Designing Reactive Power Control Rules for Smart Inverters using Machine Learning

Aditie Garg

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Electrical Engineering

Vassilis Kekatos, Chair

Virgilio A. Centeno

Jaime De La Reelopez

May 1, 2018

Blacksburg, Virginia

Keywords: smart inverters; support vector machines; kernel-based learning; voltage regulation; power loss minimization; linearized distribution flow model

Designing Reactive Power Control Rules for Smart Inverters using Machine Learning

Aditie Garg

(ABSTRACT)

Due to increasing penetration of solar power generation, distribution grids are facing a number of challenges. Frequent reverse active power flows can result in rapid fluctuations in voltage magnitudes. However, with the revised IEEE 1547 standard, smart inverters can actively control their reactive power injection to minimize voltage deviations and power losses in the grid. Reactive power control and globally optimal inverter coordination in real-time is computationally and communication-wise demanding, whereas the local Volt-VAR or Watt-VAR control rules are subpar for enhanced grid services. This thesis uses machine learning tools and poses reactive power control as a kernel-based regression task to learn policies and evaluate the reactive power injections in real-time. This novel approach performs inverter coordination through non-linear control policies centrally designed by the operator on a slower timescale using anticipated scenarios for load and generation. In real-time, the inverters feed locally and/or globally collected grid data to the customized control rules. The developed models are highly adjustable to the available computation and communication resources. The developed control scheme is tested on the IEEE 123-bus system and is seen to efficiently minimize losses and regulate voltage within the permissible limits.

Designing Reactive Power Control Rules for Smart Inverters using Machine Learning

Aditie Garg

(GENERAL AUDIENCE ABSTRACT)

The increasing integration of solar photovoltaic (PV) systems poses both opportunities and technical challenges for the electrical distribution grid. Although PV systems provide more power to the grid but, can also lead to problems in the operation of the grid like overvoltages and voltage fluctuations. These variations can lead to overheating and burning of electrical devices and equipment malfunction. Since the solar generation is highly dependent on weather and geographical location, they are uncertain in their output. The uncertainity in the solar irradiance can not be handled with the existing voltage control devices as they need to operate more frequently than usual which can cause recurring maintenance needs for these devices.

Thus, to make solar PV more flexible and grid-friendly, smart inverters are being developed. Smart inverters have the capability of advanced sensing, communication, and controllability which can be utilized for voltage control. The research discusses how the inverters can be used to improve the grid profile by providing reactive power support to reduce the power losses and maintain voltages in their limits for a safer operation.

Acknowledgments

I would first like to thank my thesis advisor Dr. Vasillis Kekatos for his patience, motivation, enthusiasm, and immense knowledge. The door to his office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently steered me in the right the direction whenever he thought I needed it.

I wish to thank Dr. Virgilio Centeno and Dr. Jaime De La Ree for creating a nurturing environment at the power systems laboratory. Dr. Centeno additionally, has always motivated me to pursue work in the direction of my passion and I wish to stand by this inspiration for the rest of my life. I would also like to express my appreciation towards all members of the power systems laboratory at Virginia Tech for their constant support and healthy interactions.

I would like to extend my heartfelt thanks to my dear friends, Megha Ravishankar Rao and Manik Aggarwal. All our fun get-togethers revitalized me. Finally, I would like to thank my parents Dinesh Kumar and Alka Gupta without whose sincere efforts and constant guidance I would not have been able to experience my present life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Electricity is an inevitable and integral part of life in today's world. In the past carbon fossil fuels were the main conventional sources of producing electricity. These sources were not environmental friendly, involved emission of greenhouse gases and depletion of non-renewable sources such as coal, petroleum and natural gas which have limited availability. But, with increased awareness of the harmful effects of the gases, governments in many countries are assigning topmost priority to improve energy efficiency and reduce our carbon footprint on the planet [3].

In addition, to limited sources of generation, there is an increase in peak consumption with the increase of demand on basis of rising per capita kWh consumption [4]. Thus,

the gap between generation and consumption capacities of the grid is reducing, leading to increased pressure on the grid [4]. Increase in demand and limited amount of generation resources makes the grid highly dynamic and vulnerable to voltage fluctuations [5]. Thus, finding new sources of energy has become very essential. Renewable generation is one of the key solutions used to provide electricity and stability to the grid.

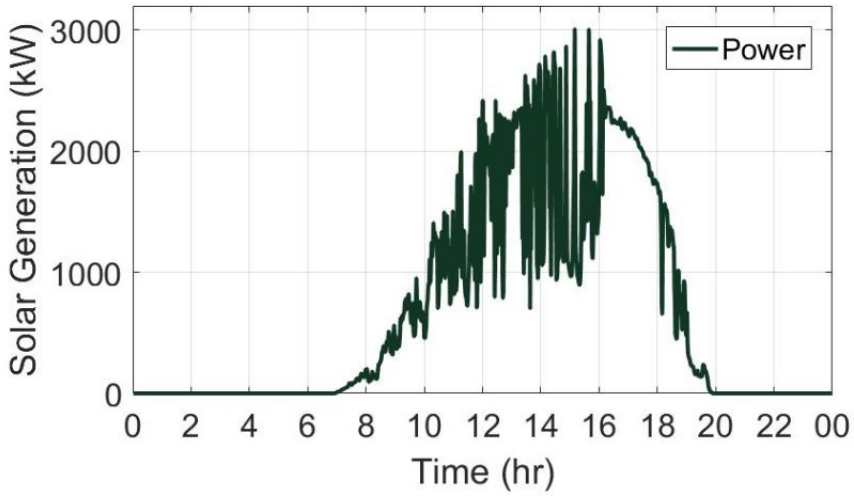


Figure 1.1: Solar generation from Pecan street data [1].

The most common renewable energy resources used in the present world are photovoltaic and wind energy. Increased penetration of renewable sources into the grid provides a logical solution to handle the environmental issues. Since solar energy is locally available near the load, it is easy to harness. Nevertheless, with increasing quantum of these renewable sources results in more problems for the grid, such as reverse power flow, voltage regulation etc. For example, a solar farm connected at the end of a long rural feeder can introduce voltage regulation problems to all the residential buses along the feeder. Moreover, frequent power flow reversals due to remote injection of renewable energy, strain the apparent power

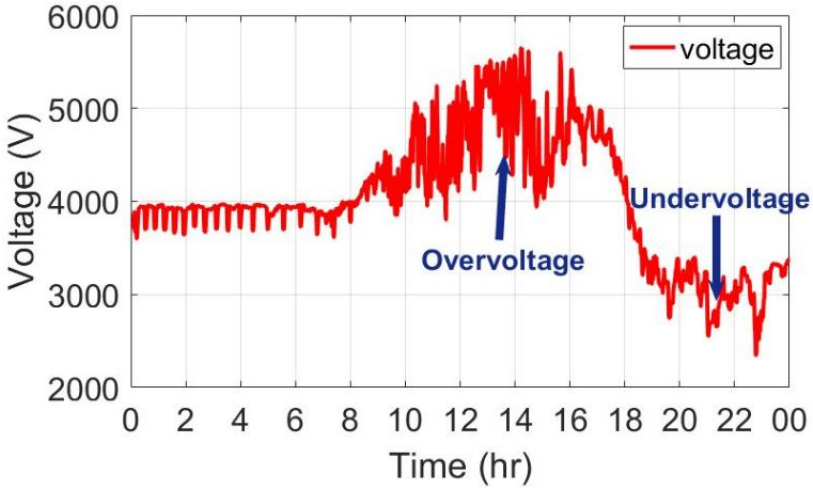


Figure 1.2: Bus voltage profile.

capabilities of substation transformers [6]. Solar generation from residential photovoltaics (PVs) can fluctuate by up to 15% of the PV nameplate rating within one-minute intervals bringing in the issue of uncertainty in generation too. Active power generated over a day taken from Pecan street [1] is shown in Fig. 1.1. Although the active solar generation is seen to peak during the daytime when there is maximum sunlight, there are lot of variations in the real power output from the solar cell. This uncertainty depends mostly on climate and geographical location of the solar cells. Hence, making the real power generation highly dynamic and unpredictable. Moreover, in Fig. 1.2 the voltage profile is shown where the consumption and generation data is taken from pecan street data [1]. Figure. 1.2 depicts that voltage fluctuates rapidly during the day, overvoltage is seen during daytime due to excess of solar power generation but undervoltage is seen during the night due to lack of enough generation. Thus, to avoid these fluctuations and have a stable grid operation, voltage regulation is required.

Traditionally, voltage regulation is carried out using on-load tap changing in substation transformers, switching of capacitor banks, and step voltage regulators. These equipment have an operational delay of 30-90 sec [7]. due to their control techniques. The voltage regulators need to carry out switching operations frequently due to high frequency of PV-induced voltage fluctuations. This reduces the longevity of the voltage regulator as it depends on the life of its switch. Since the real power generation and loads are increasingly becoming more uncertain.Increased uncertainty leads to frequent switching actions, perhaps more installations might be needed, thus critically challenging reactive power control in distribution grids.

These problems can be alleviated through the use of smart inverter technologies. Presently, PV units interface with the inverters to provide advanced communication, metering, and control functionalities [8]. These inverters provide smart multi-unit control through control of real power limit, controlled ramp rate for real power limit, control of reactive power output or power factor, ride-through capability for specific grid disturbances, bi-directional power flow capability, and alternatives to conventional transfer trip schemes [8]. Use of smart inverters for reactive power control provides a fast responding solution for various grid objectives such as power loss minimization and voltage regulation [6]. With the increasing penetration level of PV into the grid more sophisticated rules for interconnection are emerging too. Intelligent solutions to the problems present in the grid harnessing the inverter control capabilities will be key to the successful implementation of large-scale PV generation in distribution grid. In future, with a large range of control functions incorporated into newer PV inverter designs,

will allow them to play an important role in the operation of the distribution grid [8].

According to the amended IEEE 1547 standard [9], inverters are allowed to operate at non-unit power factors, giving inverters the freedom to improve the grid voltage profile. Studies have shown that inverters can improve the grid voltage profile, or even displace utility-owned voltage regulating equipment at more than 50% solar penetration [10]. With hundreds of inverters in the grid, it must be noted that coordination of each inverter needs to be taken into consideration to keep the grid stable.

1.2 Prior Works

In a typical distribution grid setup, the instantaneous loads and solar generation from each node are communicated to a central utility controller; the controller minimizes ohmic losses subject to voltage regulation constraints; and the computed setpoints for reactive power injections are sent back to each of the inverters. The problem of finding the optimal reactive injection setpoints for inverters is an instance of the optimal power flow (OPF) task, which is non-convex in general. In radial networks, the OPF can be relaxed into a second-order cone program (SOCP) via polar coordinates [11], where the problem of power loss and voltage deviation minimization is solved. To alleviate the complexity of the involved optimization problems, approximate grid models have been employed in [12], [13], [14].

Reactive power control problem can be solved using centralized, decentralized or local techniques [15]. Centralized approaches need good communication setup as global infor-

mation is needed for control actions. While, decentralized methods require local or neighbouring inputs like voltage or power values for evaluating the control settings for single- and unbalanced multiphase grids [16, 17]. Purely localized schemes provide reactive power support using only local measurements. Thus, centralized schemes incur high computational complexity as large sets of data have to be communicated between the controller and each inverter; decentralized solvers require multiple communication exchanges among inverters; and local schemes have no performance guarantees as the control set point only depends on these local inputs making the system highly unreliable to disturbances happening on other nodes.

Reactive power control policies can be mapped as linear or non linear policies with respect to its input features. Linear policies, however, are restricted to capturing linear relations between the features and dependent variables, and very often only capture second-order statistical relations. Such limitations call for extensions to nonlinear and higher order algorithms. To overcome this, kernel methods are adopted in this research. In this research reactive power control policies are modeled by creatively cross-pollinating ideas from machine learning and using the powerful tool of kernel-based learning which are practically feasible.

1.3 Thesis Outline

The content in this thesis has been organized as follows. The theory and application behind the reactive power injection control using kernel methods is described in Chapter 2. Chapter

3 describes the grid model and various forms of problems formulations for reactive power control. Chapter 4 explains the prior control strategies, points out the shortcoming in each method and motivates the problem utilized in this research. Chapter 5 presents our research contribution towards developing kernel based policies for reactive power control. The novel approach developed is tested on the IEEE 123-bus system and the results are shown in Chapter 6. The thesis is concluded in Chapter 7.

1.4 Notation

Regarding *notation*, lower- (upper-) case boldface letters denote column vectors (matrices), with the exception of the line complex power flow vector $\mathbf{S}$. Calligraphic symbols are reserved for sets. Symbol $^\top$ stands for transposition. Vectors $\mathbf{0}$ and $\mathbf{1}$ are the all-zeros and all-ones vectors, while $\mathbf{e}_n$ is the n-th canonical vector. Maxtrix $\mathbf{I}_N$ represent a $N \times N$ identity matrix. $\otimes$ denotes the kronecker product. Symbol $\|\mathbf{x}\|_2$ denotes the ℓ_2-norm of $\mathbf{x}$ and $\mathrm{dg}(\mathbf{x})$ defines a diagonal matrix having $\mathbf{x}$ on its diagonal. A symmetric positive (semi)definite matrix is denoted as $\mathbf{X} \succ \mathbf{0}$ ($\mathbf{X} \succeq \mathbf{0}$), while $|\mathbf{X}|$ and $\mathrm{Tr}(\mathbf{X})$ are the determinant and trace of $\mathbf{X}$. The symbol $\|\cdot\|_F$ denotes the Frobenius matrix norm with $\|\mathbf{X}\|_F^2 = \mathrm{Tr}(\mathbf{X}^\top \mathbf{X})$.

Chapter 2

Linear Regression and Kernel Methods

With the increasing PV penetrations and variability in loads, the grid is becoming highly dynamic and vulnerable to changes. To minimize losses and keep voltage of the power grid within limits, reactive power control is required. The control rules can be modeled using supervised learning learning methods like linear regression. This chapter first explains linear regression method which can predict values for continuous target variables using linear functions. Next, kernel methods are explained which offer the opportunity to translate high dimensional data to a finite dimension in kernel space while still working with linear algebra. Thus, kernel methods allows us to exploit all the intuitions and properties of linear algorithms and parallely evaluating non-linear policies which are fast to compute and evaluate.

2.1 Linear Regression Model

Regression model estimates the relationship among input data set x_n with corresponding target value t_n. The goal is to predict the value of $\mathbf{t} = [t_1, \cdots, t_N]^\top$ for a new value of $\mathbf{x} = [x_1, \cdots, x_N]^\top$ where N is the dimension of input data. A simple linear model for regression involves linear combination of input variables [18]

$$y(\mathbf{x}, \mathbf{w}) = w_0 + w_1 x_1 + \cdots + w_N x_N \tag{2.1}$$

where $\mathbf{w} = [w_0, \cdots, w_N]^\top$ are unknown parameters or coefficients. This is known as linear regression[18]. The model shown above has a target value that is a linear combination parameters $w_0, \cdots, w_N$. Also, linear methods can be applied to transformations of the input data sets. These transformations can be extended as a linear combination of fixed non-linear functions $\phi(\mathbf{x})$ known as basis functions of the input variables. Putting these functions into (2.1) the linear regression model can be reformulated as

$$y(\mathbf{x}, \mathbf{w}) = \sum_{i=0}^{M-1} w_i \phi_i(\mathbf{x}) = \mathbf{w}^\top \phi(\mathbf{x}) \tag{2.2}$$

where $\mathbf{w} = [w_0 \ \cdots \ w_{M-1}]^\top$ and $\boldsymbol{\phi} = [\phi_0 \ \cdots \ \phi_{M-1}]^\top$. Thus, the total number of parameters in this model is given by M. w_0 is the bias parameter to provide a fixed offset in data. ϕ_0 is a dummy *basis function* $\phi_0(\mathbf{x}) = 1$ [18]. By adding a nonlinear basis function the function $y(\mathbf{x}, \mathbf{w})$ becomes non linear in $\mathbf{x}$. The non linear basis functions ϕ_x can be chosen from a variety of family of functions such as polynomial $\phi_i(x) = x^i$, exponential family $\phi_i(x) = \exp \frac{-(x-\mu_i)^2}{2s^2}$ where μ_i governs the locations of basis functions in input space and s

governs the spatial space, sigmoidal basis function $\phi_i(x) = \sigma\frac{x-\mu_i}{s}$ where $\sigma(a)$ is the logistic sigmoid function given by $\sigma(a) = \frac{1}{1+\exp(-a)}$, Fourier basis etc.

