Fall 2015

Risk-averse design and operation of renewable energy power grids

Bo Sun
University of Iowa

RISK-AVERSE DESIGN AND OPERATION OF RENEWABLE ENERGY POWER GRIDS

by

Bo Sun

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Industrial Engineering
in the Graduate College of
The University of Iowa

December 2015

Thesis Supervisors: Associate Professor Pavlo Krokhmal
Associate Professor Yong Chen

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Bo Sun

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Industrial Engineering at the December 2015 graduation.

Thesis Committee: _____
Pavlo Krokhmal, Thesis Supervisor

_____
Yong Chen, Thesis Supervisor

_____
Andrew Kusiak

_____
Amaury Lendasse

_____
Raghuraman Mudumbai

To my parents

ACKNOWLEDGEMENTS

I would like to express my special appreciation and thanks to my advisor, Professor Pavlo Krokhmal, for his guidance and support during my Ph.D. studies. He patiently provided excellent instructions, insightful discussions, and scientific suggestions which were not only necessary for me to complete the Ph.D. program, but more importantly beneficial for my future life. It is so grateful to have worked under his supervision during the past five years. I would also like to thank my co-advisor Professor Yong Chen, who gave me invaluable advice and consistent support. Sincere thanks to my dissertation committee members, Professors Andrew Kusiak, Amaury Lendasse and Raghuraman Mudumbai, for their feedback and comments.

I also thank my lab members, Maciej Rysz, Alexander Vinel, Yana Morenko, Dmitri Chernikov, Nathaniel Richmond, Huan Yu, Yuxing Hou, Baosheng He, and Guanglin Xu from Department of Management Sciences, for their valuable discussions and kind help in my Ph.D. life.

Many thanks to the department administrative staff, Andrea Flaherty, Karla Stout, Tara Hoadley, and Jennifer Rumping, who were always there to help. Also, I thank Deborah Hampton, who provided enormous assistance when I worked at CCAD.

This work was supported by NSF EPSCoR project, which is highly appreciated.

I would also like to express my gratitude to all my friends for their encourage-

ment and support. Special thanks to my friend, Dr. Paul Weller, who has made a profound influence on my career. It was my great fortune to have met him. He not only helped broaden the scope of my thoughts in quantitative finance, but also gave me a tremendous lift and worthwhile advice, no matter when I was looking for an internship or a full-time job.

Last but not least, I especially thank my parents and Hongye Sun, who have provided endless love, care and support that I needed to build a dream to pursue. No matter how frustrated I felt, their consolation and encouragement would always magically wipe out my frowns. Without them, I would have never made it this far.

ABSTRACT

The need for effective energy harvesting from renewable resources becomes increasingly important, especially in the light of the inevitable depletion of the fossil fuel energy sources. Among renewable energy sources, wind energy represents one of the most attractive alternatives. In this thesis, we construct several stochastic optimization models, including the traditional risk-neutral expectation based model, and risk-averse models based on linear and nonlinear coherent measures of risk, to study the strategic planning and operation of futuristic power grids where the loads are served from renewable energy sources (wind farms) through High Voltage Direct Current lines. Exact solutions algorithms that employ Benders decomposition and polyhedral approximations of nonlinear constraints have been proposed for the formulated linear and nonlinear mixed-integer optimization problems. The conducted numerical experiments illustrate the efficiency of the developed algorithms, as well as effectiveness of risk-averse models in reducing the power grid's exposure to power shortage risks when the energy is produced from renewable sources. We further extend the risk-averse models to demonstrate how energy storage devices may impact the risk profile of power shortages in the renewable energy power grid. Additionally, we consider convex relaxations of optimal power flow problem over radial networks, that allow for solving mixed-integer optimization problems in traditional alternating current distribution networks. Exactness of a specific second-order cone programming relaxation has been discussed. We finally propose an "extended" optimal power

flow problem and prove its second-order cone programming relaxation to be exact theoretically and empirically.

PUBLIC ABSTRACT

As renewable energy resources increasingly contribute to the global energy consumption and gradually replace traditional fossil fuels, how to harvest and utilize renewables in an effective and efficient way becomes crucial. The advantages of renewable energy as a clean, plentiful and commonly available source of energy usually come at a cost of uncertainty or intermittency, which would impact on the stability of power grids and even cause mismatch between the power supply and load (considered as power shortage risk) in some scenarios when using or integrating renewable energy into transmission and distribution networks. Therefore, how to mitigate the negative impact from the intermittency of renewable energy and hedge the power shortage risk, is the main purpose of this study.

In this thesis, we focus on wind energy and consider the wind farm location problem. Several stochastic optimization models are proposed in terms of various criteria including risk-neutral expectation, linear and nonlinear coherent measures of risk, which aim to demonstrate the importance of strategic planning of wind farms and in which way power shortage risk can be reduced. Furthermore, we also investigate the impact of energy storage devices on the risk profile of power shortage by extending the risk-averse models. Lastly, the convexification of optimal power flow problem over radial networks is studied, which is conducive to solving mixed-integer optimization problems in traditional alternating current distribution networks.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ALGORITHMS

Algorithm

CHAPTER 1
INTRODUCTION

1.1 Wind Energy

1.1.1 Global Overview

Wind is a free and clean source of renewable energy. As an alternative to traditional fossil fuels, wind energy not only significantly mitigates negative impact on the environment (such as acid rain and greenhouse effect), but also is plentiful and widely distributed, contributing to both social and economic development. According to [8], the potential of global wind power as a source of electricity was assessed to be 72 terawatts (TW) based on an analysis of data for the year 2000, which has been found to account for about 40 times the global electricity demand of the year 2001. In addition, the price of wind energy has decreased over 80% since 1980 and is competitive with or even cheaper than other possible resources [4], and more importantly this trend is expected to be kept.

Due to various advantages of wind energy, the wind energy industry has experienced substantial growth in the last few years globally. In 2014, the global wind power installed capacity has reached an estimated 336,327 megawatts (MW), which can satisfy around 4% of the global electricity demand [131]. According to [52] and [131], the wind power installed capacity of some countries is listed in Table 1.1. From the table, we find that the annualized growth rate of wind power installed capacity at global level is over 26% during the last fifteen years, which results in the cumulative

installed capacity by the end of 2013 is more than 32 times that by the end of 1998. This exponential growth trend is further described in Figure 1.1 based on the data from [48]. Almost 90% of the current global installed capacity is added during the last 10 years (2004-2013).

Country	End 2013 (MW)	End 1998 (MW)	Annualized Growth Rate
Canada	7,698	83	35.25 %
China	91,413	224	49.30 %
Denmark	4,772	1,450	8.27 %
India	20,150	968	22.43 %
Ireland	2,037	63	26.08 %
Italy	8,551	180	29.35 %
Germany	34,660	2,874	18.06 %
Netherlands	2,693	363	14.29 %
Portugal	4,724	60	33.79 %
Spain	22,959	834	24.73 %
Sweden	4,470	150	25.40 %
U.K.	10,531	334	25.87 %
U.S.	61,108	1,952	25.81 %
Other	42,746	304	39.06 %
Total	318,530	9,839	26.09 %

Table 1.1: Wind power installed capacity comparison

The Global Wind Energy Council (GWEC) has envisioned that by the year 2030, the wind power production could reach 2,000 gigawatts (GW) which roughly amounts to 17-19% of the world's electricity demand. This number will increase up to 20-30% by the year 2050.

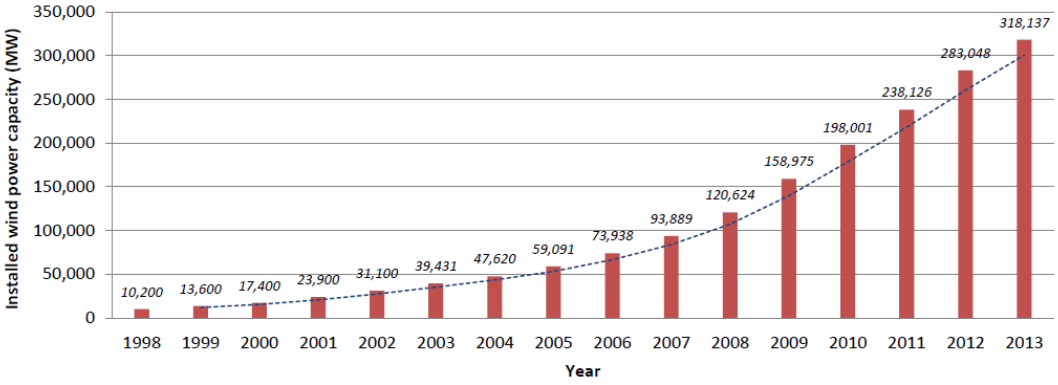

Figure 1.1: Global cumulative wind power installed capacity (MW)

1.1.2 The United States Overview

In the United States, the wind energy industry has been one of the fastest growing sectors of economy in the last several years. For the year 2013, the electricity produced from wind power in the U.S. accounted for 4.13% of all generated electrical energy, meaning that it became the fifth largest electricity source according to the data from the Department of Energy's Energy Information Administration (EIA). As of the end of the year 2013, the cumulative installed wind power capacity in the U.S. has reached 61,108 MW as the third largest wind power producer (almost 20% global share). A detailed description with regard to states' installed wind power capacity is given in Figure 1.2.

As of the end of the year 2013, Texas has the highest installed capacity of 12,355 MW, followed by California (5,830 MW), Iowa (5,178 MW), Illinois (3,568 MW), and Oregon (3,153 MW). Based on the percentage of energy generated by wind in 2013, the top five states are Iowa (27.4%), South Dakota (26.0%), Kansas (19.4%),

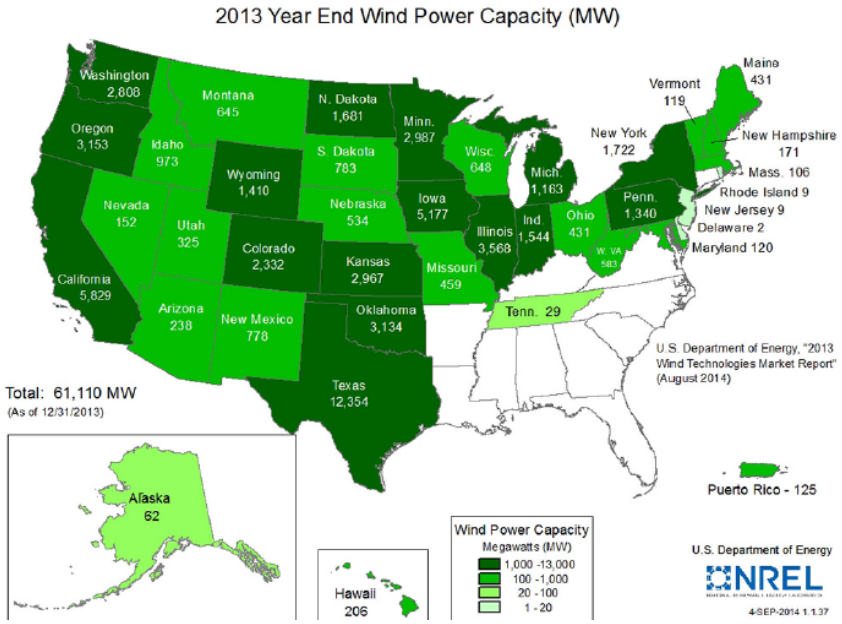

Figure 1.2: Current wind power installed capacity in the U.S. as of 12/31/2013 (MW)

Idaho(16.2%), and Minnesota (15.7%). Annual installed wind power capacity in five main wind power states of the U.S. is shown in Figure 1.3. The wind power in Texas has developed fastest, with almost 35% annualized growth rate, especially during the period (2004-2009). During the same period, the growth rate for Iowa was only lower than that for Texas, and, as a result, the cumulative installed wind power capacity exceeded the corresponding metrics for California in 2008 and kept ranking second in cumulative installed capacity until 2012. Moreover, we may find that in some specific years, the growth rates were close to zero, which corresponded to flat segments in Figure 1.3, such as the period (2012-2013). One reasonable explanation for this may be due to the expiry of the Federal Production Tax Credit (PTC), which has been playing a crucial role in wind power development since being introduced in 1992. For

instance, sharp drops of 77% and 91% in annual installations were observed in 2003 and 2012 following the expiry of PTC in 2004 and 2013.

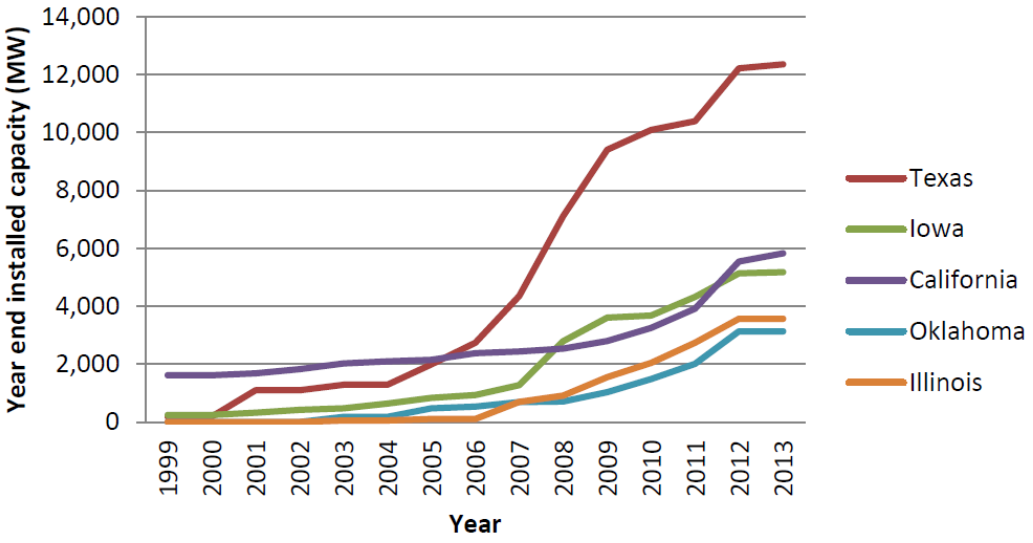

Figure 1.3: Installed capacity growth of five main wind power states in the U.S.

