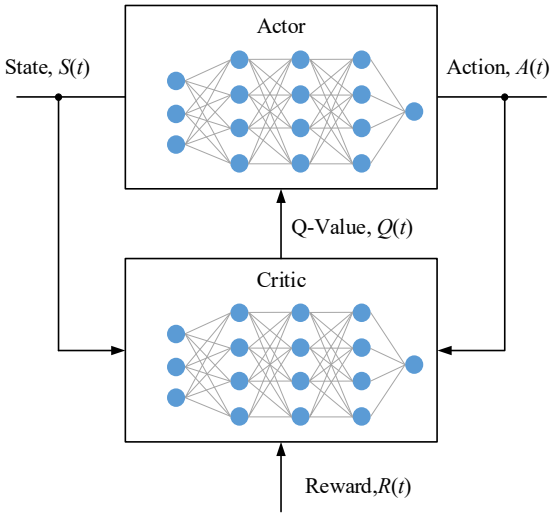


Fig. 2.5: Actor-Critic-Method structure

By "Actor-Critic" algorithms, not only a policy $\pi(a|s)$, but also a value function V can be determined. A TD algorithm is used to the "Critic" to determine a state score for the current policy. The temporal difference estimates whether the corresponding action was good or bad. A good action means that the action to a state presents a better-than-expected value. This response from the "Critic" updates the "Actor" through policy gradient [75].

Deep Q-Learning

Normally it is a big challenge for RL to find a suitable control with the input in high dimensions [76]. However, the development of deep learning has made it possible to extract features from data in high dimensions. The DeepMind team researchers applied the RL Q-learning algorithm to estimate and update the state evaluation by an ANN. This is called "Deep Q-Learning", abbreviated to "DQN" [77].

Conventionally in Q-learning, all state evaluations are made into tables. However, for complex environments with a large number of states, huge memory space is needed to insert state evaluations. Therefore, Q-learning with tables is impractical in this case. An approximation function estimates state evaluations to solve the mentioned problem.

$$Q(s,a,\theta) \approx Q^*(s,a) \tag{2.54}$$

A function can represent continuous states compared to Q-learning, which can hold the rewards for discrete states of limited quantity. Typically, a linear function can serve as a function approximator. But with the help of the ANN, it is possible to use a non-linear function approximator. This ANN, by which Q-values are estimated, is called the Q-network. In Q-learning, Q-values are directly renewed by (2.40), while the parameters (or named as weight) θ of the Q-network are trained by minimizing the value from the loss functions $L_i(\theta_i)$ to make the output from the Q-network approximate the corresponding Q-value at each state:

$$L_i(\theta_i) = \mathbb{E}_{s,a\sim\rho(\cdot)}[(y_i - Q(s,a;\theta_i))^2] \tag{2.55}$$

Here $y_i = \mathbb{E}_{s'\sim\varepsilon}[r + \gamma \max_{a'} Q(s',a';\theta_{i-1}) \mid s,a]$ is the target for iteration in the environment ε, and $\rho(s,a)$ is the probability distribution over the state s and the action a. The parameters of the last iteration θ_{t-1} are recorded in the optimization of the loss function. Deriving the loss function with weight, the following gradient is obtained:

$$\nabla_{\theta_i} L_i(\theta_i) = \mathbb{E}_{s,a\sim\rho(\cdot);s'\sim\varepsilon}\left[\left(r + \gamma \max_{a'} Q(s',a';\theta_{i-1}) - Q(s,a;\theta_i)\right)\nabla_{\theta_i} Q(s,a;\theta_i)\right] \tag{2.56}$$

Usually, the loss function is optimized by stochastic gradient descent instead of calculating the full expectation in (2.56). This algorithm is no different from the Q-learning algorithm when the weights are renewed after each time step and the expectation is replaced by a sample of the distribution ρ and the environment ε.

Remarkably, the algorithm is also model-free, which achieves the learning goal simply by sampling from the environment. Moreover, it is also off-policy, i.e., the targeted policy is a "greedy" policy $a = \max_a Q(s,a;\theta)$. The action is chosen stochastically according to a probability distribution to explore the environment.

In the following research, the algorithm of DQN was improved in different ways. One well-known improvement is the construction of the target network. This new NN has the same structure as the primary network but is not updated as frequently as it is. The new network updates the current Q value to reach the target Q value proposed by the target network. Its loss function is as follows.

$$L_i(\theta_i) = \mathbb{E}_{(s,a,r,s')\sim U(D)}\left[\left(r + \gamma \max_{a'} Q(s',a';\theta_t^-) - Q(s,a;\theta_i)\right)^2\right] \tag{2.57}$$

Here θ_i is the parameter of the primary network, θ_t^- is the parameter presented by the target network, which only renews itself at every C steps ($C \geq 1$). This method also reduces the correlation in the sequence of state observations and the influence of the change of Q-values on the policy. The DQN algorithm is more effective than the best RL algorithms in the last tweeny years [78].

Deep Deterministic Policy Gradient

Although with the algorithm of DQN mentioned above, the tasks with continuous-time observation in high dimensions can be solved, in this case, the continuous-time action is not yet possible, so the DQN is not applied to this case.

One idea is the application of "DPG", the abbreviation of "deterministic policy gradient", a kind of "actor-critic" algorithm. An optimal policy $\pi^*(a \mid s)$ is determined by a policy gradient. If it is a stochastic policy, it is a probability distribution $\pi(a \mid s)$ over the available actions at each state. When the policy is deterministic, each state corresponds to a particular action. The deterministic policy is denoted as $\mu_\theta : \mathcal{S} \to \mathcal{A}$ with the parameter $\theta \in \mathbb{R}^n$. In this case, the Q-values are learned through such as Q-learning.

Due to the success of the DQN algorithm, the DeepMind team applied deep NN to DPG, so the DDPG algorithm was developed, which is the abbreviation of Deep Deterministic Policy Gradient. The algorithm of DDPG is a combination of the algorithm of "Actor-Critic" and the algorithm of "DQN", that is, the agent has two components, one is the "Actor" that outputs the action to the environment, the other is the "Critic" that is responsible for the estimation and actualization of the Q value, but both components are KNN, like the Q network in DQN. By training and updating, the "Actor" can output the optimal action, and the "Critic" can represent the actual Q-value, which approximates the targeted Q-value.

Like the algorithm of DQN, experience recurrence and target network are applied in DDPG to reduce the dependence on different states and increase training stability. But compared to the target network update in DQN, the target network update in DDPG is "softer", which means the parameters from the target network renew very slowly: $\theta' = \tau\theta + (1-\tau)\theta'$ with $\tau \ll 1$, so that the learning process is stabilized. To avoid the divergence of the trainees from both target network μ' and Q', stable targets are required, this can lead to the slowing down of the learning process because the propagation of the value estimation is delayed, but this is helpful for the stability of the learning process.

Another feature of DDPG is batch normalization. Normally, state observation in high dimensions is related to different physical units and scopes of different physical quantities that differ from each other, which is difficult for the learning process. Stack normalization is used to solve this problem. The principle is that each sample dimension is normalized in a small stack to obtain uniform mean and variance. In addition, a running average of the mean and variance is maintained, which is used for normalization during testing. With batch normalization, it is possible to perform effective learning on various tasks without artificially adjusting the data size [79].

2.2 Interactions of the converters dominated network

The significant increase in the integration of converter-interfaced renewable energy generators (CI-REG) in power systems results in different stability issues. These issues occur due to the system inertia reduction, the low short-circuit strength at remotely sited CI-REGs during faults, reduced reactive power reserve, and the poorly tuned CI-REG control strategy [80, 81].

2.2.1 Stability issues resulted from converters

New types of stability issues arise in the converters dominated network because, compared to the conventional synchronous generators, there is different dynamic behavior in CI-REGs [82]. Fig. 2.6 shows the dynamic timescales associated with CI-REG (electromagnetic) and conventional generation (electromechanical).

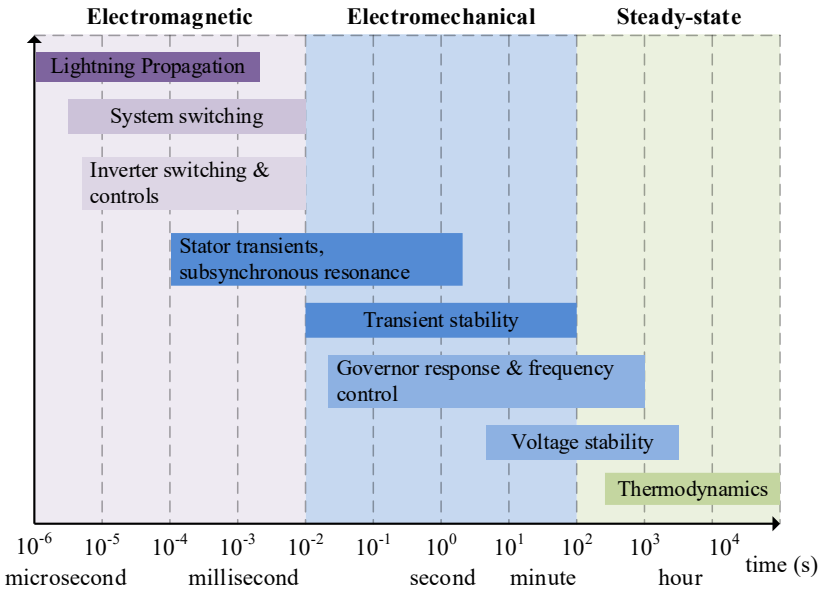


Fig. 2.6: Power system dynamics in different timescales

With the massive integration of CI-REG, which is based on power electronics converters for conversion between AC and DC, the dynamic response of the power system becomes more dependent on fast-responding power electronics. Fig. 2.6 illustrates that with the CI-REG, the timeframe of dynamics is in the 10^{-6} to 10^{-2} s timescale due to the dynamics associated with the converter switching and controls. Compared to conventional power system dynamics, whose dynamic phenomena are typically analyzed in the time range of 10^{-2} to 10^{3} s, faster dynamics will become more prominent in the dynamic analysis of the converters dominated power system. As a result, new stability issues arise, and stability analytics should be shifted to the microsecond timescale. [80, 83]

The following stability issues are affected when there is a high penetration of CI-REGs in power systems [80, 83, 84]:

- Rotor angle stability issue
- Frequency stability issue
- Voltage stability issue
- Resonance stability issue
- Converter-driven stability issue

The following will describe and analyze each of the above five types of stability issues affected by CI-REG.

Rotor angle stability issue / Synchronization issue

The power system is required for all synchronous generators in an interconnected network to remain synchronized under normal operating conditions and after a disturbance. In the event of a

synchronized failure, the generators will become unstable or lose stability, which can quickly propagate throughout the network and lead to the shutdown of the entire system [85, 86]. The ability of the synchronous generators to maintain synchronization after disturbances is called rotor angle stability [87]. According to [88], the stability of the rotor angle in the presence of large disturbances is commonly referred to as transient stability. Before a large amount of renewable energy integration through CI-REG, rotor angle stability was considered one of the critical aspects of power systems. Lack of sufficient synchronizing torque in the network would lead to transient rotor angle instability, whereas lack of damping torque would result in small-disturbance rotor angle instability [83].

The change in the rotor angle of the synchronous generator after disturbances is evaluated using the following swing equation:

$$M\frac{d^2\delta}{dt^2}+D\frac{d\delta}{dt}=P_{\mathrm{m}}-P_{\mathrm{e}} \tag{2.58}$$

where M is the inertia coefficient in $\mathrm{MVAs^2/rad}$, D is the mechanical damping coefficient in Nms, δ is the power angle in rad, P_{m} is the input mechanical power in MW, and P_{e} is the output electrical power in MW.

The generator electrical power P_{e} could be influenced by the terminal voltage magnitude and its angle. Due to the dominance of CI-REG in the power system, the trajectory of the terminal voltage is mainly driven by the source at the converter interface [89]. Since the CI-REG has more flexibility in controlling active and reactive power, this capability could improve transient stability. The voltage angle can be affected by varying the active power while changing the reactive power can help improve the voltage amplitude and, thus, the transient stability [90].

Study results in [91] show that small-disturbance rotor angle instability is improved by integrating solar-PV generation, except in the remote fault scenarios where a generator with power system stabilizer (PSS) is replaced with solar-PV generation. Transient rotor angle stability has worsened due to increased solar-PV penetration when a fault occurs at critical points. However, transient stability has improved when a fault occurs at less critical points; hence, fault proximity to solar PV generation is an important determinant of transient stability.

In [92], the analysis and simulation results show that under specific fault conditions, increasing the penetration level, such as replacing synchronous generators with DFIG-based wind farms in the case study, leads to a tendency to increase the transient stability of the specific test power system. However, it should also be noted that under certain fault conditions, the transient stability of the test power system decreases to some extent as the wind power penetration increases.

The impact of CI-REG on rotor angle stability depends on several factors, including the number of CI-REGs in the system, the location of their installation, the type of control applied, the load conditions in the system, the type of disturbance and its location concerning the CI-REGs and the large power plants, and so on [83, 93].

Frequency stability issue

Frequency stability reflects the balance of load demand and generation in the power system [88]. It is tightly associated with rotor angle stability because the rotor speed of a synchronous generator is synchronized with the system frequency. By rearranging the swing equation (2.58) with the

inertial coefficient $M = 2H / w_{sys}$ and system frequency $w = d\delta / dt$, the following equation is derived:

$$\frac{dw}{dt} = \frac{w_{sys}}{2H}\left(P_m - P_e - D \cdot w\right) \tag{2.59}$$

where H is the inertial constant in MVAs and w_{sys} is the power system rate frequency in rad/s. It shows that the frequency dynamics are inversely proportional to the inertia H.

Due to the large number of CI-REGs replacing the synchronous generators, the inertia of the system is reduced. This means that in such systems, the frequency response has faster dynamics [94]. Therefore, the potential for instability to occur earlier is increasing. This phenomenon poses a new challenge for system frequency control [80]. In addition, renewable energy like wind and solar resources are stochastic, resulting in fluctuations in output active power. This introduces uncertainty and further challenges to the system frequency stability [95].

Much research has been carried out to improve the frequency stability in the converters dominated system. [96] presents a method to re-introduce inertia by creating artificial or virtual inertia using a photovoltaic system. The results show that the virtual inertia leads the frequency dynamics in the expected direction, thus improving the system frequency stability. The "virtual synchronous generator" frequency control can also improve frequency stability. This strategy enables the converter to have the characteristics and the primary and secondary frequency adjustment mechanisms of traditional synchronous generators [97]. To smooth the fluctuations of combined power output from multiple PV systems, a central and local coordinated control has been studied and proven effective[98]. A positive contribution of CI-REGs with grid-forming control strategies to frequency stability is emphasized in [99], which implements an AC limiting scheme to stabilize the grid-forming techniques.

Voltage stability issue

Voltage stability is defined as the ability of the power system to recover and maintain the acceptable voltage ranges at all the buses in this system during the normal operating condition as well as after any disturbances being subjected to the network [88]. The voltage stability issues posed by CI-REG could mainly result from the lack of reactive power reserve due to replacing conventional synchronous generators [100].

In CI-REG-based networks, many studies on reactive power management have been carried out to overcome the problem of insufficient reactive power reserves. A consensus-based distributed voltage control is proposed in [101] for reactive power distribution. With this control, reactive power sharing is achieved in inverter-based microgrids, so the system's voltage stability is improved. A trust region framework is proposed for coordinating the reactive power output of variable-speed wind generators with other reactive sources to improve voltage stability. The results indicate a significant improvement in the load capability margin and the steady-state voltage stability margin [102].
The effect of CI-REGs on power system static voltage stability is demonstrated in [103]. Compared to other control strategies, this work shows that the static voltage stability can be improved by operating the renewable generator in voltage control mode. It is also worth noting that doubly-fed induction generator-based wind turbines improve system voltage stability compared to solar PV generators [103].

The placement of wind farms based on variable-speed wind turbines at different locations in the power system can have different effects on the steady/state voltage stability [104]. By integrating voltage-controlled CI-REGs from the weak busbars into the network, they can significantly improve the V-Q stability margin compared to integrating them from the strong busbars.

The influence on the voltage stability of different operating modes of the wind farm has been verified in [83]. In the test system, a synchronous generator is replaced by a DFIG wind farm with a voltage control mode and has a better static voltage stability margin compared to the same generator replaced by a DFIG wind farm with a power factor control mode.

Resonance stability issue

In this stability issue, resonance refers mainly to subsynchronous resonance (SSR), which is a dynamic phenomenon of interest in power systems [105]. The first SSR events occurred at the Mohave Generating Station in southern Nevada in 1970 and 1971 [106]. At that time, it was in a power system with only conventional turbine generators. However, the potential risk of SSR can also occur in current large-scale integration of induction generator-based wind farms into power systems consisting of series compensated lines [107]. The first SSR event for DFIG-based wind power was detected at the Zorillo Gulf wind farm (Texas) in October 2009 [108].

A system that has undergone SSR manifests dynamic oscillations at frequencies below the system fundamental frequency. The SSR can be split into two categories [84]:

(i) *Torsional Resonance*: Due to resonance between the series compensation and the mechanical torsion frequency of the turbine generator shaft.

(ii) *Electrical Resonance*: Due to resonance between the series compensation and the electrical characteristics of the generator.

A torsional resonance is excited when the frequency of a torque harmonic matches the natural frequency of the system [109]. This type of resonance is usually discussed in the context of power systems with turbine generators and series compensated transmission lines. However, the interaction of a turbine generator with fast-acting controllers of the power system components can also lead to a torsional resonance [110].

In the case of power systems with only conventional turbine generators, purely electrical resonances are essentially unobservable. However, the above-mentioned SSR phenomenon observed in Texas is electrical resonance. It is caused by the fact that the variable speed induction generator of the DFIG wind turbine generator is directly connected to the grid, which makes this electrical resonance between the generator and the series compensation possible. The resulting resonances mainly lead to large current and voltage oscillations that can damage electrical equipment within the generator and on the transmission system. The mitigation of such resonances can be achieved by means of SSR damping controllers with various FACTS devices and converter controllers of DFIG [111]

converter-driven stability issue

CI-REGs are usually embedded with a multiple timescale control system to regulate the current and power exchanged with the power grid [112]. The wide timescale control system of converters can lead to cross-coupling with the electromechanical dynamics of electrical machines and the electromagnetic transients (EMT) of other converters connected nearby and the power networks. Therefore, oscillations can occur across a wide frequency range in CI-REGs dominant network [113].

Fig. 2.7 maps the relationship between the harmonic frequency range and the control loops in the CI-REG, including the outer control loops for the DC voltage and the reactive power, the PLL for synchronizing the CI-REG to the grid, and the inner current control loop. The control loops designed with different bandwidths interact with the grid impedance, resulting in harmonic instability ranging from sub-synchronous frequency to switching frequency [114]. It is clearly shown in Fig. 2.7 that the fast inner current loops of CI-REG may cause high-frequency oscillations, typically in the range of hundreds of hertz to several kilohertz, while the outer power and voltage control loops and the PLL can result in unstable low-frequency oscillations, typically less than 10 Hz [115].

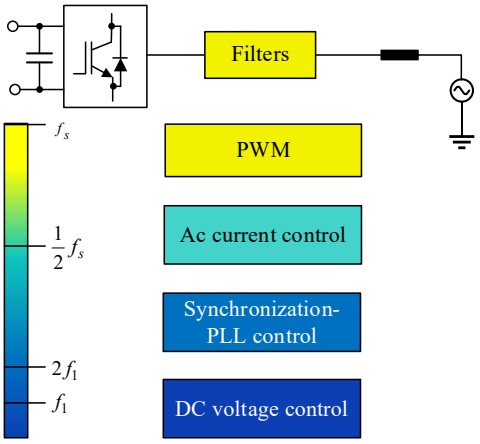


Fig. 2.7: Control loops in the CI-REG and the corresponding frequency range [114]

CI-REGs require filters at the network terminal to attenuate high-frequency content created by pulse-width modulation (PWM) [116]. As discussed above, the control loops of the CI-REGs could trigger the resonance of these filters. Compared to the single CI-REG case, the parallel CI-REGs pose a greater challenge because the converter interactions excite more complex resonances at different frequencies [117]. Interaction between the control loops of parallel CI-REGs may also cause high-frequency oscillations [118].

2.2.2 Control structure of the converters

Grid-connected converters are the main form of integrating renewable energy resources into the electrical network [1]. Grid-following control has been widely used for grid-connected converters with some providing grid supporting functionality [119, 120]. They are mostly current-controlled voltage-source converters where, given a voltage at the point of common coupling (PCC), the current is controlled accordingly to provide the desired active and reactive power [116].

The model of a power converter and its controller structure is presented in Fig. 2.8. A grid-following converter is equipped with subordinate current control, cascade voltage, and a reactive power control loop [121]. A complete control structure should include the phase-lock-loop (PLL) and abc-to-dq transformations [122]. Locking the network fundamental frequency voltage by PLL keeps the converter synchronized with the grid. The PLL and the cascade control loop perform a significant role in the dynamic behavior of the converter-based system. An LC low-pass filter is applied at the AC side to reduce the output voltage harmonics generated by the PWM [123].

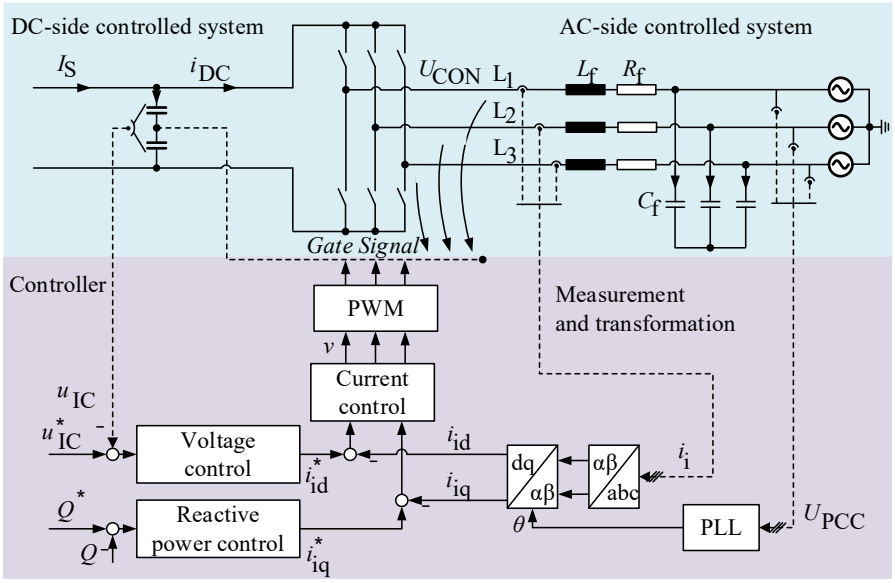


Fig. 2.8 Control structure of a grid-following CON

Voltage control loop

Voltage and reactive power control convert the voltage and reactive power setpoints into the current setpoints based on the voltage at the connection point.

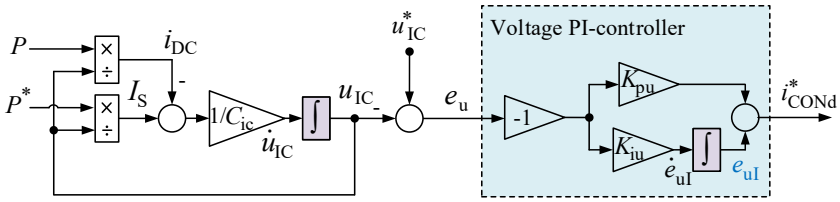


Fig. 2.9: Block diagram of the voltage control

Reactive power control loop

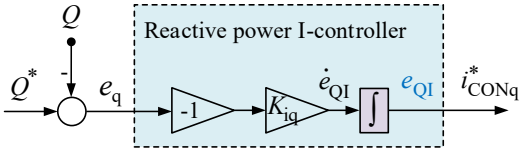


Fig. 2.10: Block diagram of the reactive power control

Current control loop

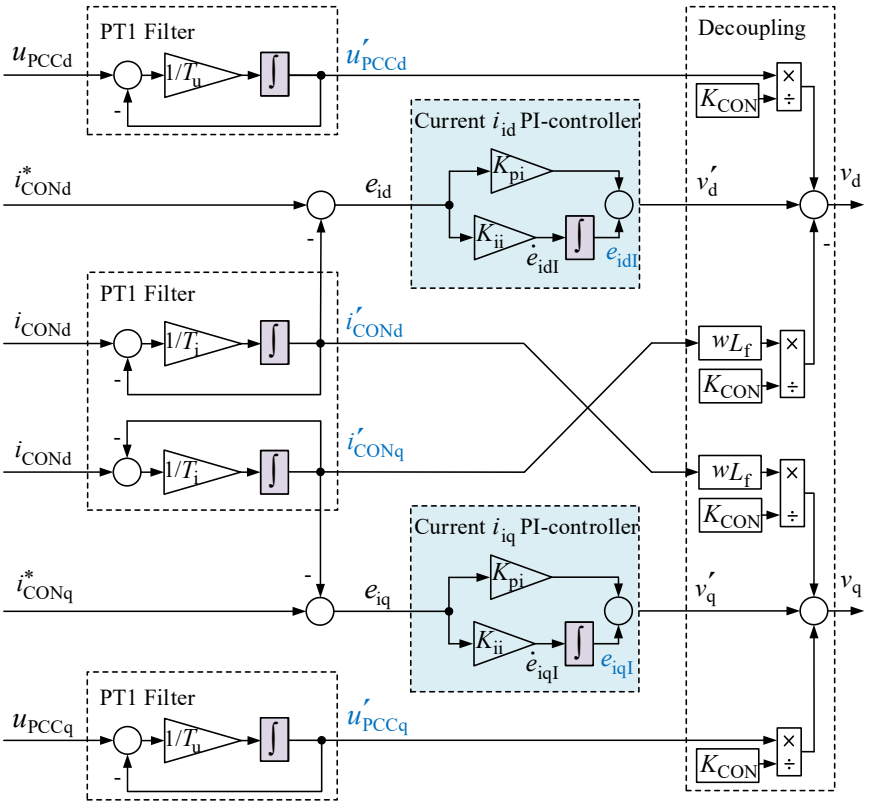


Fig. 2.11: Block diagram of the current control

PLL control:

A CI-REG is synchronized with the grid via a PLL. A commonly utilized PLL as shown in Fig. 2.12 is based on the alignment of the dq transformation angle to achieve the connected point PCC voltage without a q-axis component [124]. The PI controller calculates the filtered grid voltage angular frequency change Δw, which is then added to the nominal angular frequency w_N. By integrating the resulting angular frequency in time, the grid voltage phase angle θ is then obtained.

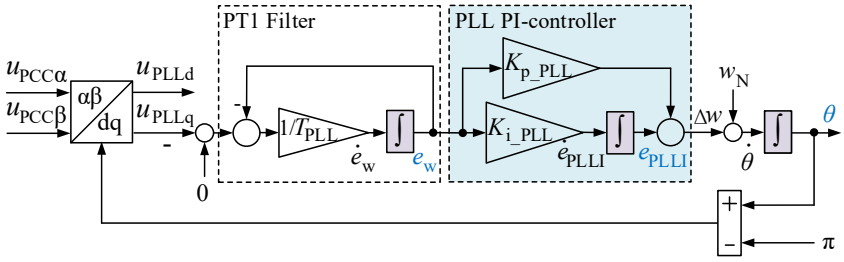


Fig. 2.12: Block diagram of PLL

2.2.3 Stability criteria

The small signal dynamic stability is an essential aspect of the power system, associated with the performance of the small perturbation around an operating point under steady-state conditions [125]. The system's small signal stability can be investigated in the time domain as well as in the Laplace domain.

A linearized system can be expressed in the small signal model (SSM) as follows.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t),\ x(t_0) = 0 \\ y(t) &= Cx(t) \end{aligned} \tag{2.60}$$

where, $A \in \mathbb{R}^{n \times n}$ is the state matrix, $B \in \mathbb{R}^{n \times p}$ represents the input matrix, and $C \in \mathbb{R}^{q \times n}$ is the output matrix. Accordingly, $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^p$ represents the system inputs and $y \in \mathbb{R}^q$ donates the system outputs.

The system inputs, states and outputs are available, with which system stability statements can be made employing stability criteria. Here, however, the unexcited "free" system is usually considered, for which $u \equiv 0$ is the essential assumption. Hereby stability statements can be made about the input-output behavior of the system. For the investigations, *n*-th order systems in the time domain according to equation (2.60) are to be considered, representing the general transfer function in the Laplace domain as shown in (2.61).

$$G(s) = \frac{Y(s)}{U(s)} = C(sI - A)^{-1} B = C \frac{\mathrm{adj}(sI - A)}{\det(sI - A)} B \tag{2.61}$$

As seen in (2.61), the poles of the transfer function $G(s)$ are equal to the eigenvalues $\lambda_i, i \in \{1, 2, ..., n\}$ of matrix *A*. The denominator polynomial of the transfer function is also called the characteristic polynomial. The Lyapunov stability in the Laplace domain can be analyzed through the eigenvalue's location.

Eigenvalues

The dynamic behavior of a system can be represented by the eigenvalue of the system. If a second-order system has a complex eigenvalue $p = \alpha \pm \mathrm{j}\beta$, then the oscillation frequency of the system under excitation can be calculated from (2.62).

$$f = \frac{|\beta|}{2\pi} \tag{2.62}$$

Further information can be obtained from the location of the eigenvalues, e.g., system oscillation behavior concerning its overshoot, the time constant, and the settling time, see Fig. 2.13.

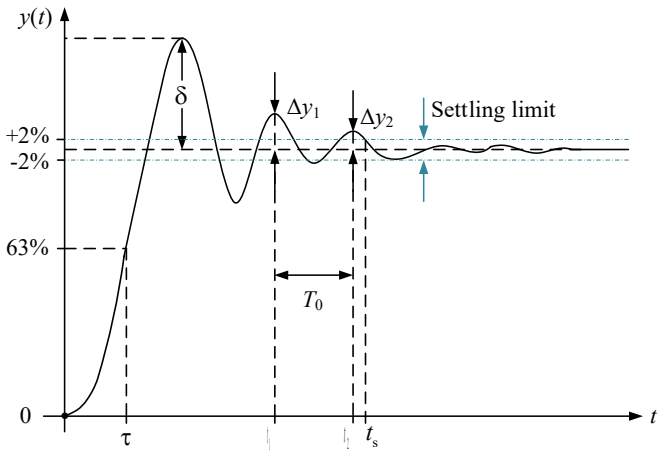


Fig. 2.13: System behavior profile with characteristic features [126]

Overshoot is also referred to as maximum deviation, which is the difference between the first peak value of the signal and the steady-state value. Most often, overshoot δ is expressed as a percentage of the change in the setpoint value [127]. The mathematical relationship between the overshoot δ and the damping ratio ζ is shown by equation (2.63).

$$\delta = e^{\frac{-\pi \cdot \zeta}{\sqrt{1-\zeta^2}}} \tag{2.63}$$

where the damping ratio ζ is an indicator that describes how the system oscillates after being subjected to a disturbance. It can be calculated from the real and imaginary parts of the complex eigenvalue.

$$\zeta = -\frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} \tag{2.64}$$

The response of the system $y(t)$ to a step input is characterized by a time constant τ. At $t = \tau$, the system has reached 63% of its final steady-state value [128]. This time constant can be calculated from the real part of the complex eigenvalue [129].

$$\tau = -\frac{\xi}{\alpha} = \frac{1}{\sqrt{\alpha^2 + \beta^2}} \tag{2.65}$$

The settling time t_s is considered as the end of the settling process if the system behavior deviates by no more than ± 2% from the final steady-state value [130].

$$t_s = \frac{\ln(0.02)}{\alpha} \tag{2.66}$$

The oscillation period is calculated by the following equation (2.67).

$$T_0 = \frac{2\pi T}{\sqrt{1-\xi^2}} = \frac{2\pi}{\sqrt{\alpha^2+\beta^2}\sqrt{1-\left(-\frac{\alpha}{\sqrt{\alpha^2+\beta^2}}\right)^2}} = \frac{2\pi}{\beta} \tag{2.67}$$

Using the above equations, it is possible to obtain the system dynamic response features from the dominant eigenvalue, which will be illustrated in section 3.3.2.

Operating point strongly correlated eigenvalues

It is known that the stability of a linearized system can be determined by the eigenvalue of the state matrix A [83], which allows the Lyapunov stability of a system included CI-REGs to be evaluated. If all eigenvalues of A have negative real parts, then the system at the corresponding operating point is asymptotically stable. If there is one eigenvalue in the RHP of the complex plane, it is then considered as an unstable mode.

A power system constructed by the CI-REGs has a higher order, which increases exponentially for parallel inverters. The state matrix of a higher-order system has a larger dimensionality, which results in a huge number of eigenvalues. This brings certain computation difficulties to the online system stability analysis and the oscillation modes. Since the operating point strongly influences the small signal stability, the strongly correlated eigenvalues of matrix A are considered a system SI. This will be further introduced and discussed in section 3.3.

2.3 Summary

This chapter discusses the theoretical backgrounds related to the ML and the research problems of stability in distribution networks containing a large number of converters. First, three types of ML algorithms applied in this work are presented, namely the TRR for the construction of a DT of the investigated network, the ANN for network state estimation, and the RL for control optimization. In addition, five types of stability issues affected by the presence of a large number of converters are described. Subsequently, the most common grid-following control structure of converters in distribution networks, which is also the investigated structure of the converters simulated in this work, is presented. Finally, small-signal dynamic stability indicators are analyzed, which will be applied to the quantitative analysis of network stability in this work. In the next chapter, the modelling of the various network components is presented as well as the concrete implementation of a method for optimizing control parameters using RL are introduced.

3 Method design

The increased number of inverters in the network can lead to network instabilities. To increase the stability margin, a concept of automatic control of converter dominated distribution network is then proposed in this section, see Fig. 3.1. The concept consists of four parts:

1) Construction of a DT for a reference distribution MV network utilizing parameter estimation
2) State estimation using ANN
3) Formulation of an eigenvalue-based SI
4) Adaptive control with RL

These four parts will be introduced in the corresponding section 3.1 to 3.4, respectively.

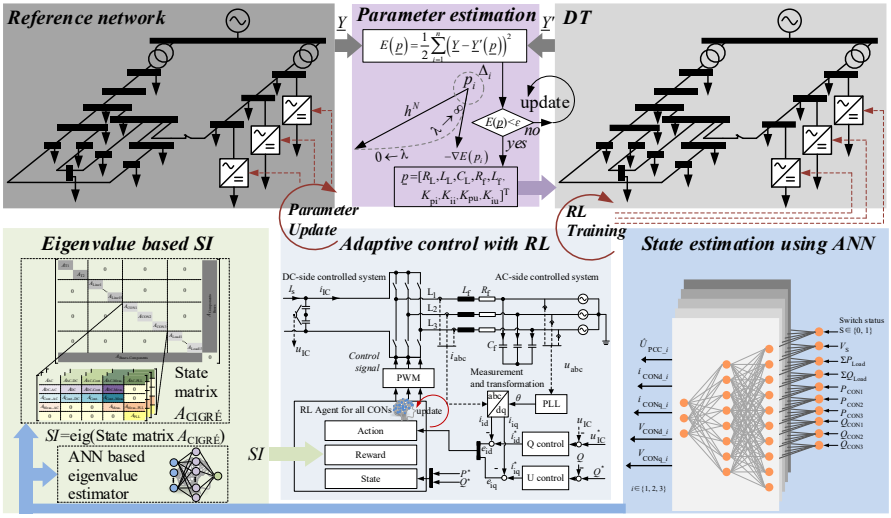


Fig. 3.1: Concept of an automatic control of converters dominated distribution network

In the first part, the injected and consumed power in the network, as well as amplitude and angle of voltages on some buses, are measured and collected from the reference network. The power data are then served as the input data sets of the constructed DT with initial parameter sets. The voltage data on the corresponding buses in DT are then collected as output data sets $\underline{Y}'$ and compared with the voltage data $\underline{Y}$ from the reference network in the parameter estimator. Through the estimator, the model parameters, such as the impedances of the lines, filter impedances of the converters as well as their controller parameters, are estimated, and then the corresponding DT of the studied network can be constructed with these parameters. The following research, which requires a large amount of data under different network conditions, is then performed in this DT.

State estimation using ANN is conducted afterward. The necessary states for SI calculation in the next step such as U_{PCCd0}, U_{PCCq0}, I_{CONd0}, I_{CONq0}, V_{d0} and V_{q0} are estimated and employed into the state matrix $\boldsymbol{A}_{\mathrm{CIGRÉ}}$. To accelerate the state estimation process, an ANN that maps the network

topology and the voltage at the measuring points to the above-mentioned states is trained with the data collected from the DT and serves as a state estimator in this work after validation.

The main goal of this work is to realize automatic control of converters and to achieve a stable distribution network. Therefore, an indicator to imply network stability is required. The indicator is significant for stability analysis and training of an RL agent in the next step. An eigenvalue-based SI is chosen to indicate the dynamic stability of the network. Thanks to the focused investigation, the real part of the operating point strongly correlated eigenvalues are determined as the SI. An ANN-based eigenvalue estimator is first trained and validated to reduce the computation effort. It is implemented in the SI calculation process afterward.

As the last part of the concept, an RL agent is utilized to parameterize the controller in converters. The SI from the last step is applied as a reward during the agent's training, whose goal is to determine a function that maps the current operating points and the controller parameters. Implementing SI as a reward can result in an optimal parameterization function to guarantee network stability under different operating points.

3.1 Digital twin of distribution networks

The electromagnetic phenomena cannot be ignored in power networks with a significant share of CI-REGs. The dynamic EMT simulation is used for this work to represent the dynamic response of power electronic converters and to conduct a comprehensive analysis of fast transients. As mentioned in section 2.1.1, a DT is understood as a real-time representation of components based on measurement data. It is also assumed that the investigated reference network has a known topology and partial, uncertain parameters, such as the impedance of transmission lines and controller parameters in CI-REGs.

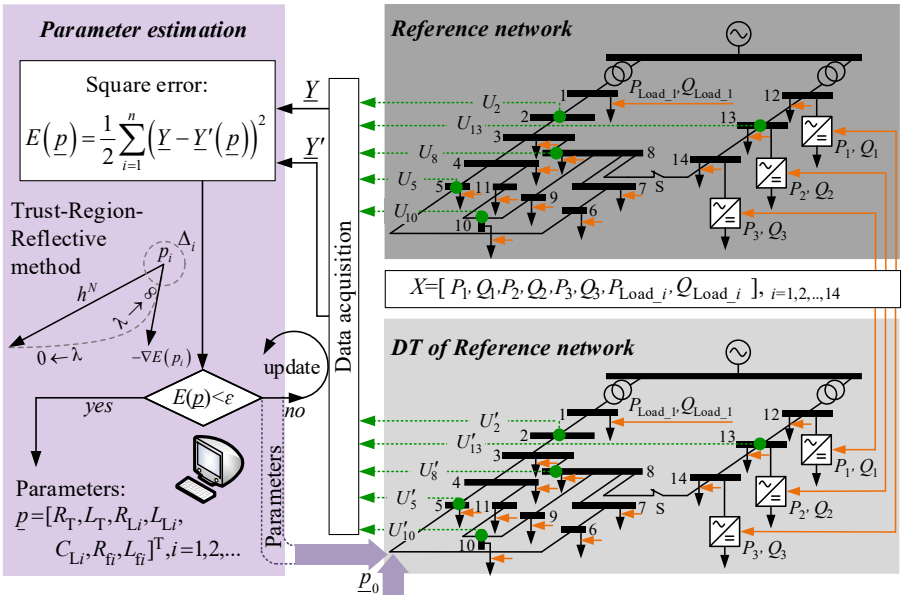


Fig. 3.2: Process of creating a DT of a reference network

Fig. 3.2 shows the DT creation process. To create the DT, an EMT model is first built according to the network topology with an initial parameter set $\underline{p}_0$. The network modeling based on EMT simulation will be introduced in section 3.1.1. Afterward, the uncertain network parameters are then estimated and optimized by the TRR method. The iteration of the TRR algorithm ends when the square error E between the measured output voltage magnitude and angle from the reference network and its DT is smaller as a predefined value ε. Consequently, the estimated model parameters $\underline{p} = [R_T, L_T, R_{Li}, L_{Li}, C_{Li}, R_{fi}, L_{fi}]^T, i = 1,2,...$ are obtained from the output of the algorithm. The estimation process will be presented in section 3.1.2. The DT model is then carried out with these optimized parameters, and further investigations such as network state estimation, automatic stability testing, and adaptive parameterization in converters will be performed on it.

3.1.1 Network modeling

A theoretical EMT model based on a set of differential equations is first constructed as a first step of creating the DT model. The CIGRÉ European medium-voltage (MV) benchmark is used as the reference model in this work [131]. Its network topology is shown in Fig. 3.3. The basic components in this network are the transformer, transmission line, converter, and load. This section deals with the modeling of these components. For the converter, only the modeling of its AC side is introduced.

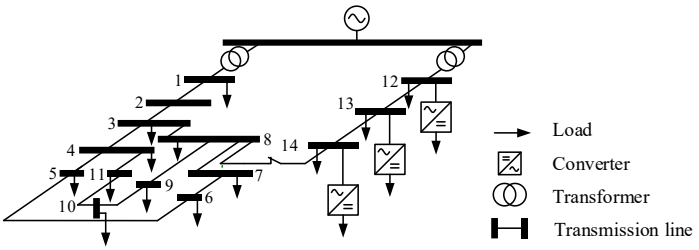


Fig. 3.3: CIGRÉ European MV benchmark [131]

The model of the aforementioned components in an αβ-coordinate system is illustrated in the following Fig. 3.4. The αβ-coordinate system is a two-axis orthogonal stationary reference frame, which is mathematically translated from the three-phase frame, is employed to simplify the analysis of three-phase circuits [132]. The differential equations, which describe the voltage across a capacitor and the current across an inductor, for each component, are derived from this model.

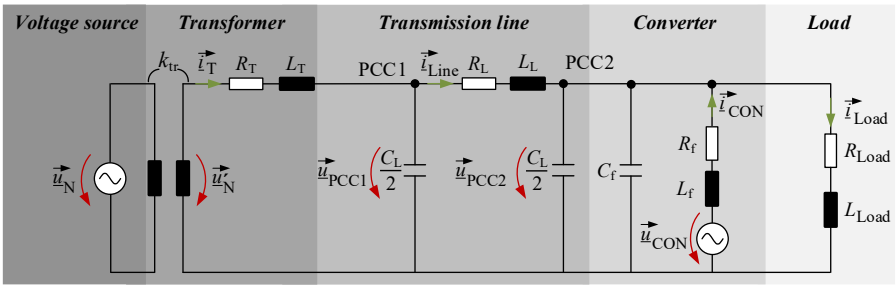


Fig. 3.4: Model of basic components

Transformer

In modeling of the CIGRE network, the shunt magnetization branches representing the iron and magnetizing losses are neglected in the transformer. Only the combined winding resistance R_T and leakage reactance X_T are considered. Therefore, an ideal transformer with a transmission ratio k_{tr} can be represented using a first-order model, which has the following differential equation.

$$\frac{d\vec{\underline{i}}_T}{dt} = \frac{1}{L_T}\left(\vec{\underline{u}}'_N - \vec{\underline{u}}_{PCC1} - R_T \cdot \vec{\underline{i}}_T\right) \tag{3.1}$$

where $\vec{\underline{u}}'_N = \frac{\vec{\underline{u}}_N}{k_{tr}}$ is the voltage on the secondary side, $\vec{\underline{i}}_T$ is the secondary current, and the calculated winding resistance R_T , and leakage inductance $L_T = X_T / w$ also refer to the secondary side.

Transmission line

The typical model for an MV transmission line uses a π model, whose characteristic is that the series impedance of the line is concentrated in the center, and the shunt capacitance of each line is divided into two equal parts, as shown in Fig. 3.4. The modeling of the line can therefore be described as follows:

$$\frac{d\vec{\underline{i}}_{Line}}{dt} = \frac{1}{L_L}\left(\vec{\underline{u}}_{PCC1} - \vec{\underline{u}}_{PCC2} - R_L \cdot \vec{\underline{i}}_{Line}\right) \tag{3.2}$$

where $\vec{\underline{i}}_{Line}$ is the transmission line current, $\vec{\underline{u}}_{PCC1}$ and $\vec{\underline{u}}_{PCC2}$ are bus voltages on both sides, R_L and L_L are the series resistance and series inductance, respectively.

Converter

An average value modeling of the converters (type 6) [133] is utilized, which is efficient for simulation and sufficient for the control and stability analysis studies. The grid-following converter is chosen for modeling and further investigation in this work. This type of converter is mostly considered as a current-controlled voltage source, whose current is controlled at the given PCC voltage to provide the desired active and reactive power.

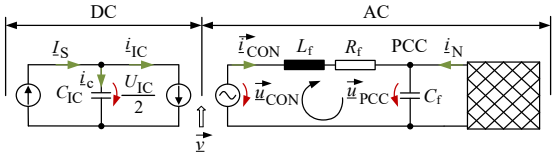


Fig. 3.5: Equivalent circuit diagram of a converter with DC (left) and AC system (right) [134]

To perform modeling and controller design in converters, the synchronous rotating dq reference frame is introduced, which is converted from the stationary αβ reference frame with the help of the park transformation. Unlike the stationary αβ reference frame, which is mostly associated with AC variables, the synchronous dq reference frame is associated with DC variables and controllers.

As shown in Fig. 3.5, the converter is described as a current-controlled voltage source with an LC filter. The following differential equation for the AC side converter system can be derived:

$$\frac{d\vec{\underline{i}}_{\mathrm{CON}}}{dt} = \frac{1}{L_{\mathrm{f}}}\left(\vec{\underline{u}}_{\mathrm{CON}} - \vec{\underline{u}}_{\mathrm{PCC}} - R_{\mathrm{f}} \cdot \vec{\underline{i}}_{\mathrm{CON}}\right) \tag{3.3}$$

where $\vec{\underline{i}}_{\mathrm{CON}}$ is the converter output current on the AC side, $\vec{\underline{u}}_{\mathrm{CON}}$ is the converter voltage, which is controlled by the intermediate circuit (IC) voltage on the DC side, R_{f} and L_{f} are resistance and inductance of the series low-pass filter, which is used at the output of the converter to filter high-frequency harmonics.

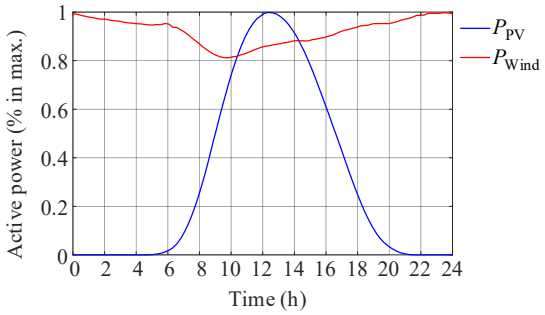


Fig. 3.6: Standard daily generation profiles for PV and offshore wind generators [131]

The representative generation profiles for both the PV array and wind turbine are shown in above Fig. 3.6. The weather conditions could rapidly change, which leads to a great change in the power generation from CI-REGs. The variation of generation powers is one of the main courses of changing the operating points of a network. The control loops could trigger the converter-driven stability issue introduced in section 2.2.1 in CI-REGs under some operating points, which is the research object in this work. To construct the DT of the studied network, the CI-REGs generation power data will be used as partial inputs of the DT model, see Fig. 3.10.

Load

The polynomial load model is commonly used to model residential, commercial, and industrial loads for analyzing and designing of power system [135, 136]. ZIP models are categorized as polynomial-type models comprising constant impedance Z, constant current I, and constant power P load (CPL) models. Most loads can be represented by a certain combination of ZIP models, with different parameters reflecting the composition [137]. A worldwide survey on load modeling has shown that 82% of utilities use CPL to represent loads in load flow and stability studies [138]. This is because the generally accepted practice of power flow analysis in electrical networks assumes that the bus voltage of the distribution network in the system keeps close to the nominal values after any initial disturbance through the tap-change transformer and voltage regulator. Therefore, loads can be treated as constant power demand. The CPL model in this work is considered as a voltage receptor and outputs current to the grid.

$$\vec{\underline{i}}_{\mathrm{Load}} = \frac{1}{R_{\mathrm{Load}} + jX_{\mathrm{Load}}}\vec{\underline{u}}_{\mathrm{PCC}} \tag{3.4}$$

where $\vec{i}_{\text{Load}}$ is the load current, $\vec{u}_{\text{PCC}}$ is the voltage at the connecting bus, and R_{Load} is the resistor and X_{Load} is the reactor corresponding to the active load power P_{Load} and reactive load power Q_{Load}, respectively.

$$Z_{\text{Load}} = \left(\frac{\frac{3}{2}\hat{U}_{\text{PCC}}{}^2}{P_{\text{Load}} + jQ_{\text{Load}}} \right)^* \tag{3.5}$$

where $\hat{U}_{\text{PCC}}$ is the nominal value of the load voltage, and the impedance $Z_{\text{Load}} = R_{\text{Load}} + jX_{\text{Load}}$.