2.1.1 Regularized Least Squares

For a successful implementation of an algorithm the model must be able to predict the results as accurately as possible with minimal assumptions. But this results in a trade off between making assumptions and having an accuracy for the model. For example, if the method is too restrictive on assumptions the model might give inaccurate results or a wrong detection of patterns due to a misfit with the learning algorithm whereas if lot of assumptions are made the fit will be good. This might happen when the data used in input might not be generated in the way it is assumed, say for example the model assumes the data to be taken from a Gaussian distribution while it was from a non Gaussian distribution. Thus resulting in misfitting the data. Whereas if model is given a lot of flexibility to assume the way data is generated or by providing a large set of hypothesis there is a high chance that it might fit the model with them. This is known as overfitting [19]. Thus, in order to control model overfitting, a regularization term is added to the error function. The regularized sum of squares error function is given by

$$\text{RSSE}(\mathbf{w}) = \text{SSE}(\mathbf{w}) + \mu(\text{RSE}(\mathbf{w})) \tag{2.3}$$

where the sum of squares of error term is given as $\text{SSE}(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\left(t_n - \mathbf{w}^\top\phi(\mathbf{x}_n)\right)^2$, μ is a regularization coefficient that controls the problem of overfitting by evaluating the relative

importance of the data dependent error SSE($\mathbf{w}$) and the regularization term RSE($\mathbf{w}$). The regularizer term for error is given by

$$\text{RSE}(\mathbf{w}) = \frac{1}{2}\mathbf{w}^\top \mathbf{w}. \tag{2.4}$$

Thus, the total error function can be written as

$$\text{RSSE}(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\left(t_n - \mathbf{w}^\top \phi(\mathbf{x}_n)\right)^2 + \mu\frac{1}{2}\mathbf{w}^\top \mathbf{w}. \tag{2.5}$$

Since the above equation is quadratic in $\mathbf{w}$ its exact minimizer can found in closed form [18]. To evaluate the minimizer the gradient (2.5) is set to zero with respect to $\mathbf{w}$. Solving for $\mathbf{w}$ as before, $\hat{\mathbf{w}}$ is obtained as

$$\hat{\mathbf{w}} = (\mu\mathbf{I} + \boldsymbol{\phi}^\top \boldsymbol{\phi})^{-1}\boldsymbol{\phi}^\top \mathbf{t}. \tag{2.6}$$

Thus, as shown above the problem of regularized least squares is solved. This type of a linear regression adjusts the cost function with a small modification of adding the regularizer to avoid overfitting of the model. The complete model complexity is dependent on the complexity of the basis functions. If the number of basis functions is larger the problem becomes more complex. This must be controlled according to the size of the data set [18]. Adding a regularizer term avoids over fitting and can be controlled by choosing effective regularization coefficient. The effective regularization coefficient can be evaluated by doing cross-validation. It must be noted, there is a trade-off in the choice of number of basis functions and evaluating the overall evaluation of the model. If one uses maximum likelihood function it might lead to over fitting or highly complex models. Thus, to decide an appropriate model complexity for a given regression model [18] kernel methods can be used.

2.2 Kernel Methods

In this section, kernel methods are explained in detail, the concept of reproducing kernel Hilbert spaces (RKHS), and the representer theorem is stated. Kernel methods provide an elegant mathematical approach to representing an infinite number of basis functions using a finite amount of computational power, which is complementary to the emerging deep learning approach. Kernel methods build upon the notion of kernel functions and RKHSs [20].

Using, kernel methods, a data set S that is defined over an input or attribute feature space $\mathcal{X}$ ($S \subseteq \mathcal{X}$) can be mapped into a higher (possibly infinite-dimensional) Hilbert space $\mathcal{H}$, also known as kernel feature space. With this in view a nonlinear algorithm is built with respect to the input data space but linear with respect to the kernel feature space. Thus, kernel processes allows the model to comprise on linear algorithms expressed in an Euclidean space and, by means of *kernelization* procedure or taking the dual it derives a nonlinear counterpart of the algorithm, hence providing a set of nonlinear properties to the linear algorithms as explained in section 2.2.1. The *kernelization* procedure involves two basic steps, firstly evaluation of an expression for kernel and secondly, substitution of the Euclidean dot product by a dot product into an RKHS [20]. This process will be explained in detail in section 2.2.2. Although the nonlinear version of linear algorithms provide an increased flexibility, when a new sample coordinates is evaluated explicitly in a high-dimensional space the computational load increases accordingly. To reduce this computation load the *kernel trick* can be used. Another advantage of using kernel methods is its ability to generalize for

all cases under consideration even if the structural complexity of a machine is too high. Since, nonlinear learning machines introduce a variety of parameters so that the model becomes extremely flexible but it falls into pit fall of overfitting the training set. Thus, the concept of regularization is also used in kernel methods explained in section 2.2.3. In summary, kernel methods allow the exploitation of a vast variety of kernel functions that can be designed and adapted to the application at hand and can be widely adopted in various types of applications.

2.2.1 Dual Representation

Kernel function for a models which span on the basis of a fixed nonlinear feature space mapping $\boldsymbol{\phi}(\mathbf{x})$, can be shown as

$$k(\mathbf{x}, \mathbf{x}') = \boldsymbol{\phi}(\mathbf{x})^\top \boldsymbol{\phi}(\mathbf{x}'). \tag{2.7}$$

From the above definition it is seen kernel function is symmetric such that $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}', \mathbf{x})$ [18]. For example of a kernel function for an identity mapping for the feature space in (2.7) such that $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{x}$ gives the kernel function as $k(\mathbf{x}, \mathbf{x}') = \mathbf{x}^\top \mathbf{x}'$. Linear models for regression can be reformulated in terms of a dual representation in which the kernel function is seen to be present and the predictions are seen to be a linear combinations of a kernel function evaluated initially at the training data points. The linear regression model for a regularized sum of squares (2.3) is considered. Solution for $\hat{\mathbf{w}}$ is evaluated by taking the gradient of

RSSE($\mathbf{w}$) with respect to $\mathbf{w}$ given as,

$$\mathbf{w} = -\frac{1}{\mu}\sum_{n=1}^{N}(\mathbf{w}^\top\phi(x_n) - t_n)\phi(x_n) = \sum_{n=1}^{N} a_n\phi(x_n) = \mathbf{\Phi}^\top\mathbf{a} \tag{2.8}$$

where $\mathbf{\Phi} = [\phi(x_1), \ \cdots \ , \phi(x_N)]^\top$ is the design matrix and the entries of vector $\mathbf{a} = [a_1, \ \cdots \ , a_N]^\top$ are defined as

$$a_n = -\frac{1}{\mu}(\mathbf{w}^\top\phi(x_n) - t_n). \tag{2.9}$$

In this expression, $\hat{\mathbf{w}}$ is seen to be a linear combinations of vectors $\phi(\mathbf{x}_n)$, with coefficients as functions of $\mathbf{w}$. The dual of problem can be taken by reformulating it in the terms of parameter vector $\mathbf{a}$. Using the results from (2.8), $\mathbf{w} = \mathbf{\Phi}^\top\mathbf{a}$, RSSE($\mathbf{a}$) takes the form

$$\text{RSSE}(\mathbf{a}) = \frac{1}{2}\mathbf{a}^\top\mathbf{\Phi}\mathbf{\Phi}^\top\mathbf{\Phi}\mathbf{\Phi}^\top\mathbf{a} - \mathbf{a}^\top\mathbf{\Phi}\mathbf{\Phi}^\top\mathbf{t} + \frac{1}{2}\mathbf{\Phi}\mathbf{\Phi}^\top\mathbf{t} + \frac{1}{2}\mathbf{t}^\top\mathbf{t} + \frac{\mu}{2}\mathbf{a}^\top\mathbf{\Phi}\mathbf{\Phi}^\top\mathbf{a} \tag{2.10}$$

where $\mathbf{t} = [t_1, \ \cdots \ , t_N]^\top$. The Gram matrix $\mathbf{K}$ is defined as

$$\mathbf{K} = \mathbf{\Phi}\mathbf{\Phi}^\top.$$

This Gram matrix is symmetric with elements as of a *kernel function* $k(x, x')$ given as

$$K_{nm} = \phi(x_n)^\top\phi(x_m) = k(x_n, x_m).$$

Rewriting (2.10) in terms of the Gram matrix, the regularized sum of squares can be written as

$$\text{RSSE}(\mathbf{a}) = \frac{1}{2}\mathbf{a}^\top\mathbf{K}\mathbf{K}\mathbf{a} - \mathbf{a}^\top\mathbf{K}\mathbf{t} + \frac{1}{2}\mathbf{t}^\top\mathbf{t} + \frac{\mu}{2}\mathbf{a}^\top\mathbf{K}\mathbf{a}. \tag{2.11}$$

Setting the gradient of RSSE($\mathbf{a}$) with respect to $\mathbf{a}$ to zero, $\hat{\mathbf{a}}$ is evaluated as

$$\hat{\mathbf{a}} = (\mathbf{K} + \mu\mathbf{I}_N)^{-1}\mathbf{t}. \tag{2.12}$$

Using the above result, output of the linear regression model for a new incoming x can be evaluated as

$$y(x) = \hat{\mathbf{w}}^\top \phi(x) = \hat{\mathbf{a}}^\top \mathbf{\Phi}\phi(x) = \mathbf{k}(x)^\top (\mathbf{K} + \mu \mathbf{I}_N)^{-1}\mathbf{t} \tag{2.13}$$

where vector $\mathbf{k}(x)$ consists of elements $k_n(x) = k(x_n, x)$. As expected the prediction at a new input x is given by a linear combination of the target values $\mathbf{t}$ from the training set. Hence, its seen that the dual formulation allows the solution to regularized least squares to be completely expressed in terms of the kernel function $k(x, x')$. Thus without explicitly getting into the feature space which can be of infinite dimensionality the model evaluation can be carried out directly in terms of kernels. The solution for $\mathbf{a}$ can be expressed as a linear combination of the elements of $\phi(x)$, and the original formulation can be recovered in terms of the parameter vector $\mathbf{w}$. This expression is known as dual formulation [18].

2.2.2 Constructing Kernel

Valid kernel functions needs to be evaluated to exploit kernel trick. If a feature space is chosen as $\phi(x)$ the corresponding kernel can be evaluated as

$$k(x, x') = \boldsymbol{\phi}(x)^\top \boldsymbol{\phi}(x') = \sum_{i=1}^{N} \phi_i(x)\phi_i(x') \tag{2.14}$$

where $\phi_i(x)$ are the basis functions. Here the kernel function is seen to be defined for a one-dimensional input space. The kernel functions can also be evaluated by finding the scalar product in some feature space which may be infinite dimensional too. For example,

$$k(\mathbf{x}, \mathbf{z}) = (\mathbf{x}^\top \mathbf{z})^2$$

where $\mathbf{x}, \mathbf{z}$ are two two input vectors. To test the validity whether a function constitutes a valid kernel or not without having to construct the function (x) explicitly one needs to evaluate the Gram matrix $\mathbf{K}$. According to [19] the Gram matrix $\mathbf{K}$, whose elements are given by $k(x_n, x_m)$ is given as

$$\mathbf{K} = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_N) \\ K(x_2, x_1) & K(x_2, x_2) & \cdots & K(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_N, x_1) & K(x_N, x_2) & \cdots & K(x_N, x_N) \end{bmatrix}.$$

Gram matrix $\mathbf{K}$ must be positive definite for all possible choices of set$\{x_n\}$. Now one needs to prove if the function gives a valid kernel $k(x, x')$ that is it is positive definite and symmetric or not and that it expresses the appropriate form of similarity between x and x' according to the intended application [18].

Thus, kernel functions help in detecting non linear relations using linear algorithms in a given feature space. The formulation helps in decoupling the design of model from its input feature space, eventually making the model modular.

2.2.3 RKHS and Representer's Theorem

As seen previously, kernel methods evaluate dot products of vectors in Hilbert space through a mapping function. If the space is complete that is every Cauchy sequence converges inside the space, then it is called a Hilbert space [20]. For given positive definite kernel $k(\cdot, \mathbf{x})$, the

corresponding space function is called a reproducing kernel Hilbert space $\mathcal{H}_K$ if the elements of $\mathcal{H}_K$ are complex or real valued functions $f(\cdot)$ defined on any set of elements $\mathbf{x}$ and for every element of $\mathbf{x}$, $f(\cdot)$ is bounded. The RKHS $\mathcal{H}_K$ for a kernel K is generated by $\mathbf{x}$ where $\mathbf{x}$ belongs to any set, is given by $K(\cdot, \mathbf{x})$. Another property of RKHS states that an RKHS contains a single reproducing kernel and a reproducing kernel has a unique RKHS [20].

Given pairs $\{(x_t, y_t)\}_{t=1}^{T}$ of features x_t belonging to a measurable space $\mathcal{X}$ and target values $y_t \in \mathbb{R}$, kernel-based learning aims at finding a mapping $f : \mathcal{X} \to \mathbb{R}$. The mapping f is constrained to lie on the linear function space [21]

$$\mathcal{H}_{\mathcal{K}} := \left\{ f(z) = \sum_{i=1}^{\infty} K(x, x_i) a_i, \ a_i \in \mathbb{R} \right\}. \tag{2.15}$$

defined by a given kernel (basis) $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and corresponding coefficients a_i. The minimization problem for a loss function $L(y, f(x)$ can be written as

$$\arg\min_{f \in \mathcal{H}} \sum_{i=1}^{N} L(y_i, f(x_i), b) \tag{2.16}$$

$$\text{s.to } \|f\|_{\mathcal{H}_K}^2 < \infty$$

where $\|f\|_{\mathcal{H}_K}$ is the induced norm by K, b is an intercept term that is also unknown. The loss function L depends on f only through the input-output pairs $\{f(x_i), y_i\}$. The loss function can be chosen from a variety of functions, typical choices for L include the least-square fit $\big(y_i - f(x_i) - b\big)^2$, or the ϵ-insensitive loss $\max\{|y_i - f(x_i) - b| - \epsilon, 0\}$ for regression; and the hinge loss $\max\left\{1 - y_i\left(f(x_i) + b\right), 0\right\}$, or the logistic function $\log\left(1 + e^{-y_i\left(f(x_i)+b\right)}\right)$ for

binary classification [21]. This can rewritten as a class of regularization problems given as

$$(\hat{f}, \hat{b}) := \arg\min_{f \in \mathcal{H}} \sum_{i=1}^{N} [L((y_i, f(x_i), b)) + \mu \|f\|^2_{\mathcal{H}_K}] \tag{2.17}$$

where the penalty function for RKHS is defined as $\text{RSE}(f) = \|f\|^2_{\mathcal{H}_K}$ and μ is the regularization coefficient. The regularizer is an increasing function of $\|f\|_{\mathcal{K}}$. This increasing function ensures that $f \in \mathcal{H}_{\mathcal{K}}$ and facilitates generalization over unseen data. The parameter $\mu > 0$ balances the two terms and is typically tuned via cross-validation [21]. According to [22] solution to (2.17) is finite dimensional given as

$$\hat{f}(x) = \sum_{i=1}^{N} K(x, x_i)\hat{a}_i. \tag{2.18}$$

For $f \in \mathcal{H}_K$, it can be seen that $\langle K(\cdot, x_i), f\rangle_{\mathcal{H}_K} = f(x_i)$[21]. The basis function $h_i(x) = K(x, x_i)$ is known as the representer of evaluation at x_i in $\mathcal{H}_K$. This theorem is known as the *representer's theorem*. It must be noted (2.18) is the sought function that is described only by N rather than infinitely many a_i's [cf. (2.15)]. Moreover, to evaluate $\hat{f}(x)$ at any $x \in \mathcal{X}$, one merely needs to know $\{\hat{a}_i\}_{i=1}^N$ and be able to evaluate the kernel function at $\{K(x, x_i)\}_{i=1}^N$. Thus the infinite dimensional optimization problem has been reduced to a finite dimensional convex problem making it computationally easier and faster.