According to a technical report from National Renewable Energy Laboratory (NREL) [73], the United States has the total estimated onshore wind energy potential of 10,955 GW, which could produce 32,784 TWh annually amounting to almost eight times of total U.S. electricity consumption in 2011. Besides, the offshore wind energy potential is estimated to be 4,150 GW [99]. By including a 4% contribution from offshore wind power, the U.S. Department of Energy projected that by 2030 wind power could generate 20% of total electricity demand in the U.S..

1.1.3 Variability and Storage of Wind Power

1.1.3.1 Variability in Wind Speed

Although we have seen the tremendous growth in wind energy industry in the past few years, the intermittency and variability of wind power generation remains a major factor affecting the growth of wind energy industry and market penetration. The power generated by wind is a variable resource, and it varies based on the availability of wind at any given point in time. That means, a wind turbine could only produce when wind is available, and this makes wind power not dispatchable as a power produced by a gas turbine, which can be scheduled to turn on and off. Wind speed varies hourly, daily and seasonally. Within a day, wind speed is usually greater during the daytime and smaller at night. This variability is illustrated in Table 1.2 [121]. Also, the variation of wind speed has strong dependence on the time of year. For example, the winds are strongest in the spring and weakest in the summer for the southern Great Plains (Kansas, Oklahoma and Texas) [52]. There also exist certain long-term variations in wind speed.

Compared with the variability of power generation by an individual wind turbine, the aggregate variability of power generation by multiple turbines in a wind farm is much smaller. And the variability could be further reduced if we consider the aggregate power generation by multiple wind farms over a large geographic area. That means, the intermittency of wind power may be mitigated to some extent by cooperation and coordination among multiple wind farms.

Time	Wind Speed (mph)	Time	Wind Speed (mph)
00:00	0.36	12:00	1.23
01:00	1.66	13:00	1.35
02:00	0.95	14:00	1.13
03:00	3.76	15:00	2.62
04:00	1.46	16:00	1.95
05:00	0.96	17:00	1.46
06:00	2.02	18:00	1.09
07:00	0.75	19:00	0.30
08:00	0.64	20:00	2.22
09:00	2.00	21:00	0.93
10:00	1.79	22:00	1.14
11:00	3.17	23:00	0.12

Table 1.2: Hourly average wind speed: Bandon, Oregon (01/01/2013)

1.1.3.2 Wind Energy Storage

Since the instantaneous energy generation and consumption in the power grid should remain in balance, the penetration of variable wind energy source into grid system may pose substantial challenges to the grid stability. Correspondingly, the cost for regulation and operating reserve would definitely increase due to the intermittency and non-dispatchable of wind energy generation. An efficient way to compensate this variability is to utilize some sort of energy storage. This means that when wind speed is high, the extra produced energy can be stored in some form, and this stored energy then can be used to meet the demands of consumers during low wind periods, which would reduce the effect of wind speed variability in wind energy production. For example, a traditional power plant is able to increase the fuel supply when the

penetration of wind farms is at a low level, and decrease it to bring energy generation into equilibrium with load when the wind speed is high. In this case, one may consider the traditional power plant as a storage device where energy storage is in the form of fossil fuels.

As the most widely used form of bulk-energy storage, pumped hydroelectric storage is undoubtedly a promising technology for store intermittent wind energy. According to the Electric Power Research Institute (EPRI), the pumped hydroelectric storage has accounted for around 127 GW, more than 99% of bulk storage capacity in the world in 2012. With 22 GW of capacity installed, pumped hydroelectric is also a common energy storage method in the United States. When the wind speed is high, the hydroelectric plant will store the extra energy by pumping water to a high storage reservoir, and use the water under the condition that the wind resource is low.

Another promising form of storage for wind energy is compressed air. The compressed air energy storage system would convert wind energy into compressed air which can be used to generate electricity when there is not enough or no wind resource. However, no matter pumped hydroelectric or compressed air storage, it has specific topography and geology requirements [122].

We can also use batteries to store the extra generated electricity from wind farms, and uncharge them to generate power during low wind periods. Although battery technology has recently been considered for most utility-scale applications, it is only widely used for small-scale energy storage.

1.2 Electric Power System

An electric power system is a real-time energy delivery network, used to generate, transmit and utilize electric power, which could be divided into the following major subsystems [38]:

- Generation Subsystem
- Transmission and Subtransmission Subsystem
- Distribution Subsystem
- Utilization Subsystem

According to [18], Figure 1.4 illustrates the basic electrical components of an electric power system, although a full scale actual power network is more complex. The power system starts with generation subsystem, where electrical energy is produced in the power plant by either fossil fuels (coal, natural gas or petroleum) or renewable energy (biomass, solar or wind) and then transferred through step-up transformers to a high voltage level, which is more suitable for long distance transmission. In the transmission subsystem, the high voltage power lines would efficiently transport electrical energy from power plants to power loads over long distances. After that, the high voltage electrical energy is transformed to a lower voltage level in the substations, which could be transmitted over distribution network. The utilization subsystem would finally transform the energy in distribution power lines into different appropriate voltage level and transport to its destinations, such as residential, commercial and industrial consumption.

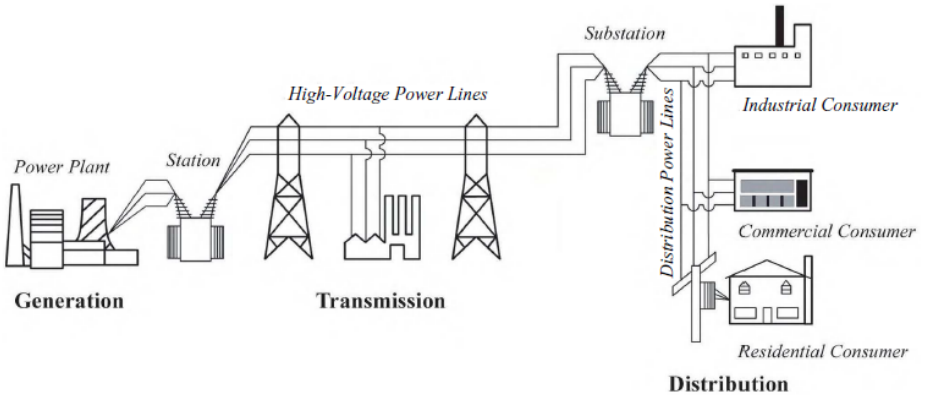

Figure 1.4: Power system structure

1.2.1 Transmission and Distribution Subsystems

The interconnected transmission and distribution network, known as "power grid" in the United States, transports and distributes the electrical energy produced in power plants to different consumers. According to [16], the generator voltage is around 11 to 30 kV, which is not appropriate for the transmission over long distances due to low energy efficiency. Therefore, step-up transformers are used to increase the voltage level. On the basis of established standard transmission voltages by American National Standards Institute (ANSI), transmission voltage lines operating at less than 230 kV are standardized at 34.5 kV, 46.0 kV, 69.0 kV, 115.0 kV, 138.0 kV, 161.0 kV, categorized into high voltage and the transmission voltages between 230.0 kV and 765.0 kV are referred to as extra-high voltage (EHV) [129]. The ultra-high voltage (UHV) represents a standardized transmission voltage of 1100.0 kV. The high voltage transmission lines are terminated in high-voltage substations, which first reduce

the voltage to some lower level. Subtransmission lines connect the high-voltage substations with the distribution substations, constituting the subtransmission network. The voltage is further reduced at the distribution substation. Note that some very large industrial consumers or railroads may be served directly from the subtransmission subsystem.

Finally, the electric energy is delivered through distribution subsystem, which carries electricity from the distribution substations to other consumers. Smaller industrial or commercial customers could be served by the primary distribution lines at the voltage levels of 4 to 34.5 kV. Residential customers normally utilize the electrical energy through the secondary distribution network which further reduces the voltage to the levels of 120/240 V.

1.2.2 High-voltage Direct Current Lines

High-voltage direct current (HVDC) lines are used to transmit bulk of electrical energy over long distances, by use of direct current (DC) in contract to the more common alternating current (AC). Since there is no need for reactive compensation, the HVDC lines would usually lose less power than equivalent AC transmission lines. This higher efficiency makes HVDC more economical than AC transmission for large amounts of power transmission over long distances, in addition to its lower transmission line costs. Moreover, HVDC transmission can improve system stability because it not only could allow the operator to quickly change the direction of power flow, but also allow the power transportation between power systems with different frequencies.

A well-known HVDC system is the Pacific HVDC Intertie, connecting southern California with Oregon. The underwater HVDC cables have been built to interconnect England and France.

Today, we are witnessing the transition from fossil fuels to renewable energy sources, and how to exploit the full potential of massive wind and solar farms is critical. The characteristic of HVDC allowing operators to control the power flow, makes HVDC more appropriate for connecting renewable sources such as wind and solar because it could mitigate the effects of intermittency and fluctuation and smooth the power outputs. Besides, the lower transmission cost by HVDC would further improve the renewable energy competence against other energy sources to reduce the demand for fossil fuels. Therefore, HVDC plays a crucial role in the development of future energy system based on renewables. In the U.S., Clean Line Energy Partners is developing four large transmission projects to transmit wind and solar energy in long distance by utilizing HVDC technology. The European Commission also advocated and supported the construction of HVDC to interconnect some countries in Europe, and to link offshore wind farms.

1.3 Review of Power Flow and Optimal Power Flow Models

1.3.1 Power Flow

Power flow, also known as load flow, is an evaluation tool to determine the steady-state complex voltages and powers at every node of the network and all the transmission lines [108]. That means, the solution to power flow problem would show

the current, voltage, real and reactive power flow of all the buses. The buses or nodes, connected to generators and loads, are interconnected by transmission lines. Due to the nonlinear relationship between voltage and current, real and reactive power at each bus, the power flow problem is a nonlinear problem. Power flow analysis is an essential part of most studies in system planning and operation.

1.3.1.1 Basic Formulation

According to Kirchhoff's Current Law (KCL), the sum of currents flowing into a node should be the same as the sum of currents flowing out of that node, hence the relationship between current and voltage at a node j can be described as followings:

$$I_j = \sum_{k=1}^{N} Y_{jk} V_k, \tag{1.1}$$

where I_j is the net injected current at node j, V_k is the voltage at node k, and Y_{jk} is an element of the admittance matrix. Since the complex power $S = VI^H$ (I^H is the Hermitian transpose of I), the power flow equations at node j are:

$$P_j + iQ_j = V_j \sum_{k=1}^{N} Y_{jk}^H V_k^H, \tag{1.2}$$

where P_j and Q_j represent the net injected active and reactive power at node j respectively, V_k^H is the Hermitian (complex conjugate) transpose of the voltage at node k, and i is the imaginary unit. If we use the polar coordinates, the complex voltage at node j can be expressed as:

$$V_j = |V_j| \angle \theta_j, \tag{1.3}$$

where $|V_j|$ and θ_j are the voltage magnitude and angle at node j respectively. Thus the mathematical formulation of power flow in polar form could be expressed as:

$$P_{G_j} - P_{L_j} = \sum_{k=1}^{N} |V_j||V_k| \left[G_{jk}\cos(\theta_j - \theta_k) + B_{jk}\sin(\theta_j - \theta_k)\right], \tag{1.4a}$$

$$Q_{G_j} - Q_{L_j} = \sum_{k=1}^{N} |V_j||V_k| \left[G_{jk}\sin(\theta_j - \theta_k) - B_{jk}\cos(\theta_j - \theta_k)\right], \tag{1.4b}$$

where P_{G_j} and Q_{G_j} denote the real and reactive power output of the generator connecting to bus j, P_{L_j} and Q_{L_j} denote the real and reactive power load at bus j, $|V_j|$, $|V_k|$, θ_j and θ_k are the voltage magnitudes and angles at bus j and k respectively, G_{jk} and B_{jk} are the conductance, susceptance between bus j and bus k. The left-hand side of equations (1.4a) and (1.4b) represent the power injections at bus j, which means $P_j = P_{G_j} - P_{L_j}$ and $Q_j = Q_{G_j} - Q_{L_j}$.

In contrast, if we express the complex voltage in rectangular form,

$$V_j = e_j + if_j, \tag{1.5}$$

where $e_j = |V_j|\cos\theta_j$ and $f_j = |V_j|\sin\theta_j$, the corresponding power flow equations in rectangular form are:

$$P_{G_j} - P_{L_j} = \sum_{k=1}^{N} \left[e_j(e_k G_{jk} - f_k B_{jk}) + f_j(f_k G_{jk} + e_k B_{jk})\right], \tag{1.6a}$$

$$Q_{G_j} - Q_{L_j} = \sum_{k=1}^{N} \left[f_j(e_k G_{jk} - f_k B_{jk}) - e_j(f_k G_{jk} + e_k B_{jk})\right]. \tag{1.6b}$$

1.3.1.2 Bus Types and Solvability

In the formulation of power flow problem, four quantities are associated with each bus: real power P, reactive power Q, voltage magnitude $|V|$ and voltage angle θ [81, 137, 51].

- Load Bus(P-Q bus): For this type of bus, there is no generator connected to it generally. The special case could be a bus connected with a generator, however the power output of which is constant. Furthermore, the real power P and reactive power Q are known from available measurements, but the voltage magnitude $|V|$ and angle θ are unknown.
- Generator Bus(P-V bus): The bus is connected with a generator. In this case, the real power P and the magnitude of voltage $|V|$ are specified, but the reactive power and the nodal voltage angle θ are computed.
- Slack Bus: This is also known as the reference bus, or the swing bus where the voltage magnitude $|V|$ and voltage angle θ are known. The real power P and reactive power Q are unspecified. Since the transmission losses are unknown in advance of the power flow calculation, this sort of bus is used to provide additional real and reactive power so as to balance the power system. Hence, it is usually a generator related bus. The voltage angle at the slack bus is generally chosen to be zero, as a reference for calculation of the other voltage angles.

The power flow problem is to determine the values of all state variables (voltage magnitude $|V|$ and voltage angle θ) by solving the above power flow equations based on specifications of other variables. Consider a system with N buses, where N^G represents the number of generator buses. Thus, we need N voltage magnitudes and N voltage angle quantities to completely describe the state of this system. However, the voltage magnitude and angle are known for the slack bus, and the voltage magnitudes of generator buses are also given. Hence the unknown state variables are $(2N - N^G -$

2). From the generator buses, we only get N^G balance equations regarding to active power injections, and there are no power flow equations for the slack bus, thus in total we also have $(2N - N^G - 2)$ equal to the number of unknowns, which serves as a necessary condition for solvability.