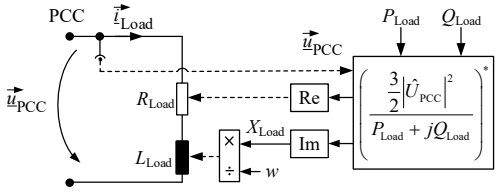


Fig. 3.7: Equivalent circuit diagram of a CPL model

As shown in Fig. 3.7, the resistance R_{Load} and the inductance $L_{\text{Load}} = X_{\text{Load}} / w$ vary with the voltage amplitude $\hat{U}_{\text{PCC}}$ to keep the active and reactive load power constant. The following equation is then derived from the diagram.

$$\frac{d\vec{i}_{\text{Load}}}{dt} = \frac{1}{L_{\text{Load}}}\left(\vec{u}_{\text{PCC}} - R_{\text{Load}} \cdot \vec{i}_{\text{Load}}\right) \tag{3.6}$$

where

$$R_{\text{Load}} = \frac{3/2 \cdot \hat{U}_{\text{PCC}}{}^2 \cdot P_{\text{Load}}}{P_{\text{Load}}{}^2 + Q_{\text{Load}}{}^2} \tag{3.7}$$

and

$$L_{\text{Load}} = \frac{X_{\text{Load}}}{w} = \frac{3/2 \cdot \hat{U}_{\text{PCC}}{}^2 \cdot Q_{\text{Load}}}{w \cdot \left(P_{\text{Load}}{}^2 + Q_{\text{Load}}{}^2\right)} \tag{3.8}$$

The above load resistance and inductance equations are derived from the impedance equation (3.5). Substitute (3.7) and (3.8) into (3.6) above, the nonlinear equation for representing the CPL model is then obtained; see the following equation.

$$\begin{aligned}\frac{d\vec{i}_{\underline{\text{Load}}}}{dt} &= \frac{1}{\frac{3/2\cdot\hat{U}_{\text{PCC}}{}^2\cdot Q_{\text{Load}}}{w\cdot\left(P_{\text{Load}}{}^2+Q_{\text{Load}}{}^2\right)}}\left(\vec{u}_{\text{PCC}}-\frac{3/2\cdot\hat{U}_{\text{PCC}}{}^2\cdot P_{\text{Load}}}{P_{\text{Load}}{}^2+Q_{\text{Load}}{}^2}\cdot\vec{i}_{\underline{\text{Load}}}\right)\\ &= \frac{w\cdot\left(P_{\text{Load}}{}^2+Q_{\text{Load}}{}^2\right)}{3/2\cdot Q_{\text{Load}}}\frac{1}{\hat{U}_{\text{PCC}}{}^2}\vec{u}_{\text{PCC}}-\frac{wP_{\text{Load}}}{Q_{\text{Load}}}\cdot\vec{i}_{\underline{\text{Load}}}\end{aligned} \tag{3.9}$$

where $\hat{U}_{\text{PCC}}{}^2 = u_{\text{PCC}\alpha}^2 + u_{\text{PCC}\beta}^2$.

The daily load profiles for residential and commercial/industrial loads are represented in Fig. 3.8, which can be considered as the average coincidence load of the CIGRÉ European MV benchmark.

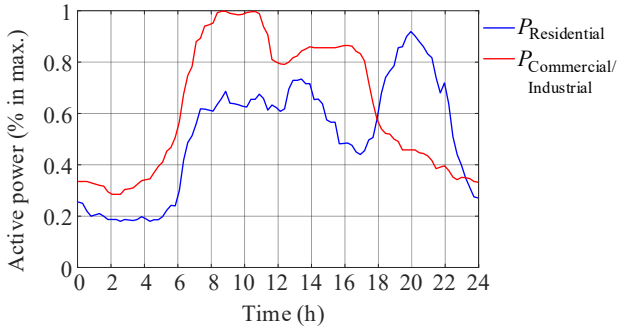


Fig. 3.8: Standard daily load profiles for residential and commercial/industrial loads [131]

The value of voltages on each bus is affected by the load power changes. For training the DT of this reference network and obtaining the estimated network parameter, the load power will also be used as partial inputs of the DT model and the voltages of a few buses can be applied as outputs in the training stage. This will be introduced in the next section 3.1.2.

Bus

Components such as transformer, transmission line, converter and load are connected to a bus. The components discussed previously are modeled as current-type sources. To connect these models via a bus model, the bus should be modeled as a voltage-type source, which means its input variables are current and output is voltage. The differential equations for voltage at both PCC buses in Fig. 3.4 are depicted in (3.10).

$$\begin{aligned}\frac{d\vec{u}_{\underline{\text{PCC1}}}}{dt} &= \frac{1}{C_{\text{L}}/2}\left(\vec{i}_{\underline{\text{T}}}-\vec{i}_{\underline{\text{Line}}}\right)\\ \frac{d\vec{u}_{\underline{\text{PCC2}}}}{dt} &= \frac{1}{C_{\text{L}}/2+C_{\text{f}}}\left(\vec{i}_{\underline{\text{Line}}}+\vec{i}_{\underline{\text{CON}}}-\vec{i}_{\underline{\text{Load}}}\right)\end{aligned} \tag{3.10}$$

where $\vec{u}_{\underline{\text{PCC1}}}$ and $\vec{u}_{\underline{\text{PCC2}}}$ are bus voltages at PCC1 and PCC2, $\vec{i}_{\underline{\text{T}}}$ is the transformer secondary current, $\vec{i}_{\underline{\text{Line}}}$ is the transmission line current, $\vec{i}_{\underline{\text{CON}}}$ is the converter output current and $\vec{i}_{\underline{\text{Load}}}$ is the load current. C_{L} and C_{f} are the line shunt capacitance and the low-pass filter capacitance for converter, respectively.

Overall representation

The block diagrams of the basic component models are represented in Fig. 3.9. It is worth defining input and output variables for each component. In this way, each component model is modularized, and it is helpful to simplify the construction of the whole network.

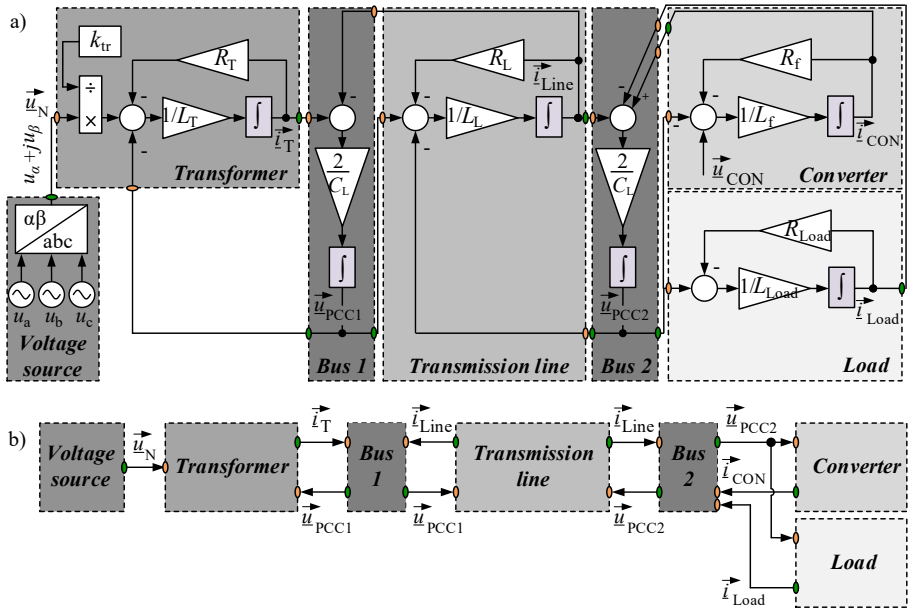


Fig. 3.9: Basic component model a) block diagram b) modularized block representation

The reference CIGRÉ benchmark and its DT model are then built with these modularized component models. For the reference model, the impedance, reactance, and capacitance parameters of transmission lines are set according to [131]. For the DT model, these parameters are minor adjusted based on the parameters from the reference network and set as the initial values for the DT, see appendix A.1.

3.1.2 Parameter estimation

As mentioned, the DT of the reference network is a digital representation of the investigated network, which can accurately describe and match the network behavior. In other words, under the same input variables set, the output from the studied network and its DT should be matched or the difference between them should be within an acceptable range. To this end, the assumed unknown parameters set $\underline{p}$ like impedance of lines and of converters' filters and the controller parameters in each converter based on the constructed network DT model from section 3.1.1 should be estimated and optimized.

Parameter estimation using the TRR method, introduced in section 2.1.1, has high accuracy and requires little effort for constrained problems. The process of parameter estimation and optimization is in fact the process of DT training. In this section, the selection of input and output variables for both training and validating of the DT model is being handled. Considering the measurement

redundancy and the possible changes on the network topology [139], voltage measuring points should be set up at buses 2, 5, 8, 10, and 13 in the CIGRÉ benchmark, see Fig. 3.2. These measured voltage data will be used for the training and validating of DT model.

As shown in the following Fig. 3.10, the consumed active and reactive power of the loads P_{Load}, Q_{Load} as well as the produced P_{CON}, Q_{CON} from the CI-REGs, which are connected with photovoltaics and wind turbine, are chosen as the input variables, while the magnitude and angle of the voltage on the selected buses are the output variables. The number of output variables needs to be constrained to train the DT effectively. In section 4.1, some scenarios are being drafted to determine the limited input and output variables while still achieving the reliable accuracy of the DT model.

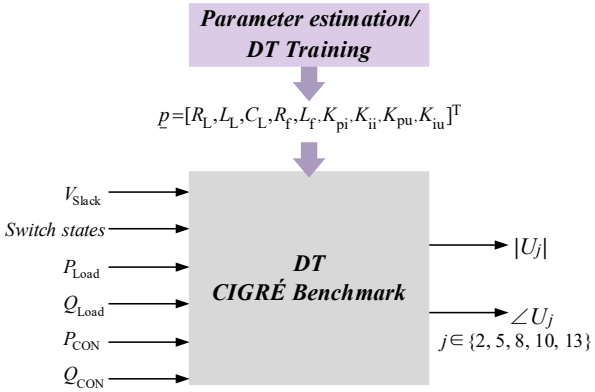


Fig. 3.10: Input and output variables of CIGRÉ benchmark DT

3.2 State estimation by means of ANN

The purpose of constructing an ANN-based state estimator is reasonably accurate and effective estimation of unmeasurable states. The process for state estimation of an electrical network can be regarded as the process of determination of functions, which capture the relationship between the measured data and the network operating states of interest. ANNs construct these functions with structures illustrated in Fig. 3.11. It is clearly to be seen that the ANNs-based state estimator consists of five ANNs for each CI-REG, which are all fully connected multilayer feedforward NNs.

Operating states in the power network vary with the changes in network topology and the power injections at load buses and generation buses. Therefore, the status of switches, which result in different network topologies, is considered one of the inputs for the estimator. The switch is located between bus 8 and bus 14; see Fig. 3.3.The voltage source that has a dominant effect on the voltage on each bus in the investigated system is also designed as one of the input variables. Measured values of active and reactive power injections at every generation bus are also selected as inputs. Unlike the generation buses, where the CI-REGs are connected, the active and reactive power consumption from all load buses is first summed and then delivered into an estimator. This input configuration is determined due to the demanded estimated states, which are PCC voltage amplitude $\hat{U}_{\mathrm{PCC0}}$, converter output currents I_{CONd0} and I_{CONq0}, as well as modulation signals V_{d0}

and V_{q0}. These states are outputs of the estimator. The mentioned PCC buses are the generation buses. Therefore, the estimated states are predominantly influenced by the generation power from each CI-REG. The load consumption may impact the other states of interest by first affecting the PCC voltages. This is the reason for using the total load power as the input of the estimator instead of using the power injections at each load bus.

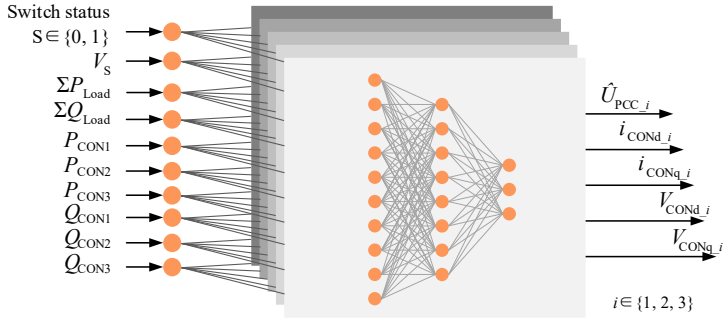


Fig. 3.11: ANNs-based state estimator

As mentioned in section 2.1.2, this work applies the SCG technique to train the ANNs-based state estimator. It is a powerful algorithm for constructing nonlinear functions between several continuous-valued inputs and one or multiple continuous-valued outputs. For a feedforward NN, the more in- and output data pairs are gathered and used for the training process, the more accurate NN model will be obtained. To avoid an under-fitting or over-fitting NN model, the *k*-folds cross-validation technique in [140] is employed in constructing the ANNs-based state estimator.

To achieve an accurate state estimator, much data needs to be gathered under different network conditions, including the status of a network switch and different loading and generation conditions.

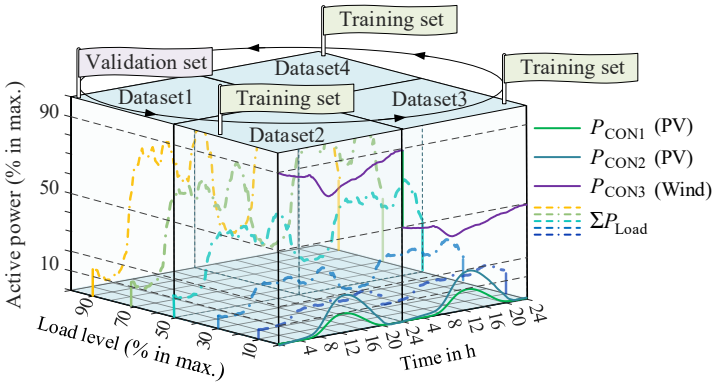


Fig. 3.12: Combinations of different load and generation levels data and representation of *k*-folds cross validation

As mentioned, the *k*-folds cross-validation technique is used in constructing the ANNs-based state estimator. Fig. 3.12 represents an example of four-fold cross-validation. The datasets are divided into four subsets known as “folds”. The one-fold dataset is used to validate the state estimator model after training it using the other three datasets. This process will repeat four times using one different fold for validation and the remaining folds as the training sets. When all the iterations have finished, the ANNs-based state estimator with a minimum Mean Squared Error (MSE) will be selected as the final state estimator and utilized in the following research.

3.3 Automatic stability detection

The stability studied in this paper focuses on the oscillatory stability of a multi-converter system. The constructed SI can be used to measure the stability margin of the steady-state equilibrium point that the system can reach in the current operation. A positive value and a higher indicator value represent that the system is more stable or robust. If the indicator has a negative value, the system is unstable in its current operating state. Based on the eigenvalues of the system's state matrix, this SI is described in detail in the following section.

3.3.1 State matrix construction

A state space representation for model differential equations is described with (3.11), in which the state matrix A can be used for the stability analysis and the function f characterizes the relationship of states $\boldsymbol{x}$ and outputs $\boldsymbol{u}$.

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+f(\boldsymbol{x},\boldsymbol{u}) \tag{3.11}$$

A complete SSM will be created for all the components, including transformers, transmission lines, loads and PLL, converter control, and the filter models of converters.

SSM of grid

In the previous section 3.1.1, the differential equations for each basic component model are represented. Before constructing the state matrix of the whole studied CIGRÉ grid, which has 15 buses, 2 transformers, 15 transmission lines, 13 loads as well as 3 converters, the state matrix of a simple grid with a structure as shown in Fig. 3.4 is first formed. Differential equations (3.1)-(3.3) as well as (3.6), whose states variables are expressed as space vectors in α/β coordinates, are represented with real and imaginary parts, and the corresponding SSM is then derived as follows. It is worth noting that for a better understanding, the following SSM regarding load is derived from the linear equation (3.6) and the part regarding converter is considered only for the AC system.

$$
\begin{bmatrix} \dot{i}_{T\alpha} \\ \dot{i}_{T\beta} \\ \dot{i}_{Line\alpha} \\ \dot{i}_{Line\beta} \\ \dot{i}_{CON\alpha} \\ \dot{i}_{CON\beta} \\ \dot{i}_{Load\alpha} \\ \dot{i}_{Load\beta} \\ \dot{u}_{PCC1\alpha} \\ \dot{u}_{PCC1\beta} \\ \dot{u}_{PCC2\alpha} \\ \dot{u}_{PCC2\beta} \end{bmatrix} =
\begin{bmatrix}
-\frac{R_T}{L_T} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{L_T} & 0 & 0 & 0 \\
0 & -\frac{R_T}{L_T} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{L_T} & 0 & 0 \\
0 & 0 & -\frac{R_L}{L_L} & 0 & 0 & 0 & 0 & 0 & \frac{1}{L_L} & 0 & -\frac{1}{L_L} & 0 \\
0 & 0 & 0 & -\frac{R_L}{L_L} & 0 & 0 & 0 & 0 & 0 & \frac{1}{L_L} & 0 & -\frac{1}{L_L} \\
0 & 0 & 0 & 0 & -\frac{R_f}{L_f} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{L_f} & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{R_f}{L_f} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{L_f} \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{R_{Load}}{L_{Load}} & 0 & 0 & 0 & \frac{1}{L_{Load}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{R_{Load}}{L_{Load}} & 0 & 0 & 0 & \frac{1}{L_{Load}} \\
\frac{2}{C_L} & 0 & -\frac{2}{C_L} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{2}{C_L} & 0 & -\frac{2}{C_L} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\frac{C_L}{2}+C_f} & 0 & \frac{1}{\frac{C_L}{2}+C_f} & 0 & -\frac{1}{\frac{C_L}{2}+C_f} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\frac{C_L}{2}+C_f} & 0 & \frac{1}{\frac{C_L}{2}+C_f} & 0 & -\frac{1}{\frac{C_L}{2}+C_f} & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} i_{T\alpha} \\ i_{T\beta} \\ i_{Line\alpha} \\ i_{Line\beta} \\ i_{CON\alpha} \\ i_{CON\beta} \\ i_{Load\alpha} \\ i_{Load\beta} \\ u_{PCC1\alpha} \\ u_{PCC1\beta} \\ u_{PCC2\alpha} \\ u_{PCC2\beta} \end{bmatrix}
$$

$$
+
\begin{bmatrix}
\frac{1}{k_{tr}L_T} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{k_{tr}L_T} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{L_f} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{L_f} & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}^{T}
\begin{bmatrix} u_{N\alpha} \\ u_{N\beta} \\ u_{CON\alpha} \\ u_{CON\beta} \end{bmatrix}
\tag{3.12}
$$

The input vector defined in the grid SSM is $\boldsymbol{u}_{GRID} = \left[u_{N\alpha}\ u_{N\beta}\ u_{CON\alpha}\ u_{CON\beta}\right]^T$ and its state-variable vector is $\boldsymbol{x}_{GRID} = \left[i_{T\alpha}\ i_{T\beta}\ i_{Line\alpha}\ i_{Line\beta}\ i_{CON\alpha}\ i_{CON\beta}\ i_{Load\alpha}\ i_{Load\beta}\ u_{PCC1\alpha}\ u_{PCC1\beta}\ u_{PCC2\alpha}\ u_{PCC2\beta}\right]^T$. From equation (3.12) the state matrix A_{GRID} can be obtained, see (3.13).

$$
\boldsymbol{A}_{GRID} =
\begin{bmatrix}
\boldsymbol{A}_T & 0 & 0 & 0 & \multirow{4}{*}{$\boldsymbol{A}_{Components\text{-}Buses}$} \\
0 & \boldsymbol{A}_{Line} & 0 & 0 & \\
0 & 0 & \boldsymbol{A}_{CON} & 0 & \\
0 & 0 & 0 & \boldsymbol{A}_{Load} & \\
\multicolumn{4}{c}{$\boldsymbol{A}_{Buses\text{-}Components}$} & 0
\end{bmatrix}
\tag{3.13}
$$

where $\boldsymbol{A}_T = \begin{bmatrix} -\frac{R_T}{L_T} & 0 \\ 0 & -\frac{R_T}{L_T} \end{bmatrix}$, $\boldsymbol{A}_{Line} = \begin{bmatrix} -\frac{R_L}{L_L} & 0 \\ 0 & -\frac{R_L}{L_L} \end{bmatrix}$, $\boldsymbol{A}_{CON} = \begin{bmatrix} -\frac{R_f}{L_f} & 0 \\ 0 & -\frac{R_f}{L_f} \end{bmatrix}$ and $\boldsymbol{A}_{Load} = \begin{bmatrix} -\frac{R_{Load}}{L_{Load}} & 0 \\ 0 & -\frac{R_{Load}}{L_{Load}} \end{bmatrix}$.

Same as the state matrixes for each component $\boldsymbol{A}_{\mathrm{T}}$, $\boldsymbol{A}_{\mathrm{Line}}$, $\boldsymbol{A}_{\mathrm{CON}}$ and $\boldsymbol{A}_{\mathrm{Load}}$, the matrixes $\boldsymbol{A}_{\mathrm{Buses\text{-}Components}}$ and $\boldsymbol{A}_{\mathrm{Components\text{-}Buses}}$ can be achieved from (3.12), which are not shown here in detail. Using these blocks' representation, the state matrix $A_{\mathrm{CIGRÉ}}$ are constructed using the following equation (3.14).

$$A_{\mathrm{CIGRÉ}} = \begin{bmatrix} A_{\mathrm{T1}} & 0 & 0 & 0 & \\ & A_{\mathrm{T2}} & & & \\ 0 & & A_{\mathrm{Line1}} & 0 & 0 \\ & 0 & & A_{\mathrm{Line15}} & \\ & & A_{\mathrm{CON1}} & & \\ 0 & 0 & & A_{\mathrm{CON2}} & 0 \\ & & & & A_{\mathrm{CON3}} \\ & & & A_{\mathrm{Load1}} & \\ 0 & 0 & 0 & & \\ & & & & A_{\mathrm{Load13}} \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} A_{\mathrm{Components\text{-}Buses}} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} A_{\mathrm{Buses\text{-}Components}} \end{matrix} \begin{matrix} 0 \end{matrix} \end{bmatrix} \tag{3.14}$$

It can be seen from equation (3.12) that the dimension of state matrixes for each component is 2×2. For each bus, it is also a matrix of dimension 2×2. Without considering the DC system and the control system for each converter, the state matrix $A_{\mathrm{CIGRÉ}}$ of the studied grid has a dimension 96×96, which comes from 14 (buses) + 2 (transformers) + 15 (transmission lines) + 13 (loads) + 3 (converters) = 47 components and 47*2=94 because each component has a matrix of dimension 2×2. The following section presents SSM for a load considering the CPL model characteristic and SSM for a converter, including its AC-, DC- and control system.

SSM of load

From section 3.1.1, the nonlinear equation (3.9) of a CPL model that describes the characteristic of the constant power is obtained. Substitute $\hat{U}_{\mathrm{PCC}}{}^2 = u_{\mathrm{PCC}\alpha}^2 + u_{\mathrm{PCC}\beta}^2$ into (3.9) and separate the equation into the real and imaginary parts, we can then gain the equations in both α and β representations, see (3.15).

$$\begin{aligned} \dot{i}_{\mathrm{Load}\alpha} &= -\frac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}} \cdot i_{\mathrm{Load}\alpha} + \frac{w \cdot \left(P_{\mathrm{Load}}{}^2 + Q_{\mathrm{Load}}{}^2\right)}{3/2 \cdot Q_{\mathrm{Load}}} \frac{1}{u_{\mathrm{PCC}\alpha}^2 + u_{\mathrm{PCC}\beta}^2} u_{\mathrm{PCC}\alpha} \\ \dot{i}_{\mathrm{Load}\beta} &= -\frac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}} \cdot i_{\mathrm{Load}\beta} + \frac{w \cdot \left(P_{\mathrm{Load}}{}^2 + Q_{\mathrm{Load}}{}^2\right)}{3/2 \cdot Q_{\mathrm{Load}}} \frac{1}{u_{\mathrm{PCC}\alpha}^2 + u_{\mathrm{PCC}\beta}^2} u_{\mathrm{PCC}\beta} \end{aligned} \tag{3.15}$$

It is clear that the equation for a CPL model is nonlinear. For the following study in this work, a linear approximation of nonlinear functions is needed. The linearized variables are represented

using $\Delta\cdot$, for example $\Delta i_{\mathrm{Load}\alpha}$. The linearization of a function is performed by deriving a first-order Taylor expansion around a point of interest [141]. Here, the point of interest is defined as the possible steady-state operating point or equilibrium point. The linear approximation of (3.15) is given by following (3.16).

$$\begin{aligned}
\Delta\dot{i}_{\mathrm{Load}\alpha} &= -\frac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}}\cdot\Delta i_{\mathrm{Load}\alpha} + \frac{w\cdot\left(P_{\mathrm{Load}}{}^2+Q_{\mathrm{Load}}{}^2\right)}{3/2\cdot Q_{\mathrm{Load}}}\left[\begin{array}{c}\dfrac{-U_{\mathrm{PCC}\alpha0}^2+U_{\mathrm{PCC}\beta0}^2}{\left(U_{\mathrm{PCC}\alpha0}^2+U_{\mathrm{PCC}\beta0}^2\right)^2}\Delta u_{\mathrm{PCC}\alpha}- \\ \dfrac{2\cdot U_{\mathrm{PCC}\alpha0}U_{\mathrm{PCC}\beta0}}{\left(U_{\mathrm{PCC}\alpha0}^2+U_{\mathrm{PCC}\beta0}^2\right)^2}\Delta u_{\mathrm{PCC}\beta}\end{array}\right] \\
\Delta\dot{i}_{\mathrm{Load}\beta} &= -\frac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}}\cdot\Delta i_{\mathrm{Load}\beta} + \frac{w\cdot\left(P_{\mathrm{Load}}{}^2+Q_{\mathrm{Load}}{}^2\right)}{3/2\cdot Q_{\mathrm{Load}}}\left[\begin{array}{c}-\dfrac{2\cdot U_{\mathrm{PCC}\alpha0}U_{\mathrm{PCC}\beta0}}{\left(U_{\mathrm{PCC}\alpha0}^2+U_{\mathrm{PCC}\beta0}^2\right)^2}\Delta u_{\mathrm{PCC}\alpha}+ \\ \dfrac{U_{\mathrm{PCC}\alpha0}^2-U_{\mathrm{PCC}\beta0}^2}{\left(U_{\mathrm{PCC}\alpha0}^2+U_{\mathrm{PCC}\beta0}^2\right)^2}\Delta u_{\mathrm{PCC}\beta}\end{array}\right]
\end{aligned} \tag{3.16}$$

where $U_{\mathrm{PCC}\alpha0}$ and $U_{\mathrm{PCC}\beta0}$ are the possible operating points of the PCC voltage. Since both $u_{\mathrm{PCC}\alpha}$ and $u_{\mathrm{PCC}\beta}$ are sine waves, it causes the coefficients of the linear functions (3.16) to change constantly, bringing continuous variation to the corresponding resulting state matrix. However, when the system reaches a steady state, the eigenvalues of the state matrix remain nearly unchanged, because they represent the characteristic of the system regarding the current state. Therefore, an assumption as shown in (3.17), which considers the orthogonal property of a variable in αß-coordinate, can be taken to simplify the equation (3.16).

$$\begin{aligned}
U_{\mathrm{PCC}\alpha0} &= \hat{U}_{\mathrm{PCC}0} \\
U_{\mathrm{PCC}\beta0} &= 0
\end{aligned} \tag{3.17}$$

Here $\hat{U}_{\mathrm{PCC}0}$ is the amplitude of PCC voltage at the equilibrium point. Substitute (3.17) into (3.16), the simplified linear functions of a CPL model are then obtained, see (3.18).

$$\begin{aligned}
\Delta\dot{i}_{\mathrm{Load}\alpha} &= -\frac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}}\cdot\Delta i_{\mathrm{Load}\alpha} - \frac{w\cdot\left(P_{\mathrm{Load}}{}^2+Q_{\mathrm{Load}}{}^2\right)}{3/2\cdot Q_{\mathrm{Load}}}\frac{1}{\hat{U}_{\mathrm{PCC}0}{}^2}\Delta u_{\mathrm{PCC}\alpha} \\
\Delta\dot{i}_{\mathrm{Load}\beta} &= -\frac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}}\cdot\Delta i_{\mathrm{Load}\beta} + \frac{w\cdot\left(P_{\mathrm{Load}}{}^2+Q_{\mathrm{Load}}{}^2\right)}{3/2\cdot Q_{\mathrm{Load}}}\frac{1}{\hat{U}_{\mathrm{PCC}0}{}^2}\Delta u_{\mathrm{PCC}\beta}
\end{aligned} \tag{3.18}$$

Considering the linearization for the load model mentioned above, the state matrix of $\boldsymbol{A}_{\mathrm{Load}}$ in (3.13) is then replaced by the following equation (3.19).

$$\boldsymbol{A}_{\mathrm{Load}} = \left[\begin{array}{cc}-\dfrac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}} & 0 \\ 0 & -\dfrac{wP_{\mathrm{Load}}}{Q_{\mathrm{Load}}}\end{array}\right] \tag{3.19}$$

In the investigated network in Fig. 3.3, there is more than one load located at different buses. Since the network impedances are relatively small, the phase differences between the bus's voltages are relatively small. Therefore, the linearized equation (3.18) that represents the relationship between the voltage, current, and powers can be used for each load at the same operating points.

SSM of converter

As shown in section 2.2.2, the type of studied converter is current-controlled voltage-source converter. Its control structure is presented in

Fig. 2.8. To construct an SSM of the converter, all differential equations for describing its AC- and DC- system model, as well as current-, voltage-, reactive power controllers and PLL model, are shown in the following.

1) AC system model

As shown in Fig. 3.5, an equivalent circuit of a converter is given in αß-coordinate. The differential equation for the AC side converter system is represented with formal (3.3). A linear approximation of nonlinear functions at an operating point is required to construct a linear SSM. A linearization at a constant operating point is preferred. Therefore, the space vector representation of αβ-coordinate (3.3) is transferred into the dq-coordinate (3.20).

$$\begin{aligned} \dot{i}_{\mathrm{CONd}} &= \frac{1}{L_{\mathrm{f}}}\left(u_{\mathrm{CONd}} - u_{\mathrm{PCCd}} + wL_{\mathrm{f}}i_{\mathrm{CONq}} - R_{\mathrm{f}}i_{\mathrm{CONd}}\right) \\ \dot{i}_{\mathrm{CONq}} &= \frac{1}{L_{\mathrm{f}}}\left(u_{\mathrm{CONq}} - u_{\mathrm{PCCq}} - wL_{\mathrm{f}}i_{\mathrm{CONd}} - R_{\mathrm{f}}i_{\mathrm{CONq}}\right) \end{aligned} \tag{3.20}$$

where u_{CONd} and u_{CONq} are the output voltage of the converter on d- and q-axis, respectively. They can be calculated by multiplying the control vector $\underline{v}$ and the IC voltage on the DC side u_{IC}.

$$\begin{aligned} u_{\mathrm{CONd}} &= v_{\mathrm{d}} \cdot \frac{u_{\mathrm{IC}}}{2} \\ u_{\mathrm{CONq}} &= v_{\mathrm{q}} \cdot \frac{u_{\mathrm{IC}}}{2} \end{aligned} \tag{3.21}$$

Consider (3.21) that is linearized at the equilibrium point U_{IC0}, V_{d0} and V_{q0}, then the linear approximation of u_{CONd} and u_{CONq} is given by (3.22).

$$\begin{aligned} \Delta u_{\mathrm{CONd}} &= \frac{V_{\mathrm{d0}}}{2} \cdot \Delta u_{\mathrm{IC}} + \frac{U_{\mathrm{IC0}}}{2}\Delta v_{\mathrm{d}} \\ \Delta u_{\mathrm{CONq}} &= \frac{V_{\mathrm{q0}}}{2} \cdot \Delta u_{\mathrm{IC}} + \frac{U_{\mathrm{IC0}}}{2}\Delta v_{\mathrm{q}} \end{aligned} \tag{3.22}$$

where U_{IC0}, V_{d0} and V_{q0} are the steady-state values of u_{IC}, v_{d} and v_{q}, respectively.

Linearization of the differential equations (3.20) with the linearized converter output voltage Δu_{CONd} and Δu_{CONq} can be obtained as follows:

$$\begin{aligned}\Delta\dot{i}_{\mathrm{CONd}} &= \frac{1}{L_{\mathrm{f}}}\left(\frac{V_{\mathrm{d0}}}{2}\cdot\Delta u_{\mathrm{IC}} + \frac{U_{\mathrm{IC0}}}{2}\Delta v_{\mathrm{d}} - \Delta u_{\mathrm{PCCd}} + wL_{\mathrm{f}}\Delta i_{\mathrm{CONq}} - R_{\mathrm{f}}\Delta i_{\mathrm{CONd}}\right)\\ \Delta\dot{i}_{\mathrm{CONq}} &= \frac{1}{L_{\mathrm{f}}}\left(\frac{V_{\mathrm{q0}}}{2}\cdot\Delta u_{\mathrm{IC}} + \frac{U_{\mathrm{IC0}}}{2}\Delta v_{\mathrm{q}} - \Delta u_{\mathrm{PCCq}} - wL_{\mathrm{f}}\Delta i_{\mathrm{CONd}} - R_{\mathrm{f}}\Delta i_{\mathrm{CONq}}\right)\end{aligned} \tag{3.23}$$

2) DC system model

The DC side system of the converter is described with (3.24).

$$\dot{u}_{\mathrm{IC}} = \frac{1}{C_{\mathrm{IC}}}\left(I_{\mathrm{S}} - i_{\mathrm{DC}}\right) = \frac{1}{C_{\mathrm{IC}}}\left(\frac{P^{*}}{u_{\mathrm{IC}}} - \frac{P}{u_{\mathrm{IC}}}\right) \tag{3.24}$$

The active power P can be calculated with (3.25).

$$P = \frac{3}{2}\left(u_{\mathrm{CONd}}i_{\mathrm{CONd}} + u_{\mathrm{CONq}}i_{\mathrm{CONq}}\right) \quad \xrightarrow[u_{\mathrm{CONq}}=v_{\mathrm{q}}\cdot\frac{u_{\mathrm{IC}}}{2}]{u_{\mathrm{CONd}}=v_{\mathrm{d}}\cdot\frac{u_{\mathrm{IC}}}{2}} = \frac{3}{2}\left(v_{\mathrm{d}}\frac{u_{\mathrm{IC}}}{2}i_{\mathrm{CONd}} + v_{\mathrm{q}}\frac{u_{\mathrm{IC}}}{2}i_{\mathrm{CONq}}\right) \tag{3.25}$$

Rewriting (3.25) into (3.26):

$$\frac{P}{u_{\mathrm{IC}}} = \frac{3}{4}\left(v_{\mathrm{d}}i_{\mathrm{CONd}} + v_{\mathrm{q}}i_{\mathrm{CONq}}\right) \tag{3.26}$$

After linearizing $\frac{P}{u_{\mathrm{IC}}}$ at the equilibrium point V_{d0}, V_{q0}, I_{CONd0} and I_{CONq0}, the following equation is obtained.

$$\Delta\left(\frac{P}{u_{\mathrm{IC}}}\right) = \frac{3}{4}\left(V_{\mathrm{d0}}\Delta i_{\mathrm{CONd}} + V_{\mathrm{q0}}\Delta i_{\mathrm{CONq}} + I_{\mathrm{CONd0}}\Delta v_{\mathrm{d}} + I_{\mathrm{CONq0}}\Delta v_{q}\right) \tag{3.27}$$

The part $\frac{P^{*}}{u_{\mathrm{IC}}}$ from (3.24) is also required to be linearized at the equilibrium point U_{IC0} and P_{0}^{*}.

$$\Delta\left(\frac{P^{*}}{u_{\mathrm{IC}}}\right) = -{U_{\mathrm{IC0}}}^{-2}P_{0}^{*}\Delta u_{\mathrm{IC}} + {U_{\mathrm{IC0}}}^{-1}\Delta P^{*} \tag{3.28}$$

With $\Delta\left(\frac{P}{u_{\mathrm{IC}}}\right)$ and $\Delta\left(\frac{P^{*}}{u_{\mathrm{IC}}}\right)$ the linearized equation for $\Delta\dot{u}_{\mathrm{IC}}$ is then driven.

$$\Delta\dot{u}_{\mathrm{DC}} = \frac{1}{C_{\mathrm{IC}}}\begin{bmatrix}-{U_{\mathrm{IC0}}}^{-2}P_{0}^{*}\Delta u_{\mathrm{IC}} + {U_{\mathrm{IC0}}}^{-1}\Delta P^{*}\\ -\frac{3}{4}\left(V_{\mathrm{d0}}\Delta i_{\mathrm{CONd}} + V_{\mathrm{q0}}\Delta i_{\mathrm{CONq}} + I_{\mathrm{CONd0}}\Delta v_{\mathrm{d}} + I_{\mathrm{CONq0}}\Delta v_{q}\right)\end{bmatrix} \tag{3.29}$$

3) converter controller models

Voltage and reactive power control convert the voltage and reactive power setpoints into the current setpoints based on the voltage at the connection point. Fig. 2.9 and Fig. 2.10 show that both controls use PI controllers.

Fig. 2.11 indicates that the current control is implemented as a pair of PI controllers along with a cross-axis decoupling term and a feedforward term for the connected point PCC voltage.

Equations to describe the dynamics of the voltage and reactive controllers are as follows.

$$\dot{e}_{\mathrm{uI}} = -K_{\mathrm{iu}} \cdot e_{\mathrm{u}} = K_{\mathrm{iu}} \left(u_{\mathrm{IC}} - u_{\mathrm{IC}}^{*} \right) \tag{3.30}$$

$$\dot{e}_{\mathrm{QI}} = -K_{\mathrm{iq}} \cdot e_{\mathrm{q}} = K_{\mathrm{iq}} \left(Q - Q^{*} \right) \tag{3.31}$$

The reactive power in (3.31) can be calculated with (3.32).

$$Q = \frac{3}{2}\left(u_{\mathrm{CONq}} i_{\mathrm{CONd}} - u_{\mathrm{CONd}} i_{\mathrm{CONq}} \right) = \frac{3}{2}\left(v_{\mathrm{q}} \frac{u_{\mathrm{IC}}}{2} i_{\mathrm{CONd}} - v_{\mathrm{d}} \frac{u_{\mathrm{IC}}}{2} i_{\mathrm{CONq}} \right) \tag{3.32}$$

Consider a linearization of the nonlinear function (3.32) at the equilibrium point V_{d0}, V_{q0}, I_{CONd0}, I_{CONq0} and U_{IC0}, then (3.33) is derived.

$$\Delta Q = \frac{3}{4}\begin{bmatrix} V_{\mathrm{q0}} U_{\mathrm{IC0}} \Delta i_{\mathrm{CONd}} + I_{\mathrm{CONd0}} \left(V_{\mathrm{q0}} \Delta u_{\mathrm{IC}} + U_{\mathrm{IC0}} \Delta v_{q} \right) \\ -V_{\mathrm{d0}} U_{\mathrm{IC0}} \Delta i_{\mathrm{CONq}} - I_{\mathrm{CONq0}} \left(V_{\mathrm{d0}} \Delta u_{\mathrm{IC}} + U_{\mathrm{IC0}} \Delta v_{\mathrm{d}} \right) \end{bmatrix} \tag{3.33}$$

Equations to describe the dynamics of the current controllers are presented as follows.

$$\dot{e}_{\mathrm{idI}} = K_{\mathrm{ii}} \cdot e_{\mathrm{id}} = K_{\mathrm{ii}} \left(i_{\mathrm{CONd}}^{*} - i'_{\mathrm{CONd}} \right) \tag{3.34}$$

$$\dot{e}_{\mathrm{iqI}} = K_{\mathrm{ii}} \cdot e_{\mathrm{iq}} = K_{\mathrm{ii}} \left(i_{\mathrm{CONq}}^{*} - i'_{\mathrm{CONq}} \right) \tag{3.35}$$

where i_{CONd}^{*} and i_{CONq}^{*} are setpoints for d- and q-axis converter output current, which comes from the voltage and reactive power controller, respectively.

$$i_{\mathrm{CONd}}^{*} = K_{\mathrm{pu}} \left(u_{\mathrm{IC}} - u_{\mathrm{IC}}^{*} \right) + e_{\mathrm{uI}} \tag{3.36}$$

$$i_{\mathrm{CONq}}^{*} = e_{\mathrm{QI}} \tag{3.37}$$

4) Measurement and decoupling for the control system

i'_{CONd} and i'_{CONq} are the filtered currents from the low-pass filter, which is added to filter the noises in the measured output current and to consider the physical time delay, as shown in Fig. 2.11. Delays are considered as first-order lags with time constant T_{i}. Equations to describe the dynamics of these filtered currents are shown in (3.38).

$$\begin{aligned} \dot{i}'_{\mathrm{CONd}} &= \frac{1}{T_{\mathrm{i}}}\left(-i'_{\mathrm{CONd}} + i_{\mathrm{CONd}} \right) \\ \dot{i}'_{\mathrm{CONq}} &= \frac{1}{T_{\mathrm{i}}}\left(-i'_{\mathrm{CONq}} + i_{\mathrm{CONq}} \right) \end{aligned} \tag{3.38}$$

The real and imaginary parts of the control vector v_{d} and v_{q} are the output-variables of the current control system, which are presented as follows.

$$\begin{aligned} v_{\mathrm{d}} &= v_{\mathrm{d}}' + \frac{1}{K_{\mathrm{CON}}} u_{\mathrm{PCCd}}' - \frac{wL_{\mathrm{f}}}{K_{\mathrm{CON}}} i_{\mathrm{CONq}}' \\ &= K_{\mathrm{pi}}\left(K_{\mathrm{pu}}\left(u_{\mathrm{IC}} - u_{\mathrm{IC}}^{*}\right) + e_{\mathrm{uI}} - i_{\mathrm{CONd}}{}'\right) + e_{\mathrm{idI}} + \frac{1}{K_{\mathrm{CON}}} u_{\mathrm{PCCd}}' - \frac{wL_{\mathrm{f}}}{K_{\mathrm{CON}}} i_{\mathrm{CONq}}' \end{aligned} \tag{3.39}$$

$$\begin{aligned} v_{\mathrm{q}} &= v_{\mathrm{q}}' + \frac{1}{K_{\mathrm{CON}}} u_{\mathrm{PCCq}}' + \frac{wL_{\mathrm{f}}}{K_{\mathrm{CON}}} i_{\mathrm{CONd}}' \\ &= K_{\mathrm{pi}}\left(e_{\mathrm{QI}} - i_{\mathrm{CONq}}'\right) + e_{\mathrm{iqI}} + \frac{1}{K_{\mathrm{CON}}} u_{\mathrm{PCCq}}' + \frac{wL_{\mathrm{f}}}{K_{\mathrm{CON}}} i_{\mathrm{CONd}}' \end{aligned} \tag{3.40}$$

where $K_{\mathrm{CON}} = \frac{U_{\mathrm{IC}}^{*}}{2}$, u_{PCCd}' and u_{PCCq}' are the filtered PCC voltages from the first-order low-pass filter, whose time constant is fixed at T_{u}. Differential equations to describe their dynamics are as follows.

$$\begin{aligned} \dot{u}_{\mathrm{PCCd}}' &= \frac{1}{T_{\mathrm{u}}}\left(-u_{\mathrm{PCCd}}' + u_{\mathrm{PCCd}}\right) \\ \dot{u}_{\mathrm{PCCq}}' &= \frac{1}{T_{\mathrm{u}}}\left(-u_{\mathrm{PCCq}}' + u_{\mathrm{PCCq}}\right) \end{aligned} \tag{3.41}$$

5) PLL model

As shown in Fig. 2.17, the utilized PLLs based on the alignment of the dq-transformation angle to achieve the connected point PCC voltage without a q-axis component [142]. The PI controller calculates the filtered grid voltage angular frequency change Δw, which is then added to the reference angular frequency w_{N}. By integrating the resulting angular frequency in time, the grid voltage phase angle θ is then obtained.

In PLL system there is three state-variable, which are dynamic are characterized by the following equations.

$$\dot{e}_{\mathrm{w}} = \frac{1}{T_{\mathrm{PLL}}}\left(-u_{\mathrm{PLLq}} - e_{\mathrm{w}}\right) \tag{3.42}$$

$$\dot{e}_{\mathrm{PLLI}} = K_{\mathrm{i\_PLL}} e_{\mathrm{w}} \tag{3.43}$$

$$\dot{\theta} = K_{\mathrm{p\_PLL}} e_{\mathrm{w}} + e_{\mathrm{PLLI}} + w_{\mathrm{N}} \tag{3.44}$$

As shown in Fig. 2.12, the equation for q-axis component of PCC voltage u_{PCCq} is represented using the nonlinear equation (3.45).

$$u_{\mathrm{PLLq}} = -\sin\left(\theta - \pi\right) \cdot u_{\mathrm{PCC\alpha}} + \cos\left(\theta - \pi\right) \cdot u_{\mathrm{PCC\beta}} = \sin\theta \cdot u_{\mathrm{PCC\alpha}} - \cos\theta \cdot u_{\mathrm{PCC\beta}} \tag{3.45}$$

A linearization of (3.45) is conducted at the equilibrium point U_{PCCd0} and the current bus voltage phase θ_0, the linear equation (3.46) is then derived.

$$\Delta u_{\mathrm{PLLq}} = \sin\theta_0 \cdot \Delta u_{\mathrm{PCC}\alpha} - \cos\theta_0 \cdot \Delta u_{\mathrm{PCC}\beta} + \underbrace{\left(U_{\mathrm{PCC}\alpha 0}\cos\theta_0 + U_{\mathrm{PCC}\beta 0}\sin\theta_0\right)}_{U_{\mathrm{PCCd0}}}\Delta\theta \tag{3.46}$$

Substitute (3.46) into (3.42), the linearized function for state-variable e_{w} is then obtained.

$$\Delta\dot{e}_{\mathrm{w}} = \frac{1}{T_{\mathrm{PLL}}}\left(-\sin\theta_0 \cdot \Delta u_{\mathrm{PCC}\alpha} + \cos\theta_0 \cdot \Delta u_{\mathrm{PCC}\beta} - U_{\mathrm{PCCd0}}\Delta\theta - \Delta e_{\mathrm{w}}\right) \tag{3.47}$$

6) Transformation

- **αß-to-dq transformation**

The αß-to-dq transformation for the PCC-voltage is represented with (3.48), in which the nonlinear sine-cosine functions are required to have a linear approximation.

$$\begin{aligned} u_{\mathrm{PCCd}} &= \cos\theta \cdot u_{\mathrm{PCC}\alpha} + \sin\theta \cdot u_{\mathrm{PCC}\beta} \\ u_{\mathrm{PCCq}} &= -\sin\theta \cdot u_{\mathrm{PCC}\alpha} + \cos\theta \cdot u_{\mathrm{PCC}\beta} \end{aligned} \tag{3.48}$$

The linearization is performed at the equilibrium operating point θ_0, U_{PCCd0} and U_{PCCq0}. PLL is as known to synchronize with the utility voltage vector phase angle θ_{e}. Therefore, at the equilibrium point $\theta_0 = \theta_{\mathrm{e}}$.

$$\begin{aligned} \Delta u_{\mathrm{PCCd}} &= \cos\theta_0 \cdot \Delta u_{\mathrm{PCC}\alpha} + \sin\theta_0 \cdot \Delta u_{\mathrm{PCC}\beta} + \underbrace{\left(-U_{\mathrm{PCC}\alpha 0}\sin\theta_0 + U_{\mathrm{PCC}\beta 0}\cos\theta_0\right)}_{U_{\mathrm{PCCq0}}}\Delta\theta \\ \Delta u_{\mathrm{PCCq}} &= -\sin\theta_0 \cdot \Delta u_{\mathrm{PCC}\alpha} + \cos\theta_0 \cdot \Delta u_{\mathrm{PCC}\beta} - \underbrace{\left(U_{\mathrm{PCC}\alpha 0}\cos\theta_0 + U_{\mathrm{PCC}\beta 0}\sin\theta_0\right)}_{U_{\mathrm{PCCd0}}}\Delta\theta \end{aligned} \tag{3.49}$$

- **dq-to-αß transformation**

The studied grid-following converter, as described, is a current-type source from the perspective of the network, which means the converter affects the connected network through its output current $\vec{i}_{\mathrm{CON}}$. As introduced, the network parts are modeled in aß-coordinate. Therefore, a dq-to-αß transformation for the converter current is required for i_{CONd} and i_{CONq}.

$$\begin{aligned} i_{\mathrm{CON}\alpha} &= \cos\theta \cdot i_{\mathrm{CONd}} - \sin\theta \cdot i_{\mathrm{CONq}} \\ i_{\mathrm{CON}\beta} &= \sin\theta \cdot i_{\mathrm{CONd}} + \cos\theta \cdot i_{\mathrm{CONq}} \end{aligned} \tag{3.50}$$

The transformation is represented with equations (3.50). These nonlinear functions are linearized at the equilibrium operating point θ_0, I_{CONd0} and I_{CONq0}; see following equation (3.51).

$$\begin{aligned} \Delta i_{\mathrm{CON}\alpha} &= \cos\theta_0 \cdot \Delta i_{\mathrm{CONd}} - \sin\theta_0 \cdot \Delta i_{\mathrm{CONq}} - \underbrace{\left(I_{\mathrm{CONd0}}\sin\theta_0 + I_{\mathrm{CONq0}}\cos\theta_0\right)}_{I_{\mathrm{CON}\beta 0}}\Delta\theta \\ \Delta i_{\mathrm{CON}\beta} &= \sin\theta_0 \cdot \Delta i_{\mathrm{CONd}} + \cos\theta_0 \cdot \Delta i_{\mathrm{CONq}} + \underbrace{\left(I_{\mathrm{CONd0}}\cos\theta_0 - I_{\mathrm{CONq0}}\sin\theta_0\right)}_{I_{\mathrm{CON}\alpha 0}}\Delta\theta \end{aligned} \tag{3.51}$$

7) Equilibrium point calculation

A linear SSM of this model must be constructed to obtain the state matrix of the converter model, including its control system. An SSM, as represented using equation (3.11), requires a linearization on the studied converter model at the equilibrium point $V_{d0}, V_{q0}, I_{CONd0}, I_{CONq0}, U_{IC0}, \theta_0, U_{PCCd0}$, and U_{PCCq0}. Due to the IC voltage controller, the voltage u_{IC} at the equilibrium point should equal to its setpoint value.

$$U_{IC0} = U_{IC}^{*} \tag{3.52}$$

Same as the IC voltage, the PCC voltages on dq axis are being controlled on the following values because of the operation of the PLL system.

$$\begin{aligned} U_{PCCd0} &= \hat{U}_{PCC} \\ U_{PCCq0} &= 0 \end{aligned} \tag{3.53}$$

where $\hat{U}_{PCC}$ is the magnitude of the PCC voltage. Its phase angle θ_e is then obtained and applied in the αß /dq transformation. In this case, the PLL output angle θ is equal to θ_e at the equilibrium point.

$$\theta_0 = \theta_e = \arctan\left(\frac{U_{PCCq0}}{U_{PCCd0}}\right) \tag{3.54}$$

Combing the following power calculation equation (3.55)

$$\begin{aligned} P^{*}|_{u_{CONd}=U_{CONd0}, u_{CONq}=U_{CONq0}, i_{CONd}=I_{CONd0}, i_{CONq}=I_{CONq0}} &= \frac{3}{2}\left(u_{CONd}\cdot i_{CONd} + u_{CONq}\cdot i_{CONq}\right) \\ Q^{*}|_{u_{CONd}=U_{CONd0}, u_{CONq}=U_{CONq0}, i_{CONd}=I_{CONd0}, i_{CONq}=I_{CONq0}} &= \frac{3}{2}\left(u_{CONq}\cdot i_{CONd} - u_{CONd}\cdot i_{CONq}\right) \end{aligned} \tag{3.55}$$

and the differential equations of output converter currents on the dq-axis (3.20) as well as the calculation equation for the output converter voltage (3.21), also considering that at the equilibrium point, both $\dot{i}_{CONd}$ and $\dot{i}_{CONq}$ are equal to zero, the remaining equilibrium point of V_{d0}, V_{q0}, I_{CONd0} and I_{CONq0} can then be deduced.