Consider a *kernel matrix* $\mathbf{K} \in \mathbb{S}^N_{++}$ having entries $[\mathbf{K}]_{i,i} := K(x_i, x_j)$, and the vector $\hat{\mathbf{f}} := [\hat{f}(x_1) \ \cdots \ \hat{f}(x_N)]^\top$ collecting the function values at $\{x_i\}_{i=1}^N$. Evaluating $\hat{f}$ at the given data based on (2.18) yields

$$\hat{\mathbf{f}} = \mathbf{K}\hat{\mathbf{a}} \tag{2.19}$$

where $\hat{\mathbf{a}} := [\hat{a}_1 \ \ldots \ \hat{a}_N]^\top$. Using properties of RKHS's, it can be shown that the function norm

in $\mathcal{H}_{\mathcal{K}}$ admits the matrix-vector expression $\|f\|_{\mathcal{K}}^2 = \hat{\mathbf{a}}^\top \mathbf{K}\hat{\mathbf{a}}$. Thus, the functional minimization in (2.17) is equivalent to finding the optimal $\hat{\mathbf{a}}$. If for example $L(f(z_t), y_t) = \big(y_t - f(z_t) - b\big)^2$ where in actually you are minimizing the error between the predicted and target values and he regularizer term is given as s^2, the functional problem (2.17) is equivalent to the finite-dimensional minimization

$$(\hat{\mathbf{a}}, \hat{b}) := \arg\min_{\mathbf{a},b} \|\mathbf{y} - \mathbf{K}\mathbf{a} - b\mathbf{1}\|_2^2 + \mu\|\mathbf{a}\|_{\mathbf{K}}^2 \tag{2.20}$$

where $\|\mathbf{a}\|_{\mathbf{K}}^2 := \mathbf{a}^\top \mathbf{K}\mathbf{a}$ and $\mathbf{y} := [y_1 \ \cdots \ y_N]^\top$.

It must be noted, the expression in (2.18) can be applied to the given data $\{x_i\}_{i=1}^N$ as well as a new point of interest $x_j \in \mathcal{X}$. Prior to evaluating at the new point of interest $\hat{f}(x_j)$ requires knowing the kernel evaluations between the point of interest and the training data, i.e., $\{K(x_j, x_i)\}_{i=1}^N$, and the minimizers $(\hat{\mathbf{a}}, \hat{b})$. These minimizers can then be fed into the validation step to evaluate the new data point be used to evaluate the $\hat{f}(x_j)$ as given below

$$\hat{f}(x_j) = \sum_{i=1}^{N} K_n(x_j, x_i)\hat{a}_i + \hat{b}_i.$$

Thus, using the kernel evaluations, minimizers, the training and new data, the functon can be evaluated for the new data point in real-time.

Chapter 3

Grid Modeling and Problem Formulation

Minimization of power losses and maintenance of voltage within limits are one of the key concerns faced by the grid operator. One solution to this problem is reactive power control generated by smart power inverters. With increasing PV installations distributed all through the grid, there is an excess of the inverter capacity available to generate or consume reactive power in grid. With the amendments in IEEE 1547 standard, question of how, when and how much reactive power must be dispatched is still significant. This section reviews an approximate grid model and formulates the reactive power control problem.

3.1 Grid Modeling

To evaluate the dispatch of reactive power by the inverters, an approximate *linearized distribution flow* (LDF) model is reviewed below. A radial single-phase grid with $N+1$ buses indexed by $n = 1, \cdots, N+1$, M branches is considered. For a radial system, the number of branches $M = N$, every bus $n = 1, \ldots, N$ is connected to a unique parent bus π_n via distribution line n, while $n = 0$ is the substation bus as shown in Fig. 3.1. The line connecting bus π_n and n is numbered as n. The grid is modeled by the branch flow equations [23].

$$s_n = \sum_{k \in \mathcal{C}_n} S_k - S_n + \ell_n(r_n + jx_n) \tag{3.1a}$$

$$v_n = v_{\pi_n} - 2\,\mathrm{Re}[(r_n - jx_n)S_n] + \ell_n(r_n^2 + x_n^2) \tag{3.1b}$$

$$|S_n|^2 = v_{\pi_n}\ell_n \tag{3.1c}$$

where for every line n the line impedance is $z_n = r_n + jx_n$; ℓ_n is the square of current magnitude in line n; $S_n = P_n + jQ_n$ is the complex power flow from sending bus π_n to bus n; $s_n = p_n + jq_n$ is the complex power injection at bus n; v_n is the squared voltage magnitude at bus n; $\mathcal{C}_n$ is the set of children buses for n; and the initial condition $s_0 = \sum_{k \in C_0} S_k$. The equations follow the notations as explained next. For all nodes $n = 1, \ldots, N$ the real power injection is collected as a vector in $\mathbf{p} := [p_1 \cdots p_N]^\top$, reactive power injection $\mathbf{q} := [q_1 \cdots q_N]^\top$, where these injections can be written as

$$\mathbf{p} = \mathbf{p}^g - \mathbf{p}^c \tag{3.2a}$$

$$\mathbf{q} = \mathbf{q}^g - \mathbf{q}^c. \tag{3.2b}$$

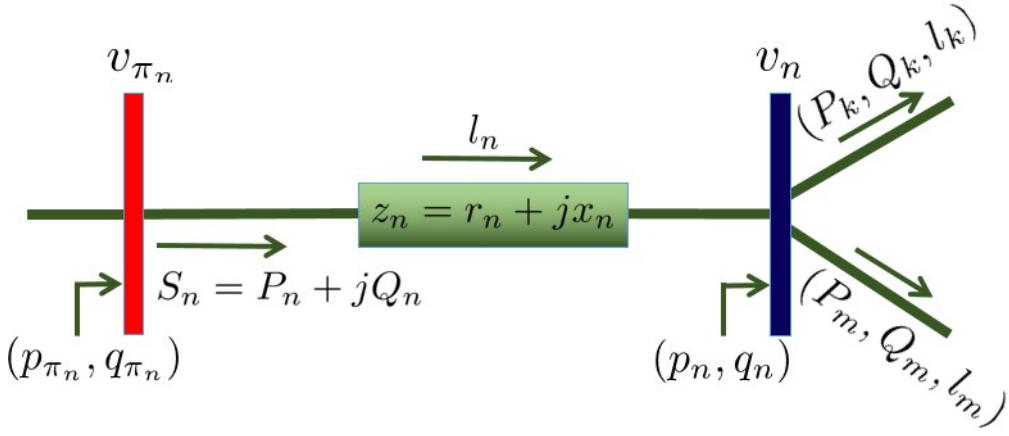


Figure 3.1: Line n feeding bus n from its parent bus π_n.

Complex nodel injections are $\mathbf{s} = \mathbf{p} + j\mathbf{q}$; squared voltage magnitudes are stacked as $\mathbf{v} := [v_1 \cdots v_N]^\top$. All lines have resistance, reactance collected together as $\mathbf{r} := [r_1 \cdots r_N]^\top$, $\mathbf{x} := [x_1 \cdots x_N]^\top$ respectively. Real and reactive line flows are defined as $\mathbf{P} := [P_1 \cdots P_N]^\top$, and $\mathbf{Q} := [Q_1 \cdots Q_N]^\top$ respectively, complex power flows are given as $\mathbf{S} = \mathbf{P} + j\mathbf{Q}$.

As per (3.1c) there exists a non linearity which complicates the power flow equations. To take care of this, the distribution grid is often remodelled as a linear model using *linear distribution flow (LDF)* model [23]. It must be noted that since the line resistance, reactance is small and its multipication with squared current magnitude will be even lesser the, for evaluating the power flow equations at flat voltage profile, the last summands in the right-hand sides of (3.1a)–(3.1b) can be dropped to formulate them as a linearized model.

Grid connectivity is captured in the branch-bus injection matrix $\tilde{\mathbf{A}}$. $\tilde{\mathbf{A}} \in \{0, \pm 1\}^{M \times (N+1)}$ it can be partioned into its first column and rest of its columns as $\tilde{\mathbf{A}} = [\mathbf{a}_0 \ \mathbf{A}]$. Thus, the reduced branch-bus injection matrix can be written as $\mathbf{A}$. $\mathbf{A}$ is a square matrix which is

invertible where $\mathbf{F} := \mathbf{A}^{-1}$ [24]. Also, $\mathbf{A}$ follows

$$\mathbf{a}_0 + \mathbf{A}\mathbf{1} = \mathbf{0}. \tag{3.3}$$

Using this connectivity matrix the LDF can rewritten as

$$\mathbf{s} = \mathbf{A}^\top \mathbf{S} \tag{3.4a}$$

$$\mathbf{A}\mathbf{v} = 2\,\mathrm{Re}[\mathrm{dg}(\mathbf{r} - j\mathbf{x})\mathbf{S}] - \mathbf{a}_0 v_0 \tag{3.4b}$$

where v_0 is the squared voltage magnitude at the substation. Using (3.3) $\mathbf{S}$ can be eliminated from (3.4) giving the squared bus voltage magnitude for all buses $n = 1, \ldots, N$ as [25]

$$\mathbf{v} \simeq 2\mathbf{R}\mathbf{p} + 2\mathbf{X}\mathbf{q} + v_0 \mathbf{1}_N \tag{3.5a}$$

where

$$\mathbf{R} := \mathbf{F}^\top \mathrm{dg}(\mathbf{r})\mathbf{F} \tag{3.5b}$$

$$\mathbf{X} := \mathbf{F}^\top \mathrm{dg}(\mathbf{x})\mathbf{F}. \tag{3.5c}$$

Since $\mathbf{F} \geq 0$, the matrices $\mathbf{R}$, $\mathbf{X}$ are also $\mathbf{R} \geq 0$, $\mathbf{X} \geq 0$. Moreover by properties of matrices, it can be easily seen that $\mathbf{R}$ and $\mathbf{X}$ are symmetric positive definite with positive entries [24]. Hence, bus voltages for all the buses in the grid are seen to increase if real or reactive power injections increase in the grid [24]. Since losses have been ignored in (3.1c) squared voltage magnitudes are an overestimate with respect to its original squared voltage magnitudes (3.5a) with the bias depending on ℓ_n's. But still, according to the numerical tests the approximation errors in voltage magnitudes is seen to be less than 0.001 pu [24, 26].

3.2 Problem Formulation

The active and reactive power injections $\mathbf{p}, \mathbf{q}$ can be decomposed into generation and inelastic load components as shown in (3.2). For known solar generation p_n^g and to comply with its apparent power limit $\bar{s}_n^g$, the reactive power injected by inverter n is constrained through the linear inequalities

$$|q_n^g| \leq \bar{q}_n^g := \sqrt{(\bar{s}_n^g)^2 - (p_n^g)^2}. \tag{3.6}$$

Moreover, to cater to the voltage regulation as mentioned in IEEE 1547 std, a linear set of inequalities can be added.

$$\underline{\mathbf{v}} \leq \mathbf{v} \leq \bar{\mathbf{v}} \tag{3.7}$$

where $\underline{\mathbf{v}}, \bar{\mathbf{v}}$ are set according to the regulation guidelines and are usually taken as $\pm(3\% - 5\%)$ about the nominal value.

To evaluate voltage deviations at each bus in the grid, let the sum of squared voltage magnitude deviations $\sum_{n=1}^{N} (v_n - v_0)^2$. Using the approximation in (3.5a) and upon dropping an inconsequential scaling factor of 4, the squared voltage deviations are

$$\Delta_s(\mathbf{q}^g) := \|\mathbf{R}\mathbf{p} + \mathbf{X}\mathbf{q}\|_2^2. \tag{3.8}$$

In addition to voltage deviation the ohmic power losses is another critical quantity in distribution grid operation. The active power losses can be expressed as $L = \sum_{n=1}^{N} r_n \ell_n$ or $\sum_{n=1}^{N} r_n \frac{P_n^2 + Q_n^2}{v_{\pi n}}$. For small voltage deviations, as advocated in [6] the power losses can be

approximated as

$$L = v_0^{-1} \left[\mathbf{P}^\top \operatorname{dg}(\mathbf{r}) \mathbf{P} + \mathbf{Q}^\top \operatorname{dg}(\mathbf{r}) \mathbf{Q} \right]. \tag{3.9}$$

Using (3.5b) and ignoring the inconsequential scaling by $v_0^{-1} \simeq 1$, the power losses can be expressed as

$$L = \mathbf{p}^\top \mathbf{R} \mathbf{p} + \mathbf{q}^\top \mathbf{R} \mathbf{q}. \tag{3.10}$$

Since $\mathbf{p}^\top \mathbf{R} \mathbf{p}$ is a constant for a given set of data, the control variable $\mathbf{q}$, which appears only the second summand is the the function of interest. The power loss function can be expressed as

$$L(\mathbf{q}^g) := \mathbf{q}^\top \mathbf{R} \mathbf{q}. \tag{3.11}$$

The positive definiteness of $\mathbf{R}$ guarantees that $L(\mathbf{q}^g)$ is a positively-valued convex quadratic function. The objectives of voltage deviations $\Delta_s(\mathbf{q}^g)$ and power loss $L(\mathbf{q}^g)$ are contradicting in general. Thus a multi objective problem can be solved to cater to these contradiction. A convex combination of the objectives can be posed to formulate the reactive control optimization problem given as

$$\begin{aligned} &\min_{\mathbf{q}^g} \ \lambda \Delta_s(\mathbf{q}^g) + (1-\lambda) L(\mathbf{q}^g) \\ &\text{s.to } \mathbf{q} \in \mathcal{Q} \end{aligned} \tag{3.12}$$

where the set $\mathcal{Q} \subseteq \mathbb{R}^N$ captures the linear constraints in (3.6) for all $n \in \mathcal{N}$. This formulation minimizes ohmic losses and voltage deviations according to the different parameter values $\lambda \in [0, 1]$, with respect to the apparent power constraints.

It must be noted, in the above formulation voltage regulation constraints are not enforced but the voltage deviations are minimized in the cost function. To enforce the voltage regulation constraints the problem can also be formulated as

$$\min_{\mathbf{q}^g} L(\mathbf{q}^g) \tag{3.13}$$

$$\text{s.to } \mathbf{q} \in \mathcal{Q}$$

$$\mathbf{v} \in \mathcal{V}$$

where the set $\mathcal{V} \subseteq \mathbb{R}^N$ captures the linear constraints in (3.7) for all $n \in \mathcal{N}$. Thus, in this formulation ohmic losses are minimized with respect to apparent power constraints and voltage regulation constraints.

Both formulations evaluate a reactive power control rule for a specified set of constraints. These formulations are exploited in this research to evaluate the reactive power support. The control rules can be centralized or decentralized depending upon the input features and model used in analysis. These are explained in detail in next chapter.

Chapter 4

Prior Works

Reactive power control can be provided for voltage regulation and minimization of ohmic losses in the power grid. The control can be provided by using the excess of PV inverter capacity to inject or reject reactive power in the distribution grid. With the amendments in IEEE 1547 standards the inverter can operate non unit factor and thus, provide reactive power support to the grid. Two approaches which can be used to provide reactive power support are centralized and decentralized (local) approaches. The uncertainty in loads and solar generation over the next reactive control period can be accounted through stochastic and robust formulations [27, 28].