Figure 1.5 gives a single line diagram of the IEEE 14-bus standard system, consisting of 14 buses and 20 branches. Buses 1 and 2 have generators, and buses 3, 6 and 8 are synchronous compensators which are only used for reactive power support. Based on the above bus classification, bus 1 serves as the slack bus, buses 2, 3, 6 and 8 are P-V buses, and the rest are P-Q buses in the system. Let us take bus 12 as an example to demonstrate the power flow at that node. As a load bus, it does not have a generator, which means the values of $P_{G_{12}}$ and $Q_{G_{12}}$ should be equal to zero, and the values of $P_{L_{12}}$ and $Q_{L_{12}}$ should be prespecified. Besides, bus 12 only connects with buses 6 and 13, and therefore the power flow equations at this node are like:

$$\begin{aligned}
-P_{L_{12}} =& |V_{12}||V_6| \left[G_{12,6}\cos(\theta_{12}-\theta_6) + B_{12,6}\sin(\theta_{12}-\theta_6)\right] + \\
& |V_{12}||V_{13}| \left[G_{12,13}\cos(\theta_{12}-\theta_{13}) + B_{12,13}\sin(\theta_{12}-\theta_{13})\right], \\
-Q_{L_{12}} =& |V_{12}||V_6| \left[G_{12,6}\sin(\theta_{12}-\theta_6) - B_{12,6}\cos(\theta_{12}-\theta_6)\right] + \\
& |V_{12}||V_{13}| \left[G_{12,13}\sin(\theta_{12}-\theta_{13}) - B_{12,13}\cos(\theta_{12}-\theta_{13})\right].
\end{aligned}$$

1.3.2 Solutions to Power Flow Problem

The basic methods to solve the nonlinear power flow equations can be categorized into:

- Gauss-Seidel Method

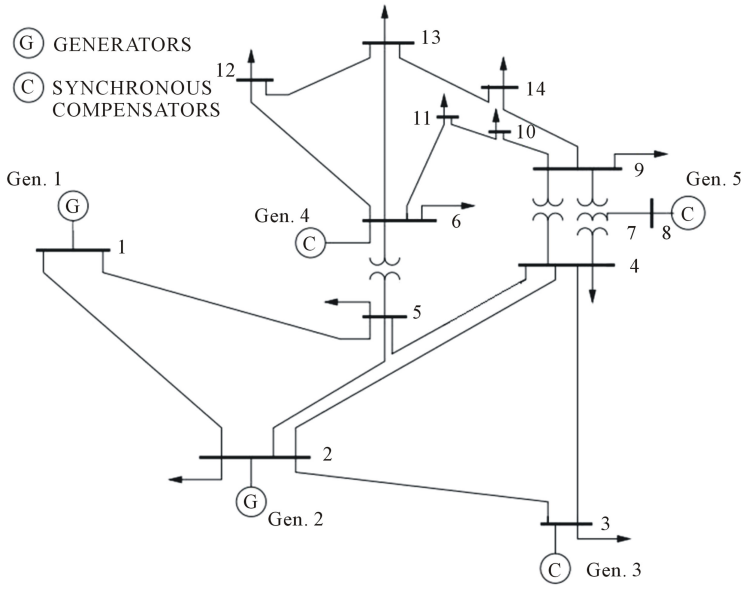

Figure 1.5: IEEE 14-bus test system

- Newton-Raphson Method
- Decoupled Method
- DC Power Flow Method

The Gauss-Seidel method is designed to find solution of a nonlinear system of equations by a series of iterations. It uses the most recent guess of solution to compute new solution values. Glimn and Stagg [47] were the first to use the Gauss-Seidel method to solve the power flow problem, followed by Brown and Tinney [20]. Although the Gauss-Seidel method requires small storage space, it exhibits poor convergence and even sometimes it does not converge despite of the existence of a solution.

The generalized Newton-Raphson method is also an iterative algorithm that is able to solve a nonlinear system of equations with the same of amount of variables.

At each iteration, each function in the system is approximated by its tangent hyperplane with Jacobian matrix. The Newton-Raphson method applied to the power flow equations were put forward and developed in [118, 36, 90, 109]. Due to its property of quadratic convergence, the Newton-Raphson method could efficiently solve the power flow problem, and become well established over the electric power system industry.

The decoupled methods was first proposed in 1970s [110, 111], and the fast decoupled version of [111] was further improved in [123, 84]. In power systems, a strong coupling between real power P and voltage angle θ, as well as between reactive power Q and voltage magnitude $|V|$ is commonly observed. However, the coupling between "$P-\theta$" and "$Q-V$" components of the problem is relatively weak [113]. Hence the essential idea of the decoupled methods is separately solving the "$P-\theta$" and "$Q-V$" problems, by taking advantage of the physical properties exhibited by electric power system being solved. As a result, the number of entries in Jacobian matrix is only half of the those used in the Newton-Raphson method, and Jacobian elements are voltage independent.

In real power dispatch or power market analysis, there is a need that power flow equations can be solved fast, however, the requirement of calculation accuracy is not very high. The DC power flow, is a simplification of the AC power flow equations presented in Section 1.3.1.1 by only restricting on MW flows, and neglecting the reactive power and voltage magnitude components. Three basic assumptions are used to derive DC power flow equations, which are listed as follows [137, 76]:

1. The resistance for each branch is ignored.

2. All the voltage magnitudes are equal to some base voltage V_0.

3. The voltage angle difference on the two ends of the branch is sufficiently small, so that $\cos(\theta_j - \theta_k) \approx 1$ and $\sin(\theta_j - \theta_k) \approx \theta_j - \theta_k$.

Give the above assumptions, the power flow equations are simplified into a linear problem, i.e., the DC power flow problem [92, 130].

1.3.3 Optimal Power Flow Problem

The optimal power flow (OPF) problem was first introduced by Carpentier [26] as an extension of conventional economic dispatch problem, to find the optimal settings of a given power system. In conventional power flow problem, the objective is to determine the steady-state complex voltages and powers of the network, with specified values of control variables. In contrast, the OPF problem is to optimize a prespecified objective while satisfying the system operating constraints, such as power flow equations, system security and physical limits, over certain control variables with unspecified values. The OPF plays a crucial role in power system operating and planning.

The general OPF formulation can be described as follows:

$$\min \quad f(\mathbf{x}, \mathbf{u}) \tag{1.8a}$$

$$\text{s.t.} \quad h(\mathbf{x}, \mathbf{u}) = 0, \tag{1.8b}$$

$$g(\mathbf{x}, \mathbf{u}) \leq 0, \tag{1.8c}$$

where $\mathbf{u}$ is the vector of control variables, such as active and reactive power generation and $\mathbf{x}$ is the vector of state variable, $f(\mathbf{x}, \mathbf{u})$ is an objective function, $h(\mathbf{x}, \mathbf{u})$ is the

power flow equations (such as (1.4a) and (1.4b)), and $g(\mathbf{x}, \mathbf{u})$ represents the limits on control variables and functional operating constraints on other variables.

There are many different ways to define an objective function for OPF, and therefore the selection of it may depend on the operating philosophy of the analyzed electric power system. As one of the most commonly used objective function in OPF, the power generation cost is generally represented by a nonlinear, second-order polynomial:

$$f(\mathbf{x}, \mathbf{u}) = \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2). \tag{1.9}$$

The inequality constraints (1.8c) may include limits on control variables, line security constraints and other associated system operating constraints. Among those, the following inequality constraints are simple and widely used:

$$\underline{P}_{G_j} \leq P_{G_j} \leq \overline{P}_{G_j},, \tag{1.10a}$$

$$\underline{Q}_{G_j} \leq Q_{G_j} \leq \overline{Q}_{G_j}, \tag{1.10b}$$

$$\underline{V}_j \leq |V_j| \leq \overline{V}_j. \tag{1.10c}$$

Thus, we can formulate a specific optimal power flow problem in polar form with objective (1.9), constraints (1.4a)-(1.4b), and constraints (1.10a)-(1.10c).

1.3.4 Solutions to Optimal Power Flow Problem

Since the OPF problem was introduced in the early 1960s, extensive research has been put on it and numerous efficient solution techniques have been developed, including linear programming, quadratic programming, nonlinear programming, Newton-based method, interior point method, heuristic methods and recent

convex relaxation methods. Extensive surveys can be found in [56, 82, 83, 88, 94, 41, 42].

Linear programming: Wells [128] first developed a linear programming approach to solve economic dispatch with security constraints by use of simplex method to the linearization of objective function and constraints. In the following years, a lot of research [114, 112, 5] was done in this category.

Quadratic programming: the sequential quadratic programming method was presented in [22]. Momoh [80] proposed a generalized quadratic model and Grudinin [54] employed successive quadratic programming method to solve a reactive power optimization.

Nonlinear programming: Dommel and Tinney [35] made a remarkable progress and first successfully solved the OPF with gradient method in 1968. Further research with nonlinear programming technique included [6, 138].

Newton-based method: an early version of Newton-based algorithm applied to OPF problem was found in [98], and an improved Newton-based method was proposed by Sun et al. [116]. Recent developments included [29, 119]

Interior point method: based on the primal-dual logarithmic barrier method, Granville [50] developed an algorithm to solve the optimal reactive power dispatch problem. The predictor-corrector interior point technique could be found in [132, 127, 120], and multiple-centrality corrections technique was employed in [25]. Min and Shengsong [78], Sousa et al. [107] presented trust-region based interior point method for OPF problems.

Heuristic methods: a genetic algorithm for solving large-scale economic dispatch problem was presented in [28], and to solve optimal power flow problem with FACTS devices, Chung and Li [30] proposed a hybrid genetic algorithm. Application of evolutionary algorithm into OPF could be found in [135, 104]. Yoshida et al. [134] applied the particle swarm optimization method for reactive power and voltage control, and Abido [2] considered a particle swarm optimization approach for OPF problem. A modified particle swarm optimization for economic dispatch problem was found in [89].

Recent convex relaxation: a common approximation method (DC OPF) to OPF problem, was found in [111, 93, 115] which linearized the power flow equations. Recently, Jabr [58] convexified the power flow problem over radial networks in the form of second order cone programming (SOCP). However, the same convexification did not hold for a meshed network. Another progress toward convexifying the OPF problem was made by Bai and Wei [10], Bai et al. [12], which first formulated the OPF problem as a semidefinite programming (SDP) problem. Although numerical experiments have been shown in [58, 12], whether or under what conditions the convexifications turn out to be exact has not been discussed. This was well studied in [68, 66, 67]. Lavaei and Low [68] demonstrated that the SDP relaxation was exact if and only if the duality gap was zero, and also showed that IEEE test bus systems with 14, 30, 57, 118 and 300 buses had no duality gap. Based on these work, Bai and Wei [11], Jabr [57] further simplified the SDP relaxation to exploit graph sparsity. Nonetheless, the exactness of SDP relaxation for mesh networks were challenged by Bukhsh et al.

[21], Lesieutre et al. [70] under certain cases. Phan [91] introduced a Lagrangian dual problem for solving and a branch-and-bound algorithm in case that the strong duality did not hold. In the basis of this work, Gopalakrishnan et al. [49] showed SDP based branch-and-bound approach was better than the Lagrangian duality based branch-and-bound.

1.3.4.1 OPF over Distribution (Tree) Network

Although OPF is traditionally solved for transmission networks, it is also becoming increasingly crucial for distribution networks. Farivar et al. [40] studied a SOCP relaxation of a multi-timescale optimal Volt/VAR control problem, which was proved to be exact provided load over satisfaction was allowed. Similar proofs for exactness of SDP or SOCP relaxation for OPF problem over radial networks under the condition of load over satisfaction could be found in [19, 136, 106], among which Sojoudi and Lavaei [106] used a simpler argument. Another type of sufficient conditions for the exactness of SOCP relaxation was discussed in [71, 44, 43], which required the upper bounds of voltage magnitudes not binding. Based on [136], Lavaei et al. [69] proved the tightness of SDP or SOCP relaxation given the voltage angles across each line was sufficiently small.

CHAPTER 2
RISK-AVERSE STRATEGIC PLANNING OF HVDC RENEWABLE ENERGY GRIDS

2.1 Introduction

The advantages of wind energy as a renewable and commonly available source of energy come at a cost of uncertainty in the amount of wind energy that can be produced during any given time period. In this respect, energy production from renewable sources differs drastically from the traditional energy production from fossil sources, whose reserves can be accurately estimated and utilized in a controlled manner. As a result, the design and operation of the existing power infrastructure, which implicitly relies on the presumption that energy production is completely controllable, may not be ideally suited for the case when a significant portion of generated and consumed energy comes from renewable sources, such as wind.

In this chapter, we consider the problem of strategic design and operation of energy grids that are based exclusively on wind energy sources, and the primary issue that we aim to elucidate is the problem of effective control of risks of power shortages due to the uncertainties in wind energy production and power demands. Specifically, the question of interest is whether risk-averse planning of energy grid can be effective in hedging the risks of power shortages due to stochastic variations in energy generation and demand. To this end, we formulate the problem of strategic design and operation of renewable energy grid as a stochastic wind farm location problem, where the risks of unsatisfied power demands due to uncertainties in power generation are

quantified by means of a class of (generally nonlinear) statistical functionals known as *coherent measures of risk* and are minimized through optimal selection of sites for wind energy harvesting and matching of energy producers and customers.

We consider our analysis to be at the *strategic level* as it pertains to planning and operation at the relatively long-term monthly scale with respect to the power generation and demand. The assumption that all energy demand within the grid is served by renewable wind energy sources implies that these sources must represent large-scale, massive wind farms. In addition, we assume that the generated electricity is transmitted to demand nodes through high-voltage direct current (HVDC) lines. HVDC transmission lines are used to transmit bulk of electrical energy over long distances by means of direct current (DC), in contrast to the more common alternating current (AC) used in most of today's electrical transmission infrastructure. Since there is no need for reactive compensation along the transmission line, the HVDC lines typically lose less power than equivalent AC transmission lines. This, in addition to lower transmission costs, makes HVDC more economical than AC transmission for large amounts of power transmitted over long distances. Moreover, HVDC transmission can improve system's stability since it allows the operator to quickly change the direction of power flow, as well as allows for the power transportation between power systems with different frequencies. These characteristics of HVDC transmission make it an appealing choice for renewable energy grids with wind or solar energy sources, as it could aid in mitigating the effects of intermittency and fluctuation and smooth the power outputs, as well as improve the economic viability of renewable energy due

to lower transmission costs.