The SSM of the converter model at the resulting equilibrium point is then constructed, and it is represented using the small signal converter model, as shown in Fig. 3.13. To show the relationship between the converter and its connected network, a network model is constructed with the following equations, which are based on the structure of the AC system model in Fig. 3.5.

$$\begin{aligned} \dot{u}_{PCC\alpha} &= \frac{1}{C_f}\left(i_{N\alpha} + i_{CON\alpha}\right) \\ \dot{u}_{PCC\beta} &= \frac{1}{C_f}\left(i_{N\beta} + i_{CON\beta}\right) \end{aligned} \tag{3.56}$$

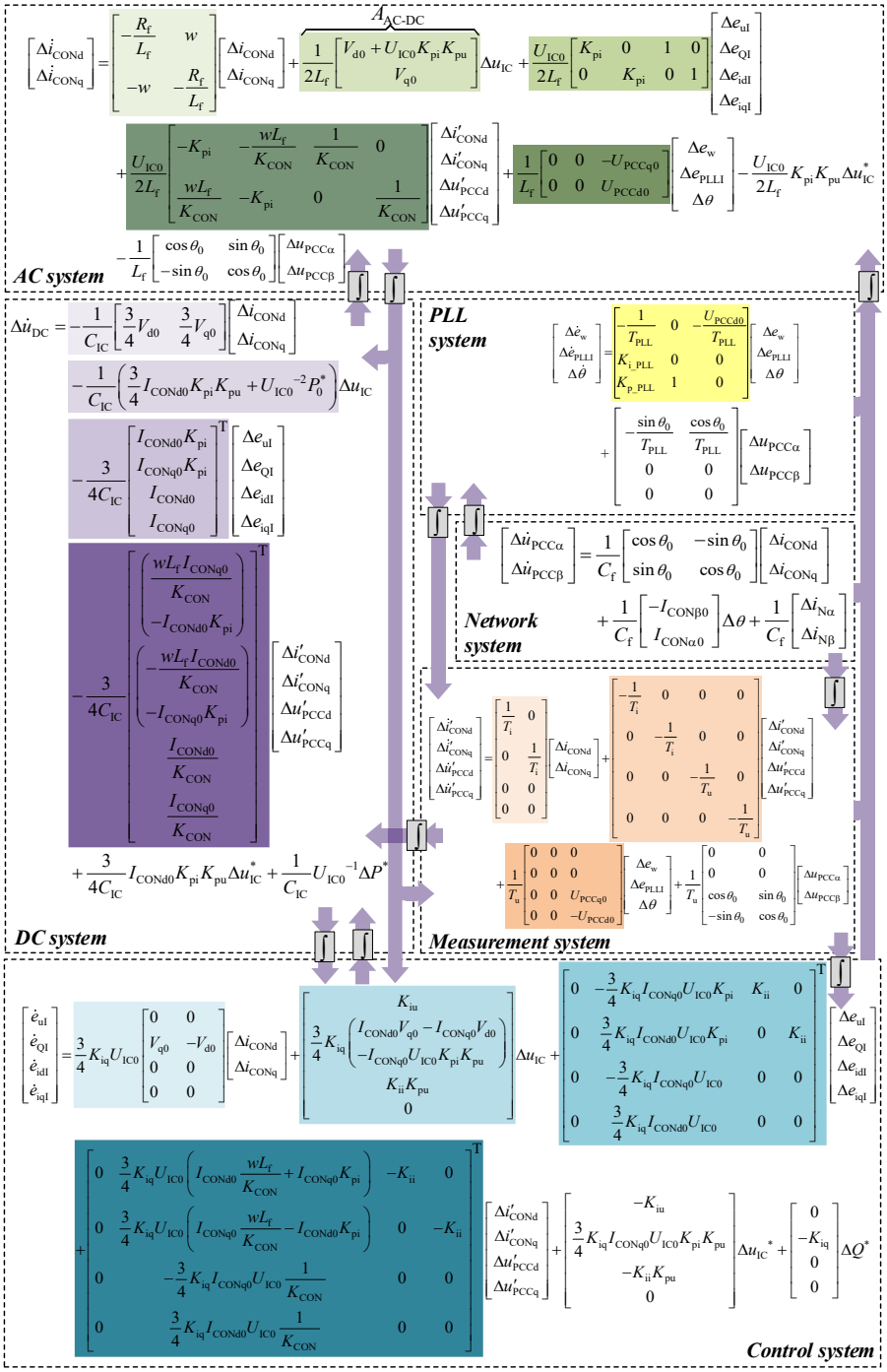


Fig. 3.13: Converter small signal model

To build the connection between the converter model and the network model with small signal, the converter currents on αß axis are represented with Δi_{CONd} and Δi_{CONq} using the equation (3.51). The small signal network model is then derived, as seen in the following equations (3.57).

$$\begin{aligned}\Delta \dot{u}_{\mathrm{PCC}\alpha} &= \frac{1}{C_{\mathrm{f}}}\left(\Delta i_{\mathrm{N}\alpha} + \cos\theta_0 \cdot \Delta i_{\mathrm{CONd}} - \sin\theta_0 \cdot \Delta i_{\mathrm{CONq}} - I_{\mathrm{CON}\beta 0}\Delta\theta\right) \\ \Delta \dot{u}_{\mathrm{PCC}\beta} &= \frac{1}{C_{\mathrm{f}}}\left(\Delta i_{\mathrm{N}\beta} + \sin\theta_0 \cdot \Delta i_{\mathrm{CONd}} + \cos\theta_0 \cdot \Delta i_{\mathrm{CONq}} + I_{\mathrm{CON}\alpha 0}\Delta\theta\right)\end{aligned} \tag{3.57}$$

From Fig. 3.13 it is clear that there are 14 state variables in a grid-following converter model using the introduced control structure, see (3.58).

$$\Delta x = \begin{bmatrix} \Delta i_{\mathrm{CONd}},\ \Delta i_{\mathrm{CONq}},\ \Delta u_{\mathrm{IC}},\ \Delta e_{\mathrm{uI}},\ \Delta e_{\mathrm{QI}},\ \Delta e_{\mathrm{idI}},\ \Delta e_{\mathrm{iqI}},\ \dots \\ \dots\ \Delta i'_{\mathrm{CONd}},\ \Delta i'_{\mathrm{CONq}},\ \Delta u'_{\mathrm{PCCd}},\ \Delta u'_{\mathrm{PCCq}},\ \Delta e_{\mathrm{w}},\ \Delta e_{\mathrm{PLLI}},\ \Delta\theta \end{bmatrix}^{\mathrm{T}} \tag{3.58}$$

Input variables for this converter model are

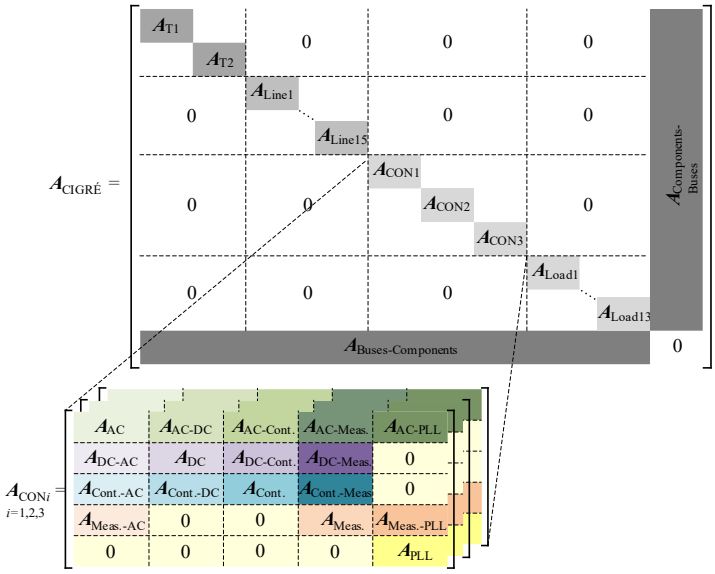
$$\Delta u = \left[\Delta u_{\mathrm{PCC}\alpha}, \Delta u_{\mathrm{PCC}\beta}, \Delta u^{*}_{\mathrm{IC}}, \Delta P^{*}, \Delta Q^{*}\right]^{\mathrm{T}} \tag{3.59}$$

where $\Delta u_{\mathrm{PCC}\alpha}$ and $\Delta u_{\mathrm{PCC}\beta}$ are the connected PCC bus voltages in aß-coordinate, $\Delta u^{*}_{\mathrm{IC}}, \Delta P^{*}, \Delta Q^{*}$ are reference value of IC voltage, active and reactive power, respectively.

To represent the complete state matrix $\boldsymbol{A}_{\mathrm{CON}}$ of a converter, $\boldsymbol{A}_{\mathrm{AC}}$, $\boldsymbol{A}_{\mathrm{DC}}$, $\boldsymbol{A}_{\mathrm{Cont.}}$, $\boldsymbol{A}_{\mathrm{Meas.}}$, and $\boldsymbol{A}_{\mathrm{PLL}}$ are used to describe the state matrixes of its AC-, DC-side system, its control system, measurement system as well as PLL system, respectively. The coupled state matrixes employ the form of $\boldsymbol{A}_{\mathrm{x\text{-}x}}$. As an example, $\boldsymbol{A}_{\mathrm{AC\text{-}DC}}$ represents the matrix, which couple the state variables from AC and DC system, as see in Fig. 3.13. The state matrix $\boldsymbol{A}_{\mathrm{CON}}$ is then represented with following equation (3.60).

$$\boldsymbol{A}_{\mathrm{CON}} = \begin{bmatrix} \boldsymbol{A}_{\mathrm{AC}} & \boldsymbol{A}_{\mathrm{AC\text{-}DC}} & \boldsymbol{A}_{\mathrm{AC\text{-}Cont.}} & \boldsymbol{A}_{\mathrm{AC\text{-}Meas.}} & \boldsymbol{A}_{\mathrm{AC\text{-}PLL}} \\ \boldsymbol{A}_{\mathrm{DC\text{-}AC}} & \boldsymbol{A}_{\mathrm{DC}} & \boldsymbol{A}_{\mathrm{DC\text{-}Cont.}} & \boldsymbol{A}_{\mathrm{DC\text{-}Meas.}} & 0 \\ \boldsymbol{A}_{\mathrm{Cont.\text{-}AC}} & \boldsymbol{A}_{\mathrm{Cont.\text{-}DC}} & \boldsymbol{A}_{\mathrm{Cont.}} & \boldsymbol{A}_{\mathrm{Cont.\text{-}Meas.}} & 0 \\ \boldsymbol{A}_{\mathrm{Meas.\text{-}AC}} & 0 & 0 & \boldsymbol{A}_{\mathrm{Meas.}} & \boldsymbol{A}_{\mathrm{Meas.\text{-}PLL}} \\ 0 & 0 & 0 & 0 & \boldsymbol{A}_{\mathrm{PLL}} \end{bmatrix} \tag{3.60}$$

There are 14 state variables in the studied CON; therefore, the state matrix $\boldsymbol{A}_{\mathrm{CON}}$ has a dimension of 14×14. This means, by considering a complete system for each converter, the state matrix $\boldsymbol{A}_{\mathrm{CIGRÉ}}$ of the studied grid, which has three converters, has a dimension 130×130. The complete matrix $\boldsymbol{A}_{\mathrm{CIGRÉ}}$ is shown in (3.61).



(3.61)

Due to the linearization of the nonlinear equations in the converter model, the state matrix $A_{\text{CIGRÉ}}$ is a matrix that varies with the system operating point. A dynamic SI for the studied network considering varying operating points is then introduced in the next section.

3.3.2 Eigenvalue based stability indicator

The small-signal-stability analysis can be conducted based on a linear SSM of a nonlinear converter-based grid model at a certain operating point. The stability of the mentioned linearized system is determined by the eigenvalue of the state matrix $A_{\text{CIGRÉ}}$, which is derived from the last section. If all eigenvalues of $A_{\text{CIGRÉ}}$ have negative real parts, then the system at the corresponding operating point is asymptotically stable. The eigenvalues also reflect the damping mode of the system [143]. Conducting the computation of eigenvalues of $A_{\text{CIGRÉ}}$ at the estimated operating points from section 3.2, the system robustness and stability margin and its dominant damping mode are then derived efficiently.

The state matrix of a higher-order system has a larger dimensionality, which results in a huge number of eigenvalues. For the state matrix $A_{\text{CIGRÉ}}$ with a dimension of 130×130, the number of corresponding eigenvalues is then 130. The operating point is a great influence factor in the distribution of the eigenvalues. However, the operating point change greatly affects only several groups of eigenvalues. Fig. 3.14 shows the complete eigenvalue locations under different active power generations from 1 MW to 5 MW of converter CON1, while the active power of the other two converters, CON2 and CON3, are kept at 3 MW, and the power factor for all converters is fixed at 0.95.

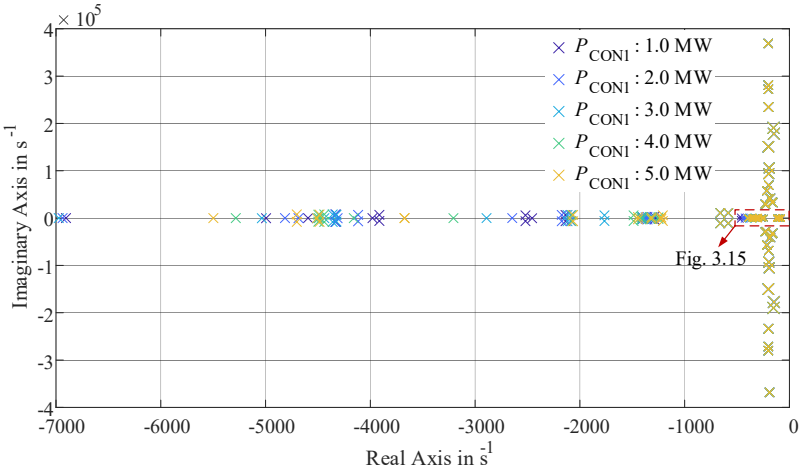


Fig. 3.14: Trajectory of complete eigenvalues for changes of converter CON1 active power

It can be clearly seen that their trajectories are considerably more variable for eigenvalues with real part values between -7000 and -1000. However, since they are far from the imaginary axis, their impact on the system stability is relatively minor. For the eigenvalues close to the imaginary axis, it can be observed that a great majority of them have only relatively small movements in location when the operating points change. By zooming in on eigenvalues, which are in the area circled by the red dotted line in Fig. 3.14,close to the imaginary axis and strongly correlated with the operating points change, the following figure Fig. 3.15 is obtained.

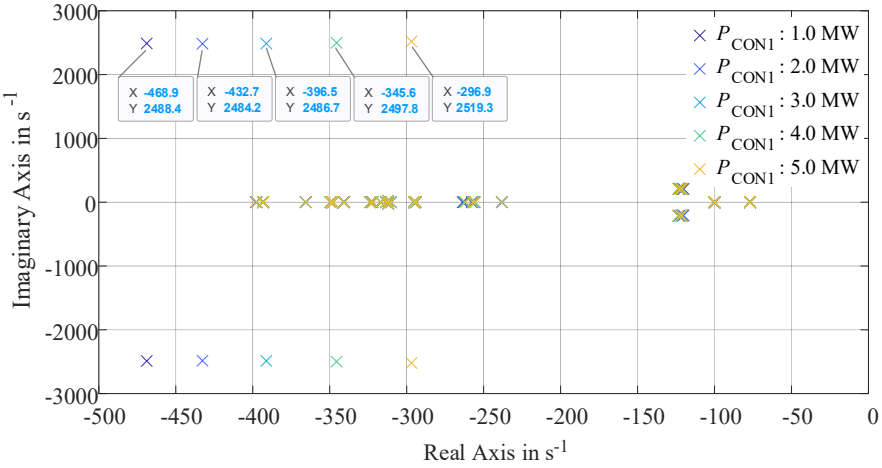


Fig. 3.15: Trajectory of partial eigenvalues for changes of converter CON1 active power

The eigenvalues marked with values in Fig. 3.15 can reflect the dominant damping mode of the system. As the active power increases from 1 MW to 5 MW, the dominant eigenvalue shifts to

the right. Using equations (2.63)-(2.67), the characteristic parameters of the system response under the dominant damping mode, such as overshoot, damping ratio, and time constant of the system, can be obtained, see the following Tab. 3.1.

Tab. 3.1 System dynamic feature from dominant eigenvalue

Active power *P* [MW]		1	2	3	4	5
Dominant eigenvalue	**Real part**	-468.87	-432.68	-396.52	-345.58	-296.94
	Imaginary part	2488.43	2484.18	2486.65	2497.79	2519.32
Overshoot		55.32%	57.86%	.61.00%	64.75%	.69.06%
Damping ratio		0.19	0.17	0.16	0.14	0.12
Time constant [μs]		394.87	396.61	397.34	396.59	394.21
Settling time t_s [ms]		8.34	9.04	10.00	11.32	13.18
Oscillation period T_O [μs]		2525.03	2529.34	2526.65	2515.51	2494.01
Damping frequency [Hz]		396.04	395.37	395.77	397.54	400.96

It is clearly shown in Tab. 3.1 that, with the increase of P_{CON1} the dominant eigenvalue, indicates a rising overshoot, a longer settling time and a decreasing oscillation period at the system response to a small disturbance signal.

Variations in the operating points can also be caused by the power factor change, changing the load in the investigated network, or by variation in the source voltage. The trajectories of the dominant eigenvalues under the mentioned variations are shown in Fig. 3.16. Compared to the other causes, the load variation has a smaller effect on the system's dominant damping mode.

The small-signal stability is studied around an operating point; therefore, an operating point strong relevant dominant eigenvalue of $\boldsymbol{A}_{\mathrm{CIGRÉ}}$ are considered as the system SI. As shown in Tab. 3.1, the overshoot, the settling time, the damping frequency as well as the damping ratio of the system response can be derived from the dominant eigenvalue. Among these response characteristics, the damping ratio is chosen as the oscillatory SI, which serves as a reward for the training of the RL agent in section 3.4.

This section presents that the dominant eigenvalue is determined from the eigenvalues movement by changing the operating points. To efficiently determine the dominant damping frequency and identify the dominant eigenvalue, the physical variables that can correspond to the dominant damping mode should be detected first. The next section introduces the detection of the mentioned variables through FFT analysis.

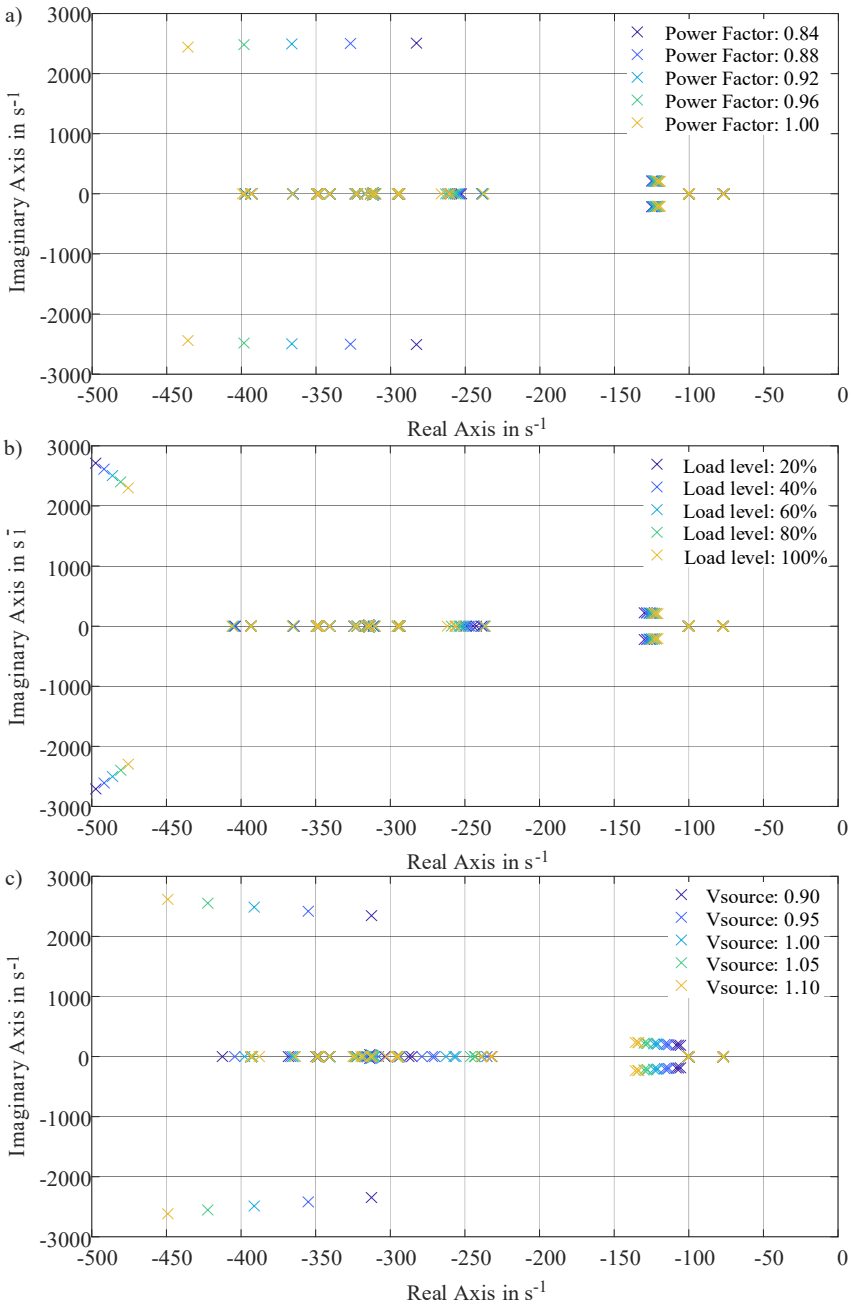


Fig. 3.16: Trajectory of partial eigenvalue for changes of a) converter power factor b) load c) voltage source

3.3.3 Variables representing dominant damping mode

From the last section 3.3.2, it is known that the damping frequency of the system's dominant dynamic feature is around 400 Hz, see Tab. 3.1. To determine which system variables response provides the best representation of the dominant oscillation mode of the system, two scenarios for small signal response are designed, and the variables for example PCC voltage, converter output current, and converter active and reactive power, are analyzed in frequency domain.

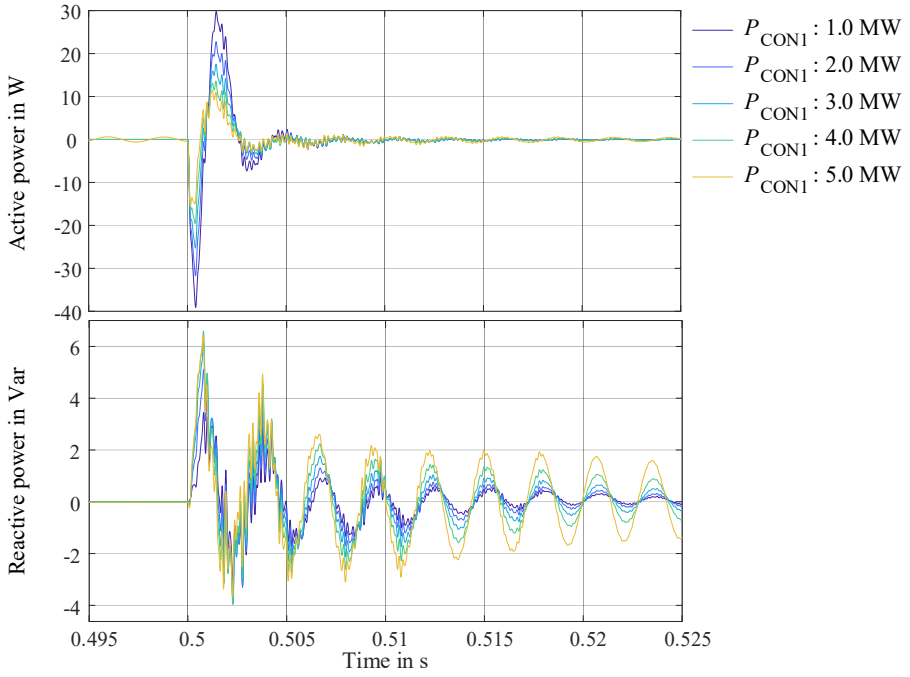


Fig. 3.17: Active power and reactive power of converter CON1 small signal response

For the first scenario, in the case of active power generations from 1 MW to 5 MW in converter CON1, a small perturbation with 1 V at 0.5 s is added to the PCC voltage at bus 12, where the converter CON1 is connected, see Fig. 3.3. The figure above Fig. 3.17 shows the variations of both active and reactive power of converter CON1 in these five cases. It is clear that the same variable has a similar oscillation mode under different operating points for the small signal. Another conclusion that can be drawn from Fig. 3.17 is that different variables have different dominant damping modes. For the active power, there is a dominant damping frequency of around 260 Hz, while for the reactive power, this frequency is around 400 Hz. To gain a clearer view of the damping mode of each signal, the FFT analysis is conducted on the time-domain signal to measure the frequency content.

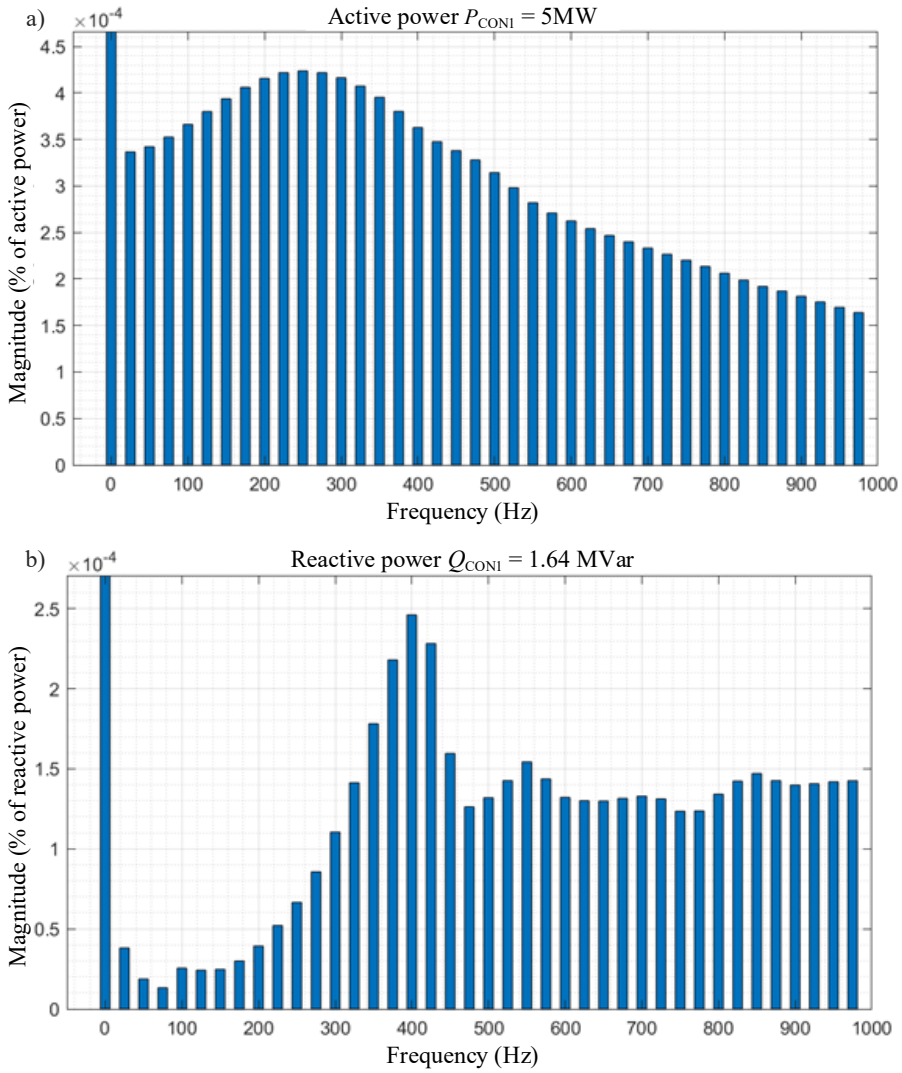


Fig. 3.18: FFT result of small signal response for a) active power and b) reactive power of converter CON1

The FFT analyses of both signals for active power generations of 5 MW are performed within a time window that starts with a small signal response of 0.5 s (see Fig. 3.17) and ends within two periods at the system frequency of 50 Hz. The results are shown in the above figure Fig. 3.18, which shows the spectrum relative to the DC component of the power signal. They are specific to the DC component because the FFT analysis is performed for the RMS value of power. It is clear that the response of active power has a dominant damping frequency of 250 Hz, while the reactive power response has a frequency of 400 Hz, which is the dominant frequency of the system damping mode.

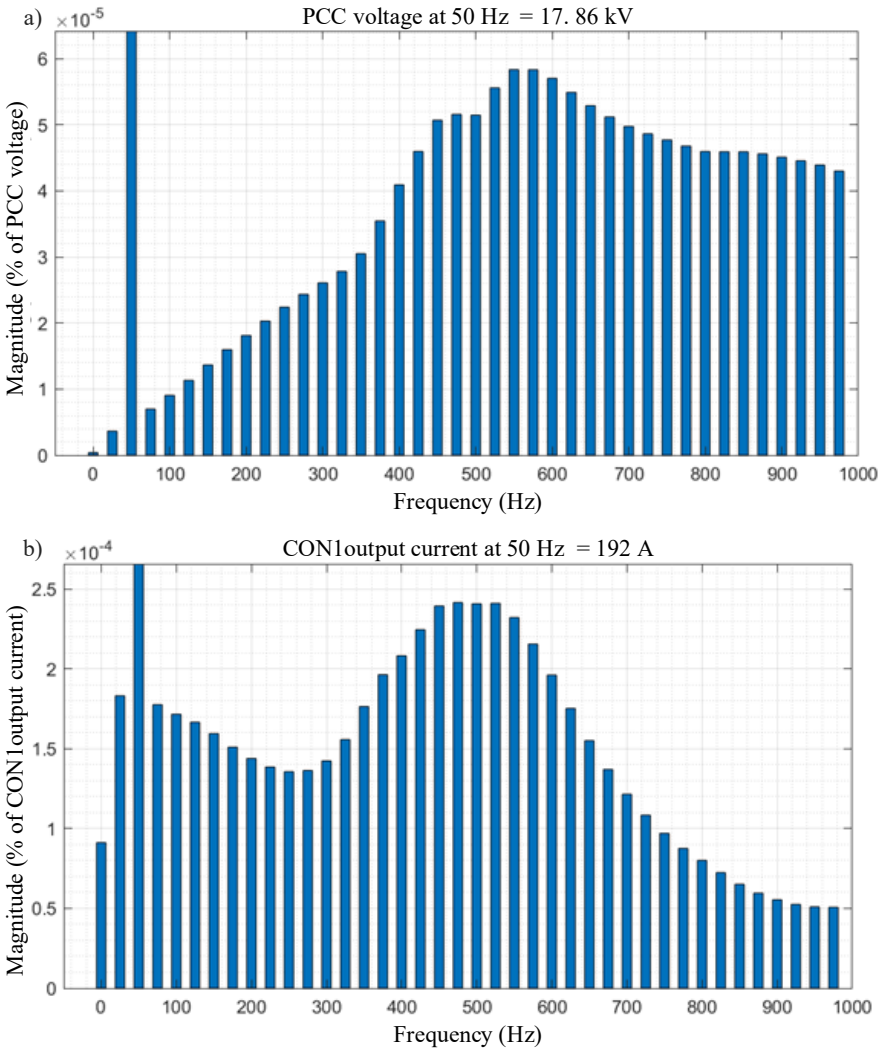


Fig. 3.19: FFT result of small signal response for a) PCC voltage and b) converter CON1 output current

The FFT analyses are conducted at both these signals to study the damping mode of the voltage at the connection bus and the converter output current. Fig. 3.19 shows that both signals have a fundamental frequency of 50 Hz. In addition, there is a dominant harmonic of about 560 Hz for voltage and about 500 Hz for current.

The results of FFT analyses of the other small signal response are shown in appendix A.3. In this scenario, a small disturbance of 1 W is injected into converter CON1 at 0.5 s while its active power is at a level ranging from 1 MW to 5 MW and its power factor keeps at 0.95, see Fig. A.13.

Same as the first scenario, the reactive power damping contains a dominant oscillation frequency of 400 Hz, which fits to the frequency of the system dominant damping mode. Therefore, a conclusion can be drawn that the reactive power response of the converter can provide information of the dominant oscillation frequency of the system. This information can easily identify the dominant eigenvalue from the system eigenvalue map.

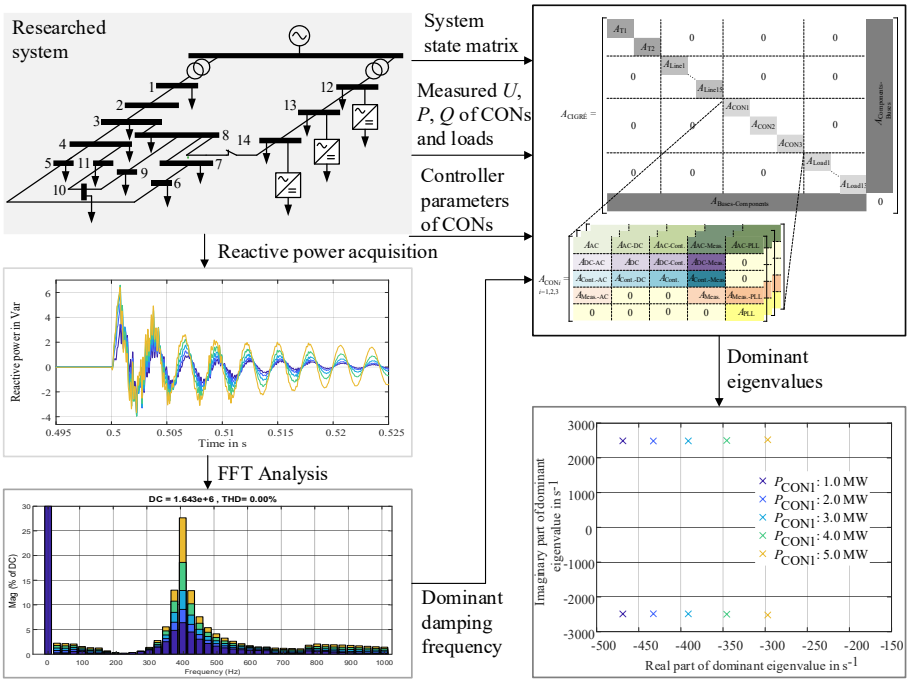


Fig. 3.20: Process of identifying dominant eigenvalues

The figure above shows the process of identifying the system's dominant eigenvalues. The measured reactive power from the studied system is extracted for FFT analysis to determine the main oscillation frequency. Through the analysis above, this frequency also characterizes the dominant oscillation frequency of the system. By deriving its system state matrix from the studied system and solving the state matrix for the eigenvalues, together with the information on the oscillation frequency, the dominant eigenvalues of the system can be quickly determined. The oscillation SI is then derived from the dominant eigenvalues.

3.3.4 ANN based stability indicator estimator

As mentioned, the eigenvalues of $\boldsymbol{A}_{\text{CIGRÉ}}$ reflecting the system damping mode. The imaginary part of the operating point strong relevant eigenvalue characterizes the main oscillation frequency of the system with varying operating points. If this eigenvalue lies in the left half-plane and away from the imaginary axis, it indicates that the system is stable with a high degree of stability. Suppose it is in the RHP that represents an unstable system. It is choosing the damping ratio of this eigenvalue as the system SI, which can indicate the stability margin as well as the instability of the system. However, calculating the system eigenvalues is time-consuming, especially for the

higher-order system. Therefore, the acquisition of the SI value is computationally intensive. This section introduces an ANN-based approach to estimate the SI value of the system, which is the basis for implementing online stability observation.

As known that ANN can generate the functional relationship between input data set X and output data set Y based on training. This ability of ANN is widely used for solving specific problems of the power system, such as load prognosis, voltage stabilization, fault classification, and the design of voltage stabilizers [144]. By training the ANN to match the desired pattern, an operating point strong relevant pattern in the designed case allows the associated eigenvalues to be located. In this work, a multi-layered feedforward network is selected as the structure of the ANN using the algorithm of supervised learning. The input neurons correspond to the current operating point of the voltage at the PCC bus, the output power and current of the converters and the loads, as well as the controller parameters of converters, while the output neurons represent the damping ratio of the system dominant eigenvalues, which is used as the SI of the system, see Fig. 3.21.

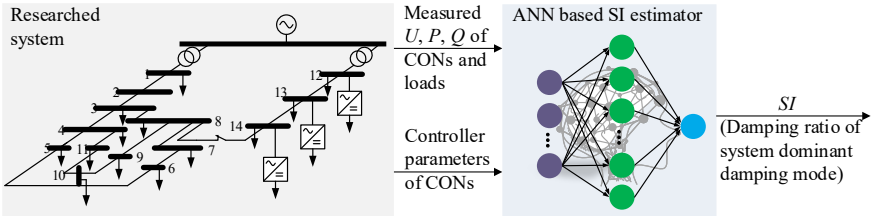


Fig. 3.21: SI estimation using ANN-based estimator

According to the state matrix construction in section 3.3.1, its eigenvalues are related to the various operating points, the controller parameters, and the components parameters. The parameter of components, for example, the inductance and resistance of the power line rarely change; therefore, these parameters are considered constants. According to the parameterization concept in this work, the controller parameters that vary with the operating points are chosen as the parts of the SI-estimator inputs. The parameterization concept, namely the adaptive control concept, will be introduced in the next section 3.4.

3.4 Adaptive control concept with reinforcement learning

As is known, the controller parameters may need to be readjusted in case the system conditions change, as the parameters optimized for a particular condition may not result in the desired response under another condition [145]. An ML method RL is implemented to retune the parameters when system condition changes to achieve desirable performance. This section introduces the training and implementation of the RL control in a multi-converters system. The following figure, Fig. 3.22 exemplifies the conceptual diagram of RL training and implementation, which will be introduced in sections 3.4.1 and 3.4.2, respectively. Employing the SI derived from the previous section as a reward in the training process and continuously varying the operating points P and Q as well as the source voltage V_s from 110kV grid to tune the control strategy for RL, an optimized strategy is finally obtained and implanted into the studied system to arrive at a system with high robustness as well as to achieve desirable performance. To ensure that the RL can perform its role of optimizing the control parameters during the operation of the network, the DT used for RL training as mentioned above requires constant updating.

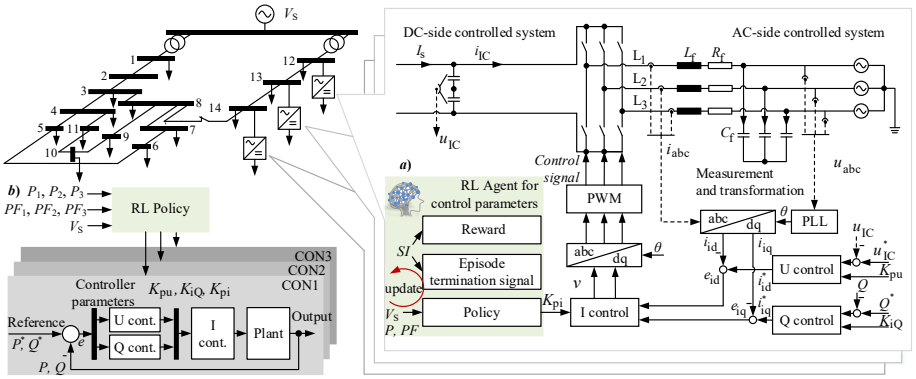


Fig. 3.22: RL control in CIGRÉ multi-converters system a) Training b) Implementation

The above figure illustrates that the outputs of the RL Agent are the parameters of the controllers in each converter, which is the way this work employs the RL to tune the parameters. In order to reduce the effort as well as the duration of RL training, the parameters that are more necessary for tuning should be filtered first. As shown in Fig. 3.23, when keeping the other control parameters constant and changing only the current PI-controller proportional parameter K_{pi} of the converter CON1, the trajectory range of the dominant eigenvalue is much larger than when changing only the integral parameter K_{ii}. To quantify the degree of influence of each control parameter on the dominant eigenvalue, the following parameter sensitivity analysis is performed using the damping ratio corresponding to the dominant eigenvalue as an indicator, which is the oscillation SI derived in the previous section 3.3.2. A selection is conducted from the parameters with high sensitivity to determine which parameters are to be adjusted using RL.

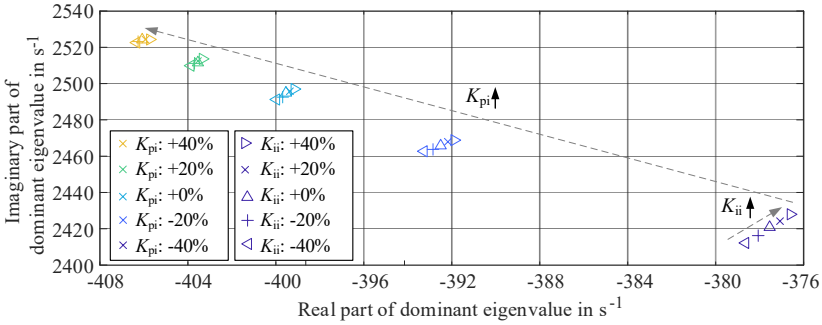


Fig. 3.23: Trajectory of dominant eigenvalues for changes of converter CON1 current controller parameters

The standard grid-following converter control structure is described in section 2.2.2. It is known that the PI-controllers are used for both the inner current and outer voltage controllers, as well as in the PLL that is used for synchronization, while the I-controller is chosen for the reactive power. Therefore, for a converter, there are a total of seven controller parameters, which is in the case that the same controller parameters are used for the d- and q-axis current. In the following, a

parameter sensitivity analysis is performed for the 21 controller parameters of the three converters in the studied system.

In order to evaluate the impact of fluctuations in the individual controller parameters on the variation of the system output, changing one parameter by a certain percentage and holding all other parameters constant. Hereby, four variations of percentages are designed, which are -40%, -20%, +20%, and +40%. The values of SI are calculated and recorded under different parameter variations, see Tab. A.5 in Appendix A.4.

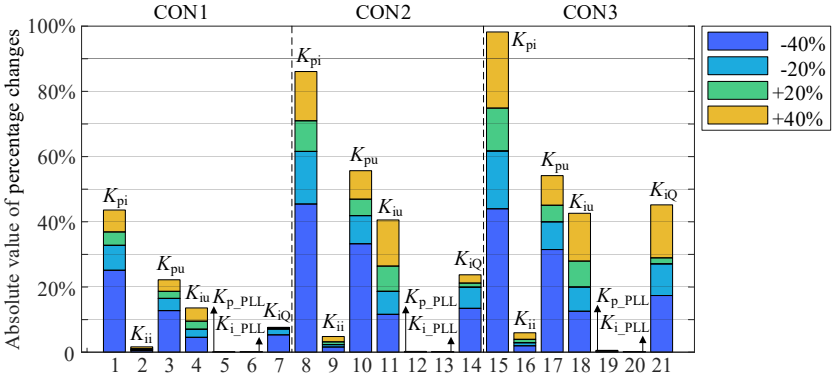


Fig. 3.24: SI absolute percentage changes under different parameter variations

When the controller parameters vary positively, some parameters cause positive changes in the indicators while others cause negative changes, see Fig. A.18 in Appendix A.4. To visualize and compare the sum of the selected SI changes caused by the four variations of each parameter, the absolute value changes of the indicator are shown in Fig. 3.24. It is observed that the variations of the indicator associated with the variation of each parameter of converter CON3 are larger compared to those of converters CON1 and CON2. This is due to the fact that CON3 is located at a farther distance from the slack bus in the studied system than the other two converters, see Fig. 3.3. The system impedance seen from the CON3 is greater than from other mentioned locations because it is mainly affected by the transmission line impedance [146]. Owing to the high system impedance, the CON3 is more vulnerable to causing system instability [147]. Therefore, its controller parameters are more sensitive for the studied SI. According to the sum of the selected SI absolute percentage changes in figure Fig. 3.24 above, parameters that cause the changes more than a predefined value are therefore chosen to be parameterized using RL.

3.4.1 RL agent training setup

As mentioned in section 2.1.3, the RL agent is trained in a simulated environment, and the parameters of each NN in the agent are continuously adjusted according to the reward changes in the process to obtain the best control strategy, i.e., the optimal agent.

Selection of training algorithm

TD3 algorithm is employed in this work, which is a method applied to the continuous action space. The application of the TD3 allows the agent to continuously change the parameters of the controller during the interaction with the environment.

The TD3 algorithm is an algebraic method based on policy gradients, which are more effective in high-dimensional action spaces or continuous actions. The actor-critic algorithm based on policy gradient theory is a widely used framework [74, 148, 149]. The framework consists of two main parts, the actor is used to update the parameter θ of the deterministic policy $\pi_\theta(s)$ through gradient ascent. It is unknown what the true action-value correspondence of the function $Q^\pi(s,a)$. However, an approximate Q-value function $Q^w(s,a)$ can be created using the parameter w. With a suitable policy calculation algorithm, it is possible to achieve this $Q^w(s,a) \approx Q^\pi(s,a)$. The policy here corresponds to the generation of the controller parameters, which outputs the action a according to the current operating points s. Here, action a represents the controller parameters derived in the previous section that greatly impact the system oscillation stability.

In the training process, the current state s and action a determine the reward r at the current moment, and as mentioned several times in the previous section, the SI is chosen as the reward. A point worth mentioning is that the Q-value is supposed to be the sum of current and future rewards, whose representation is shown in (2.40) in section 2.1.3. To estimate the steady-state state of the system and the value of the SI, the ANN is applied, and the SI is known from section 3.3.4 to be related to the current states and the controller parameters. In addition, it is expected that the RL agent achieves an optimization that outputs controller parameters change with the state so that the system always maintains its robustness and high stability margin. In other words, the RL agent is required to output the controller parameters that match it in any state, i.e., its Q-value should be maximized at the current moment. This means that the Q-value is linked only to the current reward. Therefore, the discount factor γ should be set to 0. During the training process, the output of the critic is constantly fitted to the current Q-value, and the actor selects the output action based on the maximum estimated Q-value.