Centralized approaches need global system information that is sent to the distribution network central controller (substation). The information sent requires network model and distribution network load estimation results [29]. The central controller evalautes the control

points according to the optimization technique, and sends them back to the inverters at each bus for implementation. Tap setting of transmission transformer is solved centrally in [30]. The paper evaluates the setting for a radial grid for increasing renewable penetration in the distribution grid. Another approach of model predictive control which uses time dependency is used in [31] to carry out voltage/VAR optimization with wind and PV as distributed generators. It considers the load uncertainities and captures optimal tap positions of on-load tap changer and switch statuses of capacitor banks. Additional method includes genetic algorithm used in [32]. Although, it considers forecast error of load and generation to evalaute voltage control global optimal solution is not gauranteed.

On the other hand decentralized (local) methods receives inputs from local or neighbouring and evaluates reactive power support locally. Purely localized schemes suggest having inverters implement Volt-VAR or Watt-VAR curves given only local measurements [6]. Although local Q-V rules have been analytically shown to be stable and fast-converging, their equilibria unfortunately do not coincide with the sought OPF minimizers [12, 33, 24]. In fact, there exist cases where local rules perform worse than the no-reactive support option [29].

Few prior works are explained below in detail, which motivated us to formulate the reactive power control problem as a data driven learning algorithm.

4.1 Local Rules

Local rules to evaluate reactive power control are introduced in [6]. The paper uses only local injection data to evaluate the reactive power support with respect to apparent power constraints. For the apparent power constraint (3.6) the size of PV inverter is considered to be constant. The apparent power inequality has been depicted in Fig. 4.1. As it can

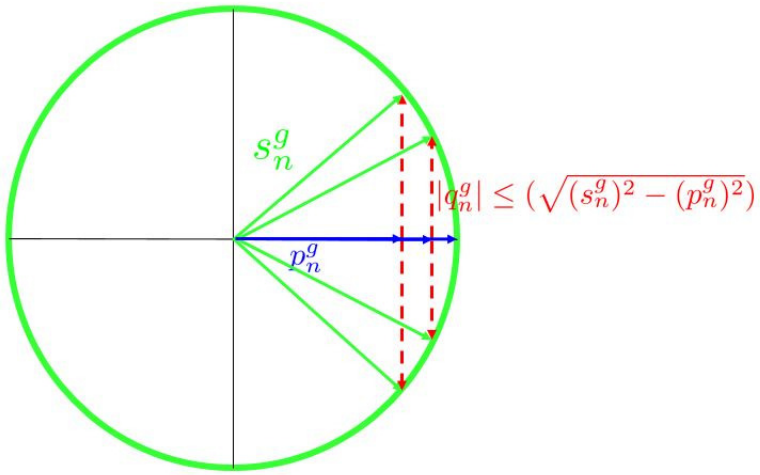


Figure 4.1: Apparent Power of the inverter.

be seen from the phasor diagram in Fig. 4.1, when the active power generation from PV is maximum that is $p_n^g = \bar{s}_n^g$, there is no room for reactive power to be dispatched and hence $q_n^g = 0$. But, if $p_n^g \leq \bar{s}_n^g$, reactive power can be dispatched according to (3.6). The paper emphasizes on developing rules which dispatch reactive power according to a scheme which reduces ohmic power losses or voltage deviations in the grid as given in (3.12). The ohmic losses are taken as (3.10). Another cost inherent in the distribution grid is voltage deviation.

The author devises a rule to minimize the voltage deviation of each bus given as

$$\delta V = \max_{n} \left| \frac{V_n - V_0}{V_0} \right| < \epsilon, \tag{4.1}$$

where δv is the maximum absolute voltage deviation with respect to substation voltage V_0, V_n is the voltage at bus n, ϵ is the voltage deviation the grid must follow for example $\epsilon \simeq (0.03 - 0.05)$. The author minimizes the voltage deviation and ohmic power losses with respect to the power flow equations as

$$\arg\min_{\mathbf{q}^g} \; [L, \delta V]^\top \tag{4.2a}$$

$$\text{s.to } (3.4)(3.6) \tag{4.2b}$$

The scheme in [6] evaluates $\mathbf{q}^g$ using only the local inputs, that is $q_n^g = F_n(p_n^g,\ p_n^c,\ q_n^c)$. A helper function is introduced in [6] to impose the apparent power constraint on the inverter. This *helper* function, Constr is given as

$$\text{Constr}(\mathbf{x}, \bar{\mathbf{x}}) = \begin{cases} \mathbf{x} & |\mathbf{x}| \leq \bar{\mathbf{x}} \\ (\mathbf{x}/|\mathbf{x}|)\bar{\mathbf{x}}, & \text{otherwise} \end{cases} \tag{4.3}$$

where $\mathbf{x}$ is the variable under minimization and $\bar{\mathbf{x}}$ is the maximum permissible value of $\mathbf{x}$. Thus, using the above helper function objective, the loss minimization can be evaluated as

$$\mathbf{q}^g_{\text{loss}} = \text{Constr}(\mathbf{q}, \bar{\mathbf{q}}) \tag{4.4}$$

where $\bar{\mathbf{q}}$ is given by (3.6). For evaluating the voltage deviation, the scheme assumes a constant ratio $\alpha = r_n/x_n$ for each line. Thus for an objective of minimizing voltage deviation,

the helper function is given as

$$\mathbf{q}^g_{\text{dev}} = \text{Constr}(\mathbf{q}^c + \frac{\mathbf{p}^c - \mathbf{p}^g}{\alpha}, \bar{\mathbf{q}}) \tag{4.5}$$

where $\bar{\mathbf{q}}$. The scheme, then evaluates a rule that can strike a balance between the voltage deviations and ohmic power losses on the basis of parameter λ given as

$$\mathbf{q}_g = \text{Constr}(\lambda \mathbf{q}^g_{\text{dev}} + (1-\lambda)\mathbf{q}^g_{\text{loss}}, \bar{\mathbf{q}}) \tag{4.6}$$

where λ is the trade off parameter, controlling the participation of each objective. Thus, this paper solves the problem (3.12) locally to provide reactive power support. Few shortcomings in this paper are: the local rule reported above evaluates the reactive power injection assuming all the lines have same α but, this is not true in the distribution grid where the ratio of r_n/x_n varies with respect to each line and cannot be considered as a constant; the scheme only depends on the local injections and fails to improve the profile of other buses for cases in which the inverter has room for reactive power injection but fails to provide it due to absence of any input for it. Although these rules provide control locally and do not require communication yet they fail to provide reactive power support optimally.

4.2 Optimal Rules

Centralized approaches require data from all the buses to evaluate reactive power support optimally. To minimize losses and voltage deviations an optimal rule can be evaluated as given in (3.12). In fact, problem (3.12) can be simplified as detailed in the ensuing lemma.

Lemma 1. *Problem* (3.12) *can be equivalently expressed as*

$$\tilde{\mathbf{q}} := \arg\min_{\mathbf{q}^g} \ \|\mathbf{C}\mathbf{q}^g + \mathbf{y}\|_2^2 \tag{4.7}$$

$$\text{s.to } \mathbf{q}^g \in \mathcal{Q}$$

where matrix $\mathbf{C} := [(1-\lambda)\mathbf{R} + \lambda\mathbf{X}^2]^{1/2}$*; vector* $\mathbf{y} := \mathbf{C}^{-1}[-(1-\lambda)\mathbf{R}\mathbf{q}^c + \lambda\mathbf{X}\mathbf{R}(\mathbf{p}^g - \mathbf{p}^c) - \lambda\mathbf{X}^2\mathbf{q}^c]$*; and the operator* $[\cdot]^{1/2}$ *represents the unique square root of a symmetric positive definite matrix.*

It is worth mentioning, $\mathbf{C}$ only depends on the distribution network, while $\mathbf{y}$ and set $\mathcal{Q}$ encompasses variable loads and solar generations collected in vector $\boldsymbol{\chi} := [(\mathbf{p}^c - \mathbf{p}^g)^\top \ \bar{\mathbf{q}}^{g\top} \ \mathbf{q}^{c\top}]^\top$.

Moreover, to minimize losses with respect to voltage regulation as constraints, the problem (4.8) can be simplified as detailed in the ensuing lemma.

Lemma 2. *Problem* (3.12) *can be equivalently expressed as*

$$\tilde{\mathbf{q}} := \arg\min_{\mathbf{q}^g} \ \mathbf{q}^\top\mathbf{R}\mathbf{q} \tag{4.8}$$

$$\text{s.to } \mathbf{q}^g \in \mathcal{Q}$$

$$\mathbf{v} \in \mathcal{V}.$$

Set $\mathcal{V}$ encompasses variable loads and solar generations collected in vector $\boldsymbol{\chi} := [(\mathbf{p}^c - \mathbf{p}^g)^\top \ \bar{\mathbf{q}}^{g\top} \ \mathbf{q}^{c\top}]^\top$.

Thus while evaluation, the model follows three specific steps to evaluate the optimal reactive power injection in the grid:

1. Each bus communicates its $(\mathbf{p}^c - \mathbf{p}^g, \bar{\mathbf{q}}^g, \mathbf{q}^c)$ to the operator.

2. The operator solves (4.7) knowing the current $\boldsymbol{\chi}$.

3. The operator sends the optimal setpoints $\tilde{\mathbf{q}}$ to inverters.

To take into consideration the uncertainity in solar generation the formulation can be extended as a process repeated on a per-minute basis or more frequently to reduce the communication and cyber overhead. As seen in the process of communicating data between nodes S1) establishes N inverter-utility communication links, and S3) requires another N utility-inverter communication links. Thus, the operator could adopt a scenario sample approach. In this rather than evaluating the problem for a long period of time, the operator may decide to issue setpoints less frequently, say every 10 minutes. But, one has to take care for the variability in $\boldsymbol{\chi}$ the feeder may encounter over this longer period. Thus, the problem (4.7) can be posed as draw a collection of anticipated $\{\boldsymbol{\chi}_t\}_{t=1}^T$; and solve the problem

$$\hat{\mathbf{q}} := \arg\min_{\mathbf{q}^g} \sum_{t=1}^{T} \|\mathbf{C}\mathbf{q}^g + \mathbf{y}_t\|_2^2 \tag{4.9}$$

$$\text{s.to } \mathbf{q}^g \in \mathcal{Q}_t, \quad t = 1, \ldots, T$$

The constrained ohmic losses minimization problem with respect to voltage regulation constraints can be solved as

$$\tilde{\mathbf{q}} := \arg\min_{\mathbf{q}^g} \sum_{t=1}^{T} \|\mathbf{R}^{1/2}(\mathbf{q}^g - \mathbf{q}_t^c)\|_2^2 \tag{4.10}$$

$$\text{s.to } \mathbf{q}^g \in \mathcal{Q}_t, \quad t = 1, \ldots, T$$

$$\mathbf{v} \in \mathcal{V}_t, \quad t = 1, \ldots, T,$$

where $(\mathbf{y}_t, \mathcal{Q}_t, \mathcal{V}_t)$ depend on $\boldsymbol{\chi}_t$ for $t = 1, \ldots, T$. One of the drawbacks of this approach is, a scenario-based problem (4.9) is solved rather than the deterministic problem in (4.7). The procedure steps S1)-S3) is solved less frequently for example 10 minutes, but the control process makes it inflexible for the setpoints $\hat{\mathbf{q}}$ to change. The control points remain unchanged over the next 10 minutes but to take into consideration the uncertainity in $\boldsymbol{\chi}_t$, affine control policies in the form of $\mathbf{q}^g(\boldsymbol{\chi}_t)$ have been suggested in [29], [34].

Although this approach solves the problem optimally, it requires frequent communication between each bus and substation making it a computationally and communication-wise challenging task.

4.3 Affine Policies

In an affine policy or linear decision rule, the solutions to the problem are restricted to be an affine function of the uncertain parameters [34]. In this section two works are explained which formulate the reactive control problem as affine policies for a robust 4.3.1 and probabilistic 4.3.2 approach. These two approaches follow different formulations wherein one minimizes a multi-objective function without any voltage constraints whilst, the other minimizes ohmic losses under voltage regulation as constraints.

4.3.1 Affinely Adjustable Robust Control Policies

A robust affinely adjustable control rule for evaluating the reactive power injection from PV inverters is introduced in [29]. The control rule focuses on minimizing the convex combination of power losses and voltage deviations in the grid. The reactive power injection is derived from an affinely adjustable robust optimization framework using a decentralized multi objective control formulation [29]. The method uses the approximate model $\mathbf{Yv} = \mathbf{s}^*$ but, for convinience and uniformity the problem is reformulated for the linear model explained in section 3.1. The paper uses a linear decision rule for evaluating the reactive power injection at each inverter. Moreover, an uncertainty budget is allocated for the robustness of the PV active generation using the dualization of constraints. To minimize power losses the epigraph trick is used, the problem can be formulated as

$$\begin{aligned} &\min_{\mathbf{t}^{loss}} \ (\mathbf{t}^{\text{loss}})^\top \mathbf{t}_{loss} + \mathbf{p}^\top \mathbf{R}\mathbf{p} \\ &\text{s.to} \quad -t_n^{loss} \leq \sqrt{r_n}\mathbf{f_n^T}\mathbf{q} \leq t_n^{loss} \ n = 1, \ldots, N \end{aligned} \tag{4.11}$$

Moreover, using (3.5a) voltage deviation with respect to slack bus voltage is given as

$$\delta v = \mathbf{v} - \mathbf{1}_N v_0 = \mathbf{R}\mathbf{p} + \mathbf{X}\mathbf{q}. \tag{4.12}$$

Using the epighraph trick to minimize voltage deviation, the problem can be formulated as

$$\begin{aligned} &\min_{\mathbf{t}^{dev}} \ \mathbf{1}_N^\top \mathbf{t}^{dev} \\ &\text{s.to} \ -\mathbf{t}^{dev} \leq \mathbf{R}\mathbf{p} + \mathbf{X}\mathbf{q} \leq \mathbf{t}^{dev} \end{aligned} \tag{4.13}$$

The apparent power constraint (3.6) in complex form can be written as $|\mathbf{p}^g + j\mathbf{q}^g| \leq \bar{\mathbf{s}}^g$. Since it's difficult to handle the above constraint, [29] uses the polyhedral formulation for $2k$ sided polygon as

$$-\bar{\mathbf{s}}^g \leq cos(\alpha_l)\mathbf{p}^g + sin(\alpha_l)\mathbf{q}^g \leq \bar{\mathbf{s}}^g \qquad l = 1, 2, ..k, \alpha_l = \frac{\pi}{k} \tag{4.14}$$

Consider a matrix which stores all the cosine and sine coefficients of $\boldsymbol{\alpha} = [\alpha_1, \alpha_2....\alpha_k]^T$, in $\mathbf{H}, \mathbf{G}$ respectively. This is defined as

$$\mathbf{H} = cos(\boldsymbol{\alpha}) \otimes \mathbf{I}_N \tag{4.15}$$

$$\mathbf{G} = sin(\boldsymbol{\alpha}) \otimes \mathbf{I}_N. \tag{4.16}$$

Consider $\mathbf{K} = [\mathbf{H}^\top \ \mathbf{G}^\top; -\mathbf{H}^\top \ -\mathbf{G}^\top]^\top$. Reformulating (4.14) the apparent power constraint can be reformulated as

$$\mathbf{K} \begin{bmatrix} \mathbf{p}^g \\ \mathbf{q}^g \end{bmatrix} \leq \begin{bmatrix} \mathbf{s}^g \\ \mathbf{s}^g \end{bmatrix}. \tag{4.17}$$

As seen earlier, the paper also solves a multi-objective problem to minimize power losses and voltage deviations. The problem is formulated as

$$\min_{\mathbf{q}, \mathbf{t}^{\text{dev}}, \mathbf{t}^{\text{loss}}} \quad (1-\lambda)((\mathbf{t}^{loss})^\top \mathbf{t} + \mathbf{p}^\top \mathbf{R}\mathbf{p}) + \lambda(\mathbf{1}_N^\top \mathbf{t}^{dev}) \tag{4.18}$$

$$\text{s.to} \quad -t_n^{\text{loss}} \leq \sqrt{r_n}\mathbf{f_n^T}\mathbf{q} \leq t_n^{\text{loss}} \quad n = 1, \ldots, N$$

$$-\mathbf{t}^{\text{dev}} \leq \mathbf{R}\mathbf{p} + \mathbf{X}\mathbf{q} \leq \mathbf{t}^{\text{dev}}$$

$$\mathbf{K}[(\mathbf{p})^\top (\mathbf{q})^\top]^\top \leq [(\mathbf{s}^g - \mathbf{H}\mathbf{p}^c - \mathbf{G}\mathbf{q}^c)^\top \ (\mathbf{s}^g - \mathbf{H}\mathbf{p}^c - \mathbf{G}\mathbf{q}^c)^\top]^\top$$

where λ is the participation factor of each objective function. The author considers uncertainty in solar generation, the active power generation belongs to the uncertainty set

$$\mathbf{p} \in [\mathbf{0}, \mathbf{p}^g_{\max}] \tag{4.19}$$

where $\mathbf{p}^g_{\max} := [p^g_{\max_1}\ p^g_{\max_1} \cdots\ p^g_{\max_N}]^\top$ is the maximum value of active power generation. A linear decision rule for the reactive power control is introduced on the reactive power control, given as

$$\mathbf{q}^g = \mathbf{q}^0 + \boldsymbol{\beta}\mathbf{p}^g. \tag{4.20}$$

When these policies are applied to reactive power generation a robust optimization problem is formulated using an affinely adjustable robust counterpart (AARC) [29]. The problem is posed as

$$\begin{aligned} \min_{\mathbf{u},\boldsymbol{\beta},\mathbf{q}^0,\boldsymbol{\theta}} \quad & \frac{1}{2}\sum_j a_j u_j^2 + \sum_j c_j u_j \\ \text{s.to} \quad & \sum_j A_{ij}u_j + \sum_j B_{ij}q_{0j} + \sum_j \theta_{ij} p^g_{max_j} \le b_i,\ \forall i \\ & \theta_{ij} \ge 0,\ \theta_{ij} \ge B_{ij\beta_j - C_{ij}},\ \forall i, \forall j \end{aligned} \tag{4.21}$$

where $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{A}, \mathbf{B}, \mathbf{C}$ are appropriate vectors and matrices as derived from (4.18),(4.19),(4.20).