The problem of configuration of power generating systems with renewable energy and storage has been extensively studied. In [102], a genetic algorithm has been proposed to determine the optimal configuration of power system in isolated island with installed renewable energy power plants. Katsigiannis and Georgilakis [61] performed tabu search to solve a combinatorial problem which aimed to optimize sizing of small isolated hybrid power systems. Similarly, Ekren and Ekren [37] applied simulated annealing method for achieving the optimal size of a PV/wind integrated hybrid energy system with battery storage. In addition to these heuristics methods, stochastic programming models have also been employed to design the energy system. Abbey and Joós [1] put forward a stochastic mixed integer programming model to optimize sizing of storage system for an existing isolated wind-diesel power system. In [65], a multi-stage stochastic mixed integer programming model has been presented for a comprehensive hybrid power system design by including renewable energy generation. More particularly, Burke and OMalley [23] considered the problem which sought to find the optimal locations to incorporate wind capacity on an existing transmission system network. A portfolio approach to decide optimal wind power deployment in Europe has been studied in [97], which endeavored to smooth out the fluctuations through geographic diversification of wind farms.

The uncertainties of classic facility location problem have been dealt with by different criteria, such as optimizing expected performance, or optimizing the worst-case performance. Daskin et al. [34] proposed a model called the α-reliable minimax

regret model, which essentially minimized the α-quantile of all regrets. Chen et al. [27] established a model called α-reliable mean-excess regret that instead minimized the expected regret of the "tail". The aim of this chapter is to develop stochastic programming models combined with various risk measures to determine the optimal locations and operation of wind farms in HVDC renewable energy grids.

The remainder of this chapter is organized as follows. In Section 2.2, we formulate three stochastic wind farm location models with different degrees of risk aversion. Branch-and-bound solution algorithms for the resulting mixed-integer linear and nonlinear programming problems, which employ Benders decomposition method and outer polyhedral approximations of nonlinear constraints are presented in Section 2.3. Section 2.4 discusses dataset generation, computational results and corresponding solution analysis.

2.2 Stochastic Wind Farm Location Models

In this section, we introduce several stochastic models that address the problem of strategic location of wind farms so as to satisfy the power demand at a given set of demand sites at a minimum cost. In all models, it is assumed that the power demand and energy generation are uncertain, or stochastic. To model these sources of uncertainty, we pursue the scenario-based approach traditional to stochastic programming, i.e., we assume that the set of random events Ω in the probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is discrete, $\Omega = \{\omega_1, \ldots, \omega_K\}$, where each elementary random event, or *scenario* ω_k has a non-zero probability $\mathsf{P}\{\omega_k\} = q_k > 0$, such that $\sum_k q_k = 1$.

Below we introduce a generic stochastic model (GS) that allows for satisfying the expected demand in each bus (demand node) of the power grid. To this end, we introduce the following notations:

i : index of demand nodes;

j : index of candidate sites of wind farms;

k : index of scenarios;

h : number of wind farms to locate;

γ : cost of power shortages;

λ : annual amortized cost per mile of HVDC transmission line built;

M : an upper bound on the number of wind turbines that can be installed at a given candidate site; for simplicity, it is assumed to be constant across all sites;

f_j : annual amortized fixed cost of wind farm site j;

c_j : annual amortized cost of per turbine purchased and installed at site j;

q_k : probability of scenario k

d_{ij} : distance from node i to candidate site j;

K_{jk} : capacity of wind turbine in candidate site j under the scenario k;

D_{ik} : demand at node i under the scenario k;

$\overline{D}_i$: expected power demand at node i.

Also, we define the following decision variables:

x_j : binary variable indicating whether wind farm site j is selected;

y_{ij} : binary variable indicating whether demand node i is connected to wind farm site j;

ζ_{ij} : number of turbines at wind farm site j serving demand node i;

p_{ijk} : power generated at wind farm site j serving demand node i under scenario k.

2.2.1 A Generic Stochastic Model for Strategic Wind Farm Location

Using the introduced above notations, a generic stochastic model for strategic wind farm location under uncertainties (GS) can be presented in the form of a mixed-integer linear programming problem:

$$\min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} \tag{2.1a}$$

$$\text{s. t.} \quad \sum_j x_j = h, \tag{2.1b}$$

$$y_{ij} \leq x_j, \qquad \forall i, j, \tag{2.1c}$$

$$\zeta_{ij} \leq M y_{ij}, \qquad \forall i, j, \tag{2.1d}$$

$$p_{ijk} \leq K_{jk} \zeta_{ij}, \qquad \forall i, j, k, \tag{2.1e}$$

$$\sum_k q_k \sum_j p_{ijk} \geq \overline{D}_i, \qquad \forall i, \tag{2.1f}$$

$$x_j, y_{ij} \in \{0, 1\},\ \zeta_{ij} \in \mathbb{Z}_+,\ p_{ijk} \in \mathbb{R}_+, \qquad \forall i, j, k. \tag{2.1g}$$

The objective function (2.1a) represents the cumulative annual cost to be minimized. Constraint (2.1b) stipulates that exactly h wind farms are to be located. Con-

straint (2.1c) states that a demand node i cannot be assigned to a wind farm j unless a wind farm is located at site j. Constraint (2.1d) limits the number of wind turbines at site j that can be assigned for serving bus i. Constraint (2.1e) ensures that power supplied by site j to bus i under scenario k is less than or equal to the total capacity of all wind turbines assigned at site j to serving bus i. Constraint (2.1f) ensures that the expected power supplied to bus i from all sites does at least meet the expected demand at bus i. Lastly, constraint (2.1g) determines the values that decision variables take, where $\mathbb{Z}_+$ and $\mathbb{R}_+$ denote the sets of non-negative integer and real numbers, respectively. In what follows, the feasible set defined by constraints (2.1b)–(2.1g) is denoted by $\mathcal{P}$.

It is easy to see that the generic model (2.1) is prone to substantial power shortages, which may occur in particular scenarios when the power demand at bus i and/or the amount of energy supplied to this bus deviate from the corresponding average figures. This is a consequence of the well-known properties of stochastic optimization models where constraints are satisfied "on average" [63]. In order to avoid power shortages, one may require that energy demands at each bus i are satisfied for every scenario $\omega_k \in \Omega$, which can be written as

$$\max_k \left\{ D_{ik} - \sum_j p_{ijk} \right\} \leq 0, \quad \forall i. \tag{2.2}$$

This method, also known as "robust optimization" approach [62], has been acknowledged in the literature as such that can often lead to overly conservative and exceedingly costly solutions [63]. In addition, enforcing the last constraint does not guarantee shortage-free power distribution in practice, since the scenario data repre-

sents only a finite sample from the generally unknown distributions of power demand and wind energy generation.

In this work, we pursue a risk-averse stochastic optimization approach which is supposed to avoid the potentially large power shortages associated with the expected-value constraint (2.1f) as well as the high costs associated with the "robust" constraints (2.2) by explicitly accounting for the *risks* of power shortages.

2.2.2 Risk-averse Models for Strategic Wind Farm Location

To quantify the risk of power shortages that may have large magnitudes but very low probabilities of occurring, we employ a class of statistical functional known as *coherent measures of risk* [9], and, more specifically, the well-known Conditional Value-at-Risk (CVaR) measure [96] and its nonlinear generalizations, Higher Moment Coherent Risk (HMCR) measures [64].

Technically, a *risk measure* is a function $\rho : \mathcal{X} \mapsto \mathbb{R}$, where $\mathcal{X}$ is an appropriately defined linear space of $\mathcal{F}$-measurable functions $X : \Omega \mapsto \mathbb{R}$. Further definition of risk measures $\rho(X)$ typically requires specifying whether the larger or smaller realizations of random element X are considered to be "risky". Here we adopt the setup common in engineering literature, where the random variable $X = X(\mathbf{x}, \omega)$ is assumed to represent the cost or loss associated with the decision $\mathbf{x}$, and thus smaller realizations of X are preferred (the alternative assumption, that X represents payoff or reward is prevalent in economics and finance domains).

Then, $\rho(X)$ as defined above is said to be a *coherent* measure of risk [9, 63] if

it satisfies the additional properties of *monotonicity*, $\rho(X_1) \leq \rho(X_2)$ for all $X_1 \leq X_2$; *sub-additivity*, $\rho(X_1+X_2) \leq \rho(X_1)+\rho(X_2)$; *positive homogeneity*, $\rho(\lambda X) = \lambda\rho(X)$ for a constant $\lambda > 0$; and *translation invariance*, $\rho(X+c) = \rho(X)+c$ for any $c \in \mathbb{R}$. The monotonicity property asserts that smaller losses bear less risk. The sub-additivity property in combination with positive homogeneity provides for *convexity* of coherent risk measures, which entails that coherent measures of risk allow for risk reduction via diversification, and, importantly, admit efficient optimization of risk via the methods of convex programming. The translation invariance property allows for efficient risk hedging, see [9] for a detailed discussion.

From this definition, it is easy to see that the risk measure defined as $\rho(X) = \mathsf{E}X$ is coherent. Hence, if one defines the stochastic cost/loss function X as the energy shortage at site i, $X_i(\omega_k) = D_{ik} - \sum_j p_{ijk}$, then constraints (2.1f) stipulating that power demand at each bus i must be satisfied on average, can equivalently be interpreted as the requirement of non-negative risk of power shortages at each bus i,

$$\rho(X_i) \leq 0, \quad \forall i. \tag{2.3}$$

Similarly, another trivial instance of coherent measures of risk is represented by the "maximum loss" measure, $\rho(X) = \max X$, which associates the risk of a stochastic loss or cost X with its largest possible realization (it is assumed here that the distribution of X has a bounded support, in the general case the maximization operator in the definition of this risk measure must be replaced with the essential supremum, $\rho(X) = \operatorname{ess\,sup} X$, see, e.g., [63] for details). Then, the conservative-but-costly approach of ensuring that power demands are satisfied at every scenario,

embodied by constraints (2.2), reduces to the risk constraints of the same form (2.3) where ρ is selected as the maximum loss measure.

In order to find, as we have proposed above, an effective – both methodologically and computationally – compromise between the "loose", risk-neutral expectation-based constraints (2.1f) and the most conservative risk-averse constraints (2.2), we will employ the well-known Conditional Value-at-Risk (CVaR) measure [96]. For a given confidence level $\alpha \in (0, 1)$, Conditional Value-at-Risk $\mathsf{CVaR}_\alpha(X)$ can be interpreted as the expected cost or loss that can occur with probability $1 - \alpha$ over the prescribed time horizon, or as the average of the $(1 - \alpha) \cdot 100\%$ of the largest (worst) realizations of the stochastic loss factor X. This interpretation is reflected in the fact that for continuously distributed X, $\mathsf{CVaR}_\alpha(X)$ can be represented in the form of the conditional expectation

$$\mathsf{CVaR}_\alpha(X) = \mathsf{E}[X \mid X \geq F_X^{-1}(\alpha)], \tag{2.4}$$

where $F_X(t)$ denotes the cumulative distribution function of X, and $F_X^{-1}(\alpha)$ is the α-quantile of X, or such a deterministic value that can be exceeded by X with probability $1-\alpha$ (in financial and risk management literature it is also known as Value-at-Risk with confidence level α, $\mathsf{VaR}_\alpha(X)$).

In the case of general distributions of X, definition (2.4) does not apply, in the sense that the corresponding conditional expectation is not guaranteed to have coherence properties [96]. It has been shown in [96] that in the general case $\mathsf{CVaR}_\alpha(X)$ can be represented as a convex combination of $F_\alpha^{-1}(X)$ and the conditional expectation of losses strictly exceeding $F_\alpha^{-1}(X)$, with weight coefficients dependent on both X and α.

A more computationally attractive definition of CVaR for general loss distributions presents it as the optimal value of the following unconstrained convex optimization problem [95, 96]:

$$\mathsf{CVaR}_\alpha(X) = \min_\eta \ \eta + (1-\alpha)^{-1}\mathsf{E}(X-\eta)_+, \tag{2.5}$$

where $X_\pm = \max\{0, \pm X\}$. Besides being a coherent measure of risk, $\mathsf{CVaR}_\alpha(X)$ possesses a number of other properties, such as, for example, continuity with respect to the confidence level α. In the context of the preceding discussion, another notable property of the CVaR measure is that, as a function of parameter α, it includes both $\rho(X) = \mathsf{E}X$ and $\rho(X) = \operatorname{ess\,sup} X$ as special cases:

$$\lim_{\alpha\to 0} \mathsf{CVaR}_\alpha(X) = \mathsf{E}X, \qquad \lim_{\alpha\to 1} \mathsf{CVaR}_\alpha(X) = \operatorname{ess\,sup} X. \tag{2.6}$$

Hence, to achieve a balance between the "risk-neutral" approach of ensuring that power shortages do not occur on average, and the "absolute risk-averse" approach requiring that power shortages never occur, one may quantify the risk of power shortages using CVaR measure with an appropriately selected value of confidence level $\alpha \in (0,1)$, whereby the shortage risk would be represented by the average of $(1-\alpha)\cdot 100\%$ largest shortages.