Observation setup

As shown in Fig. 3.22 a), the observation vector comprises $[P_{\mathrm{CON1}}\ PF_{\mathrm{CON1}}\ P_{\mathrm{CON2}}\ PF_{\mathrm{CON2}}\ P_{\mathrm{CON3}}\ PF_{\mathrm{CON3}}\ V_{\mathrm{S}}]$ the reference active and power factor of three converters. According to section 3.3.2, the variation in power and supply voltage can cause a relatively large change in the DR of the dominant oscillation mode of the system, meaning these variations significantly impact the system oscillatory stability. They are chosen as observation states for the RL-agent so that the RL-optimized controller parameters can be adjusted to these states as they change and so that the system always maintains a high stability margin.

Action setup

As mentioned, the action of the RL agent is the controller parameter of the converters. These parameters that have different effects on SI are analyzed in the previous section. Based on the analysis results, i.e., Fig. 3.24, the parameters whose effects exceed a given value v are selected for parameter optimization using RL-agent. Three different degrees of influence bounds v are chosen, and their corresponding controller parameters with influence over the threshold are listed in the table Tab. 3.2 below.

Consistent with the previous analysis, the location of converter CON3 determines the higher sensitivity of its controller parameters. In the same way, the parameters of CON2 have a greater

impact on the system stability than those of CON1. Among the four controllers in the same converter, the proportional parameters of the current controller in the inner loop are the most sensitive. Therefore, when the influence bound is chosen to be 60%, the only controller parameters that need to be optimized are the proportional parameters of the current controller in CON2 and CON3. Both parameters are then defined as the two action variables in this case.

Tab. 3.2 Three influence bounds *v* and their corresponding controller parameters with influence over the threshold

***Influence bounds v* [%]**	***Controller parameters***
40	$K_{\text{pi\_CON1}}$, $K_{\text{pi\_CON2}}$, $K_{\text{pu\_CON2}}$, $K_{\text{iu\_CON2}}$, $K_{\text{pi\_CON3}}$, $K_{\text{pu\_CON3}}$, $K_{\text{iu\_CON3}}$, $K_{\text{iQ\_CON3}}$
50	$K_{\text{pi\_CON2}}$, $K_{\text{pu\_CON2}}$, $K_{\text{pi\_CON3}}$, $K_{\text{pu\_CON3}}$
60	$K_{\text{pi\_CON2}}$, $K_{\text{pi\_CON3}}$

Rewards setup

To complete the training more efficiently, the reward is partitioned. It is known that the reward is mainly related to *SI*. When *SI* is below 0, it means that the controller parameters output by the RL agent leads the system to divergence, so a large negative reward should be given. The more negative *SI*, the larger the corresponding penalty should be, so $SI \cdot p_1$ is used as the negative reward when *SI* is less than 0. Considering that the minimum value of DR in the system is not definitely equal to SI, DR_{min} is added to the penalty. When it is less than 0, it means that the system is unstable, so a penalty of $DR_{\text{min}} \cdot p_2$ is given. In addition, there may be an eigenvalue with an imaginary part of 0 in the system, and its corresponding DR is constantly -1. Since it can also indicate system instability, the value when it is greater than 0 is also multiplied by a penalty factor p_3 as a negative reward for the RL agent.

Besides the setting of a penalty, the proper design of a positive reward will also positively impact the training. A larger SI represents a more robust system, so a positive reward should be given. To train the agent more effectively, the reward partitioning should be more elaborate when the *SI* is small, as shown in the following equation (3.62). Choosing 0.3 as a threshold value for the maximum reward lies in the fact that the maximum *SI* of the system does not exceed this value when using the traditional method of designing controller parameters.

$$reward = \begin{cases} SI \cdot p_1, & SI < 0 \\ DR_{\min} \cdot p_2, & DR_{\min} < 0 \\ -real(Eig) \cdot p_3, & real(Eig) > 0 \,\&\, imag(Eig) = 0 \\ SI \cdot 10, & 0 < SI \le 0.025 \\ SI \cdot 20, & 0.025 < SI \le 0.05 \\ SI \cdot 30, & 0.05 < SI \le 0.10 \\ SI \cdot 40, & 0.1 < SI \le 0.2 \\ SI \cdot 50, & 0.2 < SI \le 0.3 \\ SI \cdot 100, & SI > 0.3 \end{cases} \tag{3.62}$$

where p_1, p_2, and p_3 are the penalty factor for the mentioned three parts of negative reward.

Episode termination signal setup

Another possibility to improve the training efficiency is to set a suitable signal to terminate the current episode. When the action taken by the RL agent is far from the action of the optimal policy, it is highly likely to lead to a non-converging oscillatory mode in the system. Therefore, the DRs corresponding to all eigenvalues of the system are calculated and the minimum value is taken as the signal to end the current training episode, as shown in Equation (3.63).

$$signal = \min\left(DR \text{ among all the eigenvalues}\right) \tag{3.63}$$

When it is less than 0, the training of a certain episode can be ended early. A new round of training is continued based on the best control policy of the previous episode.

Actor and Critic networks setup

The actor-critic networks are implemented as shown in Fig. 3.25. Both networks with fully connected layers are initialized with small random weights before starting the training. The output of the actor-network is normalized between [-1, 1] using Relu layers. It results in better learning and convergence in the continuous action space.

Training setup

The setting of the training hyperparameters, such as the learning rate of critic and actor networks etc., are listed in the Tab. 3.3 below.

Tab. 3.3 Training hyperparameters setting

Description	***Setting***
Critic learning rate	$1e^{-03}$
Actor learning rate	$1e^{-04}$
Discount factor	0.0001
Target network updating coefficient	0.005
Activation function	Relu

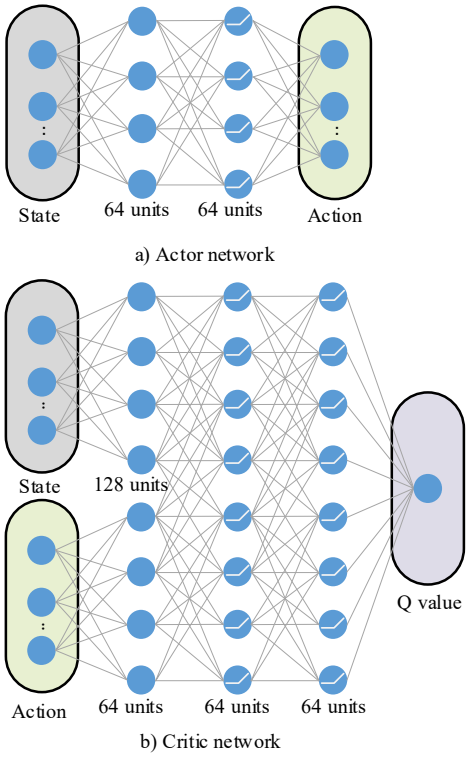


Fig. 3.25: Actor and critic networks architecture

The flowchart of the RL agent training is shown in Fig. 3.26. After initializing the parameters of the actor and critic network, the actor takes action, i.e., outputs the controller parameters under the explored system operating points. The following step is to calculate the reward under each step of training. It is calculated according to the SI, the smallest DR among all the eigenvalues as well as the smallest eigenvalue whose imaginary part is equal to 0. The accumulated reward is recorded as the reward of this episode. Actor and critic network parameters are continuously updated in the iteration, and after each update actor will output new controller parameters, thus generating a new step reward. The training in one episode keeps going until the maximum number of training steps in this episode is reached. In this work, the condition for outputting the optimal RL agent is that the reward of an episode is greater than r. The value r is decided by the steps of the episode, and the excepted SI under the explored states range during the training. A new round of training episodes starts if the episode reward is less than this value. The operating points of the system, i.e., the reference active and power factor of the three converters and the supply voltage, are changed in the new episode for a new round of exploration. The training continues until the reward of the episode is greater than the set value, and then an optimized RL agent is obtained.

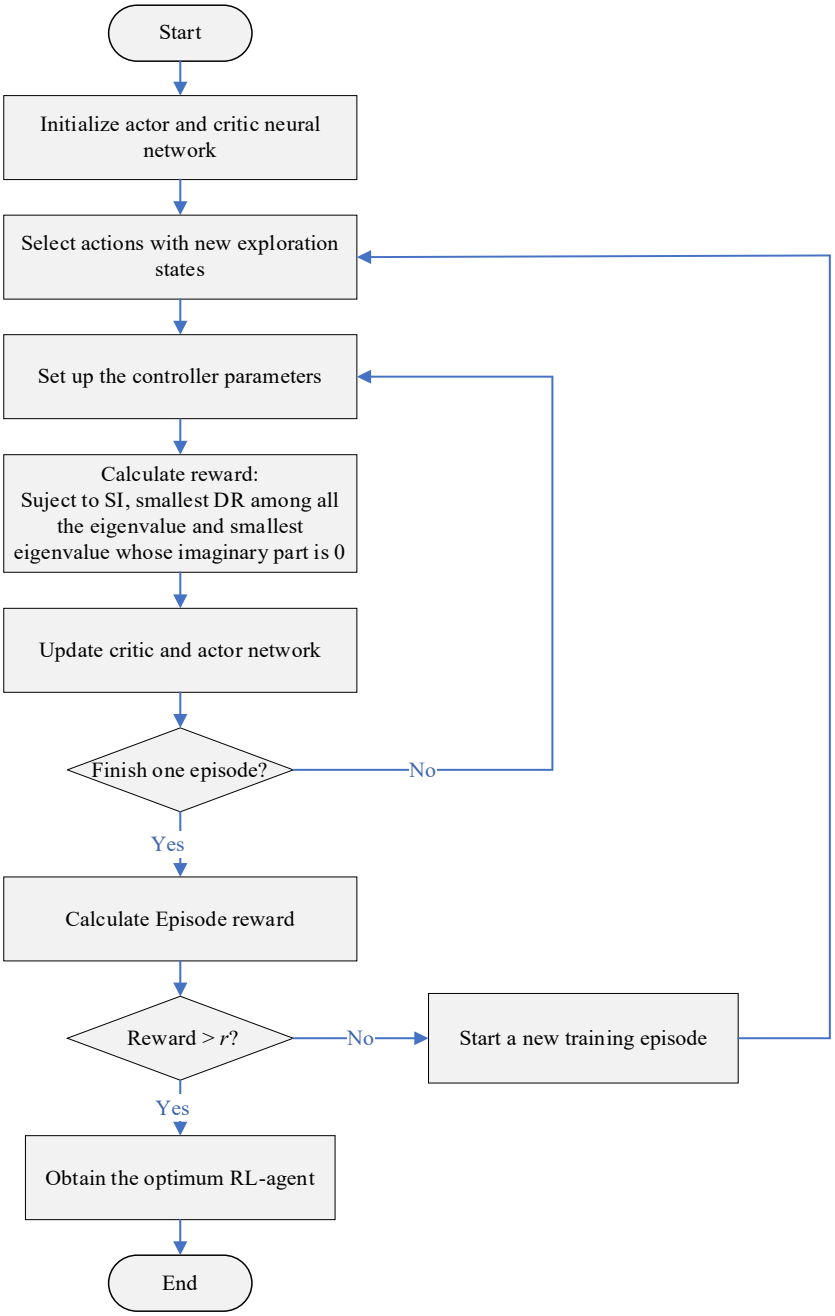


Fig. 3.26: RL agent training flowchart

In this work, three RL agents will be trained to optimize the three sets of controller parameters listed Tab. 3.2. Taking the third set of cases as an example, the TD3 algorithm is applied by

defining two actions, one for each controller gain. The RL agent is defined to have a wide search range of ±10 times the setting of method MO. The control strategy is optimized according to the SI values corresponding to the different operating states and the outputs of the different actions. With reference to the i-th action (parameter), a_i is defined over the pre-specified range $\{a_i(\min), a_i(\max)\}$. For each iteration k of the algorithm, the action $a_i(k)$ is chosen according to the Q-value function $Q_i(a_i, s, k)$. As the system learns, the Q-value function usually converges to the optimal value function Q_i^* with probability 1. The actor in the RL agent will generate desired parameter values according to the optimal Q-value function and the system states.

3.4.2 RL agent implementation

After the RL agent is trained to obtain the optimized RL policy, it will be embedded in the studied system, as shown in figure Fig. 3.22. Depending on the current operating points, the RL agent will continuously adjust the output controller parameters and implement them into the corresponding voltage, current and reactive power controllers.

As mentioned above, it is desired to train an RL agent that can bring the system to achieve the best stability at each operating point, meaning that the SI can always attain the maximum. In traditional control methods, the controller parameters are configured to maximize the SI at a given operating point. However, if the parameters are not changed at another operating point, the corresponding SI may not be the maximum value that can be reached. The following Fig. 3.27 shows the illustrative diagram of the previous description. The gray-shaded eigenvalues correspond to the maximum SI for the five different parameters at the relevant operating point. It can be seen that when $P_{\text{converter CON1}}$=1,2,3 MW, it is the case $K_{\text{pi\_converter CON1}}$ = +40% that maximizes the SI. The cases that maximize the SI when $P_{\text{converter CON1}}$ is 4 and 5 MW are $K_{\text{pi\_converter CON1}}$ = +20% and $K_{\text{pi\_converter CON1}}$ = +0%, respectively.

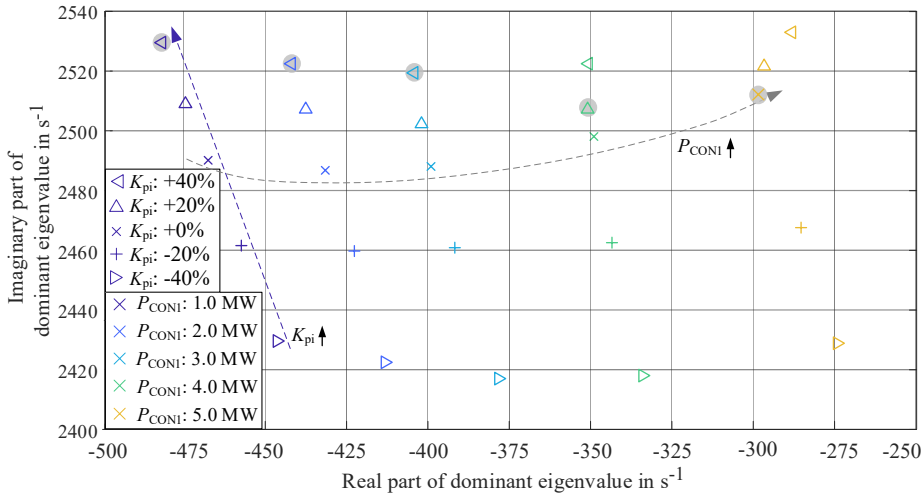


Fig. 3.27: Trajectory of dominant eigenvalue for changes of both converter active power and of controller parameter

SI can be expressed by Equation (3.64). The f in (3.64) indicates a nonlinear function, the OP stands for the operating point of the system in real-time, and the CP refers to the controller parameters.

$$SI = f\left(\boldsymbol{OP}, \boldsymbol{CP}\right) \tag{3.64}$$

The optimization goal of the RL agent is to maximize the SI globally. Therefore, the cost function is represented as the following equation (3.65).

$$J = \max\left(SI\right) = \max\left[f\left(\boldsymbol{OP}s, \boldsymbol{CP}s\right)\right] \tag{3.65}$$

When the optimized RL agent is integrated into the studied system, the SI with respect to OP and CP is given by the equation f_{RL}. In addition, the equation for the optimized relationship between OP and CP represented by the RL agent is denoted as h.

$$\boldsymbol{CP} = h\left(\boldsymbol{OP}\right) \tag{3.66}$$

Therefore, the following equation was obtained to describe the relationship between SI and OP.

$$SI = f_{\mathrm{RL}}\left[\boldsymbol{OP}, h\left(\boldsymbol{OP}\right)\right] \tag{3.67}$$

The above equation (3.67) also indicates that it is possible to obtain the index SI that characterizes the system oscillatory stability margin in relation to the current operating points.

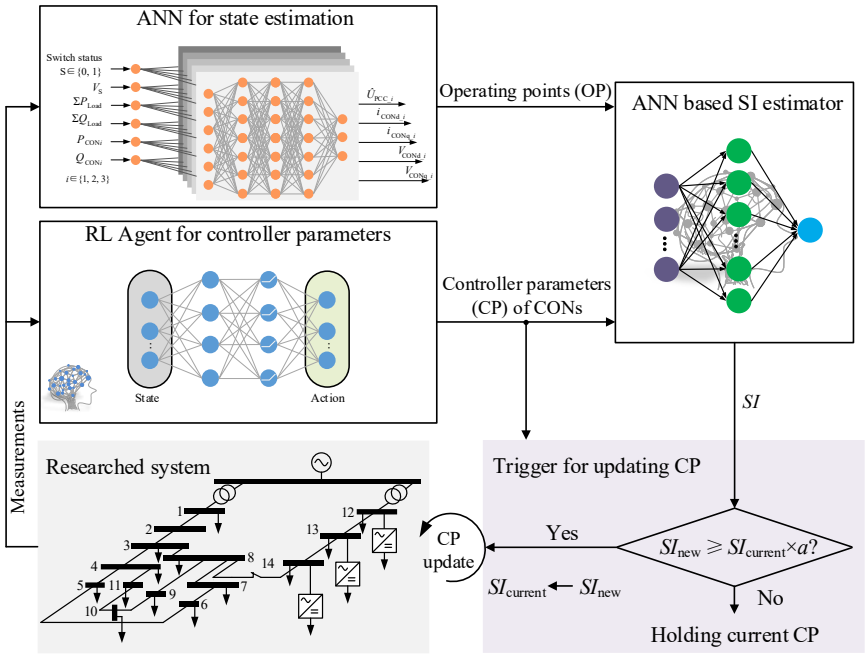


Fig. 3.28: Investigated system in operation with adaptive controller parameters using RL agent

The CP update is designed to maintain the maximum stability margin of the system. In case of a small change in the system operating points, the stability margin caused by the CP update remains unchanged. In such cases, adjusting the CP in the corresponding controller is a less meaningful process. In addition, since the CP calculation and data transmission [150] require time, it is potential that the updated CP may cause the stability margin to decrease instead of increase if the system state varies during this time. To avoid this situation, a trigger for updating the CP is added to the application of the RL agent in the investigated system.

As shown in figure Fig. 3.28 above, once the measured power of each converter and the sum of the load obtained from the measurement points of the system, the ANN responsible for the state estimation and the RL agent in charge of the controller parameters optimization calculate the corresponding operating points and parameters respectively based on the acquired inputs. The ANN for the SI estimation outputs the SI of the current system operating state depending on the above data OP and CP. The current SI is noted as SI_{current}, and the new state is noted as SI_{new}.The SI_{current} and SI_{new} are constantly compared, and when they satisfy the following relation, the CP optimized by the RL agent is updated to the corresponding controller and the value of SI_{current} is updated to SI_{new} Otherwise, the current CP is maintained, and the SI_{current} continues to be constantly compared with the new SI_{new}.

$$SI_{\text{new}} \geq SI_{\text{current}} \cdot m \tag{3.68}$$

The smaller the value of *m* in Eq. (3.68), the more frequently the CP is updated in the system. When the difference between the SI of the system before and after the CP update is small, it means that the system's stability remains almost the same margin, and there is less need for the CP update. The value of *m* is recommended to be set at 1.1, according to the study.

3.5 Summary

In this section, the concepts of this work are elaborated. First, under the assumption that the structure of the reference system is already available, but the parameters of its components are unknown, a DT of the reference system is created by parameter estimation. Through modeling of each network component, using the measured injected and consumed power in the system, as well as the voltage amplitudes and angles on some selected buses, the DT is constructed with the estimated parameters. The following studies will be performed in this DT.

An ANN-based state estimator is developed in the DT, which establishes the power flow relationship between the system states and the values of the system stability strong relevant variables, allowing a fast and accurate estimation of the operating point of the system, the output of which will be used as input to the stability detector.

A SI based on the eigenvalues of the state matrix is next constructed and implemented. To efficiently determine the dominant oscillation frequency of the system, some variables such as PCC voltage, converter output current as well as converter active and reactive power are analyzed in the frequency domain. The results show that the reactive power response of converter can provide the system with information about the dominant oscillation frequency. With this information, the dominant eigenvalues can be effectively identified from the eigenvalue map of the system. Since the calculation of the eigenvalues is time-consuming, an ANN-based SI estimator approach is

proposed to enable the eigenvalue-based SI presented above to be implanted online and indicate the current oscillatory stability margin of the system in real-time.

The values of the SI described above are utilized as RL agent reward values. The purpose is to take advantage of the ability of RL to maximize the reward and eventually obtain an optimized RL agent that mathematically relates the current steady-state operating point of the system to the parameters of the selected controllers of the converters. Thereby, the controller parameters that guarantee the maximum stability margin of the system in the changing operating point are provided.

In the next chapter, a series of numerical cases will be designed to verify the validity of RL agent-based adaptive control concepts for a converters dominated network.

4 Design of numerical case studies

The test scenarios designed for the proposed approaches of the DT of the reference network, the ANN state estimator, the automatic stability testing, and stabilization using an adaptive controller are shown in this chapter. Besides considering different influence factors, the construction of training for the DT model, the ANN-based SI estimator, and the RL agent are introduced, respectively.

4.1 Test scenarios of digital twin approach

In the reference network, it is assumed that there are only a limited number of measurement points on buses 2, 5, 8, 10, and 13, respectively, see Fig. 4.1. The measured data are the voltage amplitude and the phase angle. Based on the structure of the reference network, its DT model using the EMT modeling method is built, and the model is fed with the same power data generated and absorbed at each bus of the reference network. Afterward, through adjusting the parameters of transmission lines and the filter and controller in the converters by the selected algorithm, the difference between the voltage amplitude of a selected bus in the reference network and the DT model is reduced. Once the difference is minimized below the set value 5e-3, the parameter update ends, and the resulting model is used as the DT of the reference network.

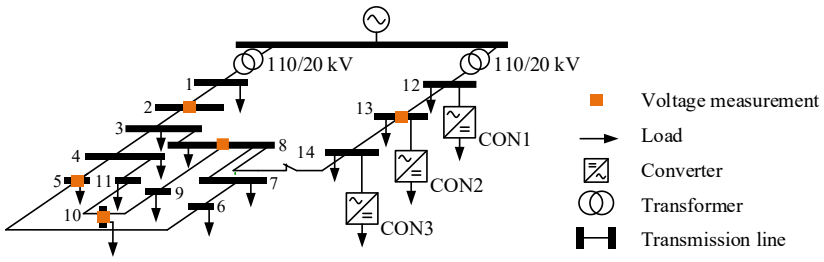


Fig. 4.1: CIGRÉ reference MV network

When the number of groups of measurement data used as reference values is large, the time required for parameter estimation grows. Therefore, only one set of voltage data from the five measurements is selected as a reference. To investigate the influence of the position of the measurement point corresponding to the reference data on the parameter estimation, each of the five measurement sets is used for estimation. In the study of this influence factor, full load is used for the load data, i.e., load level of 100%, whose data is from [131], see Tab. A.3 in Appendix A.1. For the power generator of each converter, the maximum generated powers are applied, which are defined in Tab. A.4. Combining daily production curves of PV and wind power plants in Fig. 3.6 and the maximum generation data of three converters in Tab. A.4, the daily power generation profiles of buses 12, 13, and 14 are shown in figure Fig. 4.2. In the same way, the daily consumption of each bus can be plotted separately by combining the standard load curve in Fig. 3.8 and the load parameters of each bus in Tab. A.3, see the figure Fig. 4.3 below.

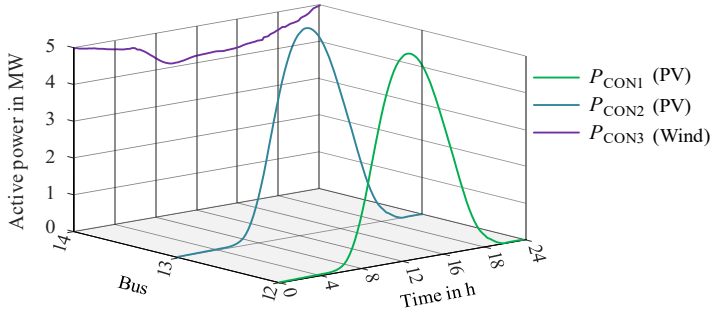


Fig. 4.2: Active power profile of power generation on bus 12, 13 and 14

To represent additional feeders served by the transformer, the load values given for buses 1 and 12 are much greater than the loads for the other buses. To allow them to be represented on a graph with other loads, these two loads are scaled down by a factor of 20.

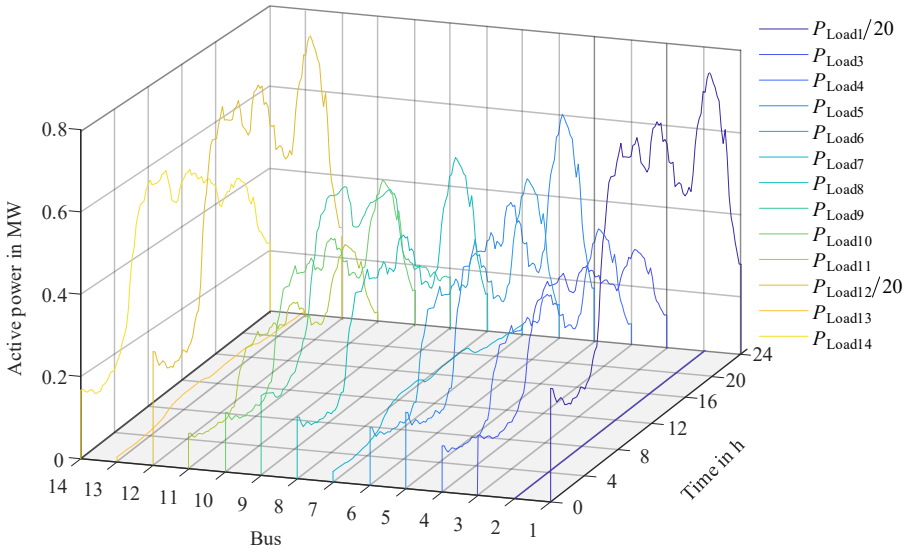


Fig. 4.3: Active power profile of load on each bus

In addition to the measurement point location, the amount of load power absorbed at the investigated bus, the presence or absence of a converter on the bus, and the amount of the generated power may influence the parameter estimation results. Therefore, the following scenarios in Tab. 4.1 for parameter estimation are designed.

Tab. 4.1 Scenarios for parameter estimation of DT

Scenario abbreviation		Description
PE_MPL		Training data obtained from different measurement point locations
Training data obtained from one measurement point → Investigating training data obtained in different network states	PE_NS_L	Investigated network connected to low and high load
	PE_ NS_CON	Selected bus connected with and without converter
	PE_ NS_G	Three converters in network with low and high generation

The daily consumption and production power data for training the DT are taken from the first 1/3 hours of the day, i.e., 0-8 h, and for its validation, the data from the remaining 16 h are used, see figure Fig. 4.4 below.

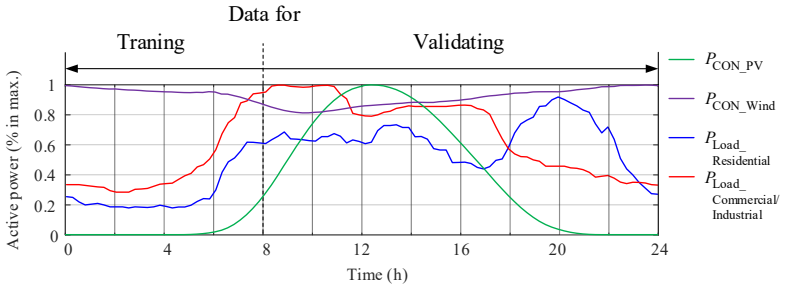


Fig. 4.4: Input variables data for training and validating the DT model

The model with the lowest error in the training results will be selected and its accuracy will be verified in the following scenarios in Tab. 4.2.

Tab. 4.2 Scenarios for validation of DT

Scenario abbreviation	Description
DT_SV	Source voltage step changes
DT_PQ	Active and reactive power of converters step changes
DT_OC	Open circuit on a selected line
DT_SC	Short circuit on a selected bus

The voltage amplitude and phase angle profiles of the selected bus of the reference network and DT model will be compared in the four test scenarios above. The corresponding mean absolute percentage error (MAPE) will be calculated. The equation of MAPE is shown in equation (4.1).

$$MAPE = \frac{100\%}{n} \sum_{k=1}^{n} \left| \frac{Ref_k - DT_k}{Ref_k} \right| \tag{4.1}$$

where Ref_k is the value of the five voltage amplitudes and five corresponding phase angles and DT_k is the value of the ten mentioned variables in the DT model. Besides, n is the sample number for each variable.

4.2 Test scenarios of ANN state estimator

The proposed NN scheme for the state estimator in this work, as shown in Fig. 3.11, has been implemented in the CIGRÉ investigated network shown in Fig. 4.1. For both training and validation of the estimator, a set of representative load daily profiles and daily generation profiles of wind and PV from CI-REG s under both switch statuses as well as their corresponding states of interest solutions are generated in the network DT from the last section. The configuration of network conditions is represented in Tab. 4.3.

Tab. 4.3 Different network conditions for training data creation

Switch	***Source voltage* [pu]**	***Loading* [%]**	***Power generations for each CI-REG* [%]**	***Time***
On; Off	0.9; 0.95; 1.00; 1.05; 1.1	10; 30; 50; 70; 90	[10 10 10]; [50 50 50]; [90 90 90]; [10 20 50]; [30 40 50]; [90 80 50]; [10 20 90]; [50 40 90]; [70 80 90]	0 – 24 h One data every 15 minutes

The topological change of the network is carried out by changing the switch S status, which is located between buses 8 and 14, see Fig. 3.3. The demand of daily load power for each bus is set for five levels from 10% to 90%. This means the total load power in the studied network ranges from 10% to 90% of the maximum values derived from the load parameters for the European MV network [131]. The power generations for each CI-REG are configurated into nine cases, as shown in Tab. 4.3. The configuration is for the three CI-REGs connected to buses 12, 13, and 14, see Fig. 3.2. The first two CI-REGs are of type PV generator and the last one is offshore wind generator. Their daily generation profiles are shown in Fig. 3.6. Both daily load and generation power data are collected every 15 minutes from 0 h to 24 h.

All training and validation data are created from the DT model of the network. Fig. 3.12 shows 10 combinations of five different load levels, from 10% to 90%, and two-generation cases, which are [10 20 90] % and [10 20 50] % of maximal power generations for each CI-REG, respectively. The maximal total load power is configurated to be 32.44 MW according to [131] and the maximal generation power for each CI-REG is set to 5 MW. The corresponding states of interest datasets are also obtained from the DT model under the different designed network conditions. These output datasets are then combined with the input datasets and used as the input-output pairs for the model training.

The ANNs structure adopted for each variable of three converters in the proposed scheme is as follows:

- $\hat{U}_{\mathrm{PCC}\_i}$: 10 input neurons - x hidden neurons, and 3 output neurons for $\hat{U}_{\mathrm{PCC}\_1,2,3}$.
- I_{CONd_i} : 10 input neurons - x hidden neurons, and 3 output neurons for $I_{\mathrm{CONd}\_1,2,3}$.
- I_{CONq_i} : 10 input neurons - x hidden neurons, and 3 output neurons for $I_{\mathrm{CONq}\_1,2,3}$.
- V_{CONd_i} : 10 input neurons - x hidden neurons, and 3 output neurons for $V_{\mathrm{CONd}\_1,2,3}$.
- V_{CONq_i} : 10 input neurons - x hidden neurons, and 3 output neurons for $V_{\mathrm{CONq}\_1,2,3}$.

In the training configurations, a sinusoidal activation function is used for all hidden layers, and the percentages of the training, validation and testing samples are set at 70%, 15% and 15%. To explore the effect of number of neurons in the hidden layer of each ANN on the accuracy of ANN state estimator, the following scenario is designed, see Tab. 4.4.

Tab. 4.4 Scenario for investigating the effect of number of neurons in the hidden layer on the accuracy of ANN state estimator

Scenario abbreviation	Description
SE_HLN_ANN	Varying number of neurons in the hidden layer of each ANN

The scenarios for testing the accuracy of the state estimation, including the variations of the operating conditions and the changes in network topology, are drawn as follows, see Tab. 4.5.

Tab. 4.5 Scenarios for testing the accuracy of the state estimation

Scenario abbreviation	Description
SE_24h_ANN	Source voltage changing for 24 hours
	Power generated and consumed changing on each bus for the same 24 hours
	Switching state of the transmission line changing during the same 24 hours

To compare the estimation results with the corresponding variables in the reference system and obtain the accuracy of the estimator, the values of the variables are then computed with the measurement using a standard load flow program. The five ANNs have been trained up to a training MSE error of 0.0008, a validation MSE error of 0.001 and a testing MSE error of 0.001. Network training and testing are performed on a 1.80 GHz Intel Core processor. In section 5.2, the performance of the proposed ANN scheme is evaluated.

4.3 Test scenarios of automatic stability testing

A SI for indicating the dynamical stability margin is designed and used to automatically test the system stability during its operation. As shown in section 3.3.2, the applied SI in this work is the

damping ratio of the dominant eigenvalue in the investigated system. The following scenarios are designed to test and validate the SI. In the tests in this section, the data of required equilibrium operating points, namely the PCC voltage $\hat{U}_{\mathrm{PCC0}}$, the converter output current I_{CONd0}, I_{CONq0} and the control vectors V_{d0}, V_{q0}, to calculate the SI are estimated using the ANN state estimator with the highest accuracy of the tested results in the previous section.

Test scenarios using eigenvalue-based SI are shown in Tab. 4.6.

Tab. 4.6 Scenarios for accuracy of stability testing using eigenvalue-based SI

Scenario abbreviation	Description
SI_P	Varying active power feeding from a selected CON
SI_Q	Varying reactive power feeding from a selected CON
SI_SV	Varying source voltage
SI_L	Varying loading conditions

In section 3.3.4 it is mentioned that for online stability detection, an ANN-based SI estimator is constructed. The different network conditions to create the training data for the SI estimator are the same from the state estimator in the previous section 3.2. Under these conditions, the corresponding eigenvalue-based SIs are calculated, and these SI values are employed as the reference output of the ANN-based SI estimator. As known that the SI is related to not only the equilibrium operating points but also the system parameters. In this work, the parameters of the network components are considered constant, while the controller parameters in each converter are adjustable and optimizable. The controller parameters are optimized with the help of the RL agent according to the operating points, see section 3.4. Therefore, this nonlinear mathematical relationship between the operating points and controller parameters is included in the SI estimator to simplify it. Using the same inputs from the ANN state estimator, namely the status of switches, the voltage source amplitude, the active and reactive powers of each converter, as well as the sum of the active and reactive power consumption, as inputs and SI as output, see Fig. 4.5, the ANN-based SI estimator is trained and validated.

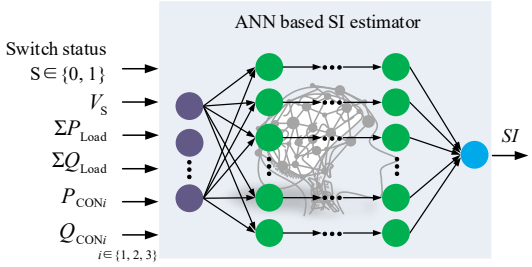


Fig. 4.5: ANN-based SI estimator

The SCG backpropagation is used as the training algorithm and the layer numbers of 10 are chosen for the ANN-based SI estimator. To verify its accuracy, scenarios designed for the estimator are the same as the test scenarios using eigenvalue-based SI.

4.4 Test scenarios of stabilization using adaptive control approach

In this subsection, as mentioned in 0, the three RL agents will be trained to optimize the three sets of controller parameters listed Tab. 3.2. Under the same critic and actor networks design (including the learning rate of both networks, the numbers and units of layers, as well as the activation function, etc., see Tab. 3.3) and the same setup of observation and reward, the results of the training are compared.

Training influence factors

- Training steps for each Episode

Initially, the training duration of each Episode is 0.25 s, the sampling time of the agent is $T_s = 50$ μs. The Episode step is obtained from the training duration divided by the sampling time, i.e., the training steps for each Episode is 5000.

A wide range of exploration was set to maximize the SI value of the RL agent at all possible operating points. It is first necessary to normalize the 7 observed variables, which are $[P_{\mathrm{CON1}}\ PF_{\mathrm{CON1}}\ P_{\mathrm{CON2}}\ PF_{\mathrm{CON2}}\ P_{\mathrm{CON3}}\ PF_{\mathrm{CON3}}\ V_{\mathrm{S}}]$. It is assumed that for each converter, the active power varies from 0 to 5 MW, and the corresponding power factor varies from 0.85 to 0.95, as shown in the following equation.

$$P_{\mathrm{CON}i} = [0,\ 5]\ \mathrm{MW},\ PF_{\mathrm{CON}i} = [0.85,\ 0.95],\ i = 1,2,3 \tag{4.2}$$

The range of its variation for the voltage source is shown in (4.3) below.

$$V_{\mathrm{S}} = [0.9,\ 1.1]\ p.u. \tag{4.3}$$

Using random numbers that are uniformly distributed in the interval [0, 1] to represent the variation of these seven observed variables, their equations are rewritten as the following equation (4.4).

$$\left.\begin{aligned} P_{\mathrm{CON}i} &= [0,\ 1]\cdot 5\ \mathrm{MW} \\ PF_{\mathrm{CON}i} &= 0.85 + [0,\ 1]\cdot 0.1 \end{aligned}\right\} i = 1,2,3 \qquad V_{\mathrm{S}} = 0.9 + [0,\ 1]\cdot 0.2\ \ p.u. \tag{4.4}$$

The observed variables are kept constant within the same episode. Since the RL agent is known to be reward-oriented, it continuously optimizes the actor and critic network parameters at each training step and records the accumulated rewards at the end of one episode. The larger the number of steps set for each episode, the higher the probability of getting an agent optimized for the corresponding observation, namely the system operating point. However, the training time will be increased. Therefore, the step of the episode is considered as an influential factor in agent training.

- Penalty factors

There are positive and negative rewards in the RL agent training process; generally, the positive rewards are referred to as rewards, and the negative ones are referred to as penalties. For the trained system, the rewards are set for the stability margin when the system is stable, while the penalty is for the instability of the system. As introduced in section 3.4.1, the rewards are only subjected to the positive SI. Besides, the corresponding reward factors are different when the SI is in different ranges, as shown in Equation (3.62). Unlike the rewards, the set penalties consist of

three different items: the SI, the smallest DR among all the eigenvalues, and the smallest eigenvalue whose imaginary part is equal to 0. The p_1, p_2, and p_3 are the penalty factor for the mentioned three penalty items. Since the third penalty is likely to differ from the first two penalties by order of magnitude, the designed penalty factors are to be tuned during the RL training so that the penalty values do not differ significantly, leading to excessive differences in the optimization of the RL agent in the former and the latter stages of the training.

- Action variables

The numbers of the action variables are given by the set of controller parameters listed in Tab. 3.2. For the first set [K_{pi_CON1}, K_{pi_CON2}, K_{pu_CON2}, K_{iu_CON2}, K_{pi_CON3}, K_{pu_CON3}, K_{iu_CON3}, K_{iQ_CON3}], the RL agent should conduct 8 actions for the 8 parameters. The other two sets of parameters [K_{pi_CON2}, K_{pu_CON2}, K_{pi_CON3}, K_{pu_CON3}] and [K_{pi_CON2}, K_{pi_CON3}], the number of action variables are set to be 4 and 2, respectively. The corresponding optimized agent is denoted as RL agent$_i$, where i =1,2,3. The greater the number of actions that need to be produced, the longer the training may also take. However, the more parameters are adjusted simultaneously, the faster the set training stop value may be reached.

The RL agent will be trained under the three influence factors listed above. The detailed variation settings are shown in the table below. The penalty factors are set in consideration of the parameters corresponding to the three factors. For p_1 and p_2, the corresponding parameters are DR, whose absolute values are between [0,1], while for p_3, the parameter is the maximum pure real eigenvalue of the system, whose values are above several thousand according to the observations.

Tab. 4.7 Variation setting of training steps, penalty factors and action variables

Description		***Setting***
Training steps		[50, 100, 500, 2000]
Penalty factors	p_1	[1000, 2000, 3000]
	p_2	$p_2 = p_1 = p$
	p_3	1
Action variables	RL agent$_1$	8 actions: [K_{pi_CON1}, K_{pi_CON2}, K_{pu_CON2}, K_{iu_CON2}, K_{pi_CON3}, K_{pu_CON3}, K_{iu_CON3}, K_{iQ_CON3}]
	RL agent$_2$	4 actions: [K_{pi_CON2}, K_{pu_CON2}, K_{pi_CON3}, K_{pu_CON3}]
	RL agent$_3$	2 actions: [K_{pi_CON2}, K_{pi_CON3}]

As mentioned in section 3.4.1, the reward for the determination to end the training is defined by the steps of the episode and the excepted SI under the range of states explored during the training. The ranges of the exploration are shown in equation (4.4). The average SI for these possible operating points is expected above 0.175. Therefore, for the four training steps setting as listed in

Tab. 4.7, the average episode reward r is designed to be $0.175 \cdot 50 \cdot [50, 100, 500, 2000]$ as the training termination signal, respectively.

To investigate the effect of training steps, penalty factors for three different action variables on the rewards, following scenarios in Tab. 4.8 with the various settings shown in Tab. 4.7 are designed.

Tab. 4.8 Scenarios for investigating the effect of training steps, penalty factors for three different action variables on rewards

Scenario abbreviation	Description
RL_A1_S	Varying training steps for RL agent$_1$
RL_A1_P	Varying penalty factors for RL agent$_1$
RL_A2_S	Varying training steps for RL agent$_2$
RL_A2_P	Varying penalty factors for RL agent$_2$
RL_A3_S	Varying training steps for RL agent$_3$
RL_A3_P	Varying penalty factors for RL agent$_3$

Under the fixed action variables, i.e., the output parameters are determined, the case that causes the maximum reward from the variations of the first two factors is selected, and the optimized agents are then obtained.

Validation of the RL agent in adaptive controller parameterization

To validate the improvement of the system stability with RL, the three optimized RL agent$_i$, i =1,2,3, are implemented into the investigated system. The following test scenarios are drafted to compare the system behaviour, including the measured power curve and the value of the SI under different operating points between three RL agents and the general controller parameterization methods. The initial system states are listed below.

- Active power of three converters: $P_{\mathrm{CON}i} = 3$ MW, i=1, 2, 3
- Power factor of three converters: $PF_{\mathrm{CON}i} = 0.9$, i=1, 2, 3
- Voltage source: $V_{\mathrm{s}} = 1$ p.u.
- Load level: 50%

The results for the following test scenarios in Tab. 4.9 will be shown in section 5.4.

Tab. 4.9 Scenarios for validation of the RL agent

Scenario abbreviation	Description
RL_SV	Varying source voltage
RL_P	Varying active power feeding from a selected CON
RL_Q	Varying reactive power feeding from a selected CON
RL_P_T	Varying active power feeding from three converters
RL_L	Varying loading conditions

5 Discussion of the simulation results

In this section, the simulation results of the DT approach, the ANN state estimation, the automatic stability testing, as well as the stabilization using the adaptive control approach are shown.

5.1 Results of Digital twin approach

Before conducting the parameter estimation, the voltage amplitude and phase on the bus 2, 5, 8, 10, and 13 from 0-8 h during a day, which are obtained from the model with initial network parameters, are compared with the reference measurement data, see Fig. 5.1 and Fig. 5.2, respectively. The initial parameters are listed in Tab. A.1. The initial capacitances of the overhead lines and underground cables as well as the filter of the converters are set to 2 times the reference parameters, while their inductances are set to 0.5 times and the resistances are set to 0.1 times. In addition to these parameters, the controller parameters of the voltage, current and reactive power of the individual converters are also the subject of parameter estimation. The reference parameters of the corresponding controllers are identical for all three converters. For the proportional parameter K_{pi} of the inner loop current controllers, the integral parameters K_{iu} and K_{iQ} of the outer loop voltage and reactive power controllers, their initial values are set to twice the reference parameter. The integral parameter K_{ii} of the current controller and the proportional parameter K_{pu} of the voltage controller are set to 0.5 times the corresponding reference parameters.

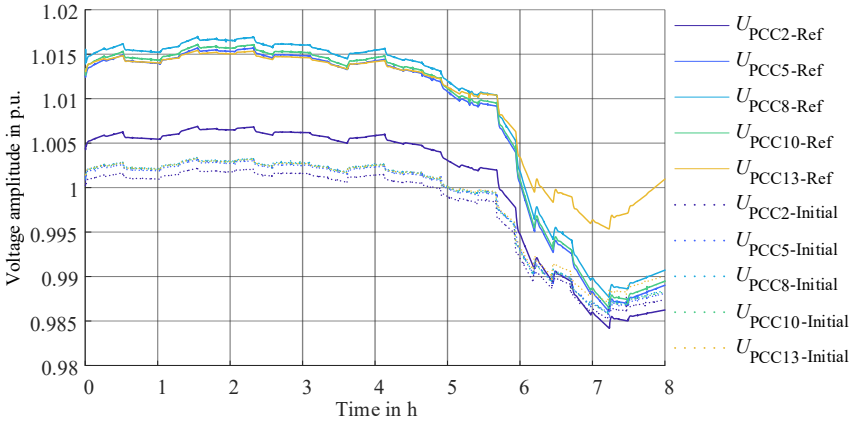


Fig. 5.1: Voltage amplitude comparison between reference network and DT model with initial parameters

As the initial values of the resistance and reactance of the lines and cables are set small, it can be seen that the differences between the voltages of the five observed buses with initial parameters are very minor. Comparing the initial values of the voltage amplitude on each bus with the reference values, the largest difference of approximately 1% occurs on bus 8. The mean absolute percentage error MAPE value between the reference and initial values was calculated for the five buses for the observed timeslot and is found to be 0.98%. Here the data are taken at a sampling time of 15 min.

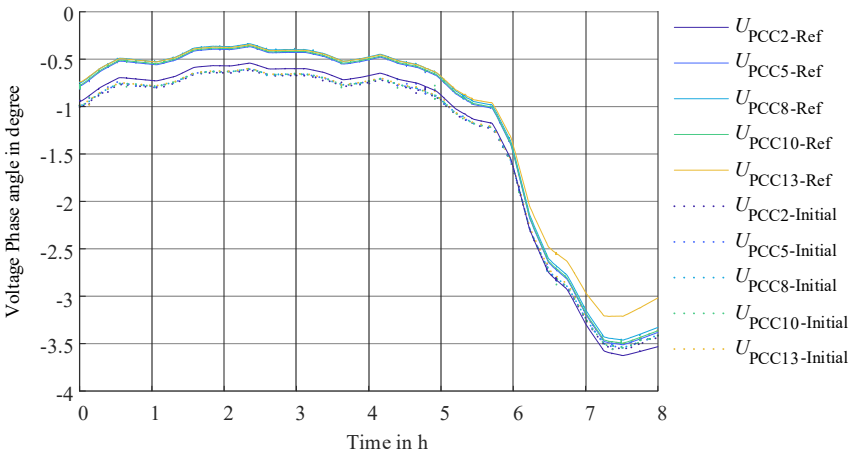


Fig. 5.2: Voltage phase angle comparison between reference network and DT model with initial parameters

Similarly, the voltage phase angle difference was calculated and the MAPE was found to be equal to 23%. The difference between the model used and the reference model is large before parameter estimation is carried out, in which case the model cannot be used as a DT.

5.1.1 Results of parameter estimation

As mentioned before, the time required for parameter estimation increases due to the large number of measurement data sets used as reference values. In order to speed up the estimation process, only the voltage in terms of its magnitude and phase angle on one bus is chosen as reference data.

- ***Scenario PE_MPL***

The measured values collected from different measurement points under different network conditions have various effects on the parameter estimation accuracy. The following starts with the factor of measurement point location. The figure below Fig. 5.3 shows the change of the factor of each parameter during the iterative process when using the voltage data of different buses as reference data.
Termination of the parameter update is when the selected training data have a MAPE value of less than 5e-3 for the voltage between the DT model and the reference model. The number of iterations required for parameter estimation is variable when different reference values are selected; for example, data from bus 8 requires only 7 iterations, while from bus13 requires 21 iterations, see Fig. 5.3. A high number of iterations corresponds to a long estimation time.

The results of the five sets of parameter estimation are used and the corresponding DT models are verified with the operation from 8 to 24 h in Fig. 4.4. The total MAPE between the voltage amplitude and phase angle of the five buses in the reference and DT models are calculated, and the results are presented in the table Tab. 5.1.