The robust formulation allows the local control to remain within its maximum capacity for all realizations of active power. This problem evaluates the coefficients for the linear decision rule and optimizes the power losses and voltage deviations according to the participation factor. Few shortcomings of this implementation are that the proposed algorithm fails to evaluate voltage deviation according to the regulation limits as it only minimizes

the voltage deviation without any regulation constraint on it; the minimum value for photo voltaic active power has been assumed to be zero which is unrealistic.

4.3.2 Probabilistic Control Policies

A probabilistic approach to evaluate reactive power control is introduced in [34]. The control rules are assumed to be affine policies of the real and reactive power consumption and active power injection. The inputs are stacked in a vector $\mathbf{w} := [\mathbf{p}^c\ \mathbf{q}^c\ \mathbf{p}^g]^\top \in \mathbb{R}^{3N}$ where, $\mathbf{w}$ stores the injection information from all the buses and could be as large as $3N$ and reactive power control variable is given as $\mathbf{q}^g$. To take care of the uncertainty in the consumption and active power injections, a Gaussian distribution is assumed for a given mean and covariance. The author evaluates the problem (3.12) for minimizing ohmic losses that is $\lambda = 0$. The voltage constrainst can be reformulated as

$$\underline{\mathbf{v}} \leq \mathbf{D}\mathbf{q}^g + \mathbf{E}\mathbf{w} + \mathbf{1}_N v_0 \leq \bar{\mathbf{v}}$$

where $\mathbf{D}, \mathbf{E}$ are appropriate matrices for input vector $\{\mathbf{q} = \mathbf{q}^g - \mathbf{q}^c, \mathbf{w}\}$, $\mathbf{1}_N$ is a vector of all 1 with N rows. Thus, the optimization problem is given as

$$\begin{aligned} \min_{\mathbf{q}^g} \quad & \mathbf{q}^{g\top}\mathbf{B}\mathbf{q}^g + \mathbf{w}^\top\mathbf{C}\mathbf{w} + \mathbf{q}^{g\top}\mathbf{D}\mathbf{w} + \mathbf{w}^\top\mathbf{D}^\top\mathbf{q}^g \\ \text{s.to} \quad & \mathbf{v} = \mathbf{D}\mathbf{q}^g + \mathbf{E}\mathbf{w} + \mathbf{1}_N v_0 \\ & -\bar{\mathbf{q}}^g \leq \mathbf{q}^g \leq \bar{\mathbf{q}}^g \\ & \mathbf{K}\mathbf{v} \leq \boldsymbol{\kappa} \end{aligned} \tag{4.22}$$

where $\boldsymbol{\kappa} = [\underline{\mathbf{v}} \ -\bar{\mathbf{v}}]^\top$. It can be seen from (3.5a) the uncertainity in $\mathbf{w}$ affects the voltages at each bus. The author utilizes a probabilistic approach for $\boldsymbol{\alpha}$ as the probability level given as

$$Prob[\mathbf{Kv} \leq \boldsymbol{\kappa}] \geq \boldsymbol{\alpha}. \tag{4.23}$$

An affine policy for the reactive power control is considered in [34] which is dependent on $\mathbf{w}$ given as

$$\mathbf{q}^g = \mathbf{Mw} + \mathbf{h} \tag{4.24}$$

where $\mathbf{M}, \mathbf{h}$ will be the optimization variables estimated for the problem. Moreover since the objective function becomes uncertain due to the uncertainty in $\mathbf{w}$ the expected value of losses will be minimized. The probabilistic constraint (4.23) can be converted to a deterministic constraint [35] by considering a probability distribution on $\mathbf{w}$. Substituting the affine policy into the constraint (4.12) it can be reformulated as

$$\text{Prob}[\mathbf{K}(\mathbf{D}(\mathbf{Mw} + \mathbf{h}) + \mathbf{Ew} + \mathbf{1}_N v_0) \leq \boldsymbol{\kappa}] \geq \boldsymbol{\alpha} \tag{4.25}$$

$$\mathbf{K}(\mathbf{DM} + \mathbf{E})\bar{\mathbf{w}} + \boldsymbol{\phi}^{-1}(\boldsymbol{\alpha})\|(\mathbf{G}^\top(\mathbf{DM} + \mathbf{E})^\top\mathbf{K})\|_2 \leq \boldsymbol{\kappa} - \mathbf{K}(\mathbf{Dh} + \mathbf{1}_N v_o) \tag{4.26}$$

where from Cholsky decomposition for distribution parameters of $\mathbf{w}$, covariance $\boldsymbol{\Sigma} = \mathbf{GG}^\top$, $\boldsymbol{\phi}^{-1}(\boldsymbol{\alpha})$ is inverse cumulative distribution function for a standard Gaussian random variable at a probability level $\boldsymbol{\alpha}$. Thus, a second order cone program (SOCP) constraint is formulated in [34] which utilizes a probalistic approach to satisfy voltage and apparent power constraints for a given distribution of $\mathbf{w}$. Few shortcomings of this approach [34] are, firstly the constraints on voltage and reactive power injection are not always satisfied and is supposed to

be satisfied only $\boldsymbol{\alpha}\%$ of the time; secondly, the affine rule needs injection inputs from all nodes making it computationally and communication-wise challenging. Thus, this approach would work well if the communication bandwidth availability is high, and the substation is able to communicate the reactive power injections to the inverters quickly.

4.4 Proposed Formulation

As seen from the prior works, existing approaches either solve problem locally or centrally considering a linear decision rule on the input parameters. Local rules suffer from sub-optimality while the central rules require heavy communication and computation. In this research a decentralized approach is developed to evaluate reactive power control policies for a multi-objective or a voltage regulation constrained problem. The policies are modeled as non-linear functions of the input feature vector. The proposed approach utilizes kernel based learning algorithm to evaluate control policies on the basis of input scenario data. Thus, the approach is practically feasible and hits the sweet spot in the performance-communication trade-off. Methodology to evaluate reactive power control is explained in the next section.

Chapter 5

Methodology

To keep the network's net power loss and voltage drop minimum, the model needs to evaluate reactive power to be injected at each node. This injection varies according to the size of the inverter, it's placement, network's topology and configuration. In this research, the reactive power injection q_n^g by inverter n can be modeled as

$$q_n^g(\mathbf{z}_n) = f_n(\mathbf{z}_n) + b_n \tag{5.1}$$

whose inputs $(f_n, \mathbf{z}_n, b_n)$ are described below.

5.1 Control Policies

Controller inputs: Vector $\mathbf{z}_n \in \mathcal{Z}_n \subseteq \mathbb{R}^{M_n}$ is given as an input to inverter to evaluate the reactive power injection at node n. Vector $\mathbf{z}_n$ can purely depend on its local values or have

few non-local or neighbouring inputs appended in the end like real power flow, squared magnitude. With only local inputs $\mathbf{z}_n := [p_n^g - p_n^c \;\; \bar{q}_n^g \;\; q_n^c]^\top$ where $\bar{q}_n^g := \sqrt{(\bar{s}_n^g)^2 - (p_n^g)^2}$. The first entry of $\mathbf{z}_n$ relates to the net active power injection at node n; second entry gives the maximum reactive injection possible at node n; and third entry relates to the reactive load at that node. This input is thus purely local to its respective node. Various variations in $\mathbf{z}_n$ can be done to improve the fit of control policies. This can be done by appending few important global inputs to the vector $\mathbf{z}_n$. It must be noted, by appending squared voltage magnitude v_n in $\mathbf{z}_n$, the stability of the resultant closed-loop control would be hard to analyze even when f_n is linear [25, 33, 24].

Since these inputs are locally available the burden on communication channels is minimal and the evaluation is fast. Ideally, if communication resources are abundant, the uncertain quantities from all buses $\{\bar{q}_n^g, p_n^c - p_n^g, q_n^c\}_{n\in\mathcal{N}}$ could be forwarded to all inverters. Thus, in this case the control inputs in $\mathbf{z}_n$ would be identical to all the inverters in the network and would be of a larger size $3N$. It must be noted that the input variables are extremely flexible and can be varied depending upon the available communication bandwidth of the network. Non-local, shared control inputs can also be added to input vector $\mathbf{z}_n$ to study the results for local inputs and few global inputs common for all inverters for example $\mathbf{z}_n := [p_n^g - p_n^c \;\; \bar{q}_n^g \;\; q_n^c \;\; P_i \;\; P_j \;\; P_k]^\top$ where i, j, k represent the line numbers and P_i is the real power flow on line i. These lines are selected on the bases of topology of each network and these inputs will be identical to each inverter.

Controller function: Next step involves the control function policy f_n. The control

function can be evaluated as a linear or non linear policy as delineated below. Using the theory of kernel based learning, the reactive power control from inverter n is postulated to lie in the RKHS.

$$\mathcal{H}_{\mathcal{K}_n} := \left\{ f_n(\mathbf{z}_n) = \sum_{t=1}^{\infty} K_n(\mathbf{z}_n, \mathbf{z}_{n,t}) a_{n,t},\ a_{n,t} \in \mathbb{R} \right\}. \tag{5.2}$$

that is uniquely determined by the kernel function $K_n : \mathcal{Z}_n \times \mathcal{Z}_n \to \mathbb{R}$.

Linear policies can be implemented by evaluating a linear kernel $K_n(\mathbf{z}_{n,t}, \mathbf{z}_{n,t'}) = \mathbf{z}_{n,t}^\top \mathbf{z}_{n,t'}$.

Nonlinear policies can be designed by selecting a polynomial kernel $K_n(\mathbf{z}_{n,t}, \mathbf{z}_{n,t'}) = \left(\mathbf{z}_{n,t}^\top \mathbf{z}_{n,t'} + \gamma\right)^\beta$, or a Gaussian kernel $K_n(\mathbf{z}_{n,t}, \mathbf{z}_{n,t'}) = \exp\left(-\|\mathbf{z}_{n,t} - \mathbf{z}_{n,t'}\|_2^2/\gamma\right)$ with design parameters β and $\gamma > 0$ or a linear combination of linear, polynomial, gaussian kernels.

Intercept: The control function also needs to evaluate an intercept value $b_n \in \mathbb{R}$ in (5.1). Although it could be incorporated into f_n, e.g., by augmenting $\mathbf{z}_n$ by a constant entry of 1, it is usually kept separate to avoid its penalization through $\|f\|_{\mathcal{K}_n}$.

5.2 Learning Policies from Scenarios

With the control function and input vector finalized, the reactive power control policies (5.1) needs to be evaluated for the input data. As per the policy the n-th entry of $\mathbf{q}^g$ for scenario t can be replaced with the policy $q_n^g(\mathbf{z}_{n,t}) = f_n(\mathbf{z}_{n,t}) + b_n$ from (5.1). Thus the algorithm evaluates the optimal function and intercept pairs $\{\hat{f}_n, \hat{b}_n\}_{n=1}^N$ which can be found via the

functional minimization

$$\min \sum_{t=1}^{T} C\left(\mathbf{y}_t, \{f_n(\mathbf{z}_{n,t})\}_n, \mathbf{b}\right) + \mu P\left(\{\|f_n\|_{\mathcal{K}_n}\}\right) \tag{5.3a}$$

$$\text{over } \{f_n \in \mathcal{H}_{\mathcal{K}_n}\}_{n=1}^{N}, \mathbf{b} \tag{5.3b}$$

$$\text{s.to } |f_n(\mathbf{z}_{n,t}) + b_n| \leq \bar{q}^g_{n,t}, \ \forall n, t \tag{5.3c}$$

$$\underline{v}_{n,t} \leq r_n(p^g_{n,t} - p^c_{n,t}) + x_n(f_n(\mathbf{z}_{n,t}) + b_n - q^c_{n,t}) \leq \bar{v}_{n,t}, \ \forall n, t \tag{5.3d}$$

where $\mathbf{b} := [b_1 \ \cdots \ b_N]^\top$, constraint (5.3c) represents the apparent power constraint, contriant (5.3d) represents the voltage regulation constraint. The regularizer $P\left(\{\|f_n\|_{\mathcal{K}_n}\}\right)$ has been added in (5.3a) to avoid overfitting of control policies to scenario data.

In a regular machine learning regression setup, the dependency between input (feature) data and output (target) data is analyzed and the closest fit is evaluated. In this formulation the grid quantities feeding each controller serves as a feature data, and the reactive injections as target values. Ideally, the designed function should behave well even for feature-target pairs not seen during the training or fitting process. In direct analogy, the inverter control policies are posed as a joint function fitting task based on scenario data. Once the functions (policies) have been designed, they can be applied to unseen feature data. While, the apparent power constraint (5.3c) is enforced for training data, it is possible for the policies obtained through (5.3) may not satisfy the apparent power limits for $\mathbf{z}_{n,t'}$'s with $t' \notin \{1, \ldots, T\}$ as policies has only been trained for the data upto scenario T in which the constraint is actively present. This limitation of kernel-based learning appears also in scenario-based and chance-constrained designs [34]. To overcome this limitation the reactive power evaluated at

t' scenario for node n can be heuristically projected within $[-\bar{q}^g_{n,t'}, +\bar{q}^g_{n,t'}]$ as

$$\left[q^g_n(\mathbf{z}_{n,t'})\right]_{\bar{q}^g_{n,t'}} := \max\left\{\min\left\{q^g_n(\mathbf{z}_{n,t'}), \bar{q}^g_{n,t'}\right\}, -\bar{q}^g_{n,t'}\right\}.$$

In this formulation, the optimal policies (functions) are evaluated for each inverter n separately as compared to single one evaluated in (2.17). But, the inverter policies are coupled through the loss term L in (2.17) as the voltage deviation and power losses are affected by each reactive power injection feeder wise. Similar multi-function setups can be found in collaborative filtering or multi-task learning [36]. In this thesis a regularizer is adopted which is separable over all the inverters given as

$$P\left(\{\|f_n\|_{\mathcal{K}_n}\}_{n=1}^N\right) = \sum_{n=1}^N \|f_n\|^2_{\mathcal{K}_n}. \tag{5.4}$$

The well known Representer's Theorem can be applied successively over n in (5.3). Thus, this ensures that

$$\hat{f}_n(\mathbf{z}_n) = \sum_{t=1}^T K_n(\mathbf{z}_n, \mathbf{z}_{n,t})\hat{a}_{n,t}.$$

still holds for all n. Thus, after the optimal policies are evaluated, the coefficients $\{\hat{a}_{n,t}\}_{n,t}$, the control policies $\{\hat{f}_n\}$ can be evaluated at any other point. As seen earlier, the inverter policy $\hat{f}_n$ over the test data $\{\mathbf{z}_{n,t}\}_{t=1}^T$ can be examined as

$$\hat{\mathbf{f}}_n = \mathbf{K}_n\hat{\mathbf{a}}_n, \quad \forall n \tag{5.5}$$

where $[\mathbf{K}_n]_{t,t'} = K_n(\mathbf{z}_{n,t}, \mathbf{z}_{n,t'})$ for $t, t' \in \{1, \ldots, T\}$, and $\hat{\mathbf{a}}_n := [\hat{a}_{n,1} \ \cdots \ \hat{a}_{n,T}]^\top$. Moreover the regualarizer term, the RKHS norms can be expressed as

$$\|f_n\|^2_{\mathcal{K}_n} = \hat{\mathbf{a}}_n^\top\mathbf{K}_n\hat{\mathbf{a}}_n, \quad \forall n. \tag{5.6}$$

Examples for various policies are delineated next.