To incorporate the quantification of risks of power shortages in the wind farm location model (2.1) via the Conditional-Value-at-Risk measure, we define the cost/loss function X as the cumulative power shortage over all buses,

$$X(\omega_k) = \sum_i \left(D_{ik} - \sum_j p_{ijk} \right)_+ . \tag{2.7}$$

In order to have an additional degree of flexibility in our model, we include $\mathsf{CVaR}_\alpha(X)$ in the objective of problem (2.1) with an appropriate weight coefficient $\gamma > 0$, which represents the cost (in millions of dollars) of 1MW of power short:

$$\begin{aligned} \min \quad & \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \gamma \mathsf{CVaR}_\alpha(X) \\ \text{s. t.} \quad & x_j, y_{ij}, \zeta_{ij}, p_{ijk} \in \mathcal{P}, \end{aligned}$$

where X is defined by (2.7). Note that the cost of shortages in the objective function is non-negative due to the fact that $X(\omega_k) \geq 0$ in (2.7). By further defining auxiliary variables U_k and η, the risk-averse, CVaR-based stochastic model (CVaRS) can be formulated as follows:

$$\min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \gamma \left(\eta + \frac{1}{1-\alpha} \sum_k q_k U_k \right) \tag{2.8a}$$

$$\text{s. t.} \quad U_k \geq \sum_i \left(D_{ik} - \sum_j p_{ijk} \right)_+ - \eta, \qquad \forall k, \tag{2.8b}$$

$$U_k \in \mathbb{R}_+, \qquad \forall k, \tag{2.8c}$$

$$x_j, y_{ij}, \zeta_{ij}, p_{ijk} \in \mathcal{P}. \tag{2.8d}$$

By means of the Conditional Value-at-Risk measure, the risk-averse formulation (2.8) accounts for the risk of power shortages as the first moment of the $(1-\alpha)$-tail of the shortages distribution. At the same time, the "risk" as a proxy for "large losses that have low probability of occurring" is commonly associated in the risk management literature with "heavy tails" of distributions, and the distributions of power shortages are well known to be heavy tailed (see, e.g., [46, 31, 77]). Therefore, it is of interest to take into account higher moments of shortage distribution in assessing

the risk of power shortages. This can be accomplished by means of the family of Higher-Moment Coherent Risk (HMCR) measures [64]. Assuming that the space $\mathcal{X}$ admits a sufficient degree of integrability, i.e., $\mathcal{X} = \mathcal{L}^p(\Omega, \mathcal{F}, \mathsf{P})$ for a given $p \geq 1$, the HMCR measures are defined as

$$\mathsf{HMCR}_{p,\alpha}(X) = \min_{\eta \in \mathbb{R}} \ \eta + (1-\alpha)^{-1} \|(X-\eta)_+\|_p, \quad p \geq 1, \quad \alpha \in (0,1), \tag{2.9}$$

where $\|X\|_p = (\mathsf{E}|X|^p)^{1/p}$. Obviously, the HMCR family contains CVaR as a special case of $p = 1$.

Similarly to CVaR-based formulation (2.8), minimization of the total cost that includes the shortage risk cost as expressed by a higher tail moment of shortage distribution is given by the following HMCR-based stochastic optimization (HMCRS) model:

$$\min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \gamma\big(\eta + (1-\alpha)^{-1} U_0\big) \tag{2.10a}$$

$$\text{s. t.} \quad q_k^{-1/p} U_k \geq \sum_i \left(D_{ik} - \sum_j p_{ijk} \right)_+ - \eta, \qquad \forall k, \tag{2.10b}$$

$$U_0 \geq \left(U_1^p + \ldots + U_K^p\right)^{1/p}, \tag{2.10c}$$

$$U_0 \geq 0; \ \ U_k \geq 0, \qquad \forall k, \tag{2.10d}$$

$$x_j, y_{ij}, \zeta_{ij}, p_{ijk} \in \mathcal{P}. \tag{2.10e}$$

Constraint (2.10c) represents a (nonlinear) p-order cone constraint, whence formulation (2.10) represents a mixed-integer p-order cone programming (MIpOCP) problem. The next section discusses the solution methods for the proposed risk-averse stochastic wind farm location models CVaRS (2.8) and HMCRS (2.10), as well as their

special case, the risk-neutral GS model (2.1).

2.3 Benders Decomposition Based Branch-and-Bound Algorithms

2.3.1 General Formulations

The discussed formulations of GS andf CVaRS models can generally be written as mixed-integer linear programming programming problems of the form

$$\begin{aligned} Z \;=\; \min \quad & \mathbf{a}^\top \mathbf{z} + \mathbf{b}^\top \mathbf{u} && \text{(MILP)} \\ \text{s. t.} \quad & \mathbf{A}\mathbf{z} + \mathbf{B}\mathbf{u} \le \mathbf{c}, \\ & \mathbf{z} \in \mathcal{Z} \subset \mathbb{Z}^n_+, \quad \mathbf{u} \in \mathbb{R}^m_+, \end{aligned}$$

where $\mathbf{z}$ and $\mathbf{u}$ represent an n-dimensional vector of integer variables and an m-dimensional vector of continuous variables, and $\mathcal{Z}$ is a bounded subset of $\mathbb{Z}^n_+$. Assume that problem MILP is bounded and feasible. Then, it can equivalently be represented as

$$\begin{aligned} Z \;=\; \min \quad & \mathbf{a}^\top \mathbf{z} + t(\mathbf{z}) && \text{(2.11a)} \\ \text{s. t.} \quad & \mathbf{z} \in \mathcal{Z}, && \text{(2.11b)} \end{aligned}$$

where for any given $\mathbf{z} \in \mathcal{Z}$, function $t(\mathbf{z})$ is defined to be the optimal objective value of the linear programming problem

$$\begin{aligned} t(\mathbf{z}) \;=\; \min \quad & \mathbf{b}^\top \mathbf{u} && \text{(2.12a)} \\ \text{s. t.} \quad & \mathbf{B}\mathbf{u} \le \mathbf{c} - \mathbf{A}\mathbf{z}, && \text{(2.12b)} \\ & \mathbf{u} \ge \mathbf{0}. && \text{(2.12c)} \end{aligned}$$

Note that since set $\mathcal{Z} \subset \mathbb{Z}_+^n$ is bounded, the unboundedness of the original problem MILP is associated with that of problem (2.12). By introducing dual variables $\boldsymbol{\xi}$, we can calculate $t(\mathbf{z})$ through solving its dual problem, under the assumption of boundedness of problem (2.12). The dual of problem (2.12) is

$$\begin{aligned} t(\mathbf{z}) \;=\; \max \quad & (\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi} \\ \text{s. t.} \quad & \mathbf{B}^\top \boldsymbol{\xi} \leq \mathbf{b}, \\ & \boldsymbol{\xi} \leq \mathbf{0}. \end{aligned} \tag{SMILP}$$

If the feasible region of problem SMILP is empty, then the primal subproblem (2.12) is either unbounded or infeasible, which implies the unboundedness or infeasibility of the original problem MILP. Otherwise, we can enumerate all extreme points $(\boldsymbol{\xi}_{\mathrm{p}}^1, \ldots, \boldsymbol{\xi}_{\mathrm{p}}^I)$, and extreme rays $(\boldsymbol{\xi}_{\mathrm{r}}^1, \ldots, \boldsymbol{\xi}_{\mathrm{r}}^J)$ of the feasible region of SMILP, where I and J denote the numbers of extreme points and extreme rays. Therefore, the dual subproblem SMILP can be rewritten as

$$t(\mathbf{z}) = \min \quad t \tag{2.13a}$$

$$\text{s. t.} \quad (\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi}_{\mathrm{r}}^j \leq 0, \quad \forall j = 1, \ldots, J, \tag{2.13b}$$

$$(\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi}_{\mathrm{p}}^i \leq t, \quad \forall i = 1, \ldots, I, \tag{2.13c}$$

$$t \in \mathbb{R}. \tag{2.13d}$$

By replacing $t(\mathbf{z})$ in problem (2.11) with that given by formulation (2.13), we obtain a reformulation of the original problem MILP:

$$\begin{aligned} \min \quad & \mathbf{a}^\top \mathbf{z} + t && \text{(RMILP)} \\ \text{s. t.} \quad & (\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi}_{\mathrm{r}}^j \leq 0, \quad \forall j = 1, \ldots, J, \\ & (\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi}_{\mathrm{p}}^i \leq t, \quad \forall i = 1, \ldots, I, \\ & \mathbf{z} \in \mathcal{Z}, \quad t \in \mathbb{R}. \end{aligned}$$

We denote problem RMILP but only with a subset of constraints (2.13b) and (2.13c) as problem MMILP, representing the master problem of mixed-integer linear programming problem MILP.

The standard Benders decomposition scheme is then invoked, which consists in solving the "relaxed" problem MMILP (as usual, the procedure is initialized by solving MMILP without any constraints (2.13b) and (2.13c) and the variable t in its objective disregarded). If it is unbounded, let $\boldsymbol{\xi}_{\mathrm{r}}^*$ be the column vector in which all the corresponding simplex multipliers are negative, after simplex terminates. Therefore, $\boldsymbol{\xi}_r^*$ is an extreme ray of the feasible region of SMILP, whence a feasibility cut

$$(\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi}_{\mathrm{r}}^* \leq 0 \tag{2.14a}$$

is added to MMILP and the problem is thus resolved until an optimal solution $(\mathbf{z}^*, t^*)$ of MMILP is obtained. Subsequently, the dual subproblem SMILP is solved for the given $\mathbf{z}^*$, and let $\boldsymbol{\xi}_{\mathrm{p}}^*$ be the corresponding optimal solution, or an extreme point of its feasible region. If $t(\mathbf{z}^*) = (\mathbf{c} - \mathbf{A}\mathbf{z}^*)^\top \boldsymbol{\xi}_{\mathrm{p}}^* > t^*$, then problem MMILP is augmented

with the optimality cut

$$(\mathbf{c} - \mathbf{A}\mathbf{z})^\top \boldsymbol{\xi}_{\mathrm{p}}^* \leq t \tag{2.14b}$$

and resolved.

The decomposition procedure stops when the condition $t(\mathbf{z}^*) = t^*$ is satisfied. During each iteration, a feasibility or optimality cut is added, and an optimal solution of RMILP is obtained in a finite number of iterations due to finiteness of I and J [15]. The following two propositions follow readily from the above discussion.

Proposition 1. *If $\tilde{\mathbf{z}} \in \mathcal{Z}$ and there is an optimal solution to the dual subproblem* SMILP *with objective value $\tilde{t} = \max\{(\mathbf{c} - \mathbf{A}\tilde{\mathbf{z}})^\top \boldsymbol{\xi} : \mathbf{B}^\top \boldsymbol{\xi} \leq \mathbf{b}, \boldsymbol{\xi} \leq \mathbf{0}\}$, then $\mathbf{a}^\top \tilde{\mathbf{z}} + \tilde{t}$ is an upper bound on the optimal solution value of problem* RMILP.

Proof. If there is an optimal solution to the dual subproblem SMILP, by strong duality

$$\min\left\{\mathbf{b}^\top \mathbf{u} : \mathbf{B}\mathbf{u} \leq \mathbf{c} - \mathbf{A}\tilde{\mathbf{z}}, \mathbf{u} \geq \mathbf{0}\right\} = \max\left\{(\mathbf{c} - \mathbf{A}\tilde{\mathbf{z}})^\top \boldsymbol{\xi} : \mathbf{B}^\top \boldsymbol{\xi} \leq \mathbf{b}, \boldsymbol{\xi} \leq \mathbf{0}\right\} = \tilde{t}.$$

Assume $\tilde{\mathbf{u}}$ is an optimal solution to the linear program $\min\{\mathbf{b}^\top \mathbf{u} : \mathbf{B}\mathbf{u} \leq \mathbf{c} - \mathbf{A}\tilde{\mathbf{z}}, \mathbf{u} \geq \mathbf{0}\}$, then $(\tilde{\mathbf{z}}, \tilde{\mathbf{u}})$ is a feasible solution to the original problem MILP. Therefore, $\mathbf{a}^\top \tilde{\mathbf{z}} + \tilde{t} = \mathbf{a}^\top \tilde{\mathbf{z}} + \mathbf{b}^\top \tilde{\mathbf{u}}$ is the objective value of a feasible solution to problem MILP. Because problem RMILP is equivalent to problem MILP, $\mathbf{a}^\top \tilde{\mathbf{z}} + \tilde{t}$ is also an upper bound on the optimal value of the problem RMILP.

Proposition 2. *Assume that $(\mathbf{z}^*, t^*)$ is an optimal solution of the master problem* MMILP. *If the optimal objective value of the corresponding problem* SMILP *is equal to t^*, i.e., $t^* = \max\{(\mathbf{c} - \mathbf{A}\mathbf{z}^*)^\top \boldsymbol{\xi} : \mathbf{B}^\top \boldsymbol{\xi} \leq \mathbf{b}, \boldsymbol{\xi} \leq \mathbf{0}\}$, then $(\mathbf{z}^*, t^*)$ is an optimal solution to the equivalent reformulation of the original problem* RMILP.

Proof. For each iteration, we solve problem MMILP, which is essentially a "relaxed" problem of problem RMILP, with a subset of constraints (2.13b) and (2.13c), so the corresponding objective value $\mathbf{a}^\top \mathbf{z}^* + t^*$ is a lower bound of problem RMILP. According to Proposition 1, since we have $\mathbf{z}^* \in \mathcal{Z}$ and the subproblem SMILP has an optimal objective value $\tilde{t}$, $\mathbf{a}^\top \mathbf{z}^* + \tilde{t}$ is an upper bound of the problem RMILP. Thus, if $\tilde{t} = t^*$, the upper bound and lower bound of problem RMILP is equal. Accordingly, problem RMILP obtains an optimal solution.

2.3.2 Benders Decomposition Based Algorithm for GS and CVaRS Models

In the following, we denote problems MMILP and RMILP with relaxed integrality constraints (namely, $\mathbf{z} \in \mathcal{Z} \subset \mathbb{Z}_+^n$ replaced by $\mathbf{z} \in \text{conv}(\mathcal{Z}) \subset \mathbb{R}_+^n$) as problem MLP and problem RLP, respectively. Furthermore, we define a node k in the branch-and-bound tree by a triple $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k) \in \mathbb{Z}_+^{2n} \times (\mathbb{R} \cup \{+\infty\})$, where $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ are the bounds on $\mathbf{z}$ at node k and $\underline{Z}^k$ is a lower bound on $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$. The problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is defined as the problem MLP with added constraints $\underline{\mathbf{z}}^k \leq \mathbf{z} \leq \bar{\mathbf{z}}^k$, and $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$ is the corresponding optimal objective value. Similarly, for any $\hat{\mathbf{z}}^k \in \mathbb{Z}_+^n$ we denote by $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$ by replacing the variable $\mathbf{z}$ with the value $\hat{\mathbf{z}}^k$ in problem SMILP, and by $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)}$ the corresponding optimal objective value. In addition, we introduce $\overline{Z}$ and $\mathcal{N}$ to denote the global upper bound on Z_{RMILP} and the set of active branch-and-bound nodes, respectively. The algorithm is described as follows (see **Algorithm 2.1** for details).