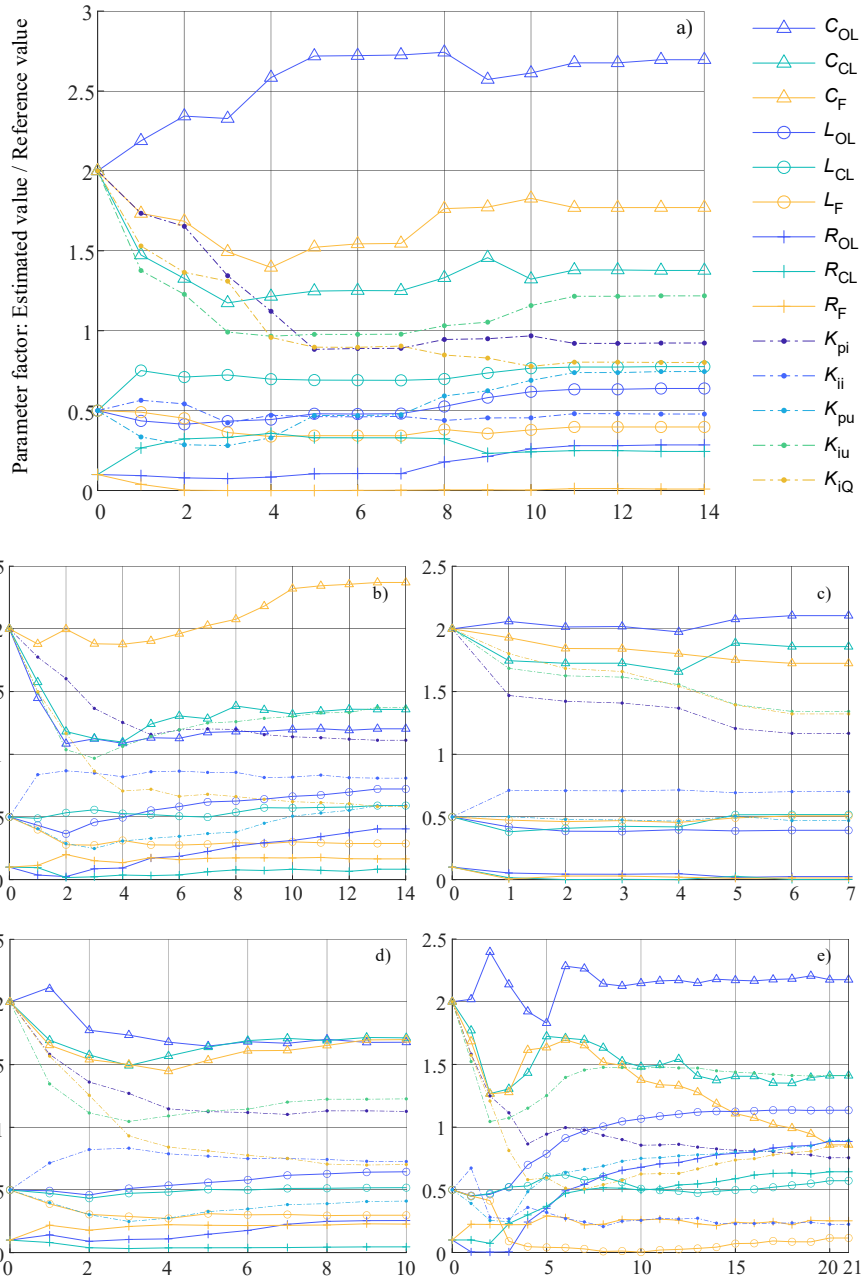


Fig. 5.3: Change of estimated parameters at each iteration step with training using data from a) bus 2; b) bus 5; c) bus 8; d) bus 10; e) bus 13

Tab. 5.1 Estimation effort and accuracy comparison with training voltage data of different buses

	Bus 2	Bus 5	Bus 8	Bus 10	Bus 13
Iteration steps	14	14	7	11	21
Convergence duration (min)	101	116	79	87	136
MAPE (%)	1.46	2.42	4.71	3.54	0.79
Accuracy (%)	98.54	97.58	95.29	96.46	99.21

The results show that the accuracy of the DT model derived from the parameter estimation using the reference data of bus 2 and 13 is higher. On the one hand, it may be related to the fact that these two buses are closer to the slack bus, and on the other hand, it may be related to the lower loads to which they are connected.

The parameter errors are listed in Tab. 5.2 are calculated using the factors of each parameter in the final iteration step minus one, whereas “one” means that the estimated parameter is the same as its reference value.

Tab. 5.2 Parameter errors with using training voltage data at different buses

Parameter	**Bus2**	**Bus5**	**Bus8**	**Bus10**	**Bus13**
Cable resistance R_{CL}	1.70	0.20	1.11	0.68	1.21
Cable inductance L_{CL}	0.38	0.36	0.86	0.71	0.40
Cable capacitance C_{CL}	0.77	1.37	0.72	0.70	-0.05
Overhead line resistance R_{OL}	-0.36	-0.28	-0.61	-0.35	0.13
Overhead line inductance L_{OL}	-0.22	-0.41	-0.48	-0.48	-0.45
Overhead line capacitance C_{OL}	-0.60	-0.71	-0.49	-0.70	-0.92
Filter resistance R_{OL}	-0.71	-0.6	-0.98	-0.74	-0.15
Filter inductance L_{OL}	-0.76	-0.92	-1.00	-0.95	-0.37
Filter capacitance C_{OL}	-0.99	-0.83	-0.99	-0.77	-0.74
Current controller K_{pi}	-0.07	0.11	0.16	0.13	-0.05
Current controller K_{ii}	-0.52	-0.19	-0.30	-0.27	-0.76
Voltage controller K_{pu}	-0.26	-0.41	-0.53	-0.59	-0.15
Voltage controller K_{iu}	0.22	0.37	0.34	0.23	0.41
Reactive power controller K_{iQ}	-0.20	-0.42	0.32	-0.30	-0.19

Fig. 5.4 visually compares the errors of the estimation results for the 14 parameters. Fig. 5.4 a) shows the error for each parameter individually, under training data with the different bus. It can be seen that the errors of the five controller parameters of converters are small compared to the errors of the parameters of the network transmission lines. Among them, the current controller parameter K_{pi} has the smallest parameter error. It can be understood that the DT model will be

more accurate when the errors of the control parameters are minor. Fig. 5.4 b) shows the sum of the errors for each parameter with the same bus training data. The absolute values of the errors are used here to prevent the positive and negative errors from eliminating each other. The comparison yields that the smallest sum of errors is caused by bus 13.

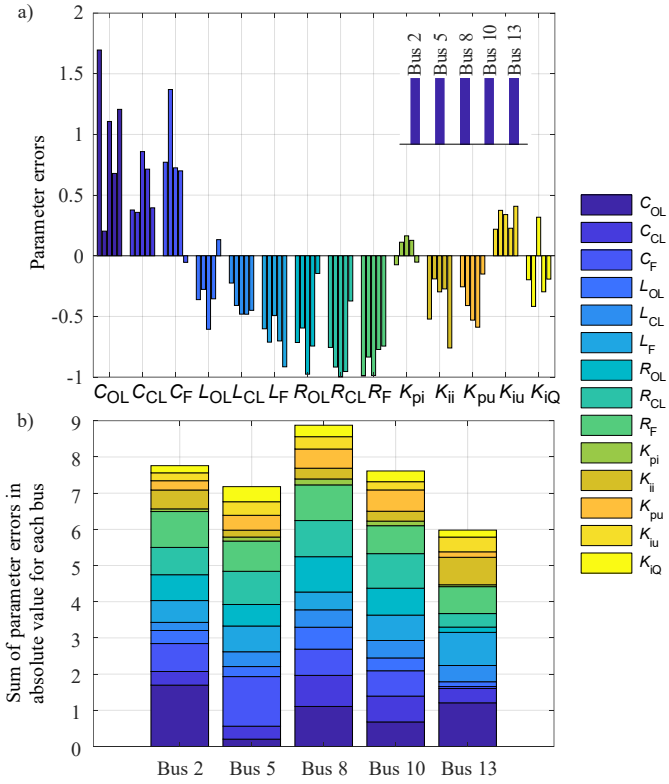


Fig. 5.4: Parameter errors comparison with different buses training data: a) Error in real value for each parameter b) Sum of errors in absolute value

Since the model with the least error is obtained using the data from bus 13, this data set will be applied in the following research. To investigate which network state measurement data as training data can obtain the DT model with high accuracy, different scenarios are set up below to collect data and the parameter estimation is then performed with these data.

- ***Scenario PE_NS_L***

The loads in the investigated network are lowed to 0.1 times and raised to 10 times the reference value and it is used as low and high load cases for collecting training data, respectively. The following Fig. 5.5 shows the changes in the estimated parameters during the training process for both cases. The parameter estimates using the low load data resulted in five parameters falling between 0.9 and 1.1 of their corresponding reference values, while only two parameters using the high load fall within this range.

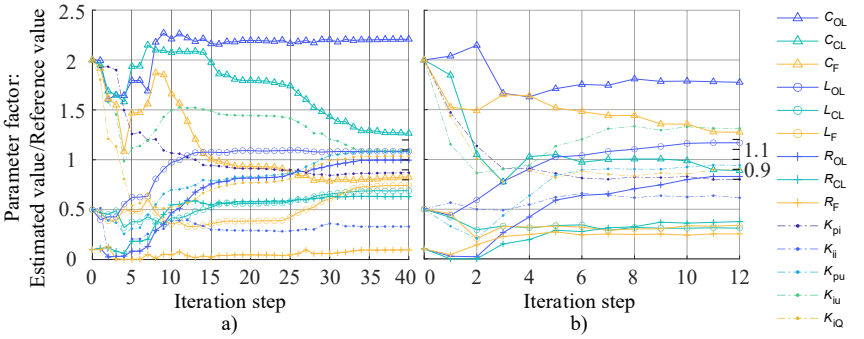


Fig. 5.5: Change of estimated parameters at each iteration step using training data with a) low load b) high load

The results of both sets of parameter estimation are utilized, and the corresponding DT models are verified with the operation from 8 to 24 h in Fig. 4.4 with low and high load, respectively. The total MAPE between the voltage amplitude and phase angle of the five buses in the reference and DT models are computed, and the results are presented in the following table Tab. 5.3. For a better comparison, the table also includes the case when the load uses the reference value.

Tab. 5.3 Effort comparison using training voltage data from bus 13 with low and high load

Load factor	0.1	1	10
Iteration steps	40	21	12
Convergence duration (min)	228	136	85
MAPE (%)	0.15	0.79	1.12
Accuracy (%)	99.85	99.21	98.88

The results show that the accuracy of the DT model derived from the parameter estimation using the lower load is higher. It is the same reason postulated in the previous section when using five sets of bus data, i.e., the network is in a low load state, and its measurement data can be utilized to obtain a more accurate DT model.

- ***Scenario PE_NS_CON***

In the following, the training data obtained from bus 13 connected and unconnected converter are compared. Parameter estimation using the data with converter has five parameters falling between 0.8 and 1.2 of their corresponding reference values, while without converter, only two parameters land within this range, see Fig. 5.6.

To compare the accuracy of the reference and DT models, the total MAPE between the voltage amplitude and phase angle of the five buses are computed. The results are presented in the following table. The training data obtained using converter produces a more accurate DT model compared to the case of training without using converter, see Tab. 5.4.

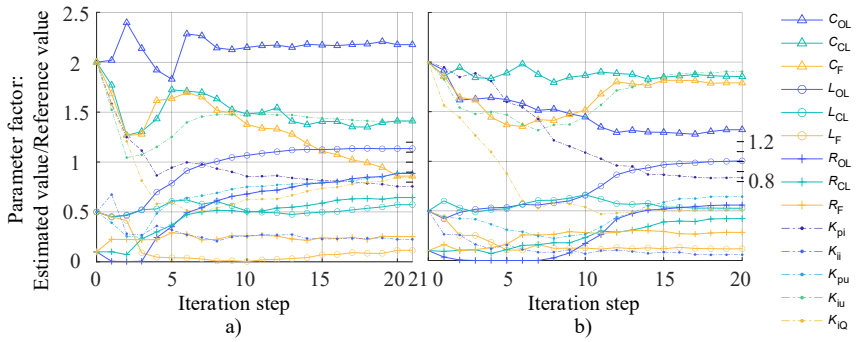


Fig. 5.6: Change of estimated parameters at each iteration step using training data from bus 13 connected a) with converter b) without converter

Tab. 5.4 Effort comparison using training voltage data at bus 13 with and without converter

CON	With CON	Without CON
Iteration steps	21	20
Convergence duration (min)	136	179
MAPE (%)	0.79	1.40
Accuracy (%)	99.21	98.60

- ***Scenario PE_NS_G***

The generated power of the three converters is set to 0.5 times and 1.5 times the reference value, and the training data are collected as low and high-generation cases, respectively.

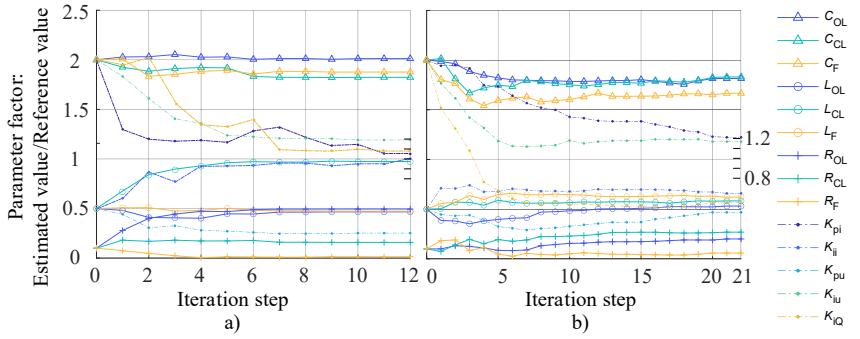


Fig. 5.7: Change of estimated parameters at each iteration step using training data with a) low generation b) high generation from three converters

The figure above shows the increase in P and Q bothFig. 5.7 shows the changes in the estimated parameters during the training process for both cases. The parameter estimates using the low generation resulted in five parameters falling between 0.8 and 1.2 of their corresponding reference values, while only one parameter using the high generation fell within this range.

Tab. 5.5 Effort comparison using training voltage data at bus 13 with low and high generation from three converters

Generation power factor	0.5	1	1.5
Iteration steps	12	21	21
Convergence duration (min)	125	136	187
MAPE (%)	0.14	0.79	1.02
Accuracy (%)	99.86	99.21	98.98

The results of both sets of parameter estimation are utilized, and the corresponding DT models are verified with the operation from 8 to 24 h in Fig. 4.4 with the low and high generation, respectively. The total MAPE between the voltage amplitude and phase angle of the five buses in the reference and DT models are computed, and the results are presented in the table above Tab. 5.5. For a better comparison, the table also includes the case when the generation uses the reference value.

- ***Summary and conclusion***

Three scenarios are set up to analyze the accuracy of the DT model when the training data for parameter estimation are sourced from different network states. According to the results, the smaller the load connected to the network, with the converter connected to the bus and the smaller the power generated by the converter, the more suitable the measurement data source for parameter estimation. The mentioned network states are, in fact, of a high degree of stability. In addition, the closer a bus is to the slack bus, the more accurate the DT model will be when the voltage data on the bus are used as training data for estimation. By taking advantage of these conditions that improve the model's accuracy, the resulting model is applied as the DT model for the reference network in this work and for subsequent stability studies. Before proceeding with further studies, the accuracy of the DT model should be verified; see the next sub-section 5.1.2.

5.1.2 Verification of the DT model

In the following, the DT model is evaluated in four different scenarios described in sub-section 4.1. Furthermore, four sub-scenarios within each scenario are set up to investigate the accuracy of the DT model for different combinations of levels of power generation and load. It is defined as high generation and heavy load at 0.9 of their nominal value and low generation and low load at 0.1 of the nominal value.

- ***Scenario DT_SV***

In this case, the voltage reduces from 1 p.u. to 0.95 p.u. at 0.1 s and raises to 1.05 p.u. at 0.2 s.

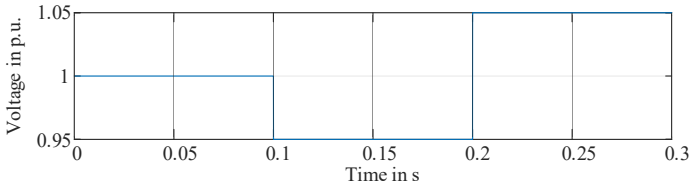


Fig. 5.8: Source voltage step changes

In response to changes in the source voltage, as shown above, the amplitude and phase angle of the five bus voltages in the DT model change, and their comparison with the corresponding reference values are shown below. Illustrated here is the case of heavy load and high power generation.

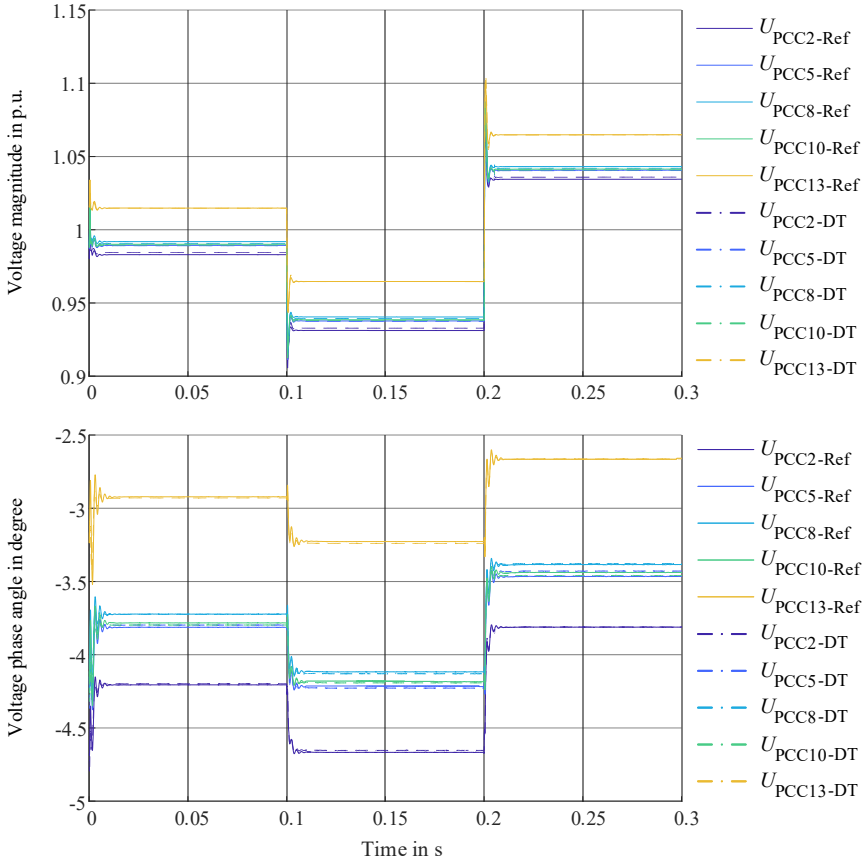


Fig. 5.9: Voltage comparison under heavy load and high generation between reference network and DT model in response to a step change of source voltage

In order to carry out a direct comparison of the errors between the DT model and the reference model for four different combinations of load and generation, the table below shows the MAPE values for each case.

Tab. 5.6 Accuracy comparison with different generation and load level

Generation leveling	0.9	0.9	0.1	0.1
Load level	0.9	0.1	0.9	0.1
MAPE (%)	0.66	0.45	0.39	0.05
Accuracy (%)	99.34	99.55	99.61	99.95

According to the table above Tab. 5.6, the MAPE is less than 1% for all four scenarios, which means that the DT models are all above 99% accurate for the response of source voltage step change. Since the DT model with the highest accuracy results from the model training under low load and low generation, its accuracy under the same configuration also presents the highest values in this validation scenario.

- ***Scenario DT_PQ***

The active power P and the reactive power Q of converter CON3 connected to bus14 change as shown below. P_{CON3} rises from 2 to 4 MW at 0.1 s and Q_{CON3} rises from 1.5 to 2.5 MVar at 0.2 s.

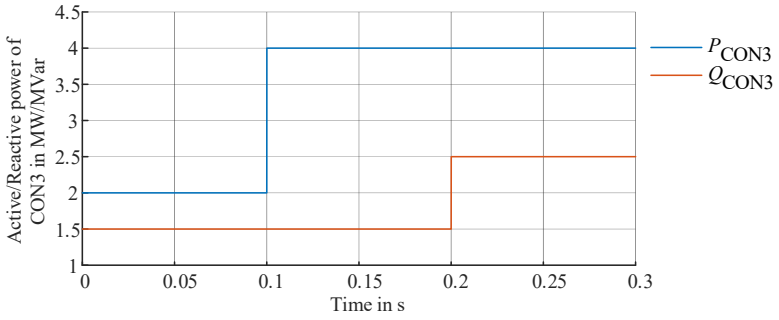


Fig. 5.10: Active and reactive power step change

The changes of the amplitude and phase angle of the five bus voltages in the DT model and their comparison with the corresponding reference values are shown in Fig. 5.11 in response to changes of an active and reactive power step change in CON3. The DT model is also evaluated in four combinations of levels of power generation and load. The power generation, in this case, excludes the generation of CON3. Fig. 5.11 illustrates the case of heavy load and high generation.

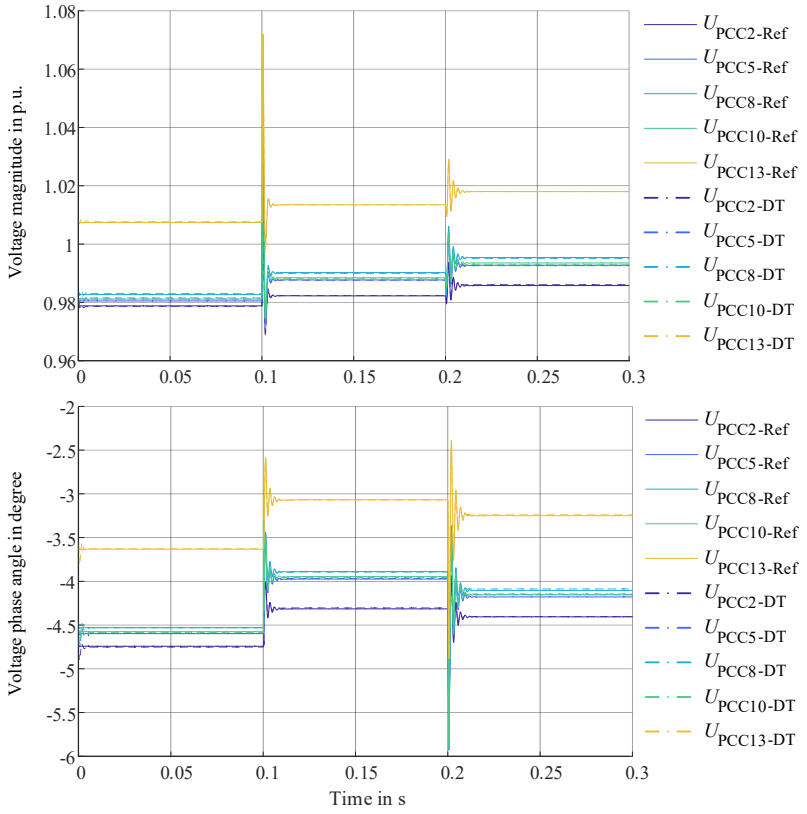


Fig. 5.11: Voltage comparison under heavy load and high generation between reference network and DT model in response to a step change of active and reactive power

The increase in P and Q both result in an increase in voltage amplitudes. Compared to $t = 0.1$ s, the period of the system oscillation increases at 0.2 s when Q increases. Since the increase in Q is based on the increase in P, the system stability is weaker at 0.2 s, so the time to reach a steady state is longer.

Tab. 5.7 Accuracy comparison with different generation and load level

Generation leveling	0.9	0.9	0.1	0.1
Load level	0.9	0.1	0.9	0.1
Error of validation (%)	0.59	0.43	1.63	0.33
Accuracy (%)	99.41	99.57	98.37	99.67

Tab. 5.7 shows the MAPE is less than 1% for the scenarios except for the case with 0.1 generation leveling and 0.9 load level, which has an error of 1.63% for the response of active and reactive power step change. Here, the accuracy under the low load and low generation scenario again shows the highest value.

- ***Scenario DT_OC***

The line between bus 8 and 14 opens at 0.1 s and closes at 0.2 s. Fig. 5.12 presents the changes in the amplitude and phase angle of the five bus voltages in the DT and the reference model in response to line switch status changes under heavy load and high generation.

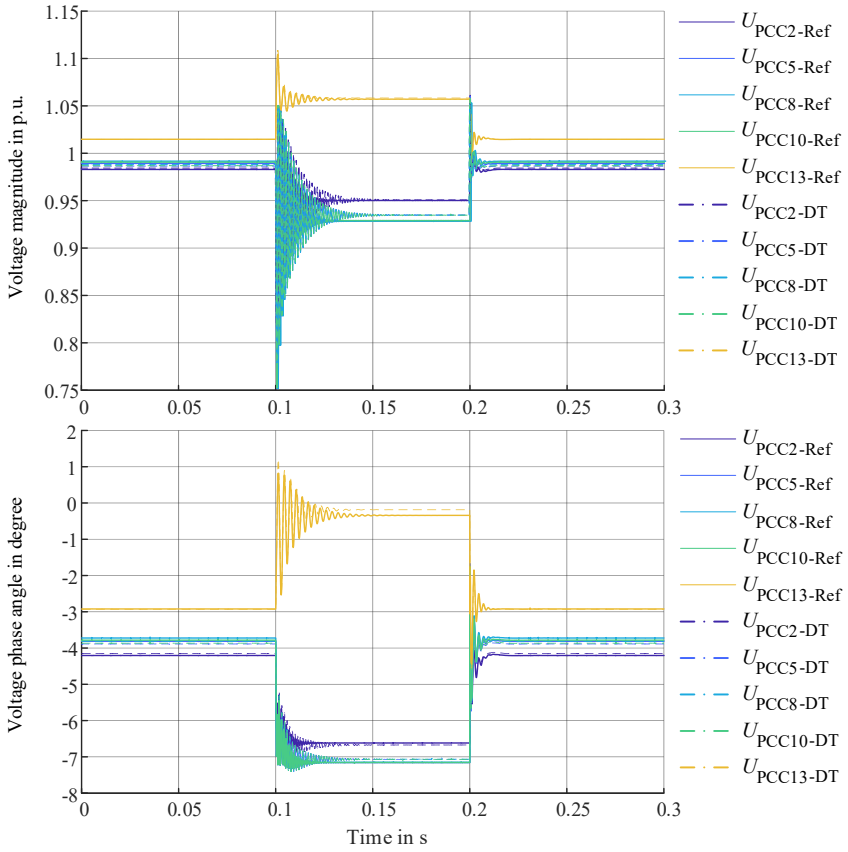


Fig. 5.12: Voltage comparison under heavy load and high generation between reference network and DT model in response to line switch status changes

The DT model is again evaluated in four combinations of levels of power generation and load. Tab. 5.8 lists the error of validation for four cases, and the maximum value of 3.17% occurs in heavy load and high generation.

Tab. 5.8 Accuracy comparison with different generation and load level

Generation leveling	0.9	0.9	0.1	0.1
Load level	0.9	0.1	0.9	0.1
Error of validation (%)	3.17	0.54	2.32	0.43
Accuracy (%)	96.83	99.46	97.68	99.57

- ***Scenario DT_SC***

A three-phase short circuit appears on bus 14 at 0.1 s and is cleared at 0.2 s. Due to the short circuit, the voltage amplitude drops significantly, see Fig. 5.13 below. Bus 2 has a smaller voltage drop than the other buses because of its spatial proximity to the upper layer high voltage network. However, bus 13, which is also close to the 110 kV network, has a more severe voltage drop than bus 2, which is directly connected to bus 14.

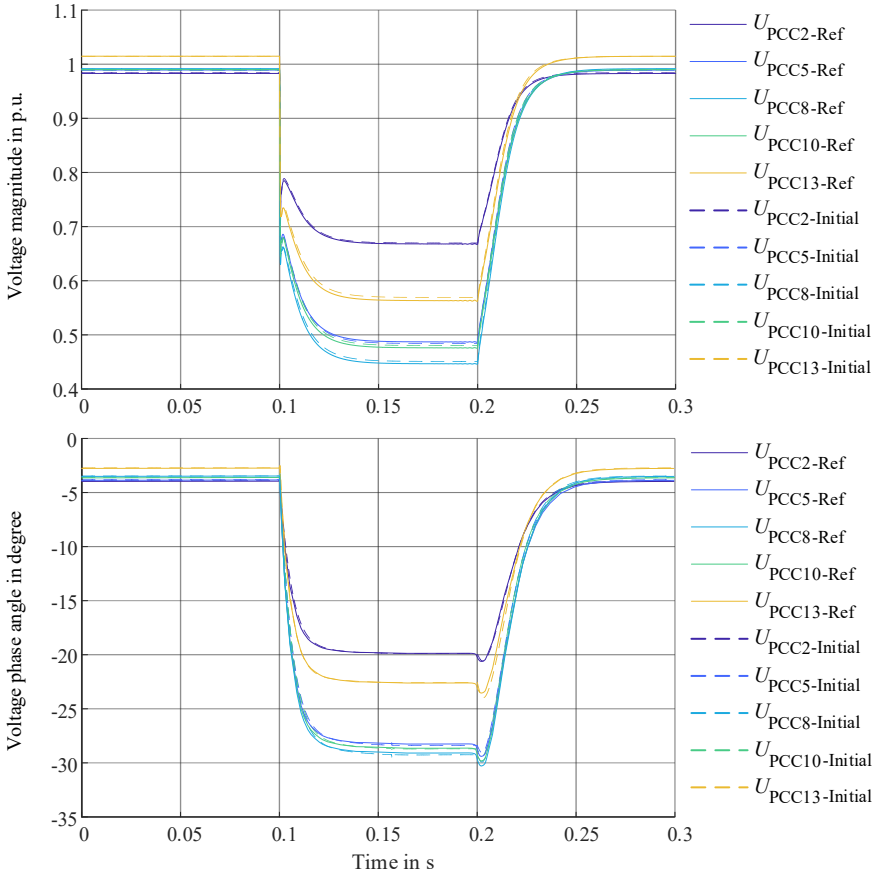


Fig. 5.13: Voltage comparison under heavy load and high generation between reference network and DT model in response to short circuit on bus 14

Tab. 5.9 Accuracy comparison with different generation and load level

Generation leveling	0.9	0.9	0.1	0.1
Load level	0.9	0.1	0.9	0.1
Error of validation (%)	2.20	1.87	1.07	0.98
Accuracy (%)	97.80	98.13	98.93	99.02

In this scenario, a maximum error value of 2.20% occurs in generation leveling of 0.9 and load level of 0.9, see Tab. 5.9. The highest accuracy value is again shown here for low load and low generation cases.

- ***Summary for the DT validation***

The DT model has been validated using the four scenarios mentioned above, i.e., step change of source voltage, active and reactive power, open circuit on the line between bus 8 and 14, and short circuit on bus 14. For each scenario, different combinations of leveling of power generation and load are set up. The amplitude and phase angle of the five bus voltages for each scenario are plotted and compared with the reference model. In addition, the corresponding error MAPE values are calculated and tabulated for comparison. In the case of low load and low generation, the DT outperforms the other combinations with the accuracy of above 99%. The minimum accuracy of 96.83% occurs in the open circuit on the line with heavy load and high generation.

5.2 Results of ANN state estimation

The number of neurons in the hidden layer of each ANN varies from 5 to 15, which are considered adequate settings in terms of training speed and convergence characteristics based on multiple training sessions. The correlation coefficient R is chosen as an indicator of ANN fitting degree to determine the optimal number of neurons for each of the five ANNs. The R, which ranges from -1 to 1, shows the correlation between the given values Y for training and the predicted values $\hat{Y}$ obtained from the ANN model. For R =1 and R = -1, the ANN model shows the strongest positive and negative fitting degrees, respectively. A coefficient R of zero indicates a random relationship between the values mentioned above.

$$R = \frac{\operatorname{cov}\left(Y, \hat{Y}\right)}{\sqrt{\operatorname{var}\left(Y\right)}\sqrt{\operatorname{var}\left(\hat{Y}\right)}} = \frac{\sum_{k=1}^{n}\left(Y_k - \overline{Y}\right)\left(\hat{Y}_k - \overline{\hat{Y}}\right)}{\sqrt{\sum_{k=1}^{n}\left(Y_k - \overline{Y}\right)^2}\sqrt{\sum_{k=1}^{n}\left(\hat{Y}_k - \overline{\hat{Y}}\right)^2}} \tag{5.1}$$

where cov(Y, $\hat{Y}$) is the covariance of Y and $\hat{Y}$, var(Y) and var($\hat{Y}$) are the variances of Y and $\hat{Y}$, n is the sample size, Y_k and $\hat{Y}_k$ are the k-th sample points, $\overline{Y}$ and $\overline{\hat{Y}}$ are the means of variables Y and $\hat{Y}$, respectively.

- ***Scenario SE_HLN_ANN***

The Fig. 5.14 below shows the change in R-value for each ANN as their number of neurons increases. For ANN1, the maximum R-value is obtained at 9 neurons. For ANN2, it is 10. For ANN3, 4, and 5, the maximum R-values are obtained with 11 neurons. Since the R values here are all greater than 0, the larger the value, the greater the degree of ANNs fit. In other words, the more accurate the ANNs can estimate the system states. Therefore, it is determined that for the five ANNs, the number of neurons in their hidden layers is 9, 10, 11, 11, and 11, respectively.

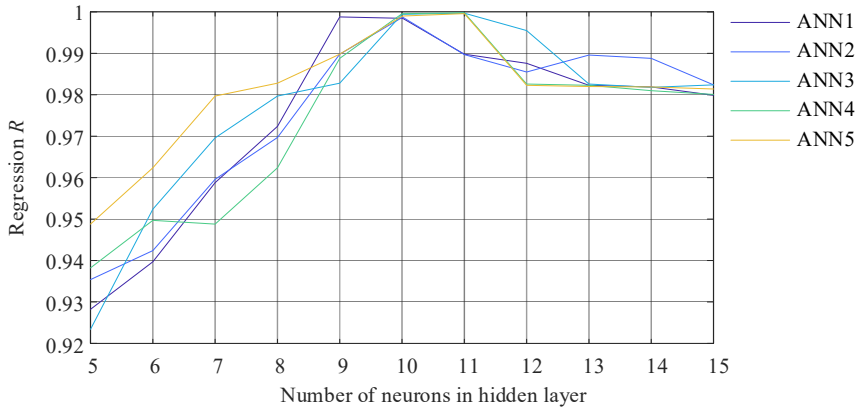


Fig. 5.14: Change in *R*-value of five ANNs with different of neurons in hidden layer

- ***Scenario SE_24h _ANN***

After training with the above setup, each ANN state estimators are verified in the following.

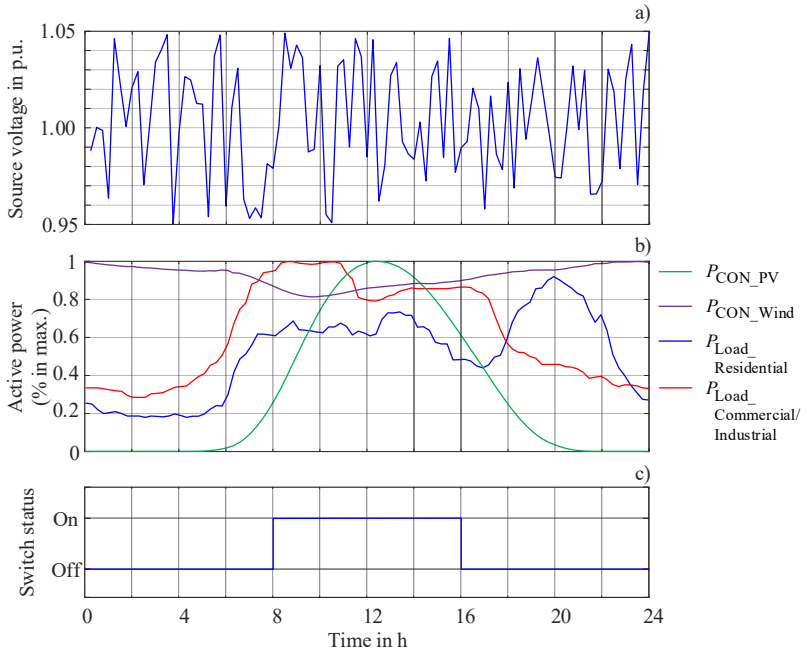


Fig. 5.15: Scenarios design for 0-24h: a) Change of source voltage b) Change of generation and consumption power c) Change of line switch status between bus 8 and 14

The data of the voltage source, the power generated and consumed on each bus during 24 hours at a measurement frequency of one data per 15 minutes, and the switching state of the transmission line are given to the ANN state estimator. Here, the voltage source varies slowly with small amplitudes between 0.95 and 1.05. The power generated by the three converters varies according to

the standard power generation of PV and wind power plants for one day, while the variation of the load power at each consumption bus is designed according to the change of the standard industrial/commercial and residential power consumption to form the typical system operating states. Besides, the line switch changes from closed to open at t=8 h and closes again at t=16 h. Fig. 5.15 shows a graphical representation of the corresponding scenario.

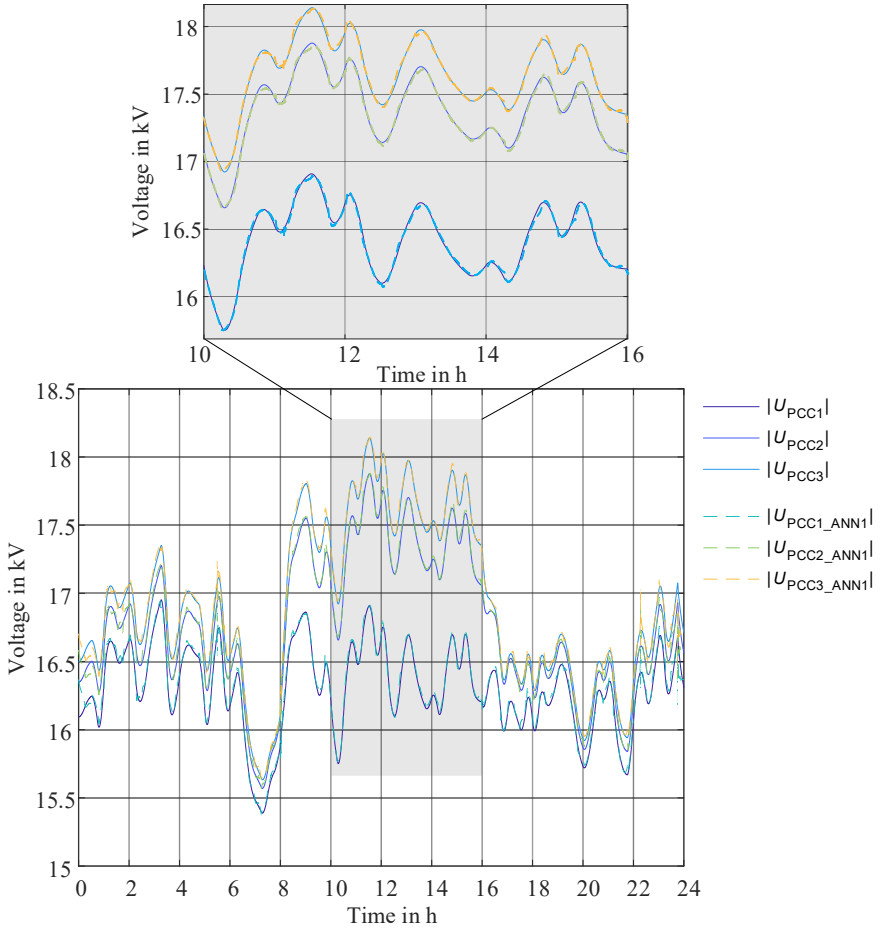


Fig. 5.16: Comparison of estimated voltage using standard load flow program and ANN1

The estimation results of each of the five ANNs during the 24 hours of the test scenario are shown below and compared with the results of a standard load flow program. Fig. 5.16 shows the results of the ANN1, which is used to estimate the voltage magnitude at the PCCs of the three connected converters. The three PCC voltages have the same trend and are most affected by the variation of the source voltage, which means that they trend along with the changes in the source voltage. The solid line shows the result using the standard power flow calculation, while the dashed line shows the result of the ANN1 estimation. As can be seen from the partially zoomed-in view, the differences between the two corresponding voltages are small.

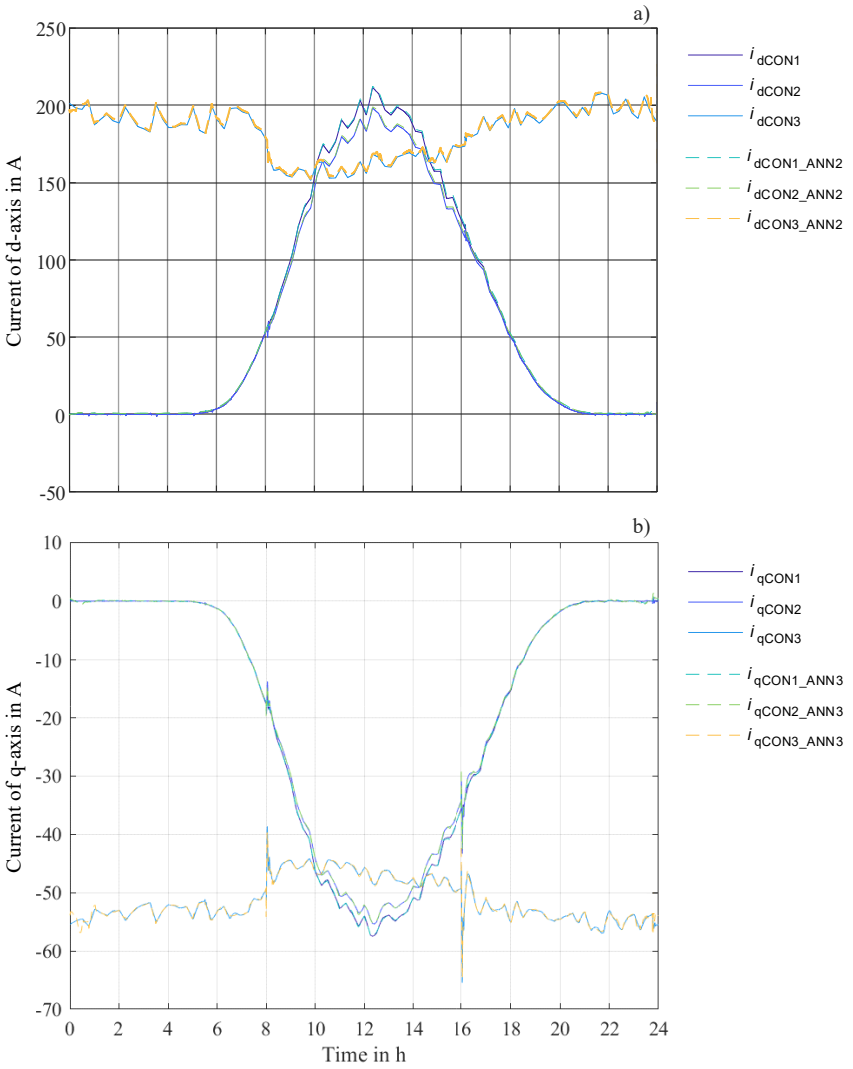


Fig. 5.17: Comparison of estimated current of: a) d-axis using standard load flow program and ANN2; b) q-axis using standard load flow program and ANN3

The trends of i_{d} and i_{q} of three converters in the absolute values are consistent with the corresponding generating active power P_{CON}. Since converter CON1 and CON2 are driven by PV, their currents are close to zero from 0 to 6 h and 20 to 24 h. The current of the wind-driven CON3 is influenced by wind strength. From the figure above, the estimated results from the ANN2 and ANN3 are also remarkably similar to the results of the standard load flow program calculations.

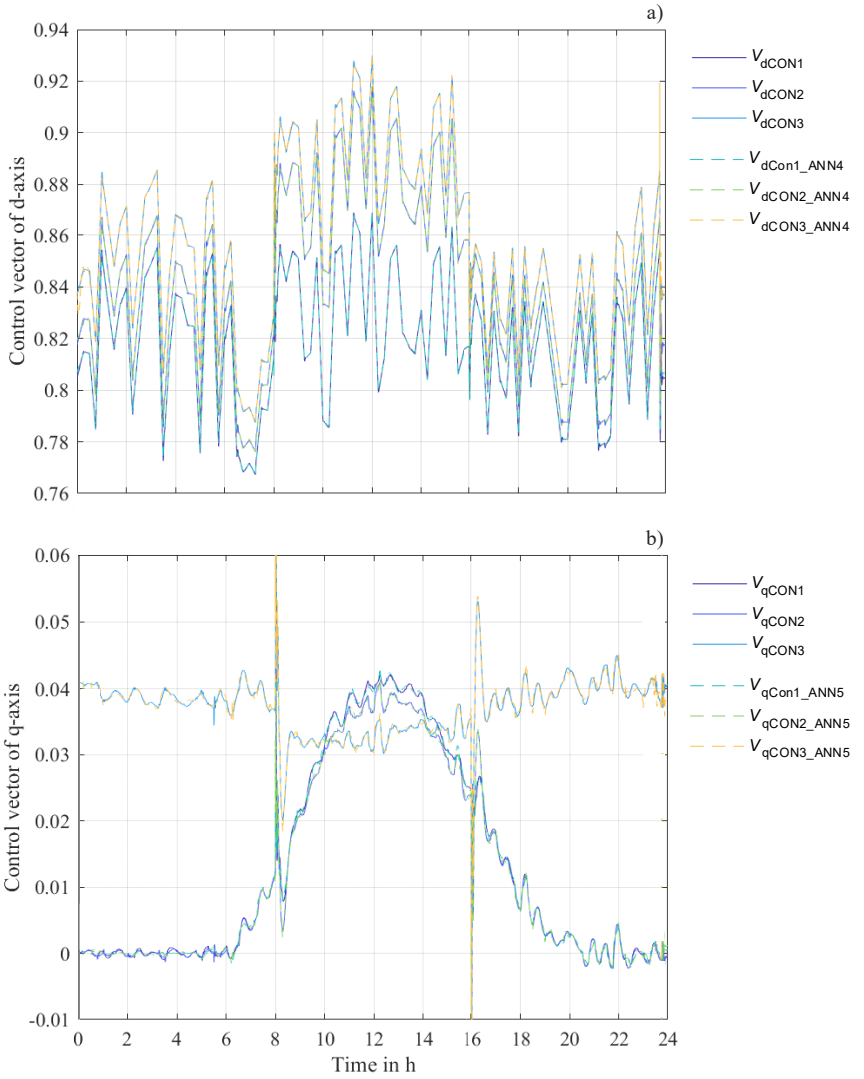


Fig. 5.18: Comparison of estimated control vector of: a) d-axis using standard load flow program and ANN4; b) q-axis using standard load flow program and ANN5

ANN4 and ANN5 are respectively utilized to estimate the control vectors of the d-axis and the q-axis of the three converters. The control vectors of the d-axis follow the same trend as the voltages of the PCCs where the corresponding converters are located, i.e., they are mainly influenced by the source voltage. On the other hand, the control vectors of the q-axis are more influenced by the generating active powers P of the corresponding converters. Fig. 5.18 b) illustrates the error in ANN5's estimation result of v_q for the period 0-6 h with respect to the results of the power flow calculations. Since its error estimates are on a scale of 0.001, they have only a small impact on the subsequent estimation of the SIs.

- ***Summary for the ANN state estimation***

During the training of the five ANNs, various network states and different network structures were set up. After extensive training, the five ANNs can accurately estimate the voltage, current and control vectors for each of the three converters. The resulting ANNs-based state estimators are then used to compute and monitor the network stability in real-time.

5.3 Results of automatic stability testing

This section shows how the SI, namely the damping ratios of the dominant eigenvalues in the system, change when varying the active and reactive power feeding of converter CON1 and the source voltage supplying as well as the loading conditions, respectively. The initial conditions for the testing are shown as follows.

- Active power of three converters: $P_{CONi} = 3$ MW, i=1, 2, 3
- Power factor of three converters: $PF_{CONi} = 0.95$, i=1, 2, 3
- Voltage source: $V_s = 1$ p.u.
- Load level: 50%

To calculate the SIs, the values of the real and imaginary parts of the dominant eigenvalues are required, as listed in the following Tab. 5.10. Along with these values, the SI can be calculated using the equation (2.64).

Tab. 5.10 System dominant eigenvalue under different network states

***Active power P* [MW]**		**1**	**2**	**3**	**4**	**5**
Dominant eigenvalue	**Real part**	-468.9	-432.7	-396.5	-345.6	-296.9
	Imaginary part	2488.4	2484.2	2486.7	2497.8	2519.3
Power factor PF		**0.84**	**0.88**	**0.92**	**0.96**	**1.00**
Dominant eigenvalue	**Real part**	-282.7	-326.9	-366.2	-398.2	-442.2
	Imaginary part	2507.8	2502.5	2495.1	2483.0	2441.2
***Vsource* V_s [p.u.]**		**0.9**	**0.95**	**1**	**1.05**	**1.1**
Dominant eigenvalue	**Real part**	-312.6	-355.0	-396.5	-422.3	-449.1
	Imaginary part	2345.7	2418.0	2486.7	2552.5	2615.6
Load level		**0.9**	**0.95**	**1**	**1.05**	**1.1**
Dominant eigenvalue	**Real part**	-472.5	-478.6	-483.4	-490.6	-498.1
	Imaginary part	2471.7	2498.9	2513.0	2535.2	2559.4

Another focus of this section is on demonstrating the accuracy of the ANN-based SI estimator for the four designed scenarios.