5.2.1 Examples

Example 1: Affine policies. If the policies in (5.1) are affine, the sought policy functions can be written as

$$f_n(\mathbf{z}_n) = \mathbf{z}_n^\top \mathbf{w}_n, \quad \forall n \tag{5.7}$$

where $\mathbf{w}_n \in \mathbb{R}^{M_n}$'s are to be determined. Given input data $\{\mathbf{z}_{n,t}\}_{n\in\mathcal{N}}$ and the desired *feeder output* $\mathbf{y}_t$ for $t = 1, \ldots, T$, the $\{\mathbf{w}_n, b_n\}_{n\in\mathcal{N}}$ needs to be found via (5.3). Let us consider the input data for inverter n is collected in the $M_n \times T$ matrix

$$\mathbf{Z}_n := \begin{bmatrix} \mathbf{z}_{n,1} & \cdots & \mathbf{z}_{n,T} \end{bmatrix}. \tag{5.8}$$

According to the Representer's Theorem, the $\hat{\mathbf{w}}_n$ corresponding to the optimal f_n can be expressed as $\hat{\mathbf{w}}_n = \mathbf{Z}_n \hat{\mathbf{a}}_n$ for some $\hat{\mathbf{a}}_n$. Then, the policy evaluated at any input $\mathbf{z}_n$ is

$$q_n(\mathbf{z}_n) = \hat{f}_n(\mathbf{z}_n) + \hat{b}_n = \mathbf{z}_n^\top \mathbf{Z}_n \hat{\mathbf{a}}_n + \hat{b}_n.$$

Evaluating the policy specifically at the input data $\{\mathbf{z}_{n,t}\}_{t=1}^T$ provides

$$\hat{\mathbf{f}}_n = \mathbf{K}_n \hat{\mathbf{a}}_n$$

where the kernel matrix now is $\mathbf{K}_n = \mathbf{Z}_n^\top \mathbf{Z}_n$. The function norm is $\|f_n\|_{\mathcal{K}_n}^2 = \|\hat{\mathbf{w}}_n\|_2^2 = \hat{\mathbf{a}}_n^\top \mathbf{Z}_n^\top \mathbf{Z}_n \hat{\mathbf{a}}_n = \hat{\mathbf{a}}_n^\top \mathbf{K}_n \hat{\mathbf{a}}_n$.

Example 2: Non-linear policies. For non-linear policies the input vector $\mathbf{z}_n$ needs to transformed into a vector $\boldsymbol{\phi}_n := \phi_n(\mathbf{z}_n)$ where $\phi_n : \mathbb{R}^{M_n} \to \mathbb{R}^{\Phi_n}$ is a non-linear mapping. We saw that the simple linear kernel contained only terms of degree one only. But, the entries

of $\boldsymbol{\phi}_n$ could be for example the first- and second-order monomials formed by the entries of $\mathbf{z}_n$. Then, the control function can evaluated as

$$f_n(\mathbf{z}_n) = \boldsymbol{\phi}_n^\top \mathbf{w}_n \tag{5.9}$$

with $\mathbf{w}_n \in \mathbb{R}^{\Phi_n}$ constitutes a *non-linear* policy in $\mathbf{z}_n$. Thus, from Example 1, Example 2 can related using $\mathbf{K}_n = \boldsymbol{\Phi}_n^\top \boldsymbol{\Phi}_n$ with $\boldsymbol{\Phi}_n := [\boldsymbol{\phi}_{n,1} \ \cdots \ \boldsymbol{\phi}_{n,T}]$. Depending on the mapping ϕ_n, the vectors $\boldsymbol{\phi}_{n,t}$ may be of finite or infinite length [21]. It must be noted, since in kernel an inner product of $\boldsymbol{\phi}_{n,t}^\top \boldsymbol{\phi}_{n,t'}$ for any t and t' is evaluated the $\hat{f}_n$ does not deal directly with $\boldsymbol{\phi}_{n,t}$'s. These products can be efficiently calculated through the kernel function as $\boldsymbol{\phi}_{n,t}^\top \boldsymbol{\phi}_{n,t'} = K_n(\mathbf{z}_{n,t}, \mathbf{z}_{n,t'})$. Apparently, selecting the kernel function K_n induces the mapping ϕ_n.

5.3 Optimal Policies

5.3.1 Voltage Drop and Power Loss Minimization

After evaluating the control policies, the optimal function needs to be evaluated for minimizing the voltage drops and power losses in the network. Thus, the cost of (4.9) and the regularizer of (5.4) in (5.3) can be posed as a linearly-constrained quadratic program as explained next.

Lemma 3. *If the data-fitting term in* (5.3) *is selected as*

$$C\left(\mathbf{y}_t, \{f_n(\mathbf{z}_{n,t})\}_n, \mathbf{b}\right) = \|\mathbf{C}\mathbf{q}_t^g + \mathbf{y}_t\|_2^2, \quad t = 1, \ldots, T$$

and the regularizing term as

$$P\left(\{\|f_n\|_{\mathcal{K}_n}\}\right) = \sum_{n=1}^{N} \|f_n\|_{\mathcal{K}_n}^2$$

the functional optimization in (5.3) *can be transformed to the vector minimization*

$$\min \quad \frac{1}{T}(\|\mathbf{CQ} + \mathbf{Y}\|_F^2 + \mu \sum_{n=1}^{N} \mathbf{a}_n^\top \mathbf{K}_n \mathbf{a}_n) \tag{5.10a}$$

$$\text{over} \quad \mathbf{Q} \in \mathbb{R}^{N\times T}, \{\mathbf{a}_n \in \mathbb{R}^T\}_{n=1}^N, \mathbf{b} \in \mathbb{R}^N \tag{5.10b}$$

$$\text{s.to} \quad \mathbf{Q}^\top = [\mathbf{K}_1\mathbf{a}_1 + b_1\mathbf{1} \ \cdots \ \mathbf{K}_N\mathbf{a}_N + b_N\mathbf{1}] \tag{5.10c}$$

$$-\bar{\mathbf{q}}_n^g \leq \mathbf{K}_n\mathbf{a}_n + b_n\mathbf{1} \leq \bar{\mathbf{q}}_n^g, \ \forall n \tag{5.10d}$$

where $\mathbf{Y} := [\mathbf{y}_1 \ \cdots \ \mathbf{y}_T]$ *and the entries of vector* $\bar{\mathbf{q}}_n^g := [\bar{q}_{n,1}^g \ \cdots \ \bar{q}_{n,T}^g]^\top$ *have been defined in* (3.6).

Proof of Lemma (4)*:* Based on (5.1) and (5.5), the reactive power injection of inverter n for scenarios $t = 1, \ldots, T$ can be expressed by the vector $\mathbf{K}_n\mathbf{a}_n + b_n\mathbf{1}$. Then, the apparent power constraint for inverter n and across all scenarios can be expressed as in (5.10d).

Consider now the first summand in (5.10a). Based on the equality in (5.10c), the t-th column of $\mathbf{Q}$ denoted by $\mathbf{q}_t^g$ contains the reactive injections from all inverters for scenario t. Because the squared Frobenius norm of a matrix equals the sum of the squared ℓ_2-norms of its columns, it follows that $\sum_{t=1}^T \|\mathbf{C}\mathbf{q}_t^g + \mathbf{y}_t\|_2^2 = \|\mathbf{CQ} + \mathbf{Y}\|_F^2$. The second summand in (5.10a) follows directly from (5.6).

5.3.2 Power Loss Minimization under Voltage Constraints

Another model implemented in this research minimizes power loss in the network with respect to apparent power and voltage regulation constraints. This model stresses on keeping the voltage under regulation limits. Thus, after the control policies is evaluated, the optimal function is evaluated for minimizing power losses with respect to the constraints as described above. Again this problem can be posed as a linearly-constrained quadratic program as explained next.

Lemma 4. *If the data-fitting term in* (5.3) *is selected to minimze losses given as*

$$C\left(\mathbf{y}_t, \{f_n(\mathbf{z}_{n,t})\}_n, \mathbf{b}\right) = \|\mathbf{R}^{1/2}(\mathbf{q}_t^g - \mathbf{q}_t^c)\|_2^2, \quad t = 1, \ldots, T$$

and the regularizing term as

$$P\left(\{\|f_n\|_{\mathcal{K}_n}\}\right) = \sum_{n=1}^{N} \|f_n\|_{\mathcal{K}_n}^2$$

the functional optimization in (5.3) *can be transformed to the vector minimization*

$$\min \quad \frac{1}{T}(\|\mathbf{R}^{1/2}(\mathbf{Q} - \mathbf{Q}^c)\|_F^2 + \mu \sum_{n=1}^{N} \mathbf{a}_n^\top \mathbf{K}_n \mathbf{a}_n) \tag{5.11a}$$

$$\text{over} \quad \mathbf{Q} \in \mathbb{R}^{N \times T}, \{\mathbf{a}_n \in \mathbb{R}^T\}_{n=1}^N, \mathbf{b} \in \mathbb{R}^N \tag{5.11b}$$

$$\text{s.to} \quad \mathbf{Q}^\top = [\mathbf{K}_1 \mathbf{a}_1 + b_1 \mathbf{1} \ \cdots \ \mathbf{K}_N \mathbf{a}_N + b_N \mathbf{1}] \tag{5.11c}$$

$$-\bar{\mathbf{q}}_n^g \leq \mathbf{K}_n \mathbf{a}_n + b_n \mathbf{1} \leq \bar{\mathbf{q}}_n^g, \ \forall n \tag{5.11d}$$

$$\mathbf{V} = \mathbf{R}(\mathbf{P}^g - \mathbf{P}^c) + \mathbf{X}(\mathbf{Q} - \mathbf{Q}^c) + v_0 \mathbf{1}_{N \times T} \tag{5.11e}$$

$$\underline{\mathbf{v}_t} \leq \mathbf{v}_t \leq \bar{\mathbf{v}}_t, \ \forall t \tag{5.11f}$$

where real and reactive power consumption vectors are stacked as columns of the $N \times T$ *matrix* $\mathbf{P}^c := [\mathbf{p}_1^c \ \cdots \ \mathbf{p}_T^c]$ *and* $\mathbf{Q}^c := [\mathbf{q}_1^c \ \cdots \ \mathbf{q}_T^c]$ *respectively. Voltage at each bus are stacked as columns of the* $N \times T$ *matrix* $\mathbf{V} := [\mathbf{v}_1 \ \cdots \ \mathbf{v}_T]$. *Similarly real power generation is* $\mathbf{P}^g := [\mathbf{p}_1^g \ \cdots \ \mathbf{p}_T^g]$ *and the entries of vector* $\underline{\mathbf{v}}_t := [\underline{\mathbf{v}}_1 \ \cdots \ \underline{\mathbf{v}}_N]^\top$, $\bar{\mathbf{v}}_t := [\bar{\mathbf{v}}_1 \ \cdots \ \bar{\mathbf{v}}_N]^\top$, $\bar{\mathbf{q}}_n^g := [\bar{q}_{n,1}^g \ \cdots \ \bar{q}_{n,T}^g]^\top$ *where the limits for voltage regulation are defined in* (3.7), *reactive power have been defined in* (3.6).

Proof of Lemma (4)*:* Based on (5.1) and (5.5), the reactive power injection of inverter n for scenarios $t = 1, \ldots, T$ can be expressed by the vector $\mathbf{K}_n \mathbf{a}_n + b_n \mathbf{1}$. Then, the apparent power constraint for inverter n and across all scenarios can be expressed as in (5.11d). From the LDF equations (3.5a) voltage constraint is expressed as shown in (5.11e). To keep the voltage within limits linear constraint is added for all buses N across all scenarios that is given in (5.11f).

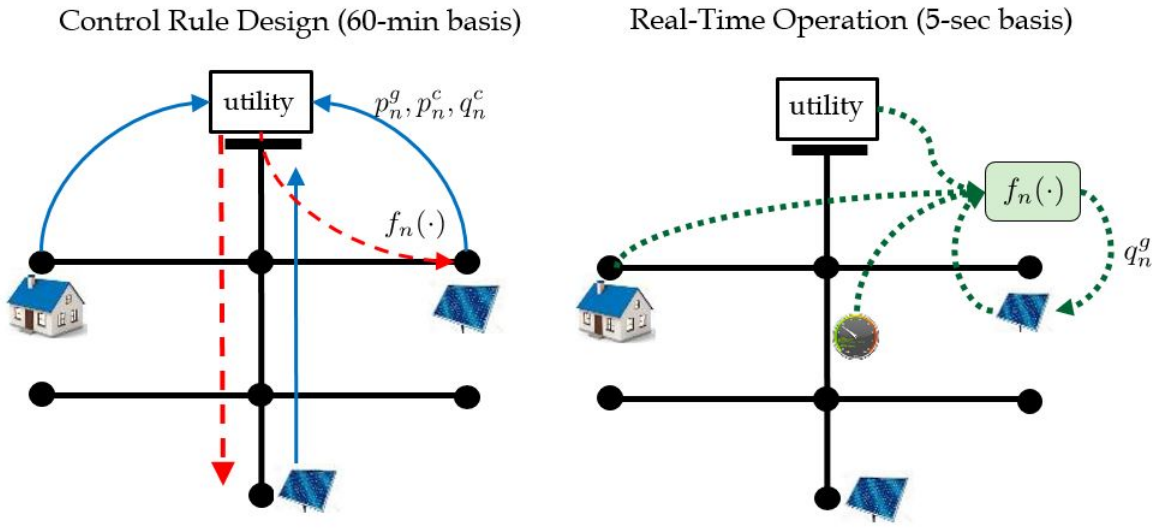


Figure 5.1: Control rules design and real time operation.

Consider now the first summand in (5.11a). Based on the equality in (5.11c), the t-th column of $\mathbf{Q}$ denoted by $\mathbf{q}_t^g$ contains the reactive injections from all inverters for scenario t. Because the squared Frobenius norm of a matrix equals the sum of the squared ℓ_2-norms of its columns, it follows that $\sum_{t=1}^{T} \|\mathbf{R}^{1/2}(\mathbf{q}_t^g - \mathbf{q}_t^c)\|_2^2 = \|\mathbf{R}^{1/2}(\mathbf{Q} - \mathbf{Q}^c)\|_F^2$. The second summand in (5.11a) follows directly from (5.6). The total cost is normalized by T.

5.4 Implementing Reactive Control Policies

After setting up the type of policy and regularizer for the problem the reactive power control policies follows the following steps:

1. The scenario data $\{\boldsymbol{\chi}_t\}_{t=1}^T$ is created by the operator for all the scenarios from $t = 1, \ldots, T$.