Step 1 Initialize the set of active branch-and-bound nodes $\mathcal{N}$ with root node defined

as $(\underline{\mathbf{z}}^0, \bar{\mathbf{z}}^0, \underline{Z}^0)$, and global upper bound $\overline{Z}$ with positive infinity.

Step 2 Select and remove a node from the set $\mathcal{N}$.

Step 3 Solve problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.

Step 4 If the solution of problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is feasible and its optimal objective value is less than the current global upper bound $\overline{Z}$, go to **Step 5**; otherwise, fathom this node and go to **Step 2**.

Step 5 Denote the optimal solution to problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ by $(\hat{\mathbf{z}}^k, \hat{t}^k)$. If the values of $\hat{\mathbf{z}}^k$ are all integers, go to **Step 6**; otherwise, branch on this node and go to **Step 2**.

Step 6 Solve the problem $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$. If its optimal objective value equal to $\hat{t}^k$ obtained in **Step 5**, then update the global upper bound $\overline{Z}$ and incumbent solution, and fathom this node; otherwise, go to **Step 7**.

Step 7 Check the solution status of problem $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$, if it is unbounded, then add a feasibility cut to problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$, go to **Step 3**; otherwise, check whether $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} > \hat{t}^k$, if it is true, then add an optimality cut to problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$, go to **Step 3**.

Algorithm 2.1 A Benders decomposition based branch-and-bound algorithm for GS and CVaRS models

Set global upper bound $\overline{Z} := +\infty$; set $\underline{Z}^0 := -\infty$.
Set $\underline{z}_i^0 := -\infty$, $\bar{z}_i^0 := +\infty$ for all $i \in \{1, \ldots, n\}$; initialize node list $\mathcal{N} := \{(\underline{\mathbf{z}}^0, \bar{\mathbf{z}}^0, \underline{Z}^0)\}$.
while $\mathcal{N} \neq \emptyset$ **do**
    Select and remove a node $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k)$ from $\mathcal{N}$.
    Solve $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.
    if $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is feasible **and** $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} < \overline{Z}$ **then**
        Let $(\hat{\mathbf{z}}^k, \hat{t}^k)$ be the optimal solution to $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.
        if $\hat{\mathbf{z}}^k \in \mathbb{Z}_+^p$ **then**
            Solve $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$.
            if $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} = \hat{t}^k$ **then**
                $\overline{Z} := Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$; update incumbent solution.
            else
                if $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$ *is unbounded* **then**
                    Add feasibility cut
                end if
                if $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} > \hat{t}^k$ **then**
                    Add optimality cut
                end if
            end if
        else

Algorithm 2.2 A Benders decomposition based branch-and-bound algorithm for GS and CVaRS models (continued)

21: Pick i_0 in $\{i \in \{1,\ldots,n\} : \hat{z}_i^k \notin \mathbb{Z}_+\}$.

22: Let $\underline{z}_i := \underline{z}_i^k$, $\bar{z}_i := \bar{z}_i^k$ for all $i \in \{1,\ldots,n\}\backslash\{i_0\}$.

23: Let $\bar{z}_{i0} := \lfloor \hat{z}_{i0}^k \rfloor$, $\underline{z}_{i0} = \lfloor \hat{z}_{i0}^k \rfloor + 1$.

24: $\mathcal{N} := \mathcal{N} \cup \left\{ \left(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}, Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}\right), \left(\underline{\mathbf{z}}, \bar{\mathbf{z}}^k, Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}\right) \right\}$.

25: **end if**

26: **end if**

27: Remove every node $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k) \in \mathcal{N}$ such that $\underline{Z}^k \geq \overline{Z}$.

28: **end while**

Proposition 3. *The Benders decomposition based branch-and-bound algorithm for* GS *and* CVaRS *models terminates with the upper bound* $\overline{Z}$ *equal to the optimal objective value of original problem* MILP.

Proof. We have shown that problem RMILP is an equivalent reformulation of the original problem MILP, thus we only need to prove that the algorithm would result in the same objective value as that of problem RMILP. In order words, all that remains to prove is that the sub-tree rooted at a fathomed node cannot contain an integer feasible solution to problem RMILP whose objective value is strictly less than the current incumbent integer solution. Denote by $\mathsf{RLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ an instance of problem RLP with additional constraints $\underline{\mathbf{z}}^k \leq \mathbf{z} \leq \bar{\mathbf{z}}^k$, and let $Z_{\mathsf{RLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$ be its optimal objective value. Note that a node is only fathomed in lines 6 and 10 in **Algorithm 2.1**. In

line 6, node $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k)$ is fathomed if $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is infeasible or if the condition $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq \overline{Z}$ is satisfied. As it was indicated above, problem MLP and problem RLP are linear relaxations of problem MMILP and problem RMILP respectively, and hence $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is a relaxation of $\mathsf{RLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$, and one has $Z_{\mathsf{RLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$. Also, if $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is infeasible, $\mathsf{RLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ will also be infeasible. Similarly, an integer feasible solution that is strictly better than the incumbent solution cannot exist in the sub-tree rooted at such a node. In line 10, the node is fathomed because the integer feasible solution to $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is also feasible to $\mathsf{RLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ according to Proposition 2, and thus it is the best integer feasible solution that can be found at the sub-tree rooted at the fathomed node.

2.3.2.1 Feasibility Cuts for GS Model

The generic stochastic model GS (2.1) can be written in the form of problem MILP, where the integer-valued vector $\mathbf{z}$ contains variables x_j, y_{ij}, and ζ_{ij}, and non-negative vector $\mathbf{u}$ contains variables p_{ijk}. Then, the corresponding problem(2.11) has the form

$$Z = \min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} \tag{2.15a}$$

$$\text{s. t.} \quad \sum_j x_j = h, \tag{2.15b}$$

$$y_{ij} \leq x_j, \qquad \forall i, j, \tag{2.15c}$$

$$\zeta_{ij} \leq M y_{ij}, \qquad \forall i, j, \tag{2.15d}$$

$$x_j, y_{ij} \in \{0, 1\},\ \zeta_{ij} \in \mathbb{Z}_+, \qquad \forall i, j, k, \tag{2.15e}$$

where function $t(\mathbf{z})$ defined as the optimal objective of subproblem (2.12) is equal to either 0 when the problem

$$t(\mathbf{z}) = \min \quad 0 \tag{2.16a}$$

$$\text{s. t.} \quad p_{ijk} \leq K_{jk}\zeta_{ij}, \qquad \forall i, j, k, \tag{2.16b}$$

$$\sum_k q_k \sum_j p_{ijk} \geq \overline{D}_i, \qquad \forall i, \tag{2.16c}$$

$$p_{ijk} \geq 0, \tag{2.16d}$$

is feasible for the given values of ζ_{ij}, or $+\infty$ when (2.16) is infeasible. Obviously, if $\hat{\mathbf{z}} = (\hat{x}_j, \hat{y}_{ij}, \hat{\zeta}_{ij})$ is an optimal solution of (2.15) such that (2.16) has a feasible $\hat{p}_{ijk}$ for the given $\hat{\zeta}_{ij}$, then $(\hat{\mathbf{z}}, \hat{\mathbf{u}}) = (\hat{x}_j, \hat{y}_{ij}, \hat{\zeta}_{ij}, \hat{p}_{ijk})$ is an optimal solution of the original problem (2.1). On the other hand, if (2.16) is infeasible, then the corresponding $\hat{\mathbf{z}}$ should be eliminated from the feasible region of the master problem (2.15). In accordance to the described above algorithm, this is accomplished by augmenting (2.15) with feasibility cuts

$$\sum_i \sum_j \sum_k K_{jk}\hat{\alpha}_{ijk}\zeta_{ij} - \sum_i \overline{D}_i\hat{\beta}_i \leq 0, \tag{2.17}$$

where $(\hat{\alpha}_{ijk}, \hat{\beta}_i)$ is an extreme ray of the dual $\mathsf{SMILP}(\hat{\mathbf{z}})$ of problem (2.16), which takes the form

$$t(\hat{\mathbf{z}}) = \max \quad \sum_i \sum_j \sum_k K_{jk}\hat{\zeta}_{ij}\alpha_{ijk} - \sum_i \overline{D}_i\beta_i \tag{2.18a}$$

$$\text{s. t.} \quad \alpha_{ijk} - q_k\beta_i \leq 0, \qquad \forall i, j, k, \tag{2.18b}$$

$$\alpha_{ijk}, \beta_i \leq 0, \qquad \forall i, j, k. \tag{2.18c}$$

Clearly, no optimality cuts are added to (2.15) since (2.16) is a "feasibility" subproblem.

2.3.2.2 Feasibility and Optimality Cuts for CVaRS Model

Analogously to above, the CVaRS model (2.8) can be reformulated in the form (2.11)–(2.12) as follows

$$Z = \min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \gamma t \tag{2.19a}$$

$$\text{s. t.} \quad \sum_j x_j = h, \tag{2.19b}$$

$$y_{ij} \leq x_j, \qquad \forall i, j, \tag{2.19c}$$

$$\zeta_{ij} \leq M y_{ij}, \qquad \forall i, j, \tag{2.19d}$$

$$x_j, y_{ij} \in \{0, 1\},\ \zeta_{ij} \in \mathbb{Z}_+,\ t \in \mathbb{R}, \tag{2.19e}$$

with the subproblem (2.16) defined as:

$$t(\hat{\mathbf{z}}) = \min \quad \eta + \frac{1}{1-\alpha} \sum_k q_k U_k \tag{2.20a}$$

$$\text{s. t.} \quad p_{ijk} \leq K_{jk} \hat{\zeta}_{ij}, \qquad \forall i, j, k, \tag{2.20b}$$

$$\sum_k q_k \sum_j p_{ijk} \geq \overline{D}_i, \qquad \forall i, \tag{2.20c}$$

$$w_{ik} \geq D_{ik} - \sum_j p_{ijk}, \qquad \forall i, k, \tag{2.20d}$$

$$U_k \geq \sum_i w_{ik} - \eta, \qquad \forall k, \tag{2.20e}$$

$$w_{ik}, p_{ijk}, U_k \geq 0, \qquad \forall i, j, k. \tag{2.20f}$$

Let α_{ijk}, μ_i, β_{ik}, and θ_k be the dual variables associated with constraints (2.20b), (2.20c), (2.20d) and (2.20e) respectively. Then, the dual of subproblem becomes

$$t(\mathbf{z}) = \max \quad \sum_i \sum_j \sum_k K_{jk}\hat{\zeta}_{ij}\alpha_{ijk} - \sum_i \sum_k D_{ik}\beta_{ik} - \sum_i \overline{D}_i\mu_i \tag{2.21a}$$

$$\text{s. t.} \quad \alpha_{ijk} - \beta_{ik} - q_k\mu_i \leq 0, \qquad \forall i, j, k, \tag{2.21b}$$

$$-\theta_k \leq \frac{1}{1-\alpha}q_k, \qquad \forall k, \tag{2.21c}$$

$$-\beta_{ik} + \theta_k \leq 0, \qquad \forall i, k, \tag{2.21d}$$

$$-\sum_k \theta_k \leq 1, \tag{2.21e}$$

$$\alpha_{ijk}, \mu_i, \beta_{ik}, \theta_k \leq 0, \qquad \forall i, j, k. \tag{2.21f}$$

If, for a given set of values $\hat{\zeta}_{ij}$, where $(\hat{x}_j, \hat{y}_{ij}, \hat{\zeta}_{ij}, \hat{t})$ is an optimal solution of (2.19), problem (2.21) is unbounded, let $(\hat{\alpha}_{ijk}, \hat{\beta}_{ik}, \hat{\mu}_i)$ be an extreme ray of the feasible region of (2.21), such that $\sum_i \sum_j \sum_k (K_{jk}\hat{\alpha}_{ijk})\hat{\zeta}_{ij} - \sum_i \sum_k D_{ik}\hat{\beta}_{ik} - \sum_i \overline{D}_i\hat{\mu}_i > 0$. Then, the feasibility cut

$$\sum_i \sum_j \sum_k K_{jk}\hat{\alpha}_{ijk}\zeta_{ij} - \sum_i \sum_k D_{ik}\hat{\beta}_{ik} - \sum_i \overline{D}_i\hat{\mu}_i \leq 0, \tag{2.22}$$

is added to the master problem. If, on the other hand, the optimal objective $t^* = t(\hat{\mathbf{z}})$ of (2.21) is finite and such that $\hat{t} < t^*$, the following optimality cut is added to the master:

$$\sum_i \sum_j \sum_k K_{jk}\hat{\alpha}_{ijk}\zeta_{ij} - \sum_i \sum_k D_{ik}\hat{\beta}_{ik} - \sum_i \overline{D}_i\hat{\mu}_i \leq t, \tag{2.23}$$

where $(\hat{\alpha}_{ijk}, \hat{\beta}_{ik}, \hat{\mu}_i)$ is an optimal solution of (2.21).

2.3.3 HMCRS Model as a Mixed-Integer p-Order Cone Programming Problem

Due to the presence of p-order cone constraint in formulation (2.10),

$$U_0 \geq (U_1^p + \ldots + U_K^p)^{1/p}, \tag{2.24}$$

the HMCRS model represents a mixed-integer p-order cone programming problem (MIpOCP). Below we propose an algorithm for the MIpOCP HMCRS problem that combines the Benders decomposition with a general branch-and-bound algorithm for solving MIpOCP problems that was discussed in [126]. The idea of this method involves solving a polyhedral approximation of the integer relaxation of MIpOCP problem at each node of the BnB tree, and is based on the work of [124] for mixed integer second order cone programming problems (MISOCP).