- ***Scenario SI_P:*** Different active power feeding from converter CON1: $P_{\text{converter CON1}}$ = 1, 2, 3, 4, 5 MW
- ***Scenario SI_Q:*** Different reactive power feeding from converter CON1, using different power factors at $P_{\text{converter CON1}}$ = 3 MW to calculate $Q_{\text{converter CON1}}$: $PF_{\text{converter CON1}}$ = 0.84, 0.88, 0.92, 0.96, 1.00

- ***Scenario SI_SV:*** Different voltage source supplying: V_s = 0.9, 0.95, 1.0, 1.05, 1.1 p.u.
- ***Scenario SI_L:*** Different loading conditions: Load level = 20%, 40%, 60%, 80%, 100%

The estimated results and the calculated reference values under varying network conditions in the four scenarios are shown in Fig. 5.19.

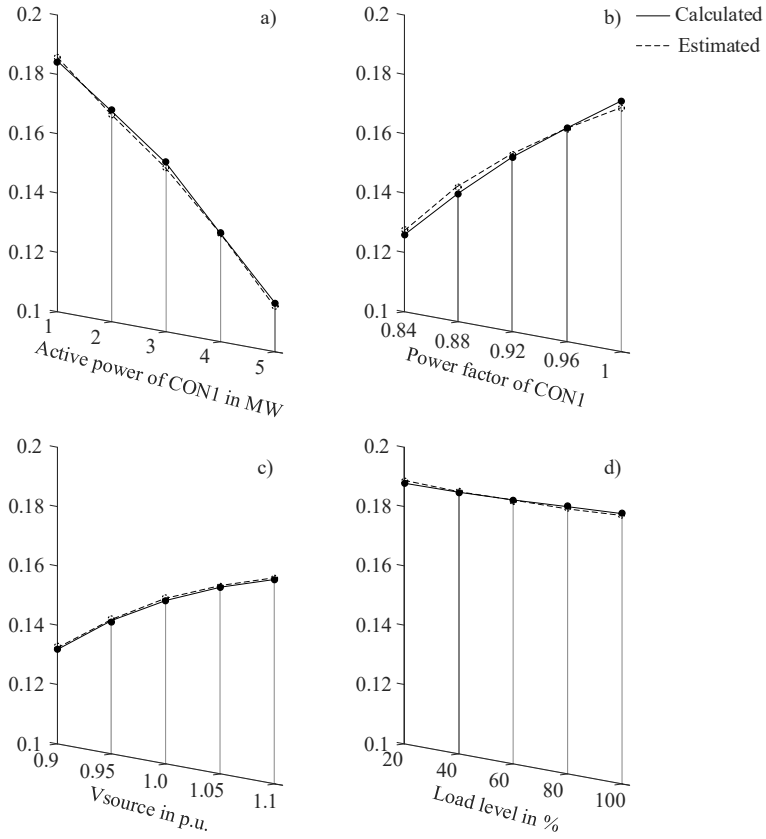


Fig. 5.19: Trajectory of SI for changes of a) Active power of converter CON1 b) Power factor of converter CON1 c) Voltage source d) Load level

It can be seen that the estimated SI is very close to the calculated value. Here, *MAPE* is used to measure the accuracy of the ANN-based SI estimator with the following equation.

$$MAPE = \frac{100\%}{n} \sum_{k=1}^{n} \left| \frac{C_k - E_k}{C_k} \right| \tag{5.2}$$

where C_k is the calculated value of the SI and E_k is the corresponding estimated value. n is the number of validation scenarios, which here will be 20.

A *MAPE* value of 1.05% is calculated, giving the SI estimator an accuracy of 98.95%. This estimator serves as an accurate and fast indication of system stability for the RL-based controller parameterization, which goal is stabilizing the system, moreover, improving its stability.

5.4 Results of stabilization using adaptive control approach

It is the key content of the adaptive control approach to train the RL agent with the resulting SI from the last chapter as a reward indicator, which enables it to optimize the stability of the system by adjusting the parameters of the converter controllers according to the system operating states.

As introduced in section 4.4, this section focuses on demonstrating the variation of rewards for three RL agents under different training settings. In addition, the agents with the maximum reward are used for validation in five different scenarios listed in section 4.4 and their system stability is compared to that under the currently prevailing controller tuning method.

Variation of rewards for three RL agents

- ***Scenario RL_A1_S and Scenario RL_A1_P***

The variation in the rewards obtained by the RL agent$_1$ for different training steps and different settings of the penalty factor is shown in Fig. 5.20.

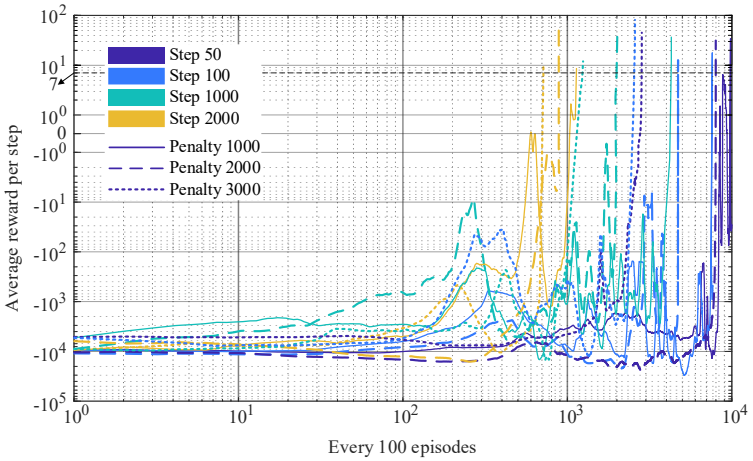


Fig. 5.20: Variation of rewards with different training steps and different penalty factors for RL agent$_1$

To better compare the effect of different settings on the reward, the average reward value per step for every 100 episodes during the evaluation episodes is used as the vertical coordinate. RL agent$_1$ outputs eight controller parameters, i.e., [K_{pi_CON1}, K_{pi_CON2}, K_{pu_CON2}, K_{iu_CON2}, K_{pi_CON3}, K_{pu_CON3}, K_{iu_CON3}, K_{iQ_CON3}]. Since these eight parameters have a large impact on the SI, and the reward of the RL agent is directly proportional to the SI, the agent that is randomly initialized in the early stage of training receives a huge negative reward, see Fig. 5.20. As the exploration proceeds, different training settings present effects on the effectiveness of optimization. The training of the RL agent stops automatically when the average reward per step reaches 7 for every 100 episodes. The value of 7 is calculated from equation (3.62) under the expectation that the average SI of all 100 random operating points of the system is 0.175. Since 0.175 falls in the interval (0.1, 0.2], the value of the reward according to (3.62) is calculated from 0.175*40 and results in 7.

At the beginning of the exploration, the RL agent$_1$ receives a penalty from the positive real part of the poles on the real axis, and in the middle of the exploration, a negative DR is added, followed

by a negative SI as a penalty to inform the RL agent$_1$ for the goal of the optimization. When the control parameters are adjusted so that the SI is positive, the SI is then used as a reward to continue optimizing the agent$_1$ within the selected operating point range until the cumulative reward reaches the set target value. Fig. 5.20 represents that the larger step can achieve a rewarding target value within a smaller number of episodes, see the yellow lines. For Step 2000, there are 2000 steps to optimize for an operating point so that the agent has a greater chance of getting optimized at that operating point than at Step 50. Moreover, a large penalty factor positively affects agent training efficiency. It is observed that the RL agent$_1$ at Penalty=3000 is always the first to reach the set optimization target value 7 for the same step setting.

- ***Scenario RL_A2_S and Scenario RL_A2_P***

The following shows the change in reward during the training of RL agent$_2$. Compared to RL agent$_1$, the agent$_2$ is only responsible for optimizing four control parameters. The following graph shows that the agent in this scenario achieves fewer training episodes than agent$_1$, because the number of parameters to be optimized is only half that of agent$_1$.

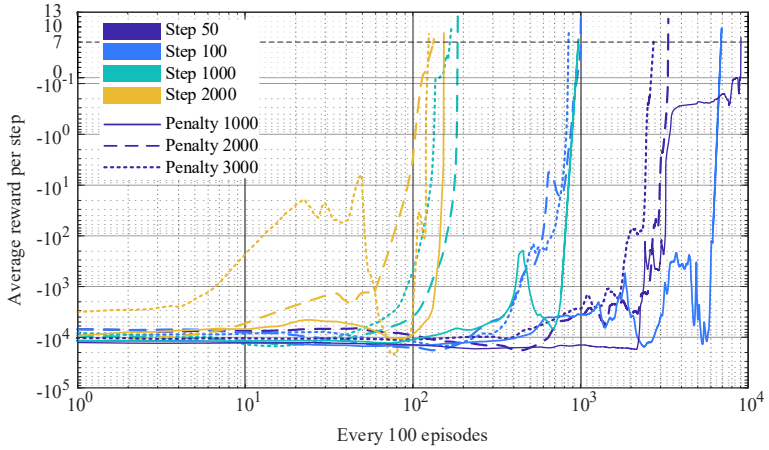


Fig. 5.21: Variation of rewards with different training steps and different penalty factors for RL agent$_2$

For the same reason, agent$_2$ achieves a smaller final reward than agent$_1$. In the previous scenario, the maximum average reward of 81.90 for RL agent$_1$ occurs at Step = 100, Penalty = 3000, whereas RL agent2 achieves the maximum value of 11.25 at Step = 50, Penalty = 2000. The different step and penalty settings have the same effect on the training efficiency as for RL agent$_1$.

- ***Scenario RL_A3_S and Scenario RL_A3_P***

The RL agent$_3$ outputs only two controller parameters i.e. K_{pi_CON1} and K_{pi_CON2}, while all the other controller parameters are set using the conventional controller configuration methods. This means that RL training will reach the predetermined reward target value 7 faster since there are only two uncertain parameters. However, the RL agent has little potential to optimize the system stability in this case. From the results, the maximum average reward value for different training settings does not exceed 10. A greater penalty is not useful for better training efficiency due to the small optimization space. Moreover, at a small step value, i.e. Step = 50, the rewards at different penal-

ties almost overlap, see purple lines in Fig. 5.22. Only at a larger step, e.g. Step = 2000, the exploration of parameter optimization by the RL agent for the same operating point is reflected, see three yellow lines on the following figure. At Penalty = 1000, the RL agent needs about 4500 episodes to reach the target reward value of optimization. While at Penalty=3000, the RL agent only needs 700 steps to complete the training.

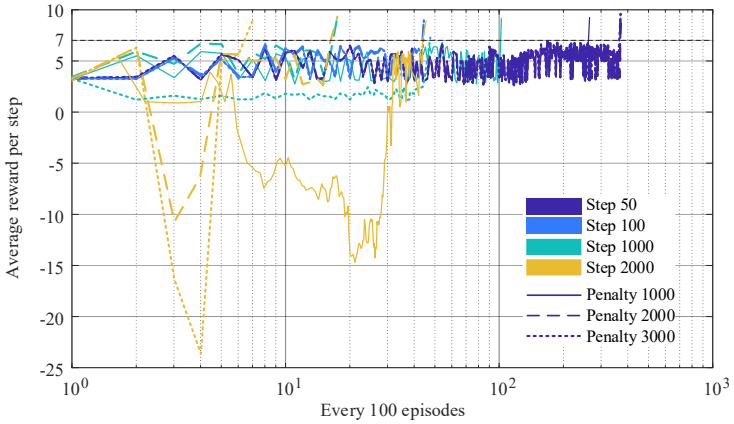


Fig. 5.22: Variation of rewards with different training steps and different penalty factors for RL agent$_3$

The following Tab. 5.11 shows all maximum average rewards, namely the rewards for the last 100 episodes of three RL agents under different step settings and penalty factors.

Tab. 5.11 Maximum average reward of three RL agents under different training settings

Average reward of step at the last episode	RL agent$_1$			RL agent$_2$			RL agent$_3$		
	Penalty factor								
Steps	1000	2000	3000	1000	2000	3000	1000	2000	3000
50	36.87	31.44	44.43	7.66	11.25	7.122	9.21	9.24	9.66
100	18.25	11.98	81.90	9.786	12.37	8.929	9.06	9.09	9.09
500	38.57	39.75	12.44	7.536	12.19	9.54	9.16	9.11	7.42
2000	8.62	51.34	9.74	8.961	7.724	8.89	8.99	9.40	8.99

The previous analysis shows that there is a general pattern of different step and penalty factor settings on the training performance of the agent. To visually analyze the effect of different training settings on the final average reward that the agent can obtain, the data in Tab. 5.11 was used to create the following figure.

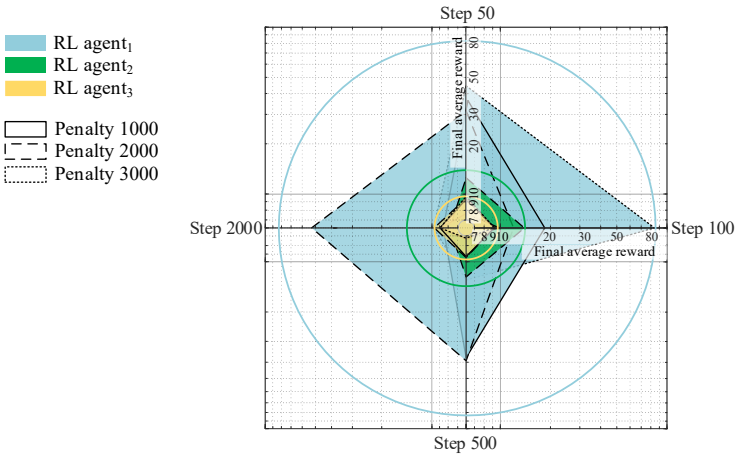


Fig. 5.23: Final average reward of three RL agents with different training steps and different penalty factors

Fig. 5.23 reflects that, for each agent, there is no general rule for the training settings when the maximum final average reward occurs. However, the area consisting of the rewards under different training settings shows that the blue area for RL agent$_1$ is much larger than the green and yellow areas for agent$_2$ and agent$_3$. The maximum final average rewards that each agent can obtain were used as radius, whose values were shown in the values marked in orange in Tab. 5.11, to draw circles. The three circles shown in the figure are therefore obtained. Compared to the green and yellow circles plotted for RL agent$_2$ and RL agent$_3$, the blue circle is obviously the largest. As discussed above, since RL agent$_1$ outputs eight parameters, it has more margin for optimization and, therefore more likely to achieve a larger reward. RL agent$_3$, however, only optimizes for two control parameters, which gives it the smallest reward among the three agents.

The three RL agents with the largest final average reward are afterward applied in the designed test scenarios in section 4.4 and compared with the system using controller parameters from the conventional parameterization methods. The trajectory of partial eigenvalues as the system changes its operating states in each of the five test scenarios will be shown in Fig. 5.24 to Fig. 5.28.

- ***Scenario RL_SV:*** Different voltage source supplying: V_s = 0.9, 1.0, 1.1 p.u.

Through the investigation, the system strength grows when the voltage of the connection point to the HV network increases. It can be observed by the trajectory of the dominant eigenvalues, which are shifted to the left for all four settings of the controller parameterization, see Fig. 5.24. Compared to RL agent$_2$ and RL agent$_3$, where only two and four control parameters are optimized, the distribution of the dominant eigenvalues of RL agent$_1$ is relatively leftward, and the frequency is between 2500 and 3000 Hz. It is relatively low compared to RL agent$_2$, whose dominant eigenvalues also lie in the range to the left of the eigenvalue map compared to the system without RL agent and with RL agen$_3$. The transient response of the system under the control of RL agent$_1$ is characterized by oscillating at a frequency of 2500 to 3000 Hz and reaching the final stable value within a short time.

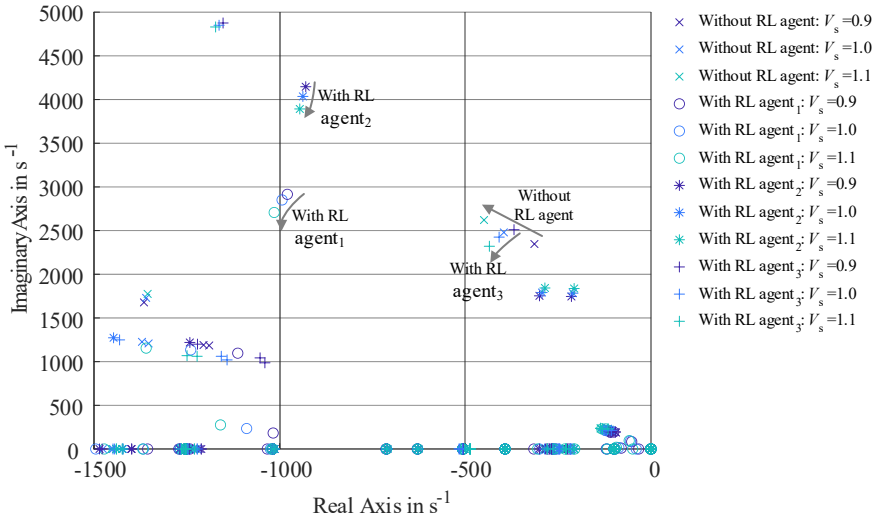


Fig. 5.24: Trajectory of partial eigenvalues for increasing the source voltage comparing between system without RL agent and with three RL agents

- ***Scenario RL_P:*** Different active power feeding from CON3: P_{CON3} = 1, 3, 5 MW, while $P_{CON1} = P_{CON2}$ = 3 MW

As the power delivered to the system via CI-REG increases, the system becomes unstable. The rightward shift of the dominant eigenvalues also reflects this trend, see Fig. 5.25.

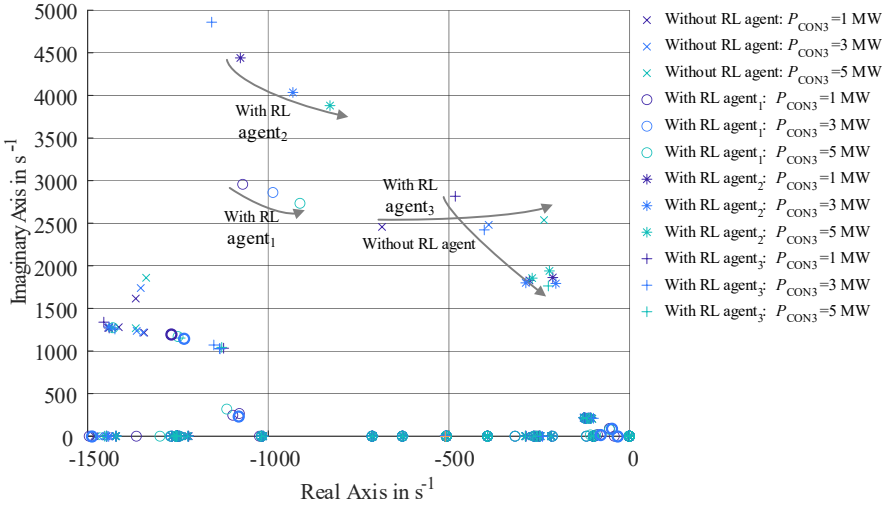


Fig. 5.25: Trajectory of partial eigenvalues for increasing the active power of CON3 comparing between system without RL agent and with three RL agents

In this scenario, the system with RL agent$_1$ has a smaller rightward shift of the dominant eigenvalues than the other three settings when the active power at CON3 increases from 1 MW to 3

MW. Moreover, the dominant eigenvalues of the system with RL agent$_1$ are also relatively concentrated on the left. Therefore, its stability margin is larger than the other settings under greater power injection by CON3.

- ***Scenario RL_Q:*** Different reactive power feeding from CON3, using different power factors at P_{CON3} = 5 MW to calculate Q_{CON3}: PF_{CON3} = 0.85, 0.9, 0.95

When the active power remains constant, an increase in the power factor implies a decrease in the reactive power. In agreement with the trend when the active power input to the MV-system from CON3 decreases, the dominant eigenvalues show a rightward shift with decreasing reactive power at different parameterization settings, see Fig. 5.26. In this scenario, the movement of the dominant eigenvalues in a system with RL agent3 shows a smaller range than in the other three settings.

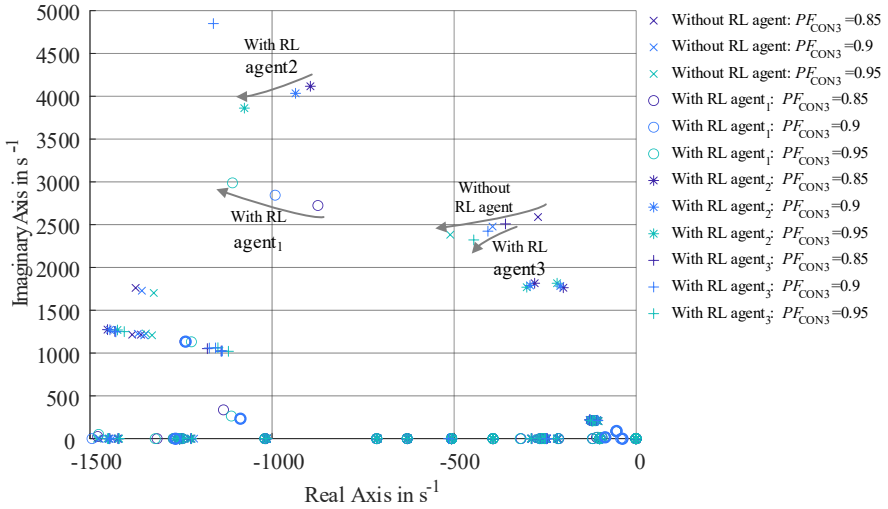


Fig. 5.26: Trajectory of partial eigenvalues for increasing the power factor of CON3 comparing between system without RL agent and with three RL agents

- ***Scenario RL_P_T:*** Different active power feeding from three converters: P_{CON1} = P_{CON2} = P_{CON3} = 1, 3, 5 MW

Fig. 5.27 shows that when the active power P delivered to the system increases significantly, a wide rightward shift of the dominant eigenvalues is expected. With a P of 5 MW for all three converters, the dominant eigenvalues of the system without an RL agent have crossed the RHP of the imaginary axis. It indicates that the system is unable to maintain a steady state. A small perturbation, such as voltage variation, can cause the system to collapse. The system with RL agent$_3$ at the same operating points is also very close to the imaginary axis, meaning that the stability threshold of the system is very small. Meanwhile, the systems with the rest settings are relatively on the left side of the eigenvalues map. However, a system with agent$_2$ has a large span of dominant eigenvalues as the total feed-in of the system changes from 3 MW up to 9 MW and then to 15 MW, with their real parts changing from approximately -1500 to close to -500. In comparison, a system with agent$_1$ only changes from close to -1200 to near -800. It is clear that this system has the largest stability margin. It can also receive more power feed-in from the converters.

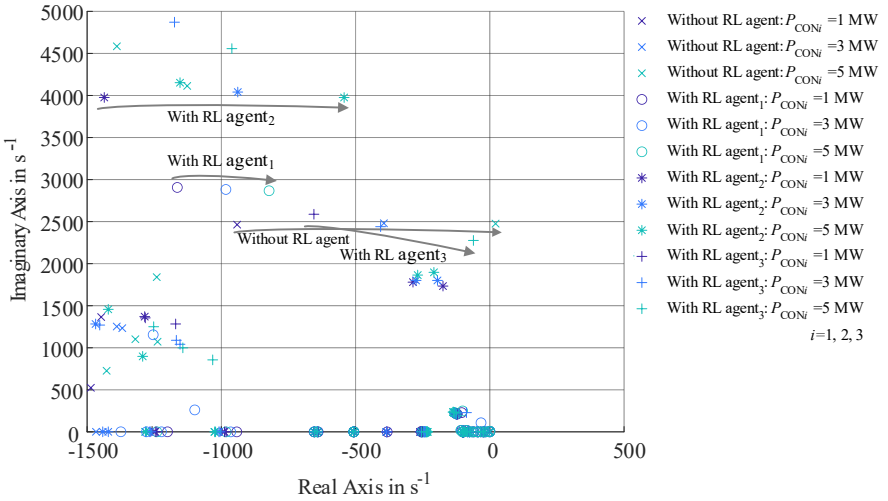


Fig. 5.27: Trajectory of partial eigenvalues for increasing the active powers of three converters comparing between system without RL agent and with three RL agents

- ***Scenario RL_L:*** Different loading conditions: Load level = 40%, 60%, 80%

The last scenario investigates the movement of the dominant eigenvalues from 40% to 80% of the load level, see Fig. 5.28. Consistent with the previous four scenarios, the dominant eigenvalues of the system with RL agent$_1$ are in relatively leftward orientation and have a small movement range. As the load level increases, the strength of the system decreases and its stability decreases. The rightward shift of the dominant eigenvalues follows the stated changes.

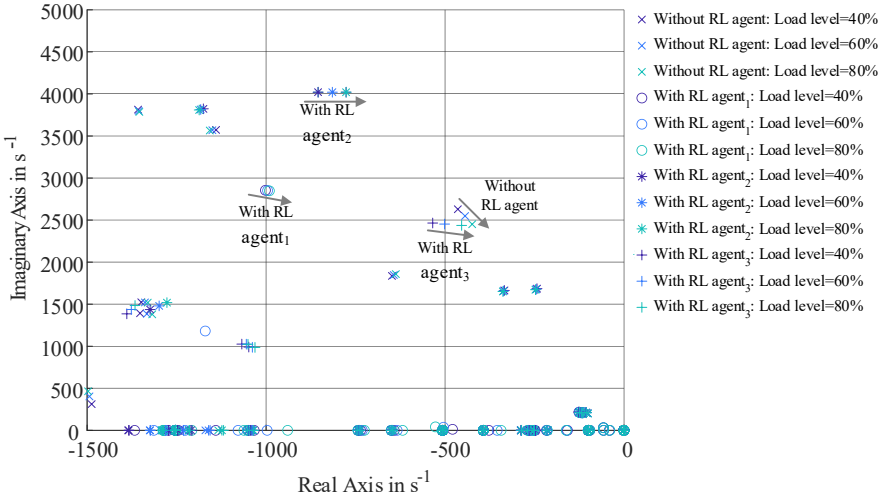


Fig. 5.28: Trajectory of partial eigenvalues for increasing the load in system comparing between system without RL agent and with three RL agents

To visually compare the stability of the system under different test scenarios, the value of SIs was calculated using equation (2.64) and is shown in Fig. 5.29 below. The values of the real and imaginary parts of the dominant eigenvalues required to calculate the SIs are listed in Tab. A.6 to Tab. A.9 in appendix A.5.

Fig. 5.29 reflects that the SI of the system with RL agent$_1$ is larger in all tested scenarios compared to the other three controller parameter configurations for the same system state. In more than 90% of the 15 system operating states, its SI is greater than 0.3. The larger value represents a higher stability margin for the system. Of all the studied system states, only the system without an RL agent has a negative SI at an active power of 5 MW for the three converters, indicating system instability. The SI of the system with RL agent$_2$ and RL agent$_3$ is in the middle of the two settings discussed earlier. The SI of the system with RL agent$_2$ is greater than that of the system with agent$_3$, which only optimizes two controller parameters for more than 85% of the scenarios.

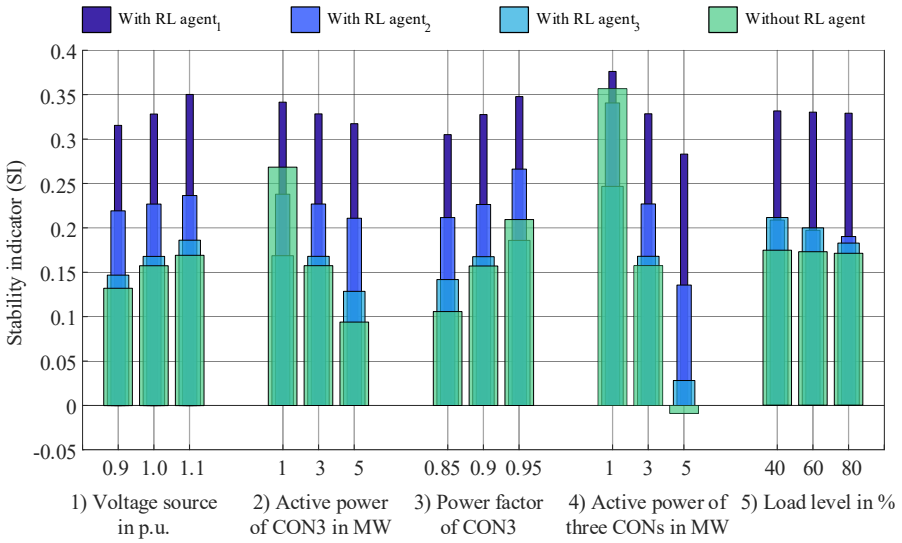


Fig. 5.29: SI comparison under five test scenarios between systems without RL agent and with three RL agents

To quantify the effect of system stability improvement, the SI for systems without RL agents and three RL agents for 15 system states were summed and averaged, respectively. The results are as follows.

$$\text{Averaged SI for system} \begin{cases} \text{with RL agent}_1 = 0.330 \\ \text{with RL agent}_2 = 0.224 \\ \text{with RL agent}_3 = 0.167 \\ \text{without RL agent} = 0.165 \end{cases} \tag{5.3}$$

Furthermore, the following equation is used to derive the stability optimization degree (SOD) of the system by the three RL agents.

$$\text{SOD} = \frac{\text{Averaged SI for system with RL agent}_i \text{ - Averaged SI for system without RL agent}}{\text{Averaged SI for system without RL agent}} \times 100\% \quad (5.4)$$

Finally, it is obtained that SOD for a system with RL agent_1 is 99.67%, for RL agent_2 is 35.84%, and for RL agent_3 is 0.96%.

- ***Summary for the stabilization using adaptive control approach***

The above results demonstrate that the stability of the system can be optimized to different degrees using RL agents. The efficiency of the RL agent in reaching the set target optimization value varies for different training settings, and a large training step and penalty factor generally allow the agent to be optimized more efficiently. In this work, the RL agent is also designed to output three different numbers of converter controller parameters, i.e., 2, 4, and 8. Among them, RL agent_1, which outputs 8 parameters, obtains the largest reward value at the end of training compared to the other two agents. It also gives the system the maximal SI at different operating points when it is put into the system for verification. It means that the system has the highest stability margin when RL agent_1 is used. According to the previous calculation, the application of RL agent_1 can improve the system stability by over 99%. Compared to the system using only the conventional converter parameterization method, the system with RL agent_2 and agent_3 also shows their positive effect on parameter optimization and leads to a certain degree of improvement in system stability.

6 Summary and outlook

6.1 Summary

In the trend of increasing renewable energy sources in the power system, there is an expansion in the number of power electronic converters for Renewable Energy sources (REs) like solar and wind to the distribution network. Their interaction with the network leads to stability problems. For the network to operate stable and with a large stability margin, this work uses reinforcement learning (RL) to optimize the parameters of the adaptive converter controller according to the system operating points.

The two main research questions and their corresponding sub-questions are answered and summarized in the following.

> *Q1: How to use the digital twin of the distribution network to increase its observability, which includes a limited number of measurement points?*

The digital twin (DT) for this work was built for one of the purposes of increasing observability. Once the DT is available, the measurements of interest can be obtained from the DT. Another important contribution of this work is the dynamic parameter adjustment method, i.e., matching the controller parameters of the converter to the current operating points of the studied network so that the system stability margin is always maintained at a relatively large value. As known that the value of the stability indicator (SI) is related to the system parameters as well as to the system operating points. Therefore, a DT model is needed to improve the observability of the converters dominated distribution network. It is known that the system has only a limited number of measurement points. These measurements are typically placed at the PCC point and measure the bus voltage and the current flowing through it. Using these limited measurements, the following sub-questions are answered.

> *Q1.1: How to generate a DT for a distribution network from a limited number of measurements?*

The DT is generated based on an EMT model with generalized element's structure and limited measurements through an optimization technique for model parameters. First, its respective theoretical EMT model based on a differential equation is constructed for each network element, such as transformer, line, load, converter, etc. They are assembled and combined according to the structure of the CIGRÉ European MV network, which has been chosen as the reference model for this work. Once the system model has been built, its parameters will be estimated by an optimization technique using non-linear least squares, which aims to minimize the error between the simulated output from the model and the measured output from the reference network. This work sets up three scenarios for parameter estimation to analyze the model's accuracy when it has training data from different network states. It turns out that the smaller the load connected to the network, in addition to the presence of a converter on a bus and the smaller the power generated by the converter, the more appropriate the measurement data source for the parameter estimation. The network states described above are the states with a higher degree of network stability. In addition, the closer the location of a bus is to the slack bus, the more accurate the model will be when the voltage data from the mentioned bus is used as training data for estimation. Using these techniques to improve the model's accuracy, its estimated parameters are finalized.

The DT model can be therefore built with a known structure and model parameters. It was validated in four test scenarios:

- Step changes in source voltage
- Step changes in converters´ active and reactive power
- An open circuit in an important line
- A short circuit at a bus to which a converter is connected

For each scenario, different combinations of generation and load levels were set. For each scenario, the amplitudes, and phases of the five bus voltages in the DT model were compared with those from the reference model. In addition, the corresponding error MAPE values were calculated as a measure of accuracy. The highest accuracy was above 99%, occurring in the source step change at low load and generation levels. At the same time, the smallest accuracy of 96.83% occurs in the open circuit case on heavily loaded and high-generation lines. The high accuracy means that the DT model can accurately simulate the behavior of the network. The DT model is therefore used with confidence in the following studies.

Q1.2: How to utilize the DT to estimate the state of the system, especially the state that is required for system stability evaluation?

An ANN-based state estimator is developed in the DT model described above and utilized to estimate the system states of interest. The estimator contains five ANNs, which are used to estimate the PCC voltages' magnitude $\hat{U}_{\mathrm{PCC0}}$, currents I_{CONd} and I_{CONq}, as well as control vectors V_{d} and V_{q} for each of the three converters in the studied network. In other words, these variables are the outputs of the ANN state estimator to be generated. They are directly related to the value of the stability degree of the small-signal system and are required for further adaptive control. Each ANN has the same variables for a total of 10 inputs, which are the switching states of the switch located at the important line, the voltage sources that have a dominant influence on the voltage at each bus, the measured values of active and reactive power injection at each generation bus and the sum of active and reactive power consumption at all load busses, respectively. All five ANNs are identically set up regarding the ANN structure with a single hidden layer. The number of neurons in the hidden layer is determined based on the maximum correlation coefficient *R* so that each ANN can minimize the estimation error and gain the greatest estimation accuracy. In the training of the five ANNs, the SCG learning algorithm was used. The training data were obtained from the previously constructed DT model. Different combinations of voltage levels, generation, and load as well as network structures were set to obtain data set for a large number of possible various network operating states.

After extensive training in ANN-based state estimators, mathematical relationships were established between the system states and the values of the system stability strong relevant variables. The ANN state estimator was validated on the following scenarios, i.e., source voltage variations in the range of 0.95 and 1.05 p.u., and the active power of the load and generation over 24 hours as well as the change of the switch status. The estimation results were compared with the corresponding variables in the reference system, which were calculated by the standard load flow program, Finally, it can be concluded that the derived state estimator based on ANNs allows a fast and accurate estimation of the operating points of the system, which is then used to calculate and monitor the stability of the network.

Q2: How to apply the converters in the distribution network to improve network stability?

Converters with appropriate control parameters demonstrate the benefits in improving network stability. The expansion of renewable energy sources has led to an increase in the number of converters in the network, and the control loop of the converters is a major factor in oscillations and converter-driven instability. The mentioned stability is sensitive to the control parameters of the control loop. In discussing improving the stability of the network, a SI is needed for the stability evaluation. Therefore, the following question is answered first.

Q2.1: Which stability indicator is suitable for detecting network stability and how can their detection be automated?

Using damping ratio of the network dominant eigenvalues as SI and developing an ANN-based SI estimator for automatic quantification of network stability are proposed and validated in this work. A small-signal model is first constructed by linearizing the DT model at the operating points of a constantly changing system. The eigenvalues are derived for the state matrix of the small signal model and the dominant eigenvalues, which vary considerably with the system voltage, load, and power fed-in for the network, are determined. Afterward, their damping ratio is used as the SI of the studied network to indicate its stability margin. To effectively determine the dominant eigenvalues of the system, determining the dominant oscillation frequency is taken as a starting point. Analysis of variables such as the PCC voltage, the converter output current, and the active and reactive power of the converter in the frequency domain reveals that the reactive power response provides information on the dominant oscillation frequency. Based on this frequency, the dominant eigenvalues can then be effectively identified from the eigenvalue map of the system. Furthermore, as the calculation of the eigenvalues is time-consuming, an ANN-based SI estimator method is proposed to enable the above eigenvalue-based SI to be implanted online and to display the current system oscillation stability margin in real-time. The ANN-based SI estimator has the same inputs as the ANN state estimator described above, namely the switching state, voltage source amplitude, active and reactive power at each converter and the sum of active and reactive power consumption, while the SI is the output. The SI was tested at different active and reactive power feeding of the converter, as well as variations in source voltage supply and load conditions. The same test scenario was used to validate the ANN-based SI estimator. As a result, the accuracy of the SI estimator was determined to be 98.95% by a MAPE value of 1.05%. The value of the output of this estimator will be used as an indicator for parameter optimization of the adaptive controller.

Q2.2: Which methods will be applied in the adaptive control of the converters?

In this work, the RL-based regulator adapts the controller parameters using the current network state for adaptive control of the converters. RL, a type of machine learning, is used in this work for the adaptive control of the converter. The system stability margin is maintained at a high level by adaptively regulating the controller parameters that greatly impact the degree of network stability under different operating points. The above value of SI is used as a reward value for the RL agent, which in this work is mathematically linked to the current steady-state operating point of the system with the selected controller parameters by using a complex NN through continuous training. RL is known to have the ability to maximize the reward it receives. Therefore, by setting the SI as a reward and training it over a large number of network operating points, an RL agent is eventually obtained, improving the network stability margin.

In this work, three RL agents were designed to output three sets of converter controller parameters. These parameters include the proportional parameters of the current PI-controller, the proportional and integral parameters of the voltage PI-controller, and the parameters of the reactive power I-controller. $Agent_1$ outputs the 8 controller parameters that have the most significant effect on the SI. The other two agents output 4 and 2 parameters respectively.

Different training set-ups were also performed to investigate the efficiency of the RL agent in reaching the target reward values. The results show that large training steps and large penalty factors lead to more efficient optimization of the agents. At the end of the training, RL $agent_1$, which outputs 8 parameters, obtained the largest reward value compared to the other two agents. This is due to the fact that RL $agent_1$ has more space for optimization. At the same time, RL $agents_2$ and $agents_3$ were only optimized for four and two control parameters, which gave them smaller rewards.

Three RL agents were put into the DT model for validation. The results demonstrate that using RL agents improves system stability compared to systems using only conventional converter parameterization methods. In particular, RL $agent_1$ is the most effective optimization, with its application improving stability by more than 99%. The use of RL for adaptive control of converters dominated distribution network to improve its stability was demonstrated in this work.

To answer the above research questions, the following scientific contributions were developed, and new insights were gained in this work:

- Investigation and evaluation of the suitable network state and measurement points location for the DT creation: Training data obtained from states with a higher degree of network stability and from measurement point closed to the slack bus resulting in a DT model with high accuracy
- Design of ANN-based state estimators allowing a fast and accurate estimation of the network operating points
- Proposed SI using damping ratio of dominant eigenvalues to indicate the stability margin of the network and develop an ANN-based SI estimator for an online implementation
- Identification of dominant oscillation frequency gained from converter reactive power response to effectively determine the dominant eigenvalues from the system eigenvalue map
- Investigation of training settings for RL, which is used to optimize the converter controller parameters
- Demonstrate the feasibility of the adaptive control using RL for improving the network stability margin

6.2 Outlook

In power systems, adaptive control is not limited to adjusting converter control parameters. The type of controller can also be adapted, for example, by using adaptive quasi-proportional resonant differential control methods to achieve zero steady-state error tracking and to extend the system bandwidth. Alternatively, different control modes can be adaptively adjusted for different grid impedances, with the grid-following converter using current-source mode under strong grid conditions and switching to voltage-source mode under weak grid conditions. Combining the advantages of both grid-following modes makes it more suitable for high penetration rates where the grid impedance fluctuates.

This work focuses on adaptively adjusting the control parameters for the converter in grid-connected generation mode so that the network stability is ensured and improved when the state changes. The adaptive control of parameters using RL is achieved by setting effective reward values as well as penalty values (negative rewards) for the RL agent. In further research, in addition to the small-signal SI, RL reward can be incorporated into another stability index, such as the static stability index for voltage stability analysis and the transient stability index for critical clearing time for a given disturbance. This setting aims to ensure and improve the stability of large-signal systems and even to adjust the control parameters of the inverter. In this way, the control parameters of the converter can be adaptively adjusted even to ensure and enhance the stability of the large-signal system.

With the help of ML's powerful data processing capabilities, adaptive control becomes even more intelligent. This work proposes that power systems can be adaptively controlled using ML, which can also assist DT models for parameter estimation as well as for estimation of system states. Potential research areas for ML in power networks are detecting the type and location of network faults, energy-saving techniques, and energy demand response, aiming to improve power system flexibility and reliability in a cost-effective manner.

In this work, the 110 kV HV network was considered a stiff network, i.e., the impact of the 110 kV network on the 20 kV MV network investigated was not considered. In future research, the 110 kV network resonance will be simulated. Subsequently, its influence is to be considered in the parameterization of the converter controller. Resonance of the 0.4 kV network is also conceivable in the adaptive control process in the future.

References

[1] *3rd Renewable Power Generation Conference (RPG 2014)*. Institution of Engineering and Technology: Institution of Engineering and Technology, Sep. 2014.

[2] Bundesministerium für Wirtschaft und Energie, *Zeitreihen zur Entwicklung der erneuerbaren Energien in Deutschland*.

[3] Bundesregierung, *Energiekonzept für eine umweltschonende, zuverlässige und bezahlbare Energieversorgung*.

[4] S. Rahmstorf and A. Ganopolski, *Long-Term Global Warming Scenarios Computed with an Efficient Coupled Climate Model, Climatic Change Ausgabe 2*. Climatic Change, Jg. 43. Zugriff am: 1999.

[5] Umweltbundesamt, *Emissionsbilanz erneuerbarer Energieträger: Bestimmung der vermiedenen Emissionen im Jahr 2018*. Dessau-Roßlau.

[6] A. Kemmler *et al., Energiewirtschaftliche Projektionen und Folgeabschätzungen 2030/2050, Dokumentation von Referenzszenario und Szenario mit Klimaschutzprogramm 2030*. PROGNOS AG, FRAUNHOFER ISI, GWS, IINAS (Hrsg.).

[7] J. Haucap und J. Meinhof, „Die Strompreise der Zukunft“, *Wirtschaftsdienst*, Jg. 102, Nr. 1, S. 53–60, 2022, doi: 10.1007/s10273-022-3175-7.

[8] Bundesministerium für Wirtschaft und Energie, „Zeitreihen zur Entwicklung der erneuerbaren Energien in Deutschland“ in.

[9] OTTI-Konferenz Zukünftige Stromnetze für Erneuerbare Energien, *2. OTTI-Konferenz Zukünftige Stromnetze für Erneuerbare Energien*. Regensburg: Ostbayerisches Technologie-Transfer-Institut, 2015.

[10] Arbeitsgemeinschaft Energiebilanzen, *Energiebilanz der Bundesrepublik Deutschland 2016 und Vorjahre*. Berlin, Germany.

[11] G. Fournier, M. Baumann, J. Buchgeister, M. Weil und R. Seign, „Elektrofahrzeuge als virtuelle Speicher zur Kompensierung volatiler erneuerbarer Energien in Deutschland – Ein neues Geschäftsmodell?“ in *Schritte in die künftige Mobilität*, H. Proff, W. Pascha, J. Schönharting und D. Schramm, Hg., Wiesbaden: Springer Fachmedien Wiesbaden, 2013, S. 253–271, doi: 10.1007/978-3-8349-4308-8_16.

[12] T. Mennel, „Das Erneuerbare-Energien-Gesetz — Erfolgsgeschichte oder Kostenfalle?“, *Wirtschaftsdienst*, Jg. 92, S1, S. 17–22, 2012, doi: 10.1007/s10273-012-1345-8.

[13] S. Heier, *Windkraftanlagen: Systemauslegung, Netzintegration und Regelung ; mit 14 Tabellen,* 4. Aufl. Stuttgart [u.a.]: Teubner, 2005.

[14] E. Rebhan, Hg., *Energiehandbuch: Gewinnung, Wandlung und Nutzung von Energie*. Berlin, Heidelberg: Springer Berlin Heidelberg; Imprint; Springer, 2002.

[15] Henning Zimmer, „Regeldynamik konventioneller Kraftwerke im Kontext veränderter Erzeugungsstrukturen“. Ph.D. dissertation, Elektrotechnik und Informationstechnik, Technischen Universität Darmstadt, Darmstadt, 2017.

[16] D. W. Felix Jenni, *Steuerverfahren für selbstgeführte Stromrichter*, 1995.

[17] M. Michel, *Leistungselektronik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[18] Yanyun Ma, „Einsatz eines Hochspannungs-Gleichstrom-Übertragung-Systems zum netzstabilisierenden Anschluss von Offshore-Windparks an das Elektroenergiesystem“. Ph.D. dissertation, 2013.

[19] D. Schroder, *Leistungselektronische schaltungen: Funktion, auslegung und anwendung*. [Place of publication not identified]: Springer, 2012.

[20] Xinya Song, „Realisierung von Umrichtermodellen und deren Validierung auf dynamisches Verhalten im Verteilernetz“. master's thesis, TU Ilmenau, Ilmenau,Germany, 2018.

[21] Michael Schärfenberg, „Untersuchung der Wechselwirkungen zwischen leistungselektronisch integrierten Netze-lementen im Verteilernetz“. master thesis, Technische Universität Ilmenau, Ilmenau, Germany, 2017.

[22] M. I. Radaideh, C. Pappas und S. Cousineau, „Real electronic signal data from particle accelerator power systems for machine learning anomaly detection“, *Data in Brief*, Jg. 43, S. 108473, 2022.

[23] J.-F. Toubeau *et al.,* „Machine learning-assisted outage planning for maintenance activities in power systems with renewables“, *Energy*, Jg. 238, S. 121993, 2022.

[24] S. Luo, „Assessment of Using Machine Learning Methods in Analyzing Data from Renewable Integrated Power Systems“, Arizona State University, 2022.

[25] S. Naderi, M. Javadi, M. Mazhari und C. Y. Chung, „A Machine Learning-Based Framework for Fast Prediction of Wide-Area Remedial Control Actions in Interconnected Power Systems“, *IEEE Trans. Power Syst.*, 2022.

[26] A. Perera und P. Kamalaruban, „Applications of reinforcement learning in energy systems“, *Renewable and Sustainable Energy Reviews*, Jg. 137, S. 110618, 2021, doi: 10.1016/j.rser.2020.110618.

[27] M. Awad und R. Khanna, „Deep Neural Networks“ in *Efficient Learning Machines*, M. Awad und R. Khanna, Hg., Berkeley, CA: Apress, 2015, S. 127–147, doi: 10.1007/978-1-4302-5990-9_7.

[28] W. Kritzinger, M. Karner, G. Traar, J. Henjes und W. Sihn, „Digital Twin in manufacturing: A categorical literature review and classification“, *IFAC-PapersOnLine*, Jg. 51, Nr. 11, S. 1016–1022, 2018, doi: 10.1016/j.ifacol.2018.08.474.

[29] General Electric Company, *GE Digital Twin: Analytic Engine for the Digital Power Plant*.

[30] Siemens AG 2018, *Siemens Electrical Digital Twin: A single source of truth to unlock the potential within a modern utility's data landscape.* [Online]. Verfügbar unter: siemens.com/electrical-digital-twin

[31] C. Brosinsky, D. Westermann und R. Krebs, „Recent and prospective developments in power system control centers: Adapting the digital twin technology for application in power system control centers“ in *2018 IEEE International Energy Conference (ENERGYCON)*, Limassol, 6/3/2018 - 6/7/2018, S. 1–6, doi: 10.1109/ENERGYCON.2018.8398846.

[32] L. Pagnier und M. Chertkov, „Embedding Power Flow into Machine Learning for Parameter and State Estimation“, 26. März 2021. [Online]. Verfügbar unter: http://arxiv.org/pdf/2103.14251v1.

[33] O. Nelles, *Nonlinear system identification: From classical approaches to neural networks and fuzzy models*. Berlin, London: Springer, 2011.

[34] R. Petrone *et al.,* „A review on model-based diagnosis methodologies for PEMFCs“, *International Journal of Hydrogen Energy*, Jg. 38, Nr. 17, S. 7077–7091, 2013, doi: 10.1016/j.ijhydene.2013.03.106.

[35] J.-R. Riba, M. Moreno-Eguilaz, S. Bogarra und A. Garcia, „Parameter Identification of DC-DC Converters under Steady-State and Transient Conditions Based on White-Box Models“, *Electronics*, Jg. 7, Nr. 12, S. 393, 2018, doi: 10.3390/electronics7120393.