2. The operator solves (5.10).

3. Each inverter n receives the optimal policy coefficients $(\hat{\mathbf{a}}_n, \hat{b}_n)$ and training data $\{\mathbf{z}_{n,t}\}_{t=1}^T$ from the operator.

4. Over the next τ min, each inverter n collects the new $\mathbf{z}_{n,t'}$ and applies its projected control policy

$$\left[q_n^g(\mathbf{z}_{n,t'})\right]_{\bar{q}_{n,t'}^g} = \left[\sum_{t=1}^{T} K_n(\mathbf{z}_{n,t'}, \mathbf{z}_{n,t})\hat{a}_{n,t} + \hat{b}_n\right]_{\bar{q}_{n,t'}^g}.$$

Figure. 5.1 shows the proposed methodology divided into a control rule design step and real time operation step. The steps are explained in detail below. Right before step *1)* the

operator collects the input data $\{p_n^c, p_n^c, q_n^c\}_{n=1}^N$ from forecasts, perturbation of feeder data, historical data. This data is collected and arranged in $\{\boldsymbol{\chi}_t\}_{t=1}^T$.

Next, the operator solves (5.10) linearly constrained quadratic program minimization problem to learn the policies. In step *3)* the operater sends each inverter n the learned parameters $(\hat{\mathbf{a}}_n, \hat{b}_n)$ and its training data $\{\mathbf{z}_{n,t}\}_{t=1}^T$. This step is repeated every τ minutes according to the operater's need and available communication bandwidth. It is worth mentioning if $\mathbf{z}_{n,t} \in \mathbb{R}^{M_n}$, the operator needs to send $(M_n + 1)T + 1$ data to inverter n. Also, the number of scenarios T affects the amount of data being communicated to each inverter. If T is increased the bandwidth needs to be large to send the data quickly.

Following this, in step *4)* the inverter applies the control policies as learned to the parameters in step *2*. Recall, as all the learned parameters $(\hat{\mathbf{a}}_n, \hat{b}_n)$ and $\{\mathbf{z}_{n,t}\}_{t=1}^T$ are already readily available to each inverter n, step *4* may run more frequently say around every 30 second as compared to τ minutes of collecting data.

If the control input $\mathbf{z}_n$ is purely local the control policy can be applied without needing any more additional information. In contrast to this, if the control input is shared or non-local, entries of real time input $\mathbf{z}_{n,t}$ needs to be sent by the operator or the origin to each inverter n. Also, if the input data is shared among inverters, bandwidth-efficient broadcasting protocols can reduce the communication overhead. For example, each $\mathbf{z}_n$ could have a common power flow input from major lines appended in the end of input data for each inverter n. Both scenarios have been implemented for evaluation and a comparison is carried out to see how the model functions for added global inputs as shown in the next chapter

Chapter 6

Numerical Tests

The developed inverter reactive control rules were tested using the recommended IEEE 123-bus benchmark feeder [2], which was converted to a single-phase grid using the procedure described in [26]. One line diagram of the 123-bus feeder system is shown in Fig. 6.1. A voltage base of 12.35 kV and a power base of 100 kVA were used. Residential load (Real power consumption) and solar generation data were generated from a Gaussian mixture model for a given mean and variance. The mean value of real power generation and real power consumption were taken as $\mathbf{p}^g_{mean} = 2.5kW$, $\mathbf{p}^c_{mean} = 10.25kW$ respectively. The variance σ was varied from $(0 - 20)\%$ in steps of 10. Reactive loads (Reactive power loads) were taken at a constant lagging power factor of 0.97. The analysis was carried out for $20\%, 50\%, 100\%$ penetration of solar power generation. The percentage of penetration represents the percentage of buses having solar power with respect to the total number of consumption buses. To allow for reactive power compensation even at peak solar irradiance, the inverters were assumed

to be oversized by 10%, yielding an apparent power capacity of $\bar{s}_n^g = 1.1\bar{p}_n^g$ for all n.

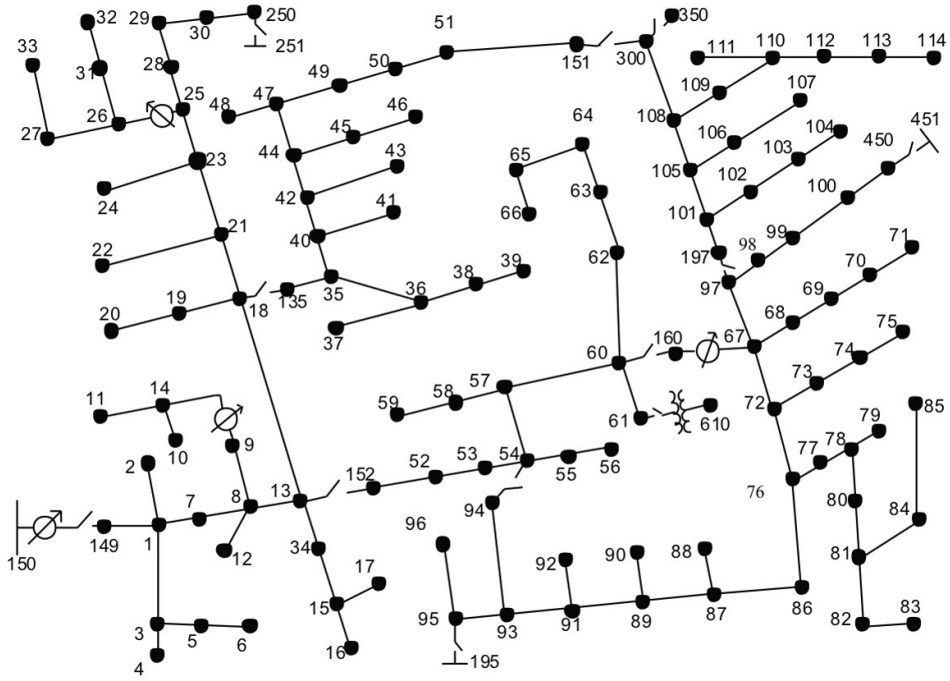


Figure 6.1: IEEE 123-bus feeder [2].

The numerical tests included five schemes: i) the unit-power factor option where inverters provide no reactive power support; ii) the fixed Watt-VAR control rules detailed in [6, Eq. (12)-(14)]; iii) the optimal reactive power setpoints obtained by solving (4.9) at each time instant; iv) the kernel-based approach of (5.10) and (5.11) for the linear; and v) the Gaussian kernel based approach for (5.10) and (5.11). The kernel-based rules were trained using the load and generation data observed during the most recent $T = 10$ scenarios, while the parameters μ and γ were decided through 5-fold cross-validation [21]. Controller n was tested for $T' = 20$ different sceanrios with only the local inputs (LI) $\mathbf{z}_n = [\bar{q}_n^g \; p_n^c - p_n^g \; q_n^c]^\top$ for each n, with global inputs (GI) of power flows $\mathbf{z}_n = [\bar{q}_n^g \; p_n^c - p_n^g \; q_n^c \; P_{15} \; P_{16} \; P_{17}]^\top$,

where line $15, 16, 17$ were chosen as the important lines which bifurcate the grid at the first level into three separate branches. The schemes iii)–v) were solved using the OSQP solver [37]. Initially, to minimize voltage deviations in the problem formulation in (5.10) the cost $\lambda\Delta_s(\mathbf{q}^g)+(1-\lambda)L(\mathbf{q}^g)$ was evaluated for $\lambda = 1$. Later results are evaluated to minimize power losses and regulate the voltage at bus for the cost $L(\mathbf{q}^g)$ was evaluated according to problem formulation (5.11). The results are shown as below.

6.1 Tests for Voltage Drop Minimization

For the problem formulation as stated in section 5.3.1 the voltage drop was minimized with respect to the apparent power constraints as shown in (5.10). Each of the test scenarios $T = 10$ was optimized by evaluating the optimal policies and tested over $T = 20$ scenarios. The evaluated control policies were compared against the five schemes. The comparison against the schemes and no reactive control is carried out for a Monte-Carlo simulation of 1000 runs. In this the input vector $\mathbf{z}_n$ is drawn from the Gaussian distribution set as explained previously, then the reactive power injection is evaluated for each of the scheme. For each run the percentage improvement is evaluated with respect to no reactive power support ($\mathbf{q}^g = \mathbf{0}$).

$$\text{Improvement in Cost(i)} = \frac{1}{T}\sum_{t=1}^{T}\frac{C|_{\mathbf{q}^g=\mathbf{0}} - C|_{\text{Scheme}(i)}}{C|_{\mathbf{q}^g=\mathbf{0}}} \times 100\% \tag{6.1}$$

where $C|_{\text{Scheme}(i)} = (\|\mathbf{Rp}+\mathbf{Xq}\|_2)|_{\text{Scheme}(i)}$ is the cost of the optimization, $i \in \{2, 3, 4, 5\}$ for each scheme as listed above. It must be noted, the local control can perform even worse as

compared to the no reactive support scheme as local scheme [6] only evaluates the reactive power control with respect to the local solar generation and load only. Thus, in few cases the percentage improvement can be seen as negative values too. These improvement percentages are enumerated in Table. 6.1.

Table 6.1: Reactive power control for voltage drop minmimization $\Delta_s(\mathbf{q}^g)$ $\lambda = 1$
Percentage improvement (%) relative to no reactive power support $\mathbf{q}^g = 0$
(LI)= With local inputs only, (GI)= With local and global inputs.

Network		Local [6]	Optimal (4.9)	Linear (5.10)		Gaussian (5.10)	
Penetration	Variance (σ)			LI	GI	LI	GI
20	0.1	4.47	99.51	95.93	98.8	95.99	99.29
20	0.2	74.25	99.98	96.7	99.31	97.77	99.42
50	0.1	88.46	99.99	97.74	99.66	97.95	99.62
50	0.2	90.92	99.99	92.12	99.5	93.1	99.56
100	0.1	69.65	99.99	94.67	99.16	94.86	99.25
100	0.2	96.35	99.99	96.12	99.1	97.04	99.21

From Table.6.1 it is seen that the optimal control techniques performes the best in comparison to the other four techniques. It evaluates the reactive power control to an optimal value with all inputs from all the buses. This is true, as the optimal control solves the problem (4.7) globally and hence, performs well. The drawback in this method is the

cyber overhead which needs a high data bandwidth to transfer data from each node to the substation at every instant making the policy slow and insecure. On the other hand, Since local rules evaluate the reactive power injection only on the basis of local inputs and considers a constant $\mathbf{x}/\mathbf{r}$ ratio for all the lines, local rules are not able to dispatch the reactive control very accurately thus giving very low percentages of improvements. Whereas the proposed kernel policy provides high improvement percentages for both linear and non-linear policies. Moreover, the performance is seen to be very close to the optimal rule as well. As seen in the table the addition of global inputs (real power flow from line $15, 16, 17$) to policies estimates a better reactive power injection and giving higher percentages of improvement as compared to only local inputs.

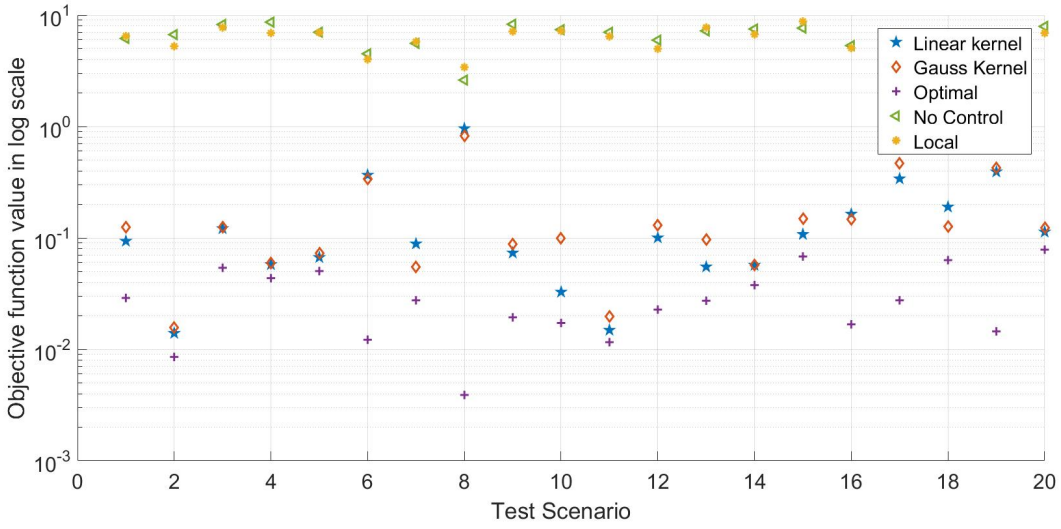


Figure 6.2: Reactive power control for $\lambda = 1$, penetration=20%, $\sigma = 10\%$ for each scheme (i-v) without global inputs.

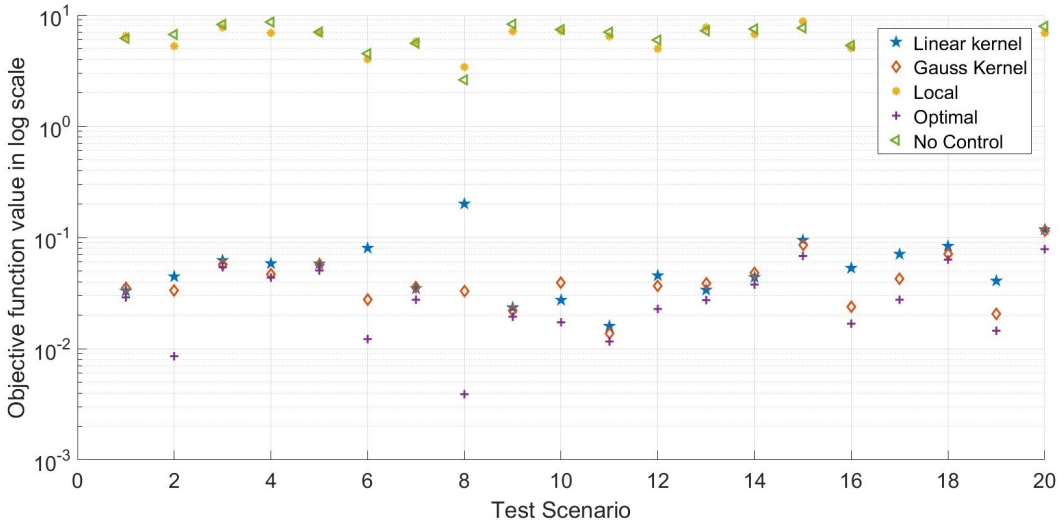


Figure 6.3: Reactive power control for $\lambda = 1$, penetration=20%, $\sigma = 10\%$ for each scheme (i-v) with global inputs (real power flows of line 15, 16, 17 as global inputs).

Validation results are shown in Fig. 6.2, 6.3. Fig. 6.2 shows the log scale value of ($\|\mathbf{Rp} + \mathbf{Xq}\|_2$) versus each scenario. In this figure it is seen that the local control behaves similar or worse to no reactive control scheme whereas the linear and non linear policies proposed in the research behaves very close to the optimal control scheme. On comparing the results Fig. 6.2, Fig. 6.3 it is seen with the addition of global inputs to the policies, the proposed method behaves even better and comes closer to the optimal control. Although adding these global inputs might lead to an increase in data communication and cyber overhead but this increase will be small if the number of added input features is kept low. In this analysis only 3 features were added and it was seen that the policies efficiency improved tremendously. Thus addition of a few global inputs can improve the kernel policies efficiency a lot. It is worth mentioning that the linear and non linear policies were seen to behave very

similarly. These results inspired us to evaluate the power loss minimization with voltage regulation constraints in which the grid is expected behave differently for linear and non-linear policies.