The polyhedral approximation of pOCP, or the relaxed MIpOCP problem, is obtained by replacing nonlinear p-order cone constraints with their polyhedral approximations. It is crucial, however, that such a polyhedral approximation is "compact" with respect to the number of facets, since a straightforward approximation of a p-cone in $\mathbb{R}^{K+1}$ by tangent hyperplanes requires $O(2^K)$ facets. To this end, a *lifted* representation of a multidimensional p-cone is used [126, 14], which expresses a p-cone in $\mathbb{R}_+^{K+1}$ as an intersection of $K-1$ three-dimensional p-cones:

$$U_{2K-1} = U_0, \quad U_{K+k} \geq (U_{2k-1}^p + U_{2k}^p)^{1/p}, \quad k = 1, \ldots, K-1. \tag{2.25}$$

Then, it is easy to see that if each of the three-dimensional p-cones is replaced by its polyhedral approximation with $O(L)$ facets, the resulting polyhedral approximation of multidimensional p-cone (2.24) will contain no more than $O(KL)$ facets. In particular,

the following gradient-based approximation of three-dimensional p-cones (2.25) in the positive orthant $\mathbb{R}^3$ was suggested in [126]:

$$U_{K+k} \geq a_j^{(p)} U_{2k-1} + b_j^{(p)} U_{2k}, \quad j = 0, \ldots, L, \tag{2.26a}$$

where

$$a_j^{(p)} = (\cos^p \theta_j + \sin^p \theta_j)^{\frac{1-p}{p}} \cos^{p-1} \theta_j, \quad b_j^{(p)} = (\cos^p \theta_j + \sin^p \theta_j)^{\frac{1-p}{p}} \sin^{p-1} \theta_j, \tag{2.26b}$$

and values θ_j, $j = 0, \ldots, L$, satisfy the condition $0 = \theta_0 < \theta_1 < \ldots < \theta_L = \frac{\pi}{2}$.

The constructed in such a way polyhedral approximation of relaxed MIpOCP problem is solved instead of the exact nonlinear pOCP formulation at every node of the BnB tree until an integer-valued solution is found. Since the employed polyhedral approximation is of outer type, its integer solution is not guaranteed to be feasible to the exact nonlinear pOCP formulation, which needs to be solved in order to verify feasibility and declare incumbent or branch further (see [124] for details).

The computational advantages of this approach come from the warm-start capabilities of LP solvers that drastically reduce computational cost of solving a polyhedral approximation of relaxed problem during BnB search in comparison to solving an exact nonlinear relaxation using an interior-point method. The computational overhead associated with the necessity of invoking an exact nonlinear relaxation for testing feasibility of the obtained solution is relatively low. It must be emphasized, however, that the effectiveness of this method is based on the premise that the employed polyhedral approximation is relatively low-dimensional. For example, Vielma et al. [124] uses a lifted polyhedral approximation of three-dimensional second-order

cones due to Ben-Tal and Nemirovski [14], whose accuracy is exponentially small in the number of facets. The accuracy of gradient approximation (2.26) of p-cones for $p \neq 2$ is only polynomially small in the number of facets, and Vinel and Krokhmal [126] introduced a fast cutting plane algorithm for solving the resulting polyhedral approximation problems. On the other hand, it has been observed in [124, 126] that the accuracy of polyhedral approximations used during the BnB process may be rather "crude" without a significant deterioration of effectiveness of the algorithm. We use this observation in the present work by employing polyhedral approximation (2.25)–(2.26) with a small number L of facets.

Finally, to solve an exact nonlinear relaxation of MIpOCP problem during the BnB algorithm, we use the fact that when $p > 1$ is a rational number, $p = r/s$, a p-order cone in $\mathbb{R}^{K+1}$ can be equivalently represented as a intersection of $O(K \log r)$ three-dimensional second-order cones [86, 13, 85]. Namely, the p-cone (2.24) is equivalent to

$$U_k^R \leq u_k^s U_0^{r-s} U_k^{R-r}, \quad u_k \geq 0, \qquad k = 1, \ldots, K, \tag{2.27a}$$

$$U_0 \geq \sum_{k=1}^{K} u_k, \tag{2.27b}$$

where $R = 2^{\lceil \log_2 r \rceil}$. Then, each nonlinear inequality (2.27) can be represented by $\lceil \log_2 r \rceil$ three-dimensional "rotated" second-order cones, see [85] for details. For example, in the case of $p = 3$, the cone (2.24) in $\mathbb{R}_+^{K+1}$ admits a representation via $2K$ quadratic cones:

$$U_0 \geq \sum_{k=1}^{K} u_k; \quad U_k^2 \leq U_0 v_k, \;\; v_k^2 \leq u_k U_k, \quad k = 1, \ldots, K.$$

2.3.3.1 Benders Decomposition Based Branch-and-Bound Algorithm for HMCRS Model

In this section, we propose an efficient method for solving the HMCRS model as a MIpOCP problem that incorporates the Benders decomposition mechanism into the branch-and-bound framework proposed in [124], and as such exploits both the mixed-integer structure and p-order cone constraints in the problem.

By employing the nomenclature introduced in Section 2.3.1, we represent the HMCRS model (2.10) in the general form of mixed-integer nonlinear programming problem (MINLP)

$$\begin{aligned} Z \;=\; \min \quad & \mathbf{a}'^{\top}\mathbf{z} + \mathbf{b}'^{\top}\mathbf{u} && \text{(MINLP)} \\ \text{s. t.} \quad & \mathbf{A}'\mathbf{z} + \mathbf{B}'\mathbf{u} \leq \mathbf{c}', \\ & \mathbf{u} \in \mathcal{K}_p, \\ & \mathbf{z} \in \mathcal{Z} \subset \mathbb{Z}^n_+, \quad \mathbf{u} \geq \mathbf{0}, \end{aligned}$$

where $\mathcal{K}_p$ is a p-order cone in an appropriate high-dimensional space, such that mixed-integer linear problem MILP is obtained from MINLP by replacing the nonlinear conic constraint with its polyhedral approximation. The integer relaxation of MINLP, obtained by replacing constraint $\mathbf{z} \in \mathcal{Z} \subset \mathbb{Z}^n_+$ by $\mathbf{z} \in \operatorname{conv}(\mathcal{Z}) \subset \mathbb{R}^n_+$, is denoted as NLP. Then, the rest of the definitions stay intact, namely problem RMILP denotes the equivalent Benders reformulation of problem MILP, and MMILP represents the corresponding master problem, or relaxation of RMILP, problem SMILP is the corresponding dual subproblem of MILP, and MLP and RLP stand for problems obtained

by relaxing the integrality constraint in problem MMILP and problem RMILP, respectively. Similarly, $\underline{\mathbf{z}}^k$, $\bar{\mathbf{z}}^k$, $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$, $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$, $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$, $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)}$ and $\mathcal{N}$ are the same as described in Section 2.3.2. In addition, we denote the problem obtained by adding constraints $\underline{\mathbf{z}}^k \leq \mathbf{z} \leq \bar{\mathbf{z}}^k$ to problem NLP for any $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k) \in \mathbb{Z}_+^{2n}$ by $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$, and the corresponding optimal objective value by $Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$. Furthermore, $\underline{Z}^k$ is a lower bound on $Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$, and $\overline{Z}$ is the global upper bound on Z_{MINLP}. The algorithm is described as follows (see **Algorithm 2.3** for details).

Step 1-5 The same as described in Section 2.3.2.

Step 6 Solve the problem $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$. If it is unbounded, then add a feasibility cut to problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ and go to **Step 3**; if its optimal objective value satisfies $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} > \hat{t}^k$, then add an optimality cut to problem $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ and go to **Step 3**. If its optimal objective value equal to $\hat{t}^k$ obtained in **Step 5**, go to **Step 7**.

Step 7 Solve problem $\mathsf{NLP}(\hat{\mathbf{z}}^k)$. If it is feasible and its optimal objective value is less than the current global upper bound $\overline{Z}$, then update the global upper bound $\overline{Z}$ and the incumbent solution.

Step 8 If the lower and upper bounds at the current node do not coincide, $\underline{\mathbf{z}}^k \neq \bar{\mathbf{z}}^k$, and $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} < \overline{Z}$, then solve $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ and go to **Step 9**; otherwise, fathom this node and go to **Step 2**.

Step 9 If the solution of problem $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is feasible and its objective value is less than the current global upper bound $\overline{Z}$, go to **Step 10**; otherwise, fathom this node and go to **Step 2**.

Step 10 Denote the optimal solution to problem $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ by $(\tilde{\mathbf{z}}^k, \tilde{\mathbf{u}}^k)$. If the values of $\tilde{\mathbf{z}}^k$ are all integers, then update the global upper bound $\overline{Z}$ and incumbent solution, fathom this node and go to **Step 2**; otherwise, branch on this node and go to **Step 2**.

Algorithm 2.3 A Benders decomposition based branch-and-bound algorithm for HMCRS model

1: Set global upper bound $\overline{Z} := +\infty$; set $\underline{Z}^0 := -\infty$.

2: Set $\underline{z}_i^0 := -\infty$, $\bar{z}_i^0 := +\infty$ for all $i \in \{1, \ldots, n\}$; initialize node list $\mathcal{N} := \left\{\left(\underline{\mathbf{z}}^0, \bar{\mathbf{z}}^0, \underline{Z}^0\right)\right\}$.

3: **while** $\mathcal{N} \neq \emptyset$ **do**

4: Select and remove a node $\left(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k\right) \in \mathcal{N}$.

5: Solve $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.

6: **if** $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ *is feasible* **and** $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} < \overline{Z}$ **then**

7: Let $(\hat{\mathbf{z}}^k, \hat{t}^k)$ be an optimal solution to $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.

8: **if** $\hat{\mathbf{z}}^k \in \mathbb{Z}_+^n$ **then**

9: Solve $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$.

10: **if** $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} = \hat{t}^k$ **then**

11: Solve $\mathsf{NLP}(\hat{\mathbf{z}}^k)$.

12: **if** $\mathsf{NLP}(\hat{\mathbf{z}}^k)$ *is feasible* **and** $Z_{\mathsf{NLP}(\hat{\mathbf{z}}^k)} < \overline{Z}$ **then**

13: $\overline{Z} := Z_{\mathsf{NLP}(\hat{\mathbf{z}}^k)}$.

14: **end if**

15: **if** $\underline{\mathbf{z}}^k \neq \bar{\mathbf{z}}^k$ **and** $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} < \overline{Z}$ **then**

16: Solve $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.

17: **if** $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ *is feasible* **and** $Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} < \overline{Z}$ **then**

18: Let $(\tilde{\mathbf{z}}^k, \tilde{\mathbf{u}}^k)$ be the optimal solution to $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$.

19: **if** $\tilde{\mathbf{z}}^k \in \mathbb{Z}_+^n$ **then**

20: $\overline{Z} := Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$.

Algorithm 2.4 A Benders decomposition based branch-and-bound algorithm for HMCRS model (continued)

else

Select i_0 in $\{i \in \{1, \ldots, n\} : \tilde{z}_i^k \notin \mathbb{Z}\}$.

Let $\underline{z}_i = \underline{z}_i^k$, $\bar{z}_i = \bar{z}_i^k$ for all $i \in \{1, \ldots, n\} \backslash \{i_0\}$.

Let $\bar{\mathbf{z}}_{i0} = \lfloor \tilde{z}_{i0}^k \rfloor$, $\underline{z}_{i0} = \lfloor \tilde{z}_{i0}^k \rfloor + 1$.

$\mathcal{N} := \mathcal{N} \cup \{(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}, Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}), (\underline{\mathbf{z}}, \bar{\mathbf{z}}^k, Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)})\}$.

end if

end if

end if

else

if $\mathsf{SMILP}(\hat{\mathbf{z}}^k)$ *is unbounded* **then**

Add a feasibility cut

end if

if $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} > \hat{t}^k$ **then**

Add an optimality cut

end if

end if

else

Select i_0 in $\{i \in \{1, \ldots, n\} : \hat{z}_i^k \notin \mathbb{Z}\}$.

Let $\underline{z}_i = \underline{z}_i^k$, $\bar{z}_i = \bar{z}_i^k$ for all $i \in \{1, \ldots, n\} \backslash \{i_0\}$.

Let $\bar{z}_{i0} = \lfloor \hat{z}_{i0}^k \rfloor$, $\underline{z}_{i0} = \lfloor \hat{z}_{i0}^k \rfloor + 1$.

Algorithm 2.5 A Benders decomposition based branch-and-bound algorithm for HMCRS model (continued)

41: $\mathcal{N} := \mathcal{N} \cup \{(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}, Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}), (\underline{\mathbf{z}}, \bar{\mathbf{z}}^k, Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)})\}$.

42: **end if**

43: **end if**

44: Remove every node $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k) \in \mathcal{N}$ such that $\underline{Z}^k \geq \overline{Z}$.

45: **end while**

Proposition 4. *The Benders decomposition based branch-and-bound algorithm for* HMCRS *model terminates with the upper bound* $\overline{Z}$ *equal to the optimal objective value of original problem* MINLP.

Proof. First, since problem MILP is an outer linear approximation of the nonlinear problem MINLP, we may regard MILP as a relaxation of MINLP. Besides, problem MMILP could be deemed as a relaxation of problem MILP because it is a relaxation of problem RMILP, which is an equivalent reformulation of problem MILP. Thus, problem MMILP is a relaxation of problem MINLP, and accordingly problem MLP is a relaxation of problem NLP.

Assuming that the polyhedral relaxation MILP is bounded, this directly implies the finiteness of this algorithm. We may encounter the issue that solution $(\hat{\mathbf{z}}^k, \hat{t}^k)$ is generated again in several nodes if we branch as lines 21–26 in **Algorithm 2.3**, however, this can only happen a finite number of times according to [124].

In the following, we will show that an integer feasible solution to problem

MINLP that has an objective value strictly less than the cost of the current incumbent integer solution cannot exist in the sub-tree rooted at a fathomed node. Note that a node is only fathomed in lines 6, 15, 17 and 19 in **Algorithm 2.3**. In line 6, we fathom the node if $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is infeasible or if the condition $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq \overline{Z}$ is satisfied. As it was indicated above, $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is a relaxation of $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$, and hence if $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is infeasible, $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ will also be infeasible. In addition, one must have $Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)}$. Therefore, an integer feasible solution which is strictly better than the incumbent solution cannot exist in the sub-tree rooted at such a node $(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k, \underline{Z}^k)$. Note that in line 10, if $Z_{\mathsf{SMILP}(\hat{\mathbf{z}}^k)} = \hat{t}^k$, then according to Proposition 2, $\hat{\mathbf{z}}$ is in fact an integer feasible solution of problem RMILP, and therefore one has to check problem NLP to make further decision. In line 15, the node is fathomed when $\underline{\mathbf{z}}^k = \bar{\mathbf{z}}^k$ or $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq \overline{Z}$. If $\underline{\mathbf{z}}^k = \bar{\mathbf{z}}^k$, then $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k) = \mathsf{NLP}(\hat{\mathbf{z}}^k)$ and hence the node k has been processed by lines 12–14. If $Z_{\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq \overline{Z}$, then $Z_{\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)} \geq \overline{Z}$ since $\mathsf{MLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$ is a relaxation of $\mathsf{NLP}(\underline{\mathbf{z}}^k, \bar{\mathbf{z}}^k)$. In line 17, the node is fathomed for the same reasons as described above with respect to line 6. The node is fathomed in line 19 because the best integer feasible solution has been found at the sub-tree rooted at the fathomed node.