[36] Frank Vanden Berghen, „Levenberg-Marquardt algorithms vs Trust Region algorithms“, IRIDIA, Universite Libre de Bruxelles, 2004.

[37] J. Nocedal und S. J. Wright, *Numerical optimization*. New York: Springer, 1999.

[38] J. J. Moré und D. C. Sorensen, „Computing a Trust Region Step“, *SIAM J. Sci. and Stat. Comput.*, Jg. 4, Nr. 3, S. 553–572, 1983, doi: 10.1137/0904038.

[39] P. Schade, *Innere-Punkte-Verfahren mit Redundanzerkennung für die Quadratische Optimierung,* 1. Aufl. Wiesbaden: Gabler Verlag / GWV Fachverlage, Wiesbaden, 2008.

[40] A. Meister, *Numerik linearer Gleichungssysteme: Eine Einführung in moderne Verfahren : mit MATLAB-Implementierung von C. Vömel,* 2. Aufl. Wiesbaden: Vieweg, 2005.

[41] J. Park, D. Kodaira, K. A. Agyeman, T. Jyung und S. Han, „Adaptive Power Flow Prediction Based on Machine Learning“, *Energies*, Jg. 14, Nr. 13, S. 3842, 2021, doi: 10.3390/en14133842.

[42] A. Karami und M. S. Mohammadi, „Radial basis function neural network for power system load-flow“, *International Journal of Electrical Power & Energy Systems*, Jg. 30, Nr. 1, S. 60–66, 2008, doi: 10.1016/j.ijepes.2007.10.004.

[43] Jayendra Krishna und L. Srivastava, „Counterpropagation Neural Network For Solving Power Flow Problem“ (en), 2008, doi: 10.5281/ZENODO.1333941.

[44] F. Schäfer, J.-H. Menke und M. Braun, „Evaluating machine learning models for the fast identification of contingency cases“, *Applied AI Letters*, Jg. 1, Nr. 2, 2020, doi: 10.1002/ail2.19.

[45] X. Hu, H. Hu, S. Verma und Z.-L. Zhang, „Physics-Guided Deep Neural Networks for Power Flow Analysis“, *IEEE Trans. Power Syst.*, S. 1, 2020, doi: 10.1109/TPWRS.2020.3029557.

[46] M. F. Møller, „A scaled conjugate gradient algorithm for fast supervised learning“, *Neural Networks*, Jg. 6, Nr. 4, S. 525–533, 1993, doi: 10.1016/S0893-6080(05)80056-5.

[47] P. E. Gill, W. Murray und M. H. Wright, *Practical optimization*. London: Academic Press, 1981.

[48] Ö. Kişi, „Streamflow Forecasting Using Different Artificial Neural Network Algorithms“, *J. Hydrol. Eng.*, Jg. 12, Nr. 5, S. 532–539, 2007, doi: 10.1061/(ASCE)1084-0699(2007)12:5(532).

[49] M. R. Hestenes und A. V. Balakrishnan, *Conjugate Direction Methods in Optimization*. New York: Springer New York, 1980.

[50] R. Vilanova und A. Visioli, Hg., *PID Control in the Third Millennium: Lessons Learned and New Approaches,* 2012. Aufl. London: Springer London, 2012.

[51] V. M. Cordero Alfaro und R. Vilanova, *Model-reference robust tuning of PID controllers,* 1. Aufl. Switzerland: Springer, 2016.

[52] A. Daraz, S. A. Malik, I. U. Haq, K. B. Khan, G. F. Laghari und F. Zafar, „Modified PID controller for automatic generation control of multi-source interconnected power system using fitness dependent optimizer algorithm“ (eng), *PloS one*, Jg. 15, Nr. 11, e0242428, 2020, doi: 10.1371/journal.pone.0242428.

[53] S. Sondhi und Y. V. Hote, „Fractional order PID controller for load frequency control“, *Energy Conversion and Management*, Jg. 85, S. 343–353, 2014, doi: 10.1016/j.enconman.2014.05.091.

[54] K. A. Al Sumarmad, N. Sulaiman, N. I. A. Wahab und H. Hizam, „Energy Management and Voltage Control in Microgrids Using Artificial Neural Networks, PID, and Fuzzy Logic Controllers“, *Energies*, Jg. 15, Nr. 1, S. 303, 2022, doi: 10.3390/en15010303.

[55] Z.-Q. Zhao, S.-J. Liu und J.-S. Pan, „A PID Parameter Tuning Method Based on the Improved QUATRE Algorithm“, *Algorithms*, Jg. 14, Nr. 6, S. 173, 2021, doi: 10.3390/a14060173.

[56] T. George und V. Ganesan, „Optimal tuning of PID controller in time delay system: a review on various optimization techniques“, *Chemical Product and Process Modeling*, Jg. 17, Nr. 1, S. 1–28, 2022, doi: 10.1515/cppm-2020-2001.

[57] J. Zhao, T. Li und J. Qian, „Application of Particle Swarm Optimization Algorithm on Robust PID Controller Tuning“ in *Lecture Notes in Computer Science, Advances in Natural Computation*, D. Hutchison et al., Hg., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, S. 948–957, doi: 10.1007/11539902_118.

[58] P. Sarkhel, N. Banerjee und N. B. Hui, „Fuzzy logic-based tuning of PID controller to control flexible manipulators“, *SN Appl. Sci.*, Jg. 2, Nr. 6, 2020, doi: 10.1007/s42452-020-2877-y.

[59] H. Xue und H. Pu, „An Improved BBO Algorithm and Its Application in PID Optimization of CFB Bed Temperature System“ in *Lecture Notes in Electrical Engineering, Proceedings of the 5th International Conference on Electrical Engineering and Automatic Control*, B. Huang und Y. Yao, Hg., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, S. 675–682, doi: 10.1007/978-3-662-48768-6_76.

[60] T. Ahmmed, I. Akhter, S. M. Rezaul Karim und F. A. Sabbir Ahamed, „Genetic Algorithm Based PID Parameter Optimization“, *ajis*, Jg. 10, Nr. 1, S. 8–13, 2020, doi: 10.5923/j.ajis.20201001.02.

[61] „Deep reinforcement learning for power system: An overview“, *CSEE JPES*, 2019, doi: 10.17775/CSEEJPES.2019.00920.

[62] A. Saenz-Aguirre, E. Zulueta, U. Fernandez-Gamiz, A. Ulazia und D. Teso-Fz-Betono, „Performance enhancement of the artificial neural network–based reinforcement learning for wind turbine yaw control“, *Wind Energy*, Jg. 23, Nr. 3, S. 676–690, 2020, doi: 10.1002/we.2451.

[63] C. Chenggang, N. Yan und C. Zhang, *An Intelligent Control Strategy for buck DC-DC Converter via Deep Reinforcement Learning*, 2020.

[64] B. V. Mbuwir, M. Kaffash und G. Deconinck, „Battery Scheduling in a Residential Multi-Carrier Energy System Using Reinforcement Learning“ in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2018, S. 1–6, doi: 10.1109/SmartGridComm.2018.8587412.

[65] Y. Qin, W. Zhang, J. Shi und L. Jinglong, *Improve PID controller through reinforcement learning*, 2018.

[66] A. I. Lakhani, M. A. Chowdhury und Q. Lu, „Stability-Preserving Automatic Tuning of PID Control with Reinforcement Learning“, 30. Dez. 2021. [Online]. Verfügbar unter: http://arxiv.org/pdf/2112.15187v2.

[67] S. Seidi Khorramabadi und A. Bakhshai, „Critic-Based Self-Tuning PI Structure for Active and Reactive Power Control of VSCs in Microgrid Systems“, *IEEE Trans. Smart Grid*, Jg. 6, Nr. 1, S. 92–103, 2015, doi: 10.1109/TSG.2014.2354651.

[68] E. Smith, D. A. Robinson und A. Agalgaonkar, „Cooperative secondary voltage control of static converters in a microgrid using model-free reinforcement learning“ in *2019 21st European Conference on Power Electronics and Applications (EPE '19 ECCE Europe)*, 2019, P.1-P.10, doi: 10.23919/EPE.2019.8914869.

[69] W. Al-Saedi, S. W. Lachowicz, D. Habibi und O. Bass, „Power flow control in grid-connected microgrid operation using Particle Swarm Optimization under variable load conditions“, *International Journal of Electrical Power & Energy Systems*, Jg. 49, S. 76–85, 2013, doi: 10.1016/j.ijepes.2012.12.017.

[70] R. Dong, S. Liu und G. Liang, „Research on Control Parameters for Voltage Source Inverter Output Controllers of Micro-Grids Based on the Fruit Fly Optimization Algorithm“, *Applied Sciences*, 9(7):1327, 2019. [Online]. Verfügbar unter: https://doi.org/10.3390/app9071327

[71] Y.-Y. Hong und W.-M. Luo, „Realization of simultaneous perturbation stochastic approximation using FPGA for on-line PID parameter tuning“, 2012.

[72] C.-C. Wong, L. Shih An und H.-Y. Wang, „Optimal PID controller design for AVR system“, *Tamkang Journal of Science and Engineering*, Jg. 12, 2009.

[73] M. Singh, R. N. Patel und R. Jhapte, „Performance comparison of optimized controller tuning techniques for voltage stability“ in *2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI)*, Kolkata, India, 2016, S. 11–15, doi: 10.1109/CMI.2016.7413701.

[74] R. S. Sutton, F. Bach und A. G. Barto, *Reinforcement Learning: An Introduction,* 2. Aufl. Massachusetts: MIT Press Ltd, 2018.

[75] U. Lorenz, *Reinforcement Learning: Aktuelle Ansätze verstehen - mit Beispielen in Java und Greenfoot,* 1. Aufl. Berlin, Heidelberg: Springer Vieweg, 2020.

[76] V. Mnih *et al.,* „Playing Atari with Deep Reinforcement Learning“, 19. Dez. 2013. [Online]. Verfügbar unter: http://arxiv.org/pdf/1312.5602v1.

[77] Sewak, *Deep Reinforcement Learning*. Springer Singapore, 2019.

[78] V. Mnih *et al.,* „Human-level control through deep reinforcement learning“ (eng), *Nature*, Jg. 518, Nr. 7540, S. 529–533, 2015, doi: 10.1038/nature14236.

[79] David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller., „Deterministic policy gradient algorithms“.

[80] N. Hatziargyriou *et al., Stability Definitions and Characterization of Dynamic Behavior in Systems with High Penetration of Power Electronic Interfaced Technologies*, 2020.

[81] R. Teodorescu, M. Liserre und P. Rodríguez, *Grid converters for photovoltaic and wind power systems*. [Piscataway, New Jersey], Chichester, West Sussex: IEEE; Wiley, 2011.

[82] A. Yazdani und R. Iravani, *Voltage-sourced converters in power systems: Modeling, control, and applications*. Hoboken, New Jersey: IEEE Press/John Wiley, 2010.

[83] L. Meegahapola, A. Sguarezi, J. S. Bryant, M. Gu, E. R. Conde D. und R. B. A. Cunha, „Power System Stability with Power-Electronic Converter Interfaced Renewable Power Generation: Present Issues and Future Trends“, *Energies*, Jg. 13, Nr. 13, S. 3441, 2020, doi: 10.3390/en13133441.

[84] N. Hatziargyriou *et al.,* „Definition and Classification of Power System Stability – Revisited & Extended“, *IEEE Trans. Power Syst.*, Jg. 36, Nr. 4, S. 3271–3281, 2021, doi: 10.1109/TPWRS.2020.3041774.

[85] *Standard approach to perform power system stability studies in oil and gas plants*, 2019. [Online]. Verfügbar unter: http://www.eaton.com.eg/content/dam/eaton/markets/oil-and-gas/knowledge-center/whitepaper/standard-approach-to-perform-power-system-stability-studies-in-oil-and-gas-plants.pdf

[86] D. P. Wadduwage, U. D. Annakkage und C. Q. Wu, „Hybrid algorithm for rotor angle security assessment in power systems“, *J. eng.*, Jg. 2015, Nr. 8, S. 241–251, 2015, doi: 10.1049/joe.2015.0033.

[87] Fetissi Selwa, Labed Djamel, Labed Imen, „The Transient Stability Study of a Synchronous Generator Based on the Rotor Angle Stability“.

[88] P. Kundur, N. J. Balu und M. G. Lauby, *Power system stability and control*. New York, London: McGraw-Hill, 1994.

[89] L. Meegahapola und D. Flynn, „Impact on transient and frequency stability for a power system at very high wind penetration“ in *Energy Society General Meeting*, Minneapolis, MN, 072010, S. 1–8, doi: 10.1109/PES.2010.5589908.

[90] A. Mitra und D. Chatterjee, „Active Power Control of DFIG-Based Wind Farm for Improvement of Transient Stability of Power Systems“, *IEEE Trans. Power Syst.*, Jg. 31, Nr. 1, S. 82–93, 2016, doi: 10.1109/TPWRS.2015.2397974.

[91] E. Munkhchuluun, L. Meegahapola und A. Vahidnia, „Impact on rotor angle stability with high solar-PV generation in power networks“ in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Torino, 92017, S. 1–6, doi: 10.1109/ISGTEurope.2017.8260229.

[92] L. Shi, S. Dai, Y. Ni, L. Yao und M. Bazargan, „Transient stability of power systems with high penetration of DFIG based wind farms“ in *Energy Society General Meeting (PES)*, Calgary, Canada, 072009, S. 1–6, doi: 10.1109/PES.2009.5276006.

[93] S. de Rijcke, H. Ergun, D. van Hertem und J. Driesen, „Grid impact of voltage control and reactive power support by wind turbines equipped with direct-drive synchronous machines“ in *2013 IEEE Power & Energy Society General Meeting*, Vancouver, BC, 2013, S. 1, doi: 10.1109/PESMG.2013.6672722.

[94] A. Ulbig, T. S. Borsche und G. Andersson, „Impact of Low Rotational Inertia on Power System Stability and Operation“, *IFAC Proceedings Volumes*, Jg. 47, Nr. 3, S. 7290–7297, 2014, doi: 10.3182/20140824-6-ZA-1003.02615.

[95] W. Li, G. Jos und C. Abbey, „Wind Power Impact on System Frequency Deviation and an ESS based Power Filtering Algorithm Solution“ in *2006 IEEE PES Power Systems Conference and Exposition*, Atlanta, Georgia, USA, 2006, S. 2077–2084, doi: 10.1109/PSCE.2006.296265.

[96] W. Farmer und A. Rix, „CURRENT METHODS FOR PV GENERATION TO PARTICIPATE IN THE INERTIAL RESPONSE OF A POWER SYSTEM“, 2018.

[97] L. Li, H. Li, M.-L. Tseng, H. Feng und A. S. F. Chiu, „Renewable Energy System on Frequency Stability Control Strategy Using Virtual Synchronous Generator“, *Symmetry*, Jg. 12, Nr. 10, S. 1697, 2020, doi: 10.3390/sym12101697.

[98] M. Datta, T. Senjyu, A. Yona, T. Funabashi und C.-H. Kim, „A Coordinated Control Method for Leveling PV Output Power Fluctuations of PV–Diesel Hybrid Systems Connected to Isolated Power Utility“, *IEEE Trans. Energy Convers.*, Jg. 24, Nr. 1, S. 153–162, 2009, doi: 10.1109/TEC.2008.2008870.

[99] A. Tayyebi, D. Gross, A. Anta, F. Kupzog und F. Dorfler, „Frequency Stability of Synchronous Machines and Grid-Forming Power Converters“, *IEEE J. Emerg. Sel. Topics Power Electron.*, Jg. 8, Nr. 2, S. 1004–1018, 2020, doi: 10.1109/JESTPE.2020.2966524.

[100] M. N. I. Sarkar, L. G. Meegahapola und M. Datta, „Reactive Power Management in Renewable Rich Power Grids: A Review of Grid-Codes, Renewable Generators, Support Devices, Control Strategies and Optimization Algorithms“, *IEEE Access*, Jg. 6, S. 41458–41489, 2018, doi: 10.1109/ACCESS.2018.2838563.

[101] J. Schiffer, T. Seel, J. Raisch und T. Sezi, „Voltage Stability and Reactive Power Sharing in Inverter-Based Microgrids With Consensus-Based Distributed Voltage Control“, *IEEE Trans. Contr. Syst. Technol.*, Jg. 24, Nr. 1, S. 96–109, 2016, doi: 10.1109/TCST.2015.2420622.

[102] V. S. S. Kumar, K. K. Reddy und D. Thukaram, „Coordination of Reactive Power in Grid-Connected Wind Farms for Voltage Stability Enhancement“, *IEEE Trans. Power Syst.*, Jg. 29, Nr. 5, S. 2381–2390, 2014, doi: 10.1109/TPWRS.2014.2300157.

[103] H. W. K. M. Amarasekara, L. Meegahapola, A. P. Agalgaonkar und S. Perera, „Impact of renewable power integration on VQ stability margin“ in *2013 Australasian Universities Power Engineering Conference (AUPEC)*, Hobart, Australia, 92013, S. 1–6, doi: 10.1109/AUPEC.2013.6725356.

[104] H. W. K. M. Amarasekara, A. P. Agalgaonkar, S. Perera und L. Meegahapola, „Placement of variable-speed wind power generators in power systems considering steady-state voltage stability“ in *2016 IEEE International Conference on Power System Technology (POWERCON)*, Wollongong, Australia, 92016, S. 1–6, doi: 10.1109/POWERCON.2016.7754029.

[105] „Index“ in *Subsynchronous Resonance in Power Systems*, P. M. Anderson, B. L. Agrawal und J. E. van Ness, Hg., IEEE, 1999, doi: 10.1109/9780470546505.index.

[106] M. C. Hall und Da Hodges, „Experience with 500 kV subsynchronous resonance and resulting turbine generator shaft damage at Mohave generating station“, *Analysis and control of subsynchronous resonance*, S. 22–29, 1976.

[107] H. U. Banna, A. Luna, S. Ying, H. Ghorbani und P. Rodriguez, „Impacts of wind energy in-feed on power system small signal stability“ in *2014 International Conference on Renewable Energy Research and Application (ICRERA)*, 2014, S. 615–622, doi: 10.1109/ICRERA.2014.7016459.

[108] H. A. Mohammadpour und E. Santi, „Analysis of subsynchronous control interactions in DFIG-based wind farms: ERCOT case study“ in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2015, S. 500–505, doi: 10.1109/ECCE.2015.7309730.

[109] D. Pejovski, A. Di Gerlando, G. M. Foglia und R. Perini, „Electrical Drive in/around Torsional Resonance Analytical Model“ in *2022 International Conference on Electrical Machines (ICEM)*, Valencia, Spain, 2022, S. 2193–2199, doi: 10.1109/ICEM51905.2022.9910627.

[110] Jan Shair, Xiaorong Xie, Luping Wang, Wei Liu, Jingbo He und Hui Liu, „Overview of emerging subsynchronous oscillations in practical wind power systems“, *Renewable and Sustainable Energy Reviews*, Jg. 99, S. 159–168, 2019, doi: 10.1016/j.rser.2018.09.047.

[111] S. Velpula, R. Thirumalaivasan und M. Janaki, „A review on SubSynchronous Resonance and its mitigation techniques in DFIG based wind farms“, *International Journal of Renewable Energy Research*, Jg. 8, S. 2275–2288, 2018.

[112] F. Blaabjerg, R. Teodorescu, M. Liserre und A. V. Timbus, „Overview of Control and Grid Synchronization for Distributed Power Generation Systems“, *IEEE Trans. Ind. Electron.*, Jg. 53, Nr. 5, S. 1398–1409, 2006, doi: 10.1109/TIE.2006.881997.

[113] X. Wang und F. Blaabjerg, „Harmonic Stability in Power Electronic-Based Power Systems: Concept, Modeling, and Analysis“, *IEEE Trans. Smart Grid*, Jg. 10, Nr. 3, S. 2858–2870, 2019, doi: 10.1109/TSG.2018.2812712.

[114] M. Paolone *et al., Fundamentals of Power Systems Modelling in the Presence of Converter-Interfaced Generation*, 2020.

[115] X. Wang, F. Blaabjerg und W. Wu, „Modeling and Analysis of Harmonic Stability in an AC Power-Electronics-Based Power System“, *IEEE Trans. Power Electron.*, Jg. 29, Nr. 12, S. 6421–6432, 2014, doi: 10.1109/TPEL.2014.2306432.

[116] J. R. Massing, G. G. Koch, A. T. Pereira, H. Pinheiro, V. F. Montagner und R. C. L. F. Oliveira, „Modeling and stability analysis of current controllers for multi-paralleled grid-connected converters with LCL-filter“ in *2017 IEEE 18th Workshop on Control and Modeling for Power Electronics (COMPEL)*, Stanford, CA, USA, 72017, S. 1–7, doi: 10.1109/COMPEL.2017.8013328.

[117] J. He, Y. W. Li, D. Bosnjak und B. Harris, „Investigation and Active Damping of Multiple Resonances in a Parallel-Inverter-Based Microgrid“, *IEEE Trans. Power Electron.*, Jg. 28, Nr. 1, S. 234–246, 2013, doi: 10.1109/TPEL.2012.2195032.

[118] C. Yoon, H. Bai, R. N. Beres, X. Wang, C. L. Bak und F. Blaabjerg, „Harmonic Stability Assessment for Multiparalleled, Grid-Connected Inverters“, *IEEE Trans. Sustain. Energy*, Jg. 7, Nr. 4, S. 1388–1397, 2016, doi: 10.1109/TSTE.2016.2551737.

[119] W. Du *et al.,* „Modeling of Grid-Forming and Grid-Following Inverters for Dynamic Simulation of Large-Scale Distribution Systems“, *IEEE Trans. Power Delivery*, Jg. 36, Nr. 4, S. 2035–2045, 2021, doi: 10.1109/TPWRD.2020.3018647.

[120] Abu Dhabi, *Grid codes for renewable powered systems*. IRENA(2022) - International Renewable Energy Agency, 2022. [Online]. Verfügbar unter: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2022/Apr/IRENA_Grid_Codes_Renewable_Systems_2022.pdf?rev=986f108cbe5e47b98d17fca93eee6c86

[121] T. Suntio, T. a. Messo und J. a. Puukko, *Power Electronic Converters: Dynamics and Control in Conventional and Renewable Energy Applictions / Teuvo Suntio, Tuomas Messo, Joonas Puukko,* 1. Aufl., 2017.

[122] M. Prodanovic und T. C. Green, „Control and filter design of three-phase inverters for high power quality grid connection“, *IEEE Trans. Power Electron.*, Jg. 18, Nr. 1, S. 373–380, 2003, doi: 10.1109/TPEL.2002.807166.

[123] J. Hamman und F. S. van der Merwe, „Voltage harmonics generated by voltage-fed inverters using PWM natural sampling“, *IEEE Trans. Power Electron.*, Jg. 3, Nr. 3, S. 297–302, 1988, doi: 10.1109/63.17947.

[124] X. He, H. Geng und S. Ma, „Transient Stability Analysis of Grid-Tied Converters Considering PLL's Nonlinearity“, *CPSS TPEA*, Jg. 4, Nr. 1, S. 40–49, 2019, doi: 10.24295/CPSSTPEA.2019.00005.

[125] P. L. Dandeno und P. Kundur, „Practical application of eigenvalue techniques in the analysis of power system dynamic stability problems“, *Can. Electr. Eng. J.*, Jg. 1, Nr. 1, S. 35–46, 1976, doi: 10.1109/CEEJ.1976.6830824.

[126] J. Lunze, *Regelungstechnik 1: Systemtheoretische Grundlagen, Analyse und Entwurf einschleifiger Regelungen,* 12. Aufl. Berlin, Heidelberg: Springer Berlin Heidelberg; Springer Vieweg, 2020.

[127] O. Föllinger und F. Dörrscheidt, *Regelungstechnik: Einführung in die Methoden und ihre Anwendung,* 10. Aufl. Heidelberg: Hüthig, 2008.

[128] *Mathematik für Ingenieure,* 8. Aufl. Wiesbaden: Vieweg+Teubner Verlag, 1990.
[129] T. Pfeifer und P. Profos, *Handbuch der industriellen Meßtechnik: Mit 135 Tabellen,* 6. Aufl. München: Oldenbourg, 1994.
[130] N. S. Nise, *Control systems engineering,* 6. Aufl. Hoboken, N.J.: Wiley, 2011.
[131] Conseil international des grands réseaux électriques. Comité d'études C6, „Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources: Task Force C6.04“ in *CIGRÉ, 2014*, S. 101.
[132] Ali Abdul und Razzaq Altahir, „Park and Clark Transformations Park and Clark Transformations: A Short Review“, 2020.
[133] C. i. g. r. Des électriques, *Guide for the development of models for HVDC converters in a HVDC grid*. Paris: CIGRÉ, 2014.
[134] T. Jiang, „Ein Beitrag zur Stabilitätsbetrachtung in Umrichter-dominierten Verteilernetzen“, Universitätsverlag Ilmenau, 2021.
[135] A. Bokhari *et al.,* „Experimental Determination of the ZIP Coefficients for Modern Residential, Commercial, and Industrial Loads“, *IEEE Trans. Power Delivery*, Jg. 29, Nr. 3, S. 1372–1381, 2014, doi: 10.1109/TPWRD.2013.2285096.
[136] E. Nashrullah und A. Halim, „Polynomial Load Model Development for Analysing Residential Electric Energy Use Behaviour“, *MATEC Web Conf.*, Jg. 218, S. 1007, 2018, doi: 10.1051/matecconf/201821801007.
[137] S. S. Ram, S. B. Daram, P. S. Venkataramu und M. S. Nagaraj, „Analysis of Exponential and Polynomial Load models using Newton-Raphson Method with Hybrid Power Flow Controller“, *International Journal of Control and Automation*, Jg. 11, Nr. 11, S. 43–56, 2018.
[138] K. Yamashita und S. Martinez Villanueva, „Initial Results of International Survey on Industrial Practice on Power System Load Modelling Conducted by CIGRE WG C4.605“, *CIGRE 2011 Bologna Symposium - The Electric Power System of the Future: Integrating Supergrids and Microgrids*, 2011.
[139] S. Singh und S. Singh, „Optimal PMU Placement in Power System Considering the Measurement Redundancy“, *Advance in Electronic and Electric Engineering*, Jg. 4, S. 593–598, 2014.
[140] G. Valente, A. L. Castellanos, L. Hausfeld, F. de Martino und E. Formisano, „Cross-validation and permutations in MVPA: Validity of permutation strategies and power of cross-validation schemes“ (eng), *NeuroImage*, Jg. 238, S. 118145, 2021, doi: 10.1016/j.neuroimage.2021.118145.
[141] C. Chen, L. Wang, R. B. Myneni und D. Li, „Attribution of Land-Use/Land-Cover Change Induced Surface Temperature Anomaly: How Accurate Is the First-Order Taylor Series Expansion?“, *Journal of Geophysical Research: Biogeosciences*, Jg. 125, Nr. 9, e2020JG005787, 2020.
[142] V. Kaura und V. Blasko, „Operation of a phase locked loop system under distorted utility conditions“, *IEEE Trans. on Ind. Applicat.*, Jg. 33, Nr. 1, S. 58–63, 1997, doi: 10.1109/28.567077.
[143] C. Chicone, *Ordinary Differential Equations with Applications,* 2. Aufl. New York, NY: Springer New York, 2006. [Online]. Verfügbar unter: http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10531325
[144] Siddhartha Khaitan, *A Survey Of Techniques for using Neural Networks in Power Systems*, 2017. [Online]. Verfügbar unter: https://hal.archives-ouvertes.fr/hal-01631454/

[145] H. Bevrani, B. François und T. Ise, *Microgrid dynamics and control,* 1. Aufl. Hoboken, New Jersey: John Wiley & Sons, 2017.

[146] R. Luhtala *et al.,* „Identification of Three-Phase Grid Impedance in the Presence of Parallel Converters“, *Energies*, Jg. 12, Nr. 14, S. 2674, 2019, doi: 10.3390/en12142674.

[147] Y. Cho, K. Hur, Y. Kang und E. Muljadi, „Impedance-Based Stability Analysis in Grid Interconnection Impact Study Owing to the Increased Adoption of Converter-Interfaced Generators“, *Energies*, Jg. 10, Nr. 9, S. 1355, 2017, doi: 10.3390/en10091355.

[148] T. Degris, M. White und R. Sutton, „Off-Policy Actor-Critic“, *Proceedings of the 29th International Conference on Machine Learning, ICML 2012*, Jg. 1, 2012.

[149] T. Degris, P. M. Pilarski und R. S. Sutton, „Model-Free reinforcement learning with continuous action in practice“ in *2012 American Control Conference - ACC 2012*, Montreal, QC, 2012, S. 2177–2182, doi: 10.1109/ACC.2012.6315022.

[150] S. Saravanan, M. Rajan, V. ANANTHAPATHMANABAN, N. Sriraam und P. Murugan, „TELEMEDICINE USING COMPUTER COMMUNICATION NETWORK“, 2022.

Appendix

A.1 Model parameters

The initial and optimized model parameters, e.g., the resistance R, the inductance L and the capacitance C of the overhead line, the cable, and the filter in the converter are given in Tab. A.1. The connections and lengths of lines and cables for the investigated system are given in Tab. A.2. The values of the peak loads for each bus of the CIGRE benchmark are shown in Tab. A.3. Tab. A.4 shows the maximum generation of three converters.

Tab. A.1 Model parameters

Component	Parameter	Initial	Reference	Optimized
Overhead line	ROL in Ω/km	0.051	0.51	0.418
	LOL in mH/km	0.5825	1.165	1.0136
	COL in µF/km	0.0101	0.0051	0.0061
Cable	RCL in Ω/km	0.0343	0.343	0.3053
	LCL in mH/km	0.436	0.872	0.672
	CCL in µF/km	0.151	0.0756	0.0862
Filter of CON	RF in Ω	0.001	0.01	0.0092
	LF in mH	6.50	13	10.66
	CF in µF	2	1	1.24
Current PI-controller in CON	Kp	0.0065	0.0032	0.0034
	Ki	0.00125	0.0025	0.0024
Voltage PI-controller in CON	Kp	0.204	0.4082	0.3756
	Ki	1020.62	510.31	581.75
Reactive power I-controller in CON	Ki	0.204	0.1021	0.0959

Tab. A.2 Connections and length of line and cable for the CIGRE benchmark system

From Bus	To Bus	Length in km	Installation	From Bus	To Bus	Length in km	Installation
1	2	2.82	Overhead line	9	10	0.77	Cable
2	3	4.42	Cable	10	11	0.33	Cable
3	4	0.61	Cable	11	4	0.49	Cable
4	5	0.56	Cable	3	8	1.30	Cable
5	6	1.54	Cable	12	13	4.89	Overhead line
6	7	0.24	Cable	13	14	2.99	Overhead line
7	8	1.67	Cable	14	8	2.00	Overhead line
8	9	0.32	Cable				

Tab. A.3 Load parameters for MV benchmark

Buss	Active Power [kW]		Power Factor	
	Residential	Residential	Residential	Commercial/ Industrial
1	14994	4998	0.98	0.98
2	-	-	-	-
3	276.45	225.25	0.97	0.85
4	431.65	-	0.97	-
5	727.5	-	0.97	-
6	548.05	-	0.97	-
7	-	76.5	-	0.85
8	586.85	-	0.97	-
9	-	573.75	-	0.85
10	475.3	68	0.97	0.85
11	329.8	-	0.97	-
12	14994	5016	0.98	0.95
13	-	34	-	0.85
14	208.55	331.5	0.97	0.85

Tab. A.4 converter parameters for MV benchmark

Bus	Power plant type	Active Power [MW]	Power Factor
12	PV	5	0.95
13	PV	5	0.95
14	Wind	5	0.95

A.2 Digital twin validation

1) Voltage step

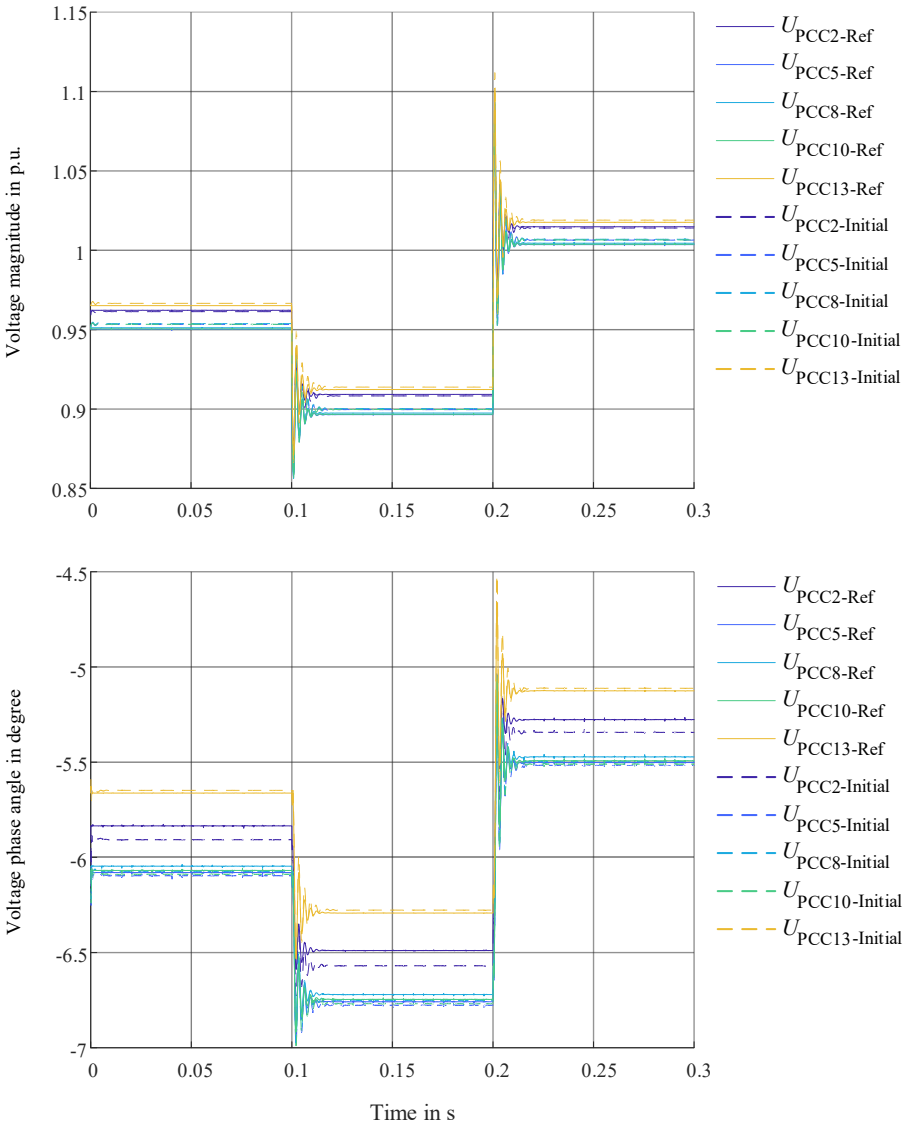


Fig. A.1: Voltage comparison under heavy load and low generation between reference network and DT model with optimized estimated parameters

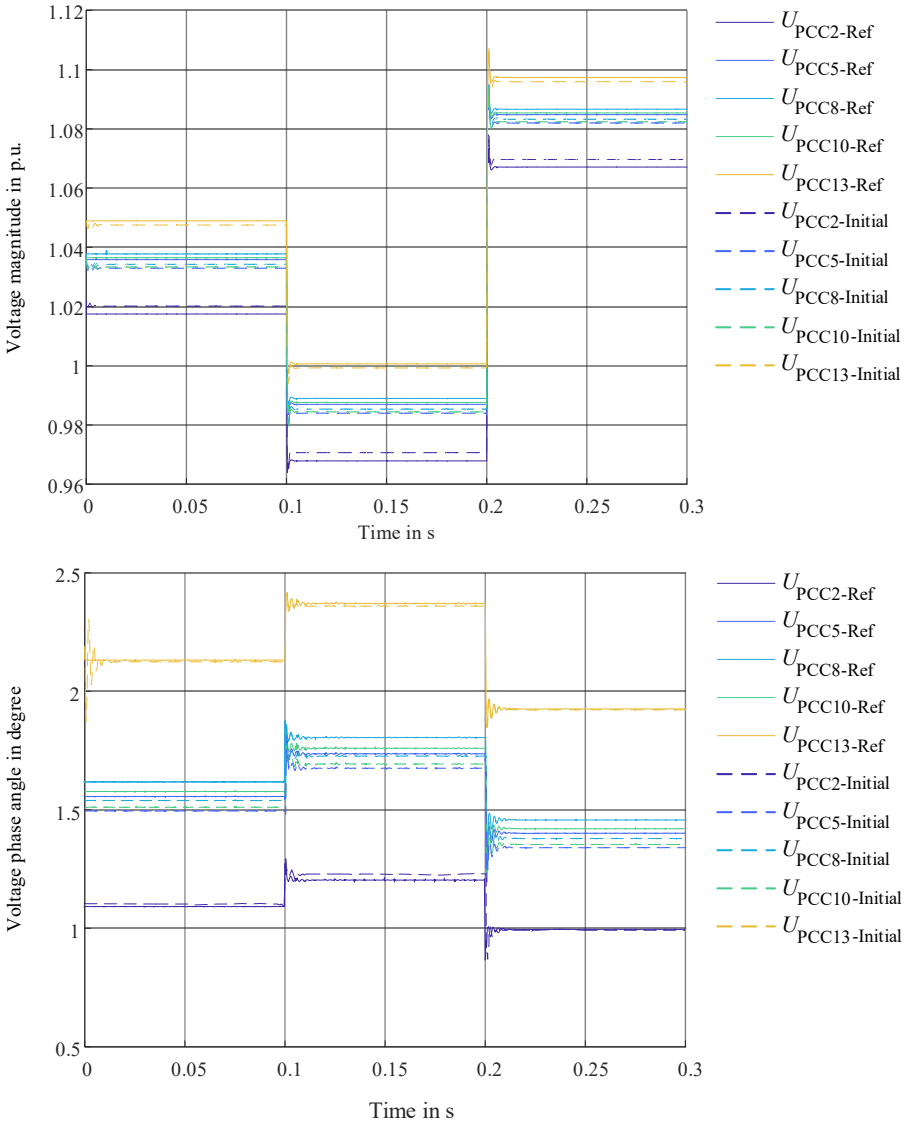


Fig. A.2: Voltage comparison under low load and high generation between reference network and DT model with optimized estimated parameters

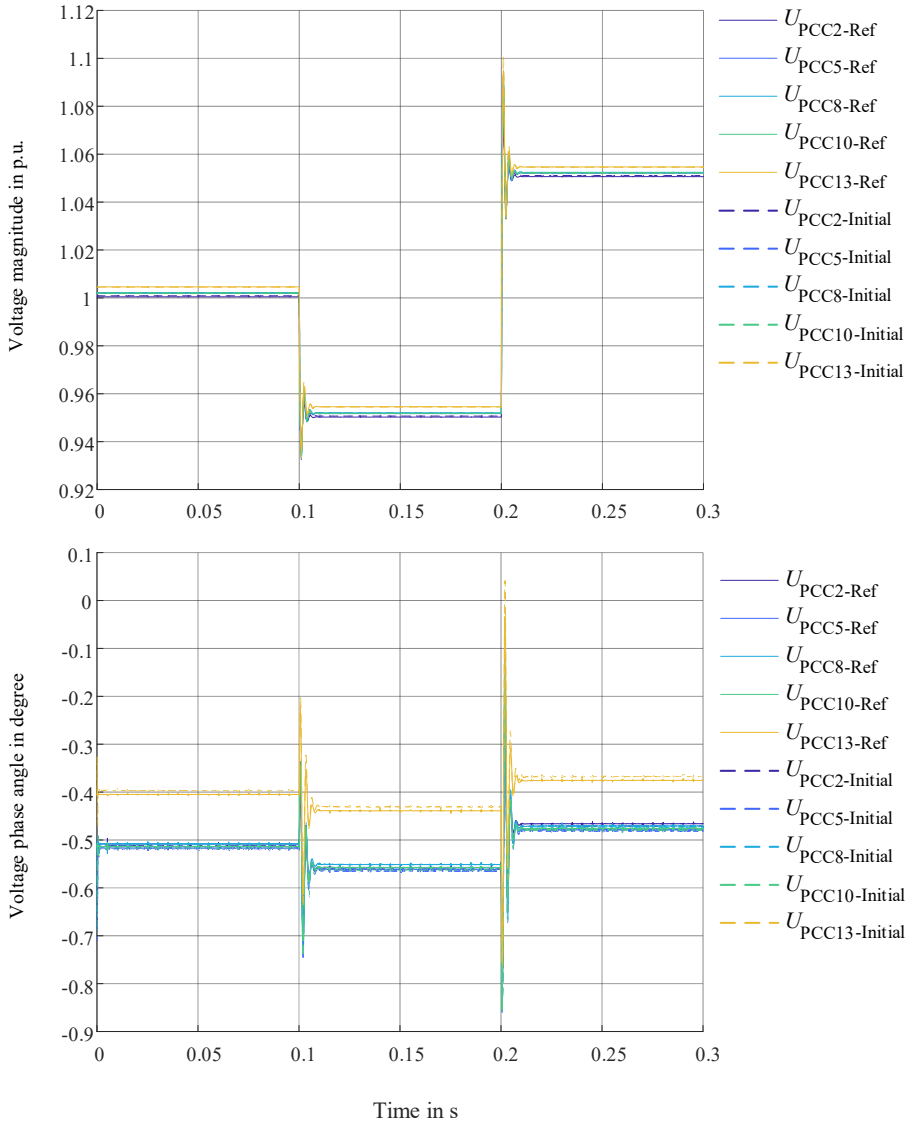


Fig. A.3: Voltage comparison under low load and low generation between reference network and DT model with optimized estimated parameters

2) P, Q step response

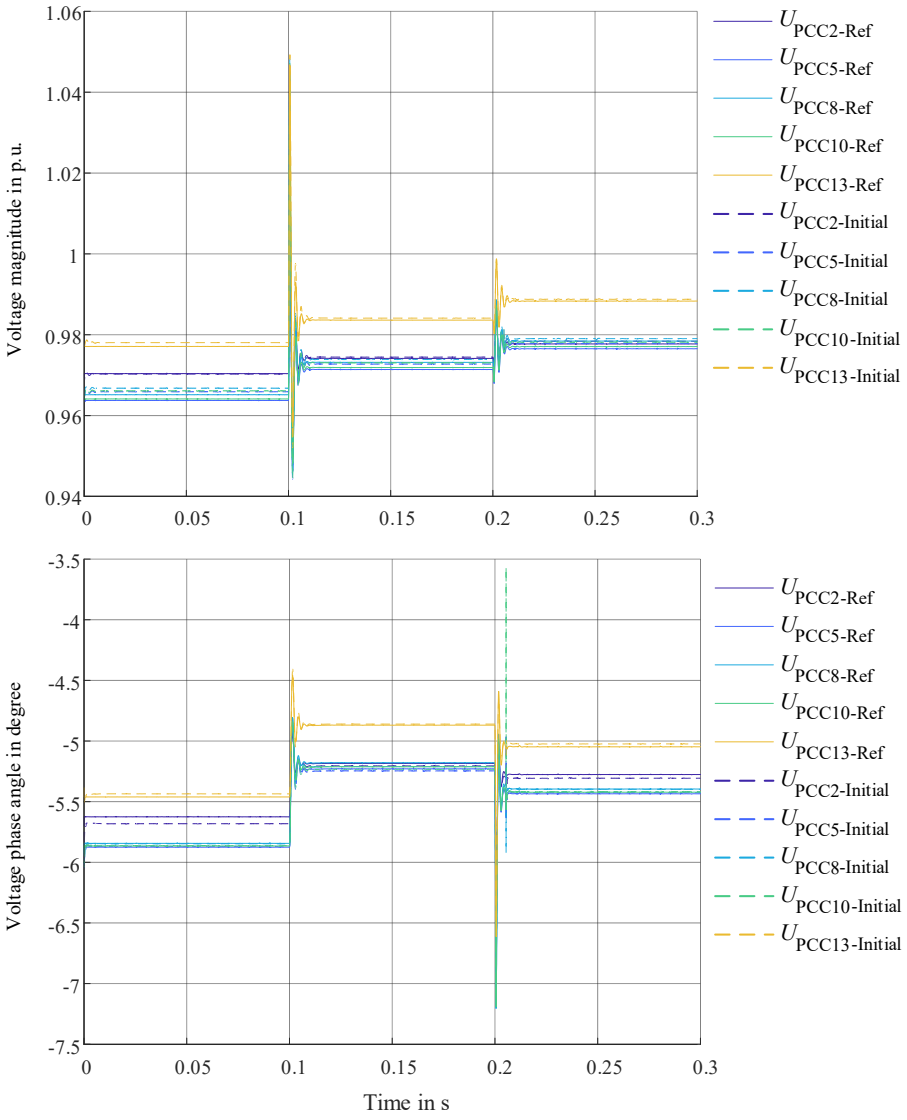


Fig. A.4: Voltage comparison under heavy load and low generation between reference network and DT model with optimized estimated parameters

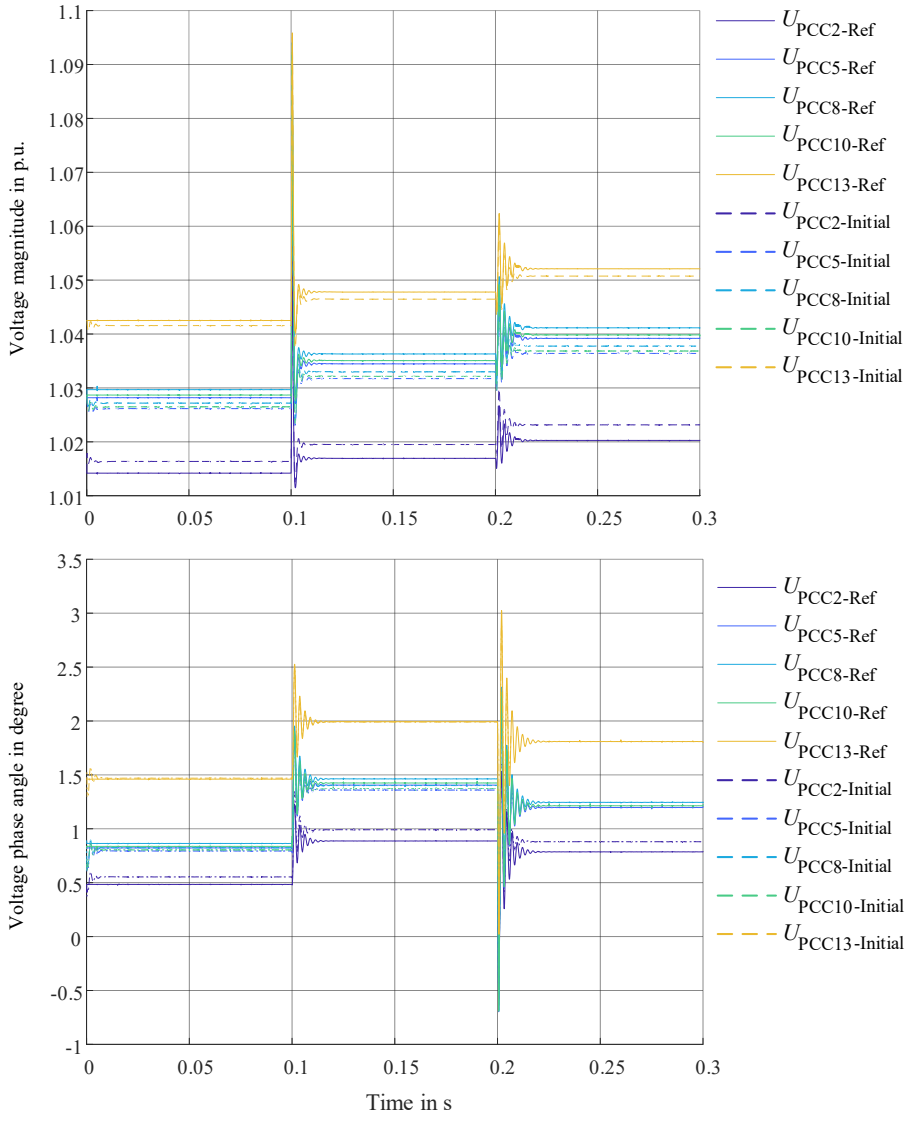


Fig. A.5: Voltage comparison under low load and high generation between reference network and DT model with optimized estimated parameters