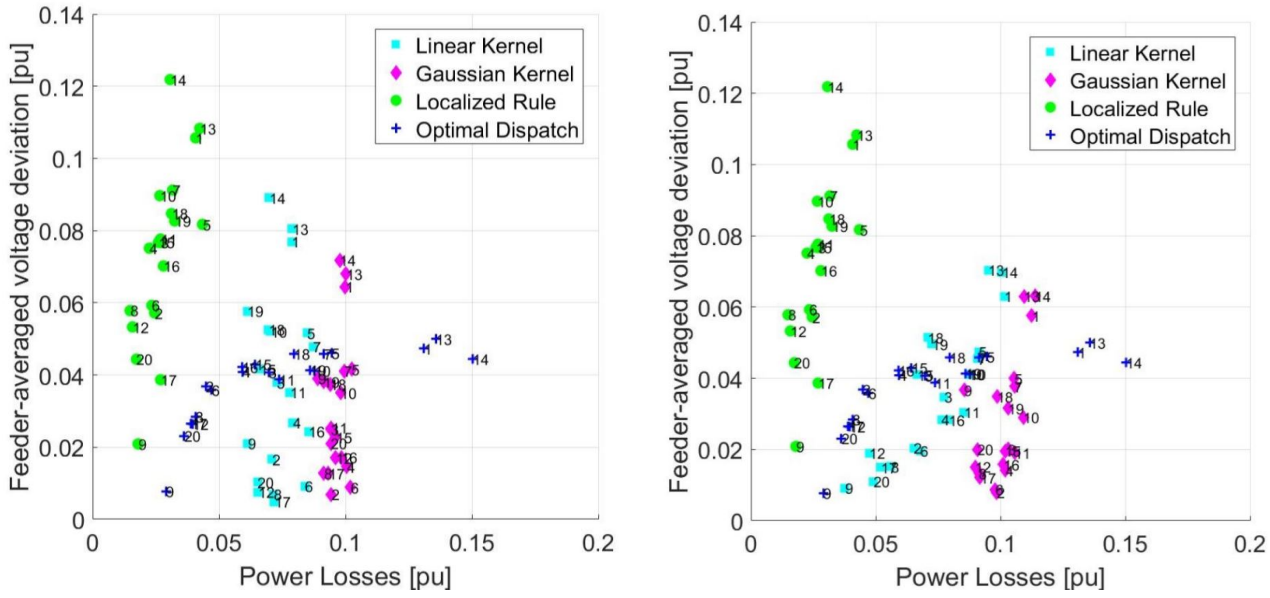
6.2 Tests for Power Loss Minimization under Voltage Constraints

The results for power loss minimization with respect to voltage and apparent power constraints (5.11) are discussed below. In this implementation, a voltage violation of 3% of base voltage value is permitted, that is $\underline{\mathbf{v}} = 0.97$ p.u., $\bar{\mathbf{v}} = 1.03$ p.u.. Each of the test scenarios $T = 10$ was optimized by evaluating the optimal policies and tested over $T = 20$ scenarios. The evaluated control policies were compared against the five schemes as described earlier. For each of the scenario, the average feeder voltage drop was evaluated as

$$\Delta\bar{v} = \frac{\|\mathbf{Rp} + \mathbf{Xq}\|_1}{N}. \tag{6.2}$$

The average voltage drop was compared against the reactive power losses. Fig. 6.4 shows the average voltage drop for each scheme with respect to the power losses for a solar penetration of 100%, solar generation and real power consumption variance of 10%. Since, for local rules $\lambda = 0$, the rule only minimizes power losses irrespective of the voltage drop as seen in (4.4), the power losses are seen to be minimal but the voltage drop is very high. Thus local rules are unable to keep the voltage within the regulation limits. Whereas, the linear and non linear

policies are seen to closely follow the optimal control scheme. Since the voltage constraint is only active during the training step, the model can violate the voltage constraints while testing. As seen in the figure these violations are not very high on an average. Moreover, comparing Fig. 6.4 (a) to Fig. 6.4 (b), it can be seen that with addition of the global inputs to the model the voltage drop reduces on an average and the model behaves more close the optimal. It is worth mentioning that in this case non-linear policies give a better regulated voltage and power losses as compared to the linear policies. In each case Fig. 6.4 (a),(b) it is seen that the non linear policies might have a little more power loss but the voltage regulation is better as compared to the linear policies. Similar trend is seen in Fig. 6.5 which shows the average voltage drop for each scheme with respect to the power losses for a solar penetration of 50% and 20%, solar generation and real power consumption variance of 10%. The figures show that the non-linear policies are perform better as compared to the linear policies. Moreover, on an average the voltages at each bus have lower voltage deviations with respect to the nominal voltages for non-linear policies as compared to linear policies.



(a) (b)

Figure 6.4: Average voltage drop vs power loss for power loss minimization under voltage constraints for 100% penetration and 0.1 variance (a) With local inputs (LI) only,(b) With local inputs and global inputs (GI).

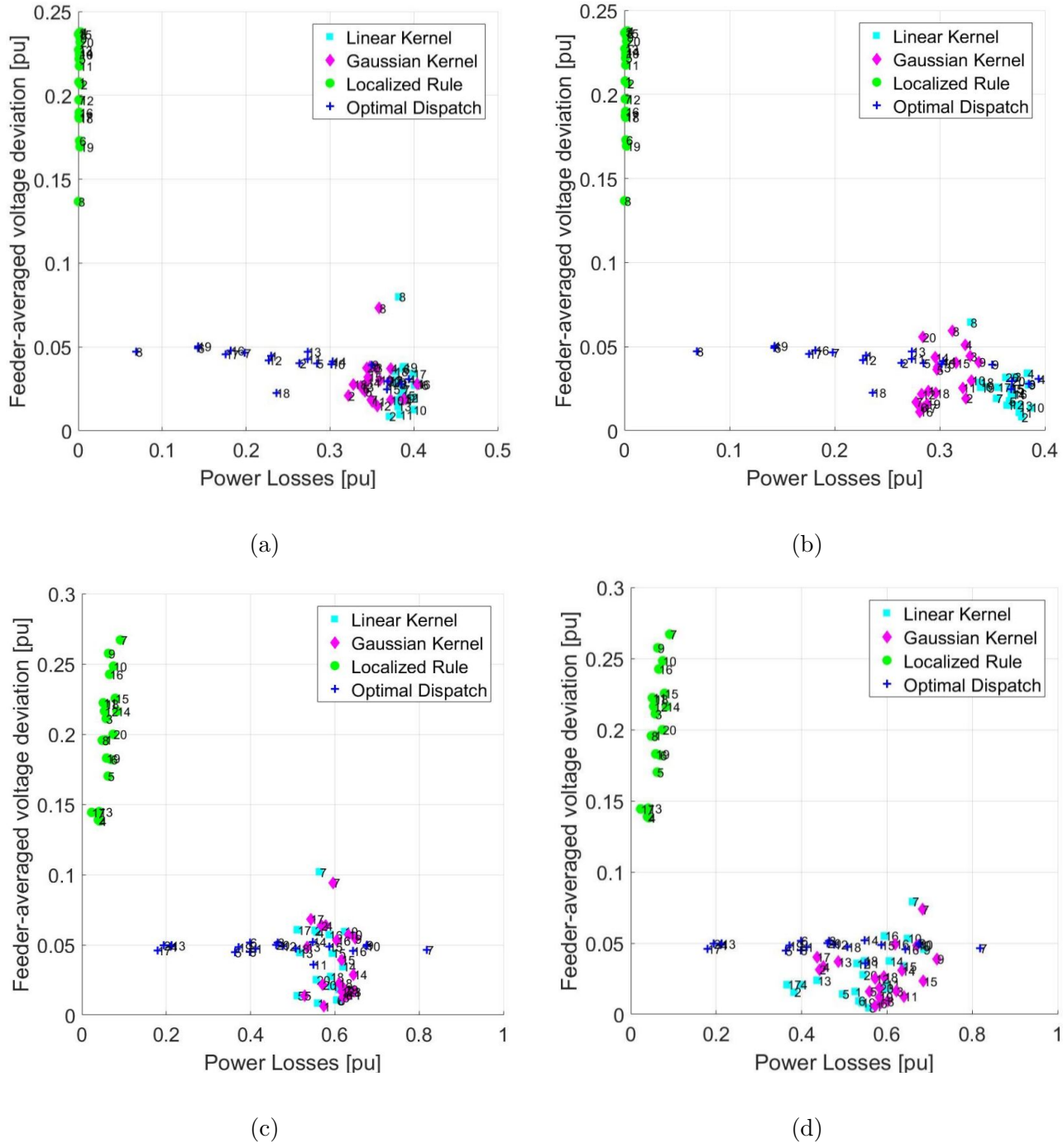


Figure 6.5: Average voltage drop vs power loss for power loss minimization under voltage constraints for scenarios, (a) 20% penetration and 0.1 variance with only Local inputs (LI),(b) 20% penetration and 0.1 variance with Local inputs and global inputs (GI), (c) 50% penetration and 0.1 variance with only local inputs (LI),(d) 50% penetration and 0.1 variance with local inputs and global inputs (GI).

Chapter 7

Conclusions

In this research we have developed a framework which helps to mitigate impacts of varying loads and high PV penetrations in the distribution grid. Since, commonly used voltage regulation devices like voltage regulators, switched capacitor banks have high response time to react to the changes in PV power generation, smart inverters have been used for reactive power control in the grid. Smart inverters can participate in feeder voltage regulation in a faster timescale and are capable to operate for various techniques and methodologies as programmed in its chipset.

Kernel-based reactive power control algorithm has been proposed in this research. The model has been evaluated for linear and non linear control policies. The policies have been designed to evaluate the control set-points for every half an hour and estimation has been done for the reactive power control in real-time using local and/or global grid data. The

model takes into consideration the cyber overhead and data communication requirements for fast and easy computation. The suggested designs are extremely flexible and adjustable to the available cyber resources. Tests have been carried out for minimization of power losses and voltage regulation. The tests corroborate that this innovative framework is able to provide reactive power support to the grid according to the objective of the framework. The model is able to achieve the desirable trade-off between reactive control performance and computational/communication requirements.

In future, this research can be extended for varying the controller inputs by adding voltages as control inputs which closes the loop of input data. Another exciting area to look into shall be designing of the kernels wherein, the kernel can be evaluated as a combination of kernels and an optimal set can be used for evaluation of the problem. This thesis is just an initialization of reactive power control problem and there are wide range of works that can be achieved using this technique as a base, like having multi-stage formulations by incorporating conventional voltage regulation equipment such as step-voltage regulators, capacitor banks or energy storage devices into the optimization problem; evaluating meshed network distribution grids for reactive power control etc.

Bibliography

[1] "Pecan street inc. dataport," 2013. [Online]. Available: https://dataport.pecanstreet.org/.

[2] W. H. Kersting, "Radial distribution test feeders," in *Proc. Power Engineering Society Winter Meeting*, vol. 2, 2001, pp. 908–912.

[3] IEA, "Worldwide trends in energy use and efficiency," 2008. [Online]. Available: https://www.iea.org/publications/freepublications/publication/Indicators_2008.pdf

[4] "Electric power consumption (kwh per capita)," [Online]. Available https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC.

[5] R. J. Bravo, "Dynamic performance of residential electronic loads during voltage deviations," *IEEE Transmission and Distribution Conference and Exposition*, 2016.

[6] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proc. IEEE*, vol. 99, no. 6, pp. 1063–1073, Jun. 2011.

[7] W. H. Kersting, *Distribution System Modeling and Analysis*. New York, NY: CRC Press, 2001.

[8] C. Schauder, "Advanced inverter technology for high penetration levels of PV generation in distribution systems," National Renewable Energy Laboratory, Tech. Rep., Mar. 2009. [Online]. Available: http://www.nrel.gov/docs/fy14osti/60737.pdf

[9] *IEEE 1547 Standard for Interconnecting Distributed Resources with Electric Power Systems*, IEEE Std., 2018. [Online]. Available: http://grouper.ieee.org/groups/scc21/1547/1547_index.html

[10] J. Bank and J. Hambrick, "Development of a high resolution, real-time, distribution-level metering system and associated visualization modeling, and data analysis functions," National Renewable Energy Laboratory, Tech. Rep. TP-5500-56610, May 2013.

[11] R. Jabr, "Radial distribution load flow using conic programming," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1458–1459, Aug. 2006.

[12] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *Proc. IEEE Power & Energy Society General Meeting*, San Diego, CA, Jul. 2012.

[13] B. Robbins, C. Hadjicostis, and A. Dominguez-Garcia, "A two-stage distributed architecture for voltage control in power distribution systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1470–1482, May 2013.

[14] K. Rogers, R. Klump, H. Khurana, A. Aquino-Lugo, and T. Overbye, "An authenticated control framework for distributed voltage support on the smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 40–47, Jun. 2010.

[15] F. Olivier, P. Aristidou, D. Ernst, and T. Van, "Cutsem active management of low-voltage networks for mitigating overvoltages due to photovoltaic units," *IEEE Trans. Smart Grid*, vol. 7, pp. 926–936, 2016.

[16] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

[17] Q. Peng and S. Low, "Distributed algorithm for optimal power flow on a radial network," in *Proc. IEEE Conf. on Decision and Control*, Venice, Italy, Dec. 2014, pp. 167–172.

[18] C. M. Bishop, *Pattern Recognition and Machine Learning.* New York, NY: Springer, 2006.

[19] S.-T. J. and N. Cristianini, *Kernel Methods for Pattern Analysis.* Cambridge University Press, 2004.

[20] J. L. Rojo-lvarez, M. Martnez-Ramn, J. Muntilde, oz Mariacute, and G. Camps-Valls, *Kernel Functions and Reproducing Kernel Hilbert Spaces*, 2018.

[21] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics, 2009.

[22] G. Wahba., "Spline models for observational data," *CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM*, vol. 59, 1990.

[23] M. Baran and F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Syst.*, vol. 4, no. 1, pp. 735–743, Jan. 1989.

[24] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Voltage regulation algorithms for multiphase power distribution grids," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3913–3923, Sep. 2016.

[25] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *Proc. IEEE Conf. on Decision and Control*, Florence, Italy, Dec. 2013, pp. 4329–4334.

[26] L. Gan, N. Li, U. Topcu, and S. Low, "On the exactness of convex relaxation for optimal power flow in tree networks," in *Proc. IEEE Conf. on Decision and Control*, Maui, HI, Dec. 2012, pp. 465–471.

[27] G. Wang, V. Kekatos, A.-J. Conejo, and G. B. Giannakis, "Ergodic energy management leveraging resource variability in distribution grids," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4765–4775, Nov. 2016.

[28] V. Kekatos, G. Wang, A. J. Conejo, and G. B. Giannakis, "Stochastic reactive power management in microgrids with renewables," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3386–3395, Nov. 2015.

[29] R. A. Jabr, "Linear decision rules for control of reactive power by distributed photovoltaic generators," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–1, 2018.

[30] A. Keane, L. F. Ochoa, E. Vittal, C. J. Dent, and G. P. Harrison, "Enhanced utilization of voltage control resources with distributed generation," *IEEE Trans. Power Syst.*, vol. 26, p. 252260, 2011.

[31] Z. Wang, J. Wang, B. Chen, M. M. Begovic, and Y. He, "MPC-based voltage/var optimization for distribution circuits with distributed generators and exponential load models," *IEEE Trans. Smart Grid*, vol. 5, p. 24122420, 2014.

[32] Z. Ziadi, M. Oshiro, T. Senjyu, A. Yona, N. Urasaki, T. Funabashi, and C. H. Kim, "Optimal voltage control using inverters interfaced with PV systems considering forecast error in a distribution system," *IEEE Trans. Sust. Energy*, vol. 5, p. 682690, 2014.

[33] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," Allerton, IL, Oct. 2014, pp. 582–588.

[34] K. S. Ayyagari, N. Gatsis, and A. F. Taha, "Chance-constrained optimization of distributed energy resources via affine policies," Montreal, Quebec, Canada, Nov. 2017.

[35] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge University Press, 2004.

[36] J. Abernethy, F. Bach, T. Evgeniou, and J.-P. Vert, "A new approach to collaborative filtering: Operator estimation with spectral regularization," *J. Machine Learning Res.*, vol. 10, pp. 803–826, 2009.

[37] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: An operator splitting solver for quadratic programs," *ArXiv e-prints*, Nov. 2017.