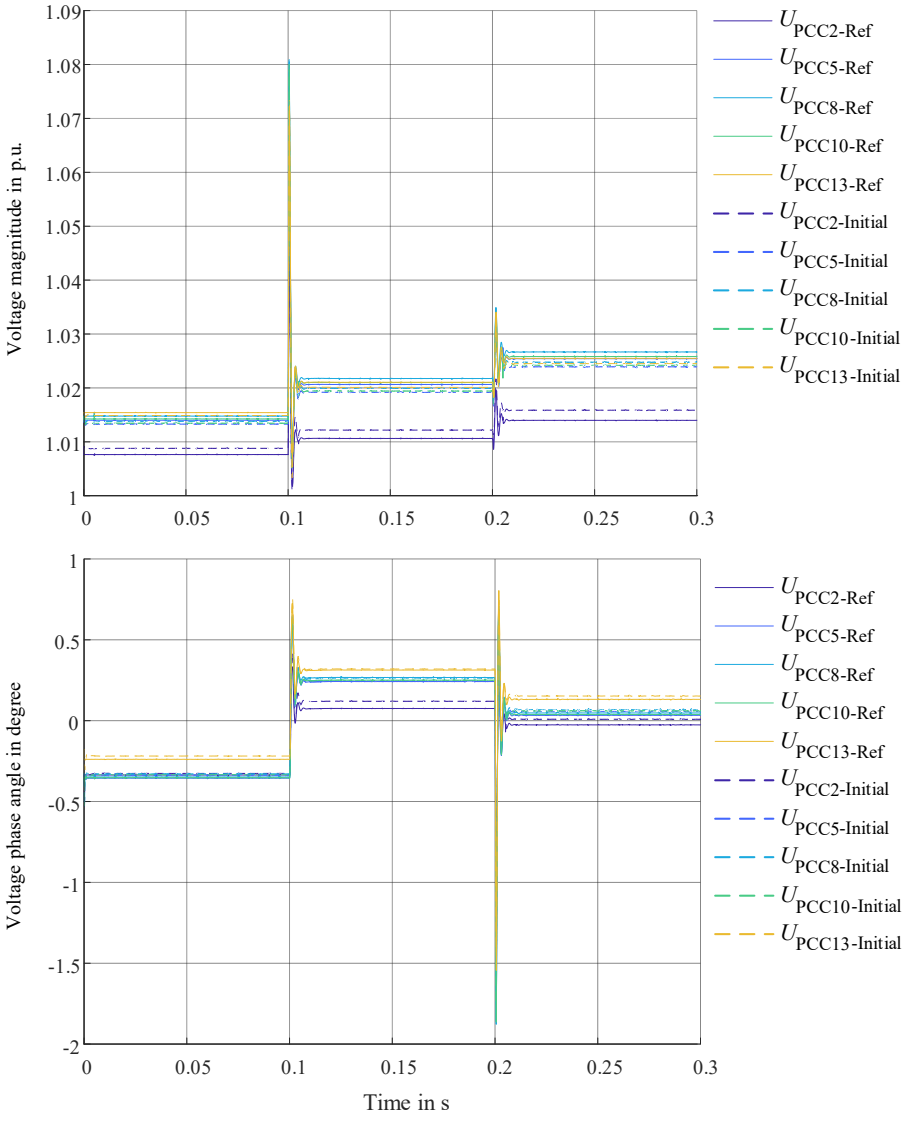


Fig. A.6: Voltage comparison under low load and low generation between reference network and DT model with optimized estimated parameters

3) Voltage step

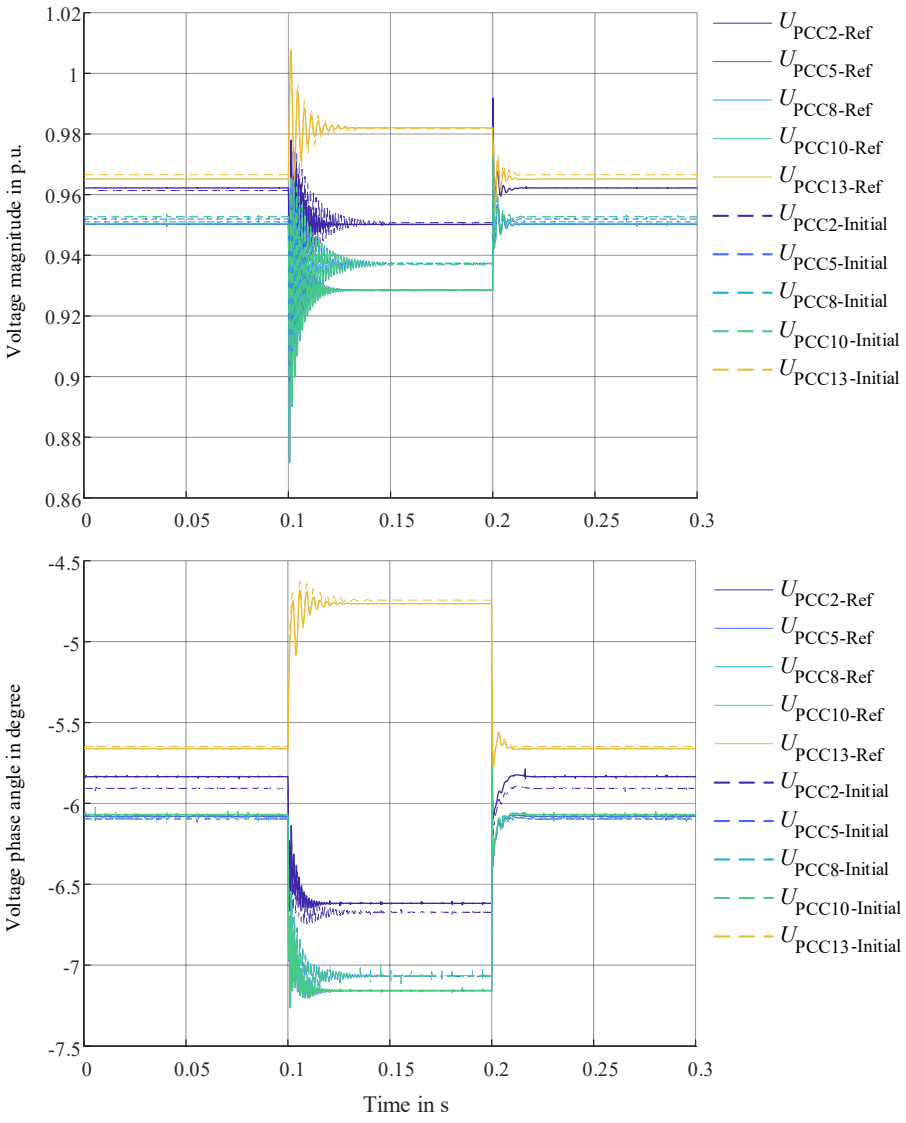


Fig. A.7: Voltage comparison under heavy load and low generation between reference network and DT model with optimized estimated parameters

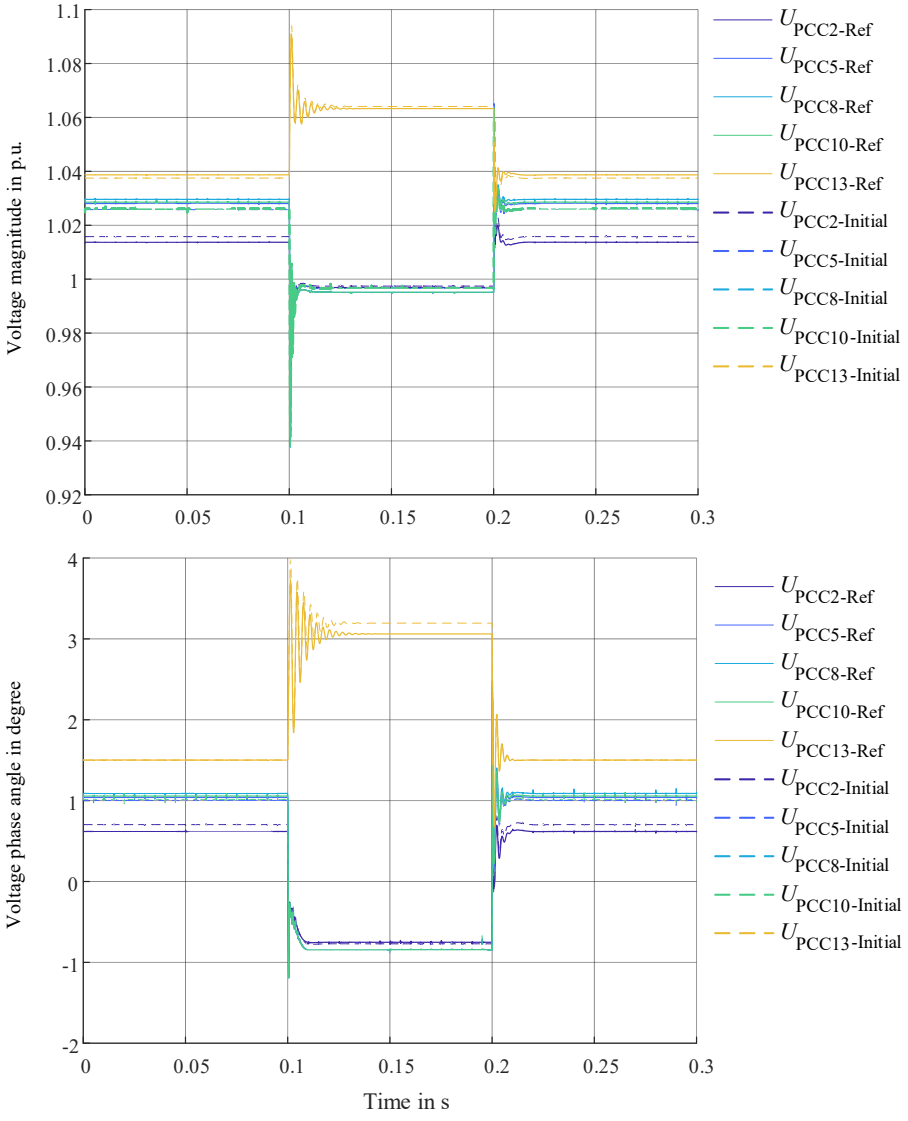


Fig. A.8: Voltage comparison under low load and high generation between reference network and DT model with optimized estimated parameters

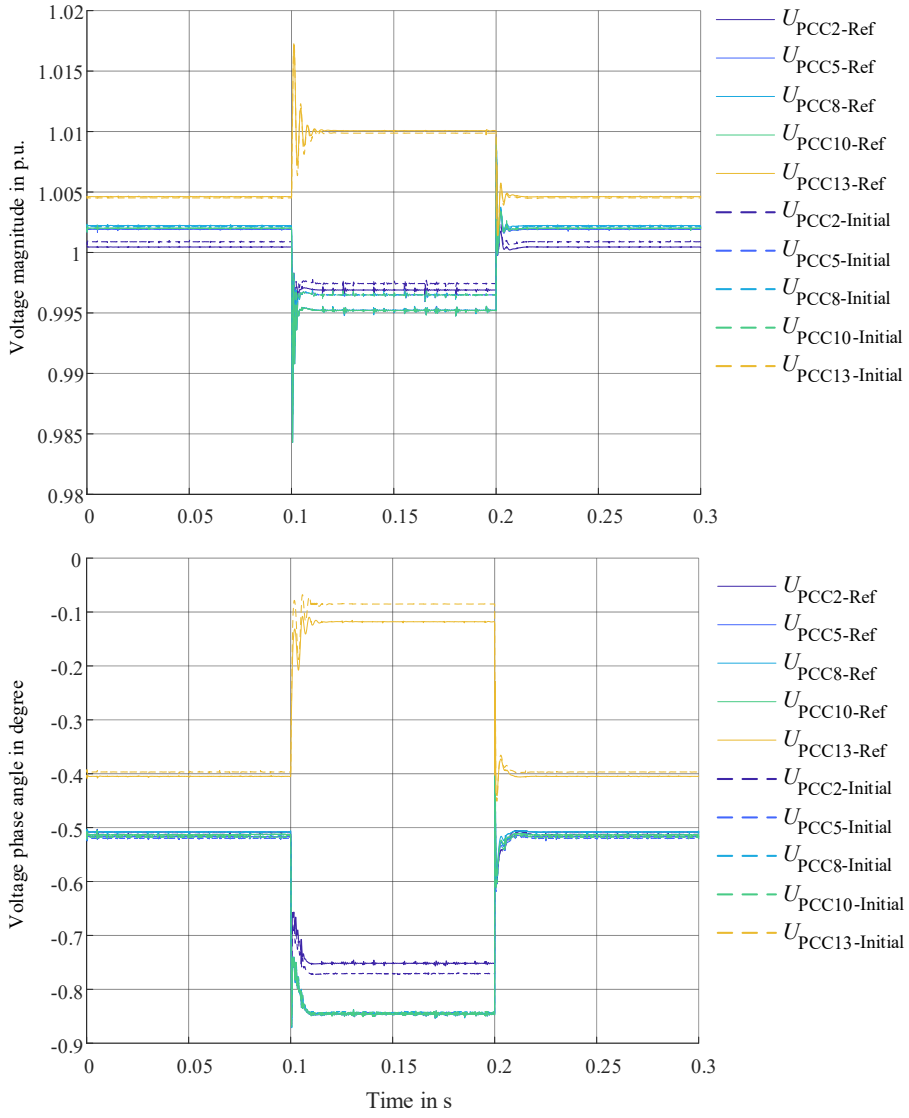


Fig. A.9: Voltage comparison under low load and low generation between reference network and DT model with optimized estimated parameters

4) Short circuit at bus 14

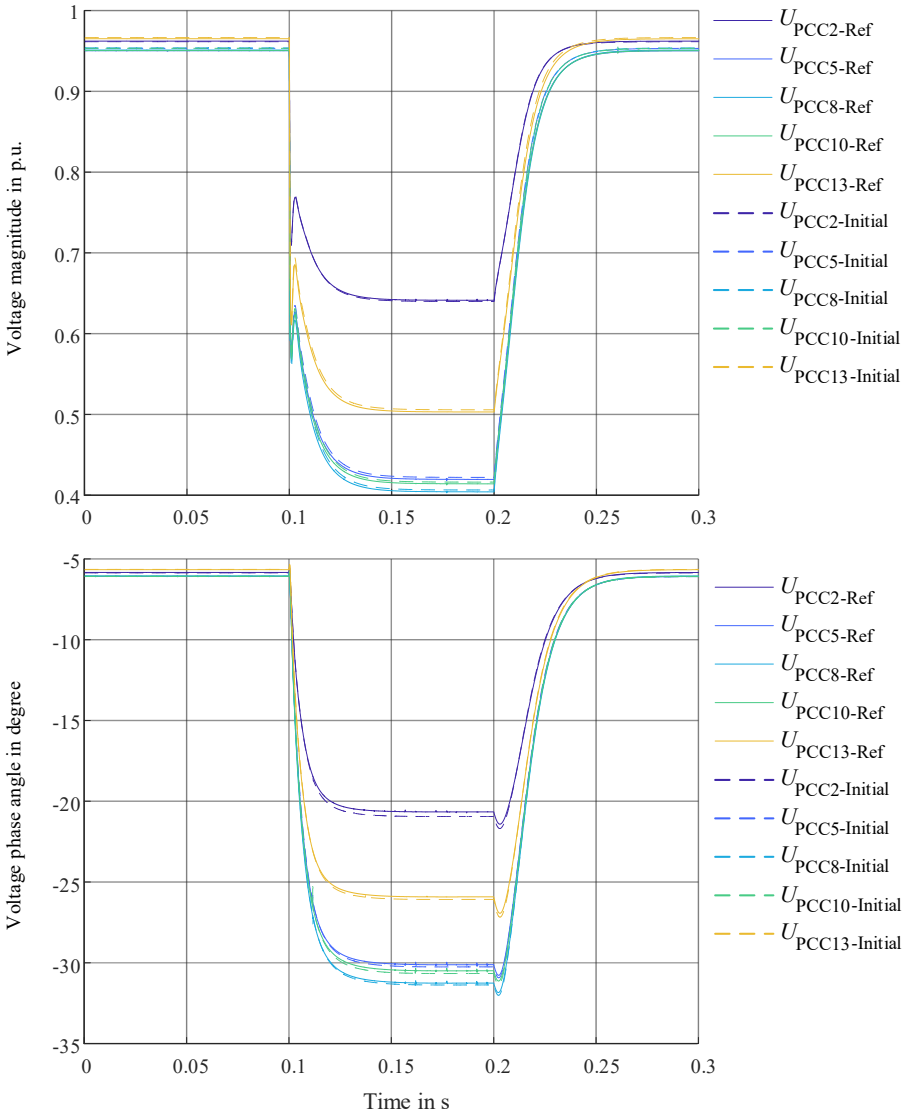


Fig. A.10: Voltage comparison under heavy load and low generation between reference network and DT model with optimized estimated parameters

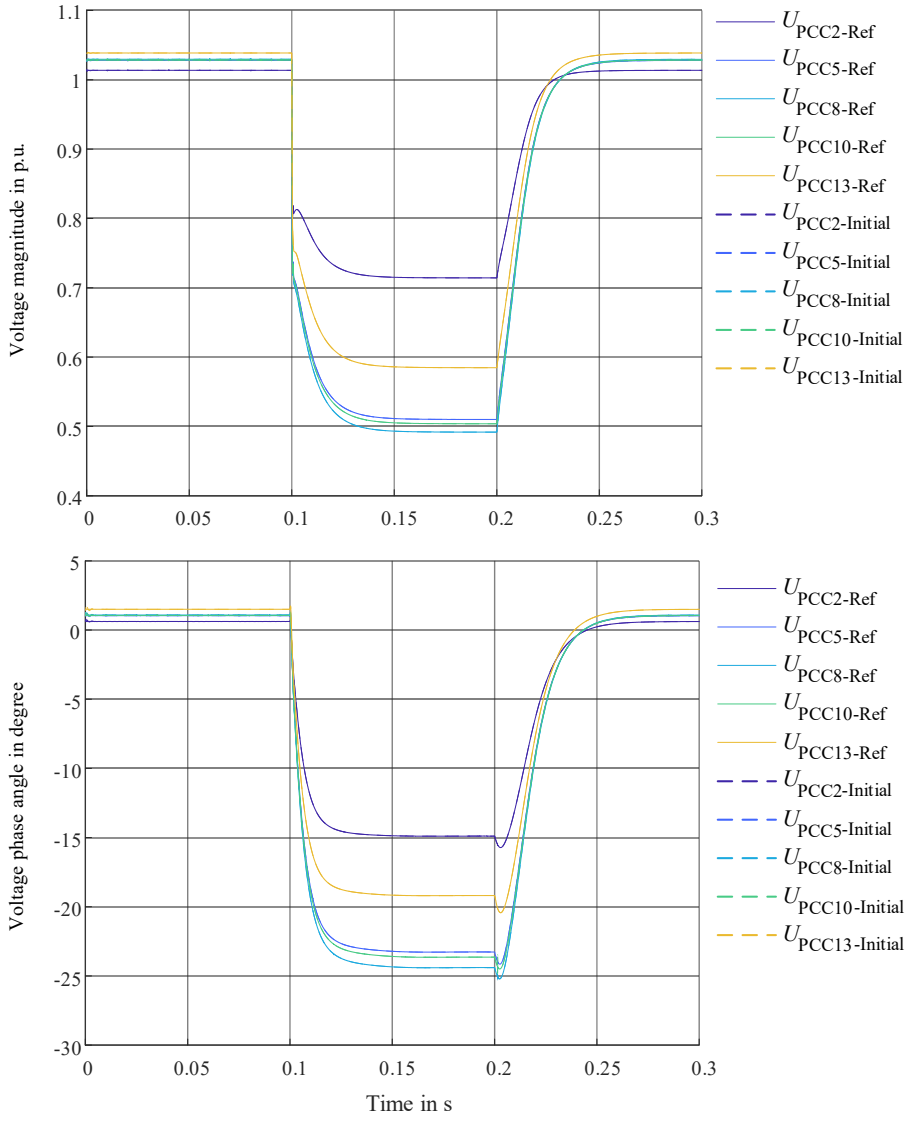


Fig. A.11: Voltage comparison under low load and high generation between reference network and DT model with optimized estimated parameters

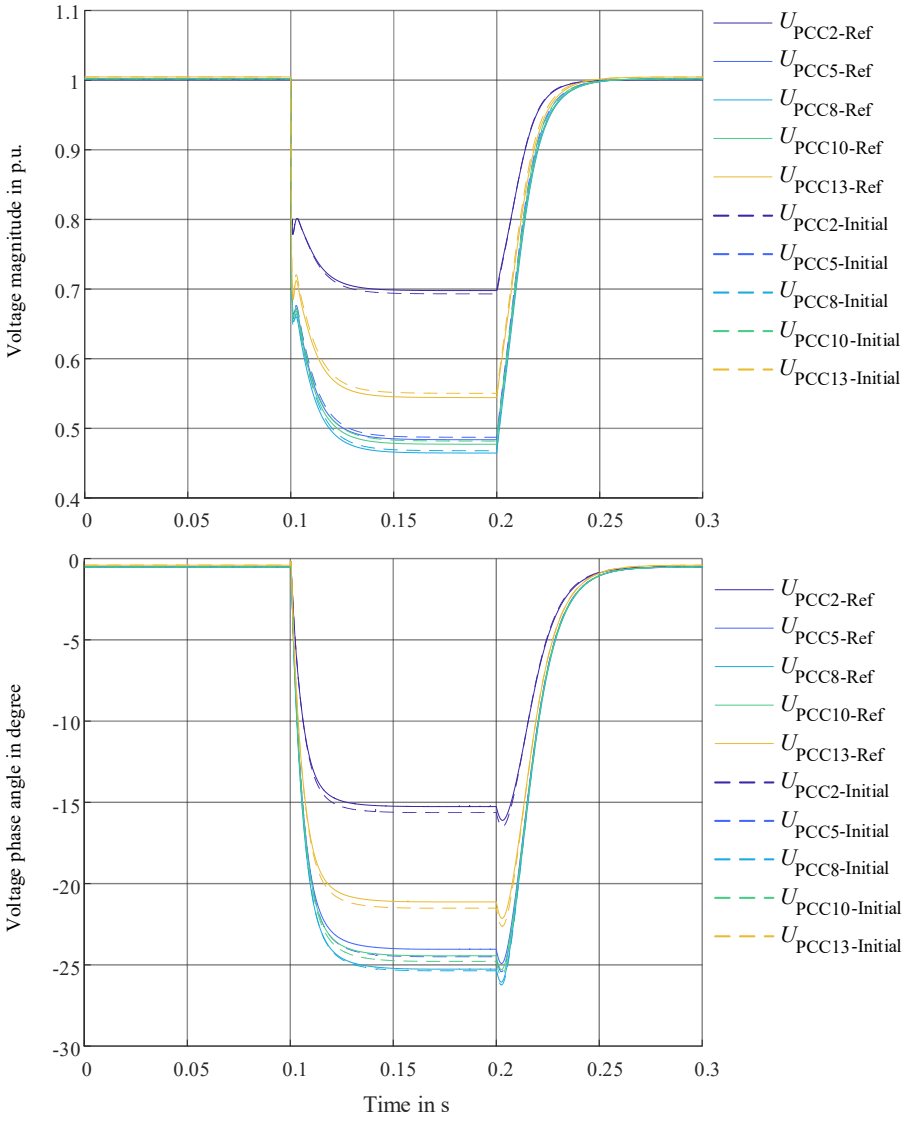


Fig. A.12: Voltage comparison under low load and low generation between reference network and DT model with optimized estimated parameters

A.3 Physical variables representing damping mode

To allow the dynamic response of active and reactive power in converter CON1 to be seen when a small disturbance of 1W is injected to a given value of the active power, the following figure Fig. A.13 shows both P and Q minus their stable values before the disturbance is injected.

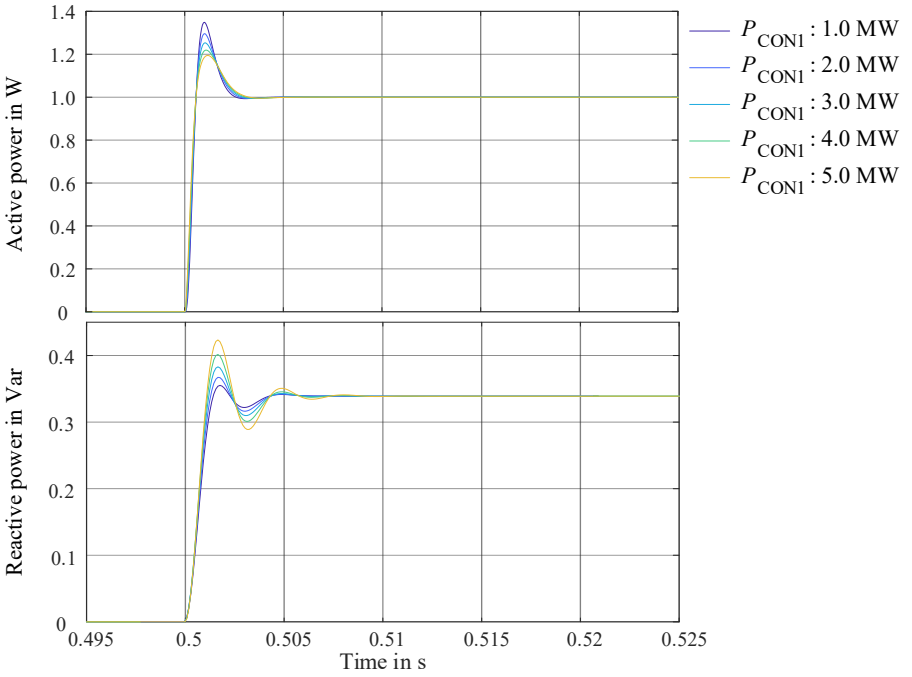


Fig. A.13: Active power and reactive power of converter CON1 for small signal response

The FFT analyses of four signals for active power generations of 3 MW are performed within a time window which starts with a small signal response of 0.5 s and ends within two time periods at the system frequency of 50 Hz. The results are shown from Fig. A.14 to Fig. A.17.

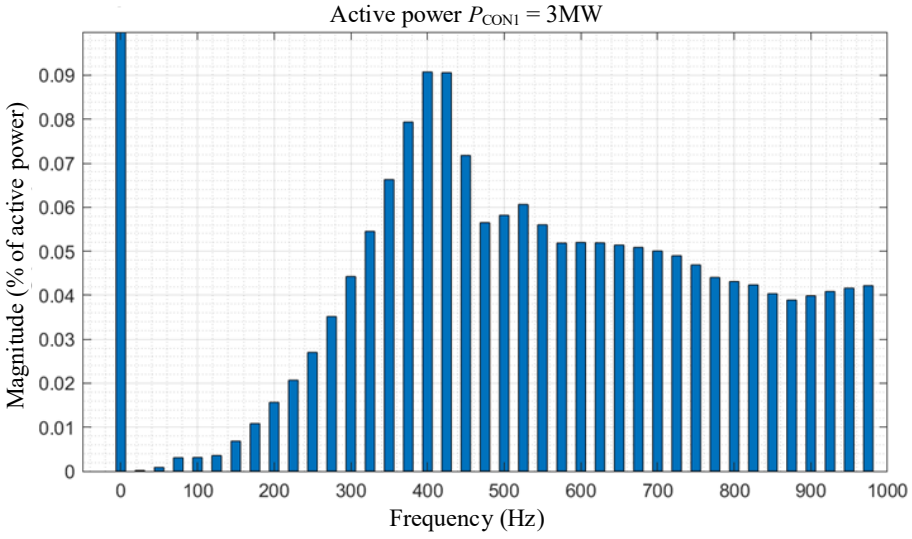


Fig. A.14: FFT result of small signal response for active power of converter CON1

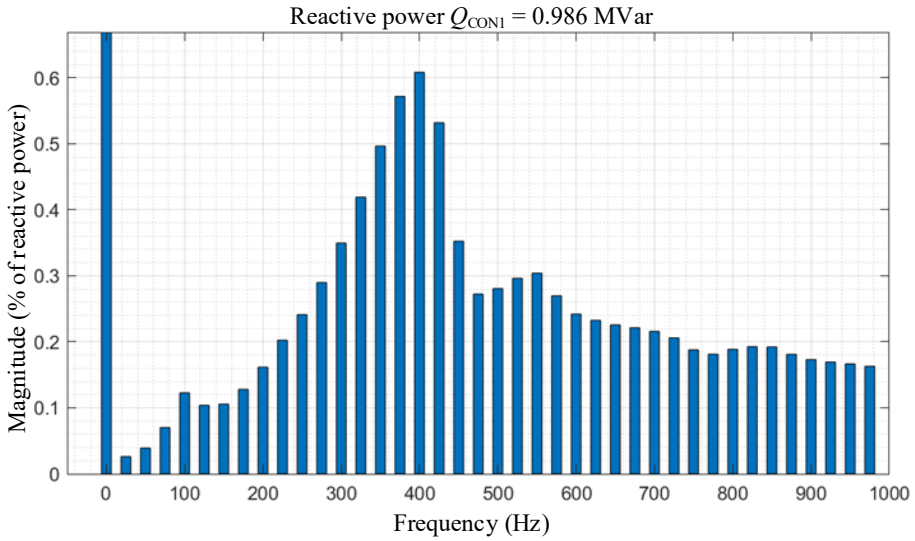


Fig. A.15: FFT result of small signal response for reactive power of converter CON1

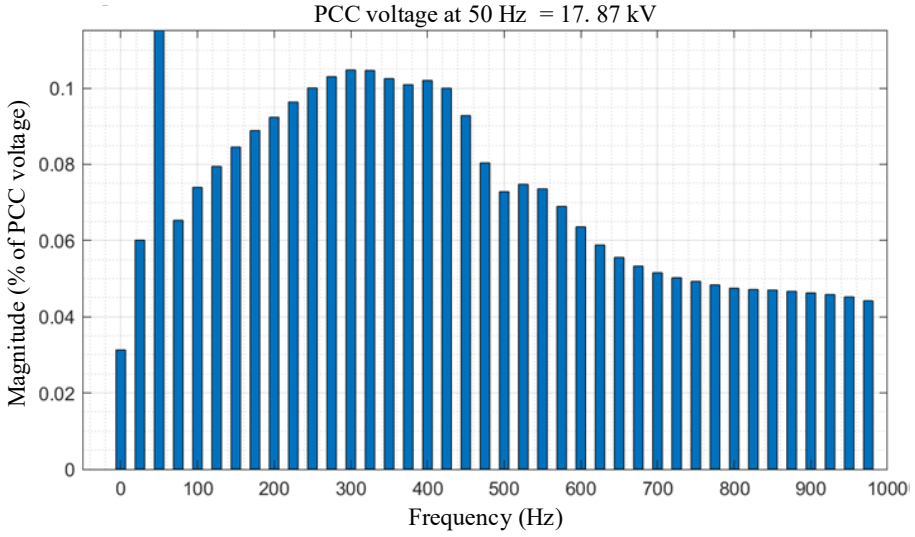


Fig. A.16: FFT result of small signal response for PCC voltage at bus 12

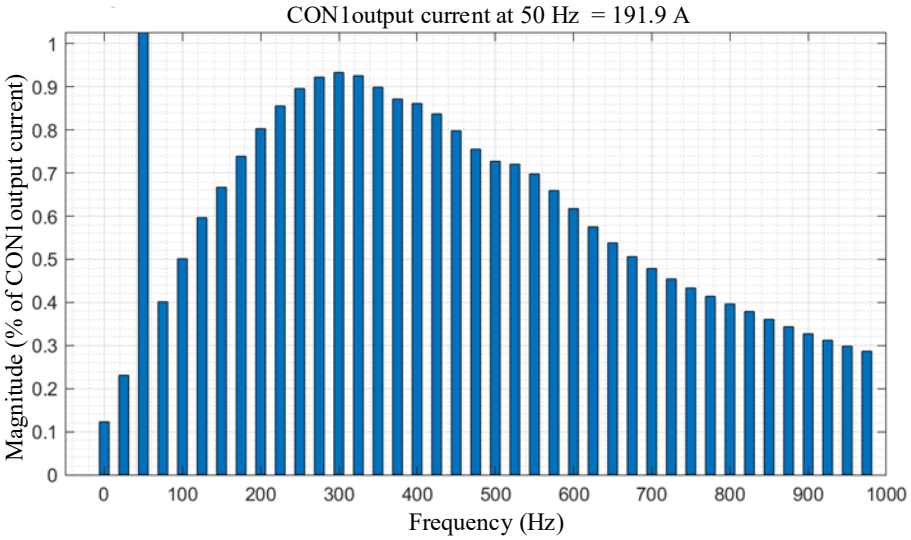


Fig. A.17: FFT result of small signal response for converter CON1 output current

A.4 Parameter sensitivity

Tab. A.5 SI values under different controller parameters setting

SI	Controller parameters	-40%	-20%	0%	+20%	+40%
CON1	K_{pi}	0.1179	0.1455	0.1575	0.1640	0.1681
	K_{ii}	0.1583	0.1579		0.1570	0.1566
	K_{pu}	0.1374	0.1516		0.1609	0.1631
	K_{iu}	0.1647	0.1614		0.1536	0.1511
	K_{p_PLL}	0.1575	0.1575		0.1574	0.1574
	K_{i_PLL}	0.1574	0.1574		0.1575	0.1575
	K_{iQ}	0.1659	0.1603		0.1570	0.1578
CON2	K_{pi}	0.0859	0.1321		0.1723	0.1813
	K_{ii}	0.1600	0.1587		0.1562	0.1549
	K_{pu}	0.1051	0.1439		0.1654	0.1713
	K_{iu}	0.1758	0.1686		0.1453	0.1352
	K_{p_PLL}	0.1576	0.1575		0.1574	0.1573
	K_{i_PLL}	0.1574	0.1574		0.1575	0.1575
	K_{iQ}	0.1787	0.1677		0.1555	0.1615
CON3	K_{pi}	0.0882	0.1296		0.1782	0.1943
	K_{ii}	0.1606	0.1590		0.1559	0.1543
	K_{pu}	0.1079	0.1441		0.1655	0.1718
	K_{iu}	0.1773	0.1692		0.1450	0.1344
	K_{p_PLL}	0.1578	0.1576		0.1573	0.1572
	K_{i_PLL}	0.1574	0.1574		0.1575	0.1575
	K_{iQ}	0.1849	0.1728		0.1604	0.1831

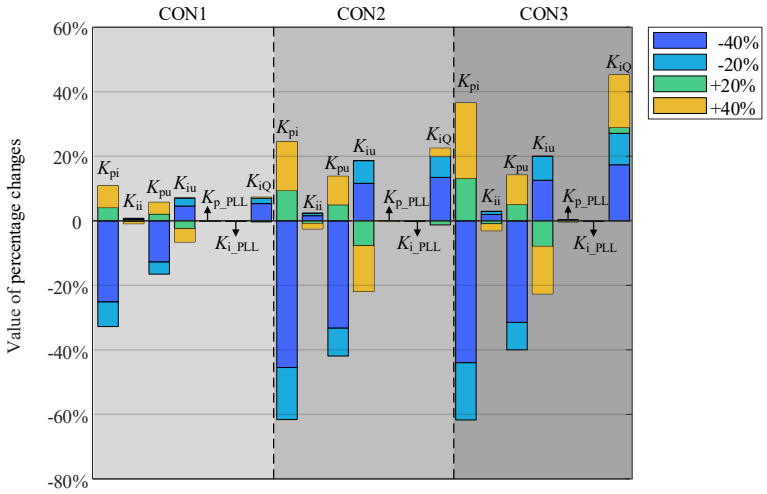


Fig. A.18: SI percentage changes under different parameter variations

A.5 Data of system dominant eigenvalue data

Tab. A.6 System dominant eigenvalue under different network states without RL agent

***Vsource V_s* [p.u.]**		**0.9**	**1**	**1.1**
Real part		-312.6	-396.5	-449.1
Imaginary part		2345.7	2486.7	2615.6
***Active power P* [MW] of CON3**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-685.2	-396.5	-236.8
	Imaginary part	2458.7	2486.7	2509.3
Power factor PF		**0.85**	**0.90**	**0.95**
Dominant eigenvalue	**Real part**	-276.4	-396.5	-512.2
	Imaginary part	2592.1	2486.7	2385.6
***Active power P* of three converters [MW]**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-943.2	-396.5	22.3
	Imaginary part	2470.7	2486.7	2482.3
Load level		**40%**	**60%**	**80%**
Dominant eigenvalue	**Real part**	-467.2	-446.8	-426.1
	Imaginary part	2635.4	2546.2	2453.9

Tab. A.7 System dominant eigenvalue under different network states with RL agent$_1$

***Vsource V_s* [p.u.]**		**0.9**	**1**	**1.1**
Real part		-972.4	-991.4	-1012.6
Imaginary part		2923.4	2850.9	2706.6
***Active power P* [MW] of CON3**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-1075.1	-991.4	-915.4
	Imaginary part	2957.0	2850.9	2734.3
Power factor PF		**0.85**	**0.90**	**0.95**
Dominant eigenvalue	**Real part**	-879.3	-991.4	-1110.5
	Imaginary part	2737.8	2850.9	2984.3
***Active power P* of three converters [MW]**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-1175.8	-991.4	-829.1
	Imaginary part	2895.9	2850.9	2809.3
Load level		**40%**	**60%**	**80%**
Dominant eigenvalue	**Real part**	-1004.4	-998.5	-993.4
	Imaginary part	2859.2	2855.6	2852.9

Tab. A.8 System dominant eigenvalue under different network states with RL agent$_2$

Vsource V_s **[p.u.]**		**0.9**	**1**	**1.1**
Real part		-930.7	-938.4	-948.2
Imaginary part		4140.8	4027.1	3895.5
Active power P **[MW] of CON3**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-1084.1	-938.4	-837.1
	Imaginary part	4424.4	4027.1	3879.9
Power factor PF		**0.85**	**0.90**	**0.95**
Dominant eigenvalue	**Real part**	-896.6	-938.4	-1075.2
	Imaginary part	4129.4	4027.1	3882.3
Active power P **of three converters [MW]**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-1435.8	-938.4	-542.4
	Imaginary part	3962.4	4027.1	3962.3
Load level		**40%**	**60%**	**80%**
Dominant eigenvalue	**Real part**	-858.4	-808.1	-778.2
	Imaginary part	4026.6	4025.3	4023.9

Tab. A.9 System dominant eigenvalue under different network states with RL agent$_3$

Vsource V_s **[p.u.]**		**0.9**	**1**	**1.1**
Real part		-373.5	-413.4	-439.4
Imaginary part		2514.7	2425.5	2317.7
Active power P **[MW] of CON3**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-481.6	-413.4	-227.9
	Imaginary part	2813.6	2425.5	1757.5
Power factor PF		**0.85**	**0.90**	**0.95**
Dominant eigenvalue	**Real part**	-359.8	-413.4	-446.9
	Imaginary part	2505.1	2425.5	2355.3
Active power P **of three converters [MW]**		**1**	**3**	**5**
Dominant eigenvalue	**Real part**	-661.0	-413.4	-63.9
	Imaginary part	2596.2	2425.5	2274.9
Load level		**40%**	**60%**	**80%**
Dominant eigenvalue	**Real part**	-534.4	-501.3	-453.6
	Imaginary part	2471.3	2459.0	2442.5

Abbreviation

AC	alternating current
ANN	artificial neural network
CI-REG	converter-interfaced renewable energy generators
CPL	constant power P load
CON	converter
DT	digital twin
DC	direct current
DT	digital twin
DZ	digitaler Zwilling
EMT	electromagnetic transient
IC	intermediate circuit
MAPE	mean absolute percentage error
ML	machine learning
MO	magnitude optimum
MDP	Markov Decision Proces
MV	medium-voltage
NN	neural network
PLL	phase-lock- loop
PCC	point of common coupling
KNN	künstliche neuronale Netz
PF	power factor
PSS	power system stabilizer
PWM	pulse-width modulation
RL	reinforcement learning
RE	renewable energies
RHP	right half-plane
SCG	scaled conjugate gradien
SI	stability indicator / Stabilitätsindikator
SOD	stability optimization degree
SSR	subsynchronous resonance
SSM	state space model
SR	Stromrichter
SO	symmetric optimum
TD	temporal difference
TRR	Trust-Region-Reflective

Notation

Notation

The notation used in this work is described below using the letter "a" as an example. All variables are represented by italic letters with serif (*a*). A matrix is marked with a bold capital letter (***A***). Vectors are represented by a bold lower case letter (***a***). Scalar variables are represented in physical quantities as upper case (*A*) and in related quantities as lower case (*a*).

List of Symbol

a	Decision parameters of the phase reserve
A	System matrix
b	Bias
B	Input matrix
c	Finite number
C	Output matrix
C_{CON}	Converter capacity
C_{IC}	Intermediate circuit capacity
C_i	Cluster
d_{ik}	Square distance between data point y_k and cluster centre v_i
D	Penetration
D_{P}	Gain of the frequency control
D_{Q}	Amplification of the voltage regulation
e	Control deviation
e_i	Unit vectors
E	Global error
E_l	Errors for the training pattern *l*
E_{qw}	Quadratic approximation for *E*
E_j	Partial derivatives
$E(\underline{p})$	Sum of the squares of errors for model parameters
$E''\left(\underline{w}_k\right)$	Hessian matrix
f	Frequency
$F\left(\underline{p}\right)$	Sum of the error squares
$\underline{g}$	Gradient of $F\left(\underline{p}\right)$
g	Activation function

g_{step}	Binary step function
g_{RELU}	"ReLU" function
$g_{Sigmoid}$	Sigmoid function
$g_{\tanh}$	Tangent hyperbolic function
g'	Derivative
G	Transfer function of the closed control loop
G_{BO}	Open loop transfer function for BO method
G_{I}	Transfer function of the closed current control loop
$G_{\mathrm{I}}^{'}$	Approximated transfer function of the closed-loop current control loop
G_{o}	Transfer function of the open control loop
G_{OI}	Transfer function of the open current control loop
G_{OU}	Transfer function of the open voltage control loop
G_{Orig}	Transfer function of the original model
G_{P}	Transfer function of the process
G_{PI}	Transfer function of the process of the current control loop
G'_{PI}	Approximated process transfer function of the current control loop
G_{PU}	Transfer function of the process of the voltage control loop
G_{R}	Transfer function of the controller
G_{RI}	Transfer function of the controller in the current control loop
G_{RU}	Transfer function of the controller in the voltage control loop
G_{SI}	Transfer function of the path in the current control loop
G_{SO}	Transfer function of the open loop for SO method
G_{SU}	Transfer function of the path in the voltage control loop
G_{t}	Transfer function of the dead time element
G_{U}	Transfer function of the closed voltage control loop
G_{W}	Recognition accuracy
G_{Z}	Assignment accuracy
$\underline{h}_i$	Iteration steps
$\underline{h}^*$	Solution of the sub-problem of the trust region
$\underline{H}$	Hessian matrix of $F\left(\underline{p}\right)$
i_{c}	Capacity current
i_{f}	Filter current

i_{IC}	Intermediate circuit current
I_{d}	Current in d-axis
I_{q}	Current in q-axis
I_{Q}	Current source on the DC side
I	Current
$\underline{i}_{\mathrm{f}}$	Filter current
$J(\Theta)$	Performance metric
k	Gain
k'	Relative gain
K_{f}	Gain factor in the voltage control loop
K_{R}	Controller gain
K_{ii}	Integral parameter of current PI-controller
K_{iu}	Integral parameter of voltage PI-controller
$K_{\mathrm{i\_PLL}}$	Integral parameter of PI-controller in PLL
K_{iQ}	Integral parameter of reactive power I-controller
K_{pi}	Proportional parameter of current PI-controller
K_{pu}	Proportional parameter of voltage PI-controller
$K_{\mathrm{p\_PLL}}$	Proportional parameter of PI-controller in PLL
L	Total number of training patterns
$L(\underline{h})$	Quadratic function
L_{f}	Filter inductance
L_{K}	Cable inductance
L_{L}	Line inductance
L_{T}	Transformer inductance
$L_i(\Theta_i)$	Loss function
m	Weighting exponent
n	Total number of output eurons
N	Number of samples
$\boldsymbol{N}$	Confidence region
o_i	Actual output of the i-th output euron at pattern l as a function of the weights
o_j	Output of j-th individual neuron

$\underline{p}$	Model parameters
$\underline{p}_i$	Iteration point
$\underline{p}_k$	Current conjugate weight vector
P	Active power
$P(s)$	Denominator polynomial
q	Action reward function
Q	Reactive power
$rRMSE$	Relativ root-mean-squared error
$RMSE$	Root-mean-squared error
R	Reward
R_{F}	Filter resistor
R_{K}	Cable resistance
R_{L}	Line resistance
R_{T}	Transformer resistance
s	Laplace operator
S	Finite state set
t	Time
t_i	Nominal output of the i-th output euron at pattern l
t_{O}	Denominator time constant
t_{s}	Settling time
T_{O}	Numerator time constant
T	Delay time constant of the control system
T_{M}	Transformation matrix
T_1	Dominant delay time constant of the process
T_{I}	Time constant of the approximated transfer function of the current control loop
T_{R}	Integral time constant
T_{RI}	Integral time constant in the current control loop
T_{RU}	Integral time constant in the voltage control loop
T_{A}	Sampling period
T_{s}	Transit time constant
T_{SI}	Time constant of the path in the current control loop
T_{SU}	Time constant of the path in the voltage control loop
T_{t}	Dead time of the current control loop
u	Input
u_{ik}	Degree of affiliation from the kth data point y_k to i-ten cluster C_i

$\underline{u}_{\mathrm{PCC}}$	Voltage at PCC point
$\underline{u}_{\mathrm{SR}}$	Complex voltage of the converter
$\underline{U}$	Partition matrix
$U(t)$	Voltage
U_{m}	Measured voltage of the power network
U_{s}	Simulated voltage of the model
$\underline{U}_{\mathrm{CON}}$	converter voltage
U_{IC}	Intermediate circuit voltage
υ	Expected value of the return function
$\underline{v}$	Control vector
v_i	Centre of the cluster *i*
ω_0	Resulting frequency of the oscillations
w	Weights
$\underline{w}$	Vector of weights
w_j	Weight vector
ω_{s}	Crossing frequency
x	Input vector
$\hat{x}_i$	State
x_{r}	Condition of the reduced model
$\underline{\vec{x}}$	Space cursor in α/β coordinates
$\underline{\vec{x}}'$	Space cursor in d/q coordinates
x_{d}	Real part of the space pointer in d/q coordinates
x_{q}	Imaginary part of the space cursor in d/q coordinates
χ	Actual value of a variable
x^*	Setpoint of a variable
$\underline{X}$	Input data
y	Output
$\underline{y}^*$	Critical points for $E_{qw}(\underline{y})$
y_j	Output target values
$\hat{y}_j$	Output at the neuron j
$\underline{Y}$	Output set of real power network
$\underline{Y}'$	Output set of DT model
$\underline{Y}_{\mathrm{DT}}$	Simulated output from DT model

$\underline{Y}_{\mathrm{m}}$	Measured output
$\underline{Y}_{\mathrm{s}}$	Simulated output
z	Disturbance
Z	Complex value
α	Real part of the pole
β	Imaginary part of the pole
δ	Overshoot of the process
δ_j	Error of backward propagation
λ_i	Eigenvalue
λ	Lagrange multiplier
λ_k	Scalar
ξ	Relative damping
ζ	Damping ratio
ε	Positive random number
η	Learning rate
τ	Time constant of eigenvalues
γ	Discount factor
ε	Environment
ρ	Distribution of the sample
π	Reinforcement learning policy
ϕ	Phase
ϕ_{r}	Phase reserve
φ_{a}	Initial phase
φ_{e}	Input phase
φ_{PLL}	Output phase of PLL
φ_{N}	Output phase
θ	Time delay of the process
θ_{O}	Time delay of the original model
σ_k	Actual learning rate
σ_i	Hankel singular values
Δ_k	Measure for approximation of $E_{qw}\left(\alpha_k \cdot \underline{p}_k\right)$ to $E\left(\underline{w}_k + \alpha_k \cdot \underline{p}_k\right)$
Δ_i	Confidence interval
Δw	Gradient
Δw_{ij}	Adaptation of weights
ρ_i	Reduction ratio

Einheiten

A	Ampere, $[I]$ = A
Hz	Hertz, $[f]$ = Hz
V	Volt, $[U]$ = V
W	Watt, $[P]$ = W
Var	Var, [Q] = Var
s	Sekunde, $[t]$ = A
s^{-1}	Hertz, $[f] = S^{-1}$

List of Figures and Tables

universitätsverlag *ilmenau*

Cai: Adaptive control of converters dominated distribution network based on digital twins

With the increasing feed-in and distribution of renewable energies in the power system, the number of converters is growing. This can lead to the destabilization of the system due to several factors, one of which is the controller parameter of the converter. This requires parameterization and optimization of the mentioned parameters. The presented work is devoted to the main research question of how the stability of a converter-dominated distribution grid is improved for varying operating points by the dynamic adaptive control method. To carry out the parameterization of the controllers, the Reinforcement Learning (RL) agent for adaptively creating proper controller parameters according to the system states to improve and ensure stability is proposed. In this work, a digital twin (DT) of the studied network is first constructed with the help of parameter estimation. The DT model is utilized to generate training data for an artificial-neural-network-based state estimator, which is dedicated to arcuately and efficiently determining the system's stable state. More-over, a small signal stability indicator (SI) using the damping ratio of the dominant eigen-value for the DT model is developed. With the SI indicating to the RL agent the system stability margin during the RL training, the agent can ultimately output optimal controller parameters for the converters. Numerical case studies are used to verify the viability of the proposed approach that network stability can be improved by the proposed adaptive control method.

ISSN 2194-2838
DOI 10.22032/dbt.59235