2.3.3.2 Feasibility and Optimality Cuts for HMCRS Model

The polyhedral approximation of the HMCRS model (2.10) can be written in the form of a mixed-integer linear problem MILP where, as before, the integer-valued vector $\mathbf{z}$ comprises variables x_j, y_{ij}, and ζ_{ij}, so that the corresponding problem (2.11)

has the form

$$Z = \min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \gamma t \tag{2.28a}$$

$$\text{s. t.} \quad \sum_j x_j = h, \tag{2.28b}$$

$$y_{ij} \le x_j, \qquad \forall i, j, \tag{2.28c}$$

$$\zeta_{ij} \le M y_{ij}, \qquad \forall i, j, \tag{2.28d}$$

$$x_j, y_{ij} \in \{0, 1\},\ \zeta_{ij} \in \mathbb{Z}_+,\ t \ge 0, \tag{2.28e}$$

and the subproblem (2.12), whose optimal objective $t(\mathbf{z})$ defines the value of variable t above, takes the form

$$t(\mathbf{z}) = \min \quad \eta + (1-\alpha)^{-1} U_{2K-1} \tag{2.29a}$$

$$\text{s. t.} \quad p_{ijk} \le K_{jk} \zeta_{ij}, \qquad \forall i, j, k, \tag{2.29b}$$

$$\sum_k q_k \sum_j p_{ijk} \ge \overline{D}_i, \qquad \forall i, \tag{2.29c}$$

$$w_{ik} \ge D_{ik} - \sum_j p_{ijk}, \qquad \forall i, k, \tag{2.29d}$$

$$U_k \ge q_k^{1/p} \left(\sum_i w_{ik} - \eta \right), \qquad \forall k, \tag{2.29e}$$

$$U_{K+r} \ge a_l^{(p)} U_{2r-1} + b_l^{(p)} U_{2r}, \qquad \forall r, l, \tag{2.29f}$$

$$w_{ik}, p_{ijk}, U_k \ge 0, \tag{2.29g}$$

where the nonnegative variables p_{ijk}, w_{ik}, and U_k comprise the vector $\mathbf{u}$, see (2.12).

Let α_{ijk}, μ_i, β_{ik}, θ_k, and ν_{lk} be the multipliers associated with constraints (2.29b), (2.29c), (2.29d), (2.29e), and (2.29f), respectively. Then, the dual of sub-

problem (2.29) can be written as

$$t(\mathbf{z}) = \max \quad \sum_i \sum_j \sum_k K_{jk}\zeta_{ij}\alpha_{ijk} - \sum_i \sum_k D_{ik}\beta_{ik} - \sum_i \overline{D}_i\mu_i \tag{2.30a}$$

$$\text{s. t.} \quad \alpha_{ijk} - \beta_{ik} - q_k\mu_i \le 0, \qquad \forall i,j,k, \tag{2.30b}$$

$$-\beta_{ik} + \theta_k \le 0, \qquad \forall i,k, \tag{2.30c}$$

$$-\sum_k \theta_k \le 1, \tag{2.30d}$$

$$-\theta_k + \sum_l a_l^{(p)} \nu_{l\lceil \frac{k}{2} \rceil} \le 0, \qquad \forall k = 1,3,\ldots,K-1, \tag{2.30e}$$

$$-\theta_k + \sum_l b_l^{(p)} \nu_{l\lceil \frac{k}{2} \rceil} \le 0, \qquad \forall k = 2,4,\ldots,K, \tag{2.30f}$$

$$\sum_l A_l^{(p)} \nu_{l(\frac{K}{2}+\lceil \frac{k}{2} \rceil)} - \sum_l \nu_{lk} \le 0, \qquad \forall k = 1,3,\ldots,K-3, \tag{2.30g}$$

$$\sum_l B_l^{(p)} \nu_{l(\frac{K}{2}+\lceil \frac{k}{2} \rceil)} - \sum_l \nu_{lk} \le 0, \qquad \forall k = 2,4,\ldots,K-2, \tag{2.30h}$$

$$-\sum_l \nu_{l(K-1)} \le \frac{1}{1-\alpha}, \tag{2.30i}$$

$$\alpha_{ijk}, \mu_i, \beta_{ik}, \theta_k, \nu_{lk} \le 0. \tag{2.30j}$$

For the sake of simplicity, we assume K is an even number. If, for a given optimal value $\hat{\mathbf{z}}$ obtained from (2.28), problem (2.30) is unbounded, let $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\mu}})$ be an extreme ray of (2.30), such that $\sum_i \sum_j \sum_k K_{jk}\hat{\alpha}_{ijk}\hat{\zeta}_{ij} - \sum_i \sum_k D_{ik}\hat{\beta}_{ik} - \sum_i \overline{D}_i\hat{\mu}_i > 0$. Then, a feasibility cut

$$\sum_i \sum_j \sum_k K_{jk}\hat{\alpha}_{ijk}\zeta_{ij} - \sum_i \sum_k D_{ik}\hat{\beta}_{ik} - \sum_i \overline{D}_i\hat{\mu}_i \le 0, \tag{2.31}$$

is added to (2.28). Conversely, let $\hat{t}$ and t^* denote an optimal value of t in the master problem (2.28), and the optimal objective function value of problem (2.30),

respectively. If $\hat{t} < t^*$, then an optimality cut

$$\sum_i \sum_j \sum_k K_{jk} \hat{\alpha}_{ijk} \zeta_{ij} - \sum_i \sum_k D_{ik} \hat{\beta}_{ik} - \sum_i \overline{D}_i \hat{\mu}_i \leq t, \tag{2.32}$$

is added to the master problem (2.28).

2.4 Computational Study

2.4.1 Parameters and Data

This section provides description and justification for the selected data sets and particular values of parameters in the three stochastic wind farm location models, GS (2.1), CVaRS (2.8), and HMCRS (2.10) considered in this study.

First, note that the choice of specific values for parameters h (the number of wind farms to locate), p (the order of tail moment in the HMCR measures of risk), and α (the parameter controlling the tail cutoff point in both CVaR and HMCR measures of risk) are at the discretion of the decision maker. It can also be argued that the set of scenario probabilities q_k, $k = 1, \ldots, K$, is in most instances also specified by the decision maker/analyst (e.g., in the case of historic scenario data, one may choose whether to adopt the "physical" probabilities or apply a change of probability measure to work in the domain of "risk-neutral" probabilities, etc).

In the case study reported below, the value of the parameter h is set at $h = 3$, implying that three wind farms are to be established on a given set of candidate locations to serve the demand nodes. The value p in HMCR measure of risk in model (2.10) is chosen as $p = 3$, and the values of α are selected at $\alpha = 0.95$ for the CVaRS model and $\alpha = 0.90$ for the HMCRS model.

The rest of the parameters can be separated into two categories: deterministic parameters, namely γ, λ, f_j, c_j and d_{ij}, which are assumed to be constant across scenarios, and stochastic parameters, specifically K_{jk} and D_{ik}, which represent the uncertainties in wind speed and consumer demand, respectively. A detailed description and rationale behind assigning specific values to these parameters follow next.

2.4.1.0.1 Deterministic Parameters

The value of the parameter γ represents the cost of power shortages, in millions of dollars per MW short. In this study, we select values of γ in the range of 0 to 0.95 with a step of 0.05 to conduct sensitivity analysis of obtained solutions with respect to γ.

We assume that λ, the estimated cost of HVDC line per mile, is 1.5 million dollars. After amortizing it by 30 years, the cost is equal to 0.05 million dollars per mile per year. To evaluate the fixed cost of building a wind farm, f_j, we refer to Kuznia et al. [65], who estimated this value at 280 million dollars. To account for variation of land prices at different locations, we randomly generated the values of parameter f_j from the uniform distribution $U(260, 300)$, and amortized them by 20 years. Next, the cost of purchasing and installing a single wind turbine is reported to be between 1 and 2 million dollars [32]. Therefore, the corresponding costs c_j have been randomly generated from the uniform $U(1, 2)$ distribution (in millions of dollars), and amortized over 20 years. The distances d_{ij} were randomly generated from the uniform $U(200, 2000)$ distribution (in miles); in addition, to model the "extreme"

situations when building a transmission line from site j to demand node i is infeasible or prohibitively expensive, some of the distances were randomly set equal to 1,000,000 miles.

2.4.1.0.2 Stochastic Parameters

The values of parameters K_{jk} and D_{ik} are obtained either directly from historical data or from Monte Carlo simulation. The corresponding scenario sets are constructed in assumption of equiprobable scenarios, i.e., $q_k = 1/K$ for all k; below we discuss the procedures used for scenario generation.

The values of parameter K_{jk} representing wind turbine power output can be obtained from wind speed data. In this study, the two methods described below were used to generate scenario sets for wind speed (and, consequently, wind power production) distribution. Importantly, we assumed that the wind speed distributions at different site locations are statistically independent.

Historical records of monthly average wind speed data for different locations were obtained through the National Climatic Data Center. Typically, the monthly average wind speed data has a smaller variance and exhibits more symmetry comparing to hourly average wind speed. In this study, we assumed that the average wind speed for each site follows a normal distribution and used maximum likelihood estimation to calculate its mean and standard deviation based on historical monthly average wind speed data. Then, scenario sets for wind speed at different sites were randomly simulated from the estimated normal distributions.

Another commonly used method for simulation of wind speed data relies on Weibull distribution ([45, 3]), whose probability density function has the form

$$f(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(\frac{x}{\lambda})^k}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

where k and λ are the shape and scale parameters, respectively. To simulate wind speed distribution, the shape parameter of Weibull distribution is often chosen as $k = 2$, and we randomly set the scale parameter as an integer from the range of $8 \leq \lambda \leq 14$.

The wind speed data can then be converted to power output K_{jk} by use of a typical power curve equation [24, 101]

$$P = \frac{1}{2}\rho A v^3 C_p,$$

where C_p is the power coefficient is set equal to 0.45, $A = \pi 50^2$ m^2 represents the area swept by the rotor blades of the wind turbine, the density of air ρ is equal to 1.225 kg/m^3, v is the wind speed in m/s. Thus, P is the power output in watts (1 W = 1 kg $\cdot$ m^2/s^3). We then scale the results to MW.

The other stochastic parameter that is considered in this case study is the demand D_{ik} at bus i under scenario k. Similarly to the wind speed data, we also employ two approaches to generating the scenario set for power demand, but, in contrast to wind speed data, we assume that demands at different locations may be correlated.

To construct scenario set for power using historical data, we used the data from Electric Reliability Council of Texas (ERCOT), which describes eight subsection's

electricity consumption in the state of Texas, and scaled it by 0.02 in consideration of current wind energy penetration level (around 4%) in the United States.

Based on [103, 72, 39, 53], a second simulated scenario set was constructed in the assumption that the power demand at each node i follows a mixed normal distribution $XY_1+(1-X)Y_2$, where X is a Bernoulli random variable with parameter $\tilde{p}$, and $Y_1 \sim N_1(\mu, \sigma^2)$ and $Y_2 \sim N_2(\mu, 100\sigma^2)$ represent the "normal" demand and "extreme/peak" demand, respectively. The value of parameter $\tilde{p}$ of the Bernoulli distribution was chosen as $\tilde{p} = 0.9$. To account for the correlation between different demand nodes, we consider a correlated multivariate distribution by additionally assuming that distributions N_1 of different nodes are correlated with each other, but N_2 are independent (i.e., one may not expect that occurrence of rare events follows a certain pattern). We use the historical data from Texas to estimate the covariance matrix of demands. The samples of the "extreme" part $N_2(\mu, 100\sigma^2)$ of the mixed normal distribution are independently generated for each node with the σ^2 estimated from the historical data of the state of Texas.

In our numerical experiments, we constructed instances of wind farm location models of two sizes, one with 7 demand nodes and 6 candidate locations, and another with 14 demand nodes and 8 candidate locations. The deterministic parameters for model of each size were generated as described above. For models of each size, the scenario sets for stochastic parameters (the wind power production and power demand) were constructed in two ways, using the historical data and simulated data in accordance to the preceding descriptions.

2.4.2 Computational Time Comparison

The GS, CVaRS, and HMCRS optimization models, introduced in Section 2.2, and the corresponding solution algorithms proposed in Section 2.3 have been implemented in C++ using CPLEX 12.5 solver. In particular, the Benders decomposition-based BnB algorithms described in Sections 2.3.2–2.3.3 were implemented using callback functionality of CPLEX, and their computational performance was compared to CPLEX's standard MIP and Barrier MIP solvers. Namely, in the following tables we denote the standard CPLEX MIP solver as "MIP", and "MIP-B" stands for the Benders decomposition (**Algorithm 2.1**) algorithm applied to GS and CVaRS models. Similarly, "MISOCP" corresponds to solving the HMCRS model using the default CPLEX MIP Barrier solver as a mixed integer second-order cone programming problem after reformulating the p-order cone constraint via a set of second-order cones [85]. We also denote the cutting-plane based branch-and-bound algorithm for mixed integer p-order cone programming problems due to Vinel and Krokhmal [125] as "BnB", and the Benders decomposition based branch-and-bound algorithm (**Algorithm 2.3**) as use "BnB-B". The computational experiments were conducted on a 3GHz PC with 4GB RAM running 32-bit Windows 7.

Tables 2.1, 2.2, and 2.3 report the computational performance of the listed algorithms for the risk-neutral (GS), linear risk-averse (CVaRS), and nonlinear risk-averse (HMCRS) problems with varying number of scenarios (K = 200, 500, 1000, and 2000), which were generated using either historical or simulated data, for models with either 7 demand nodes and 6 candidate locations or 14 demand nodes and 8

candidate locations.

	GS		CVaRS		HMCRS		
K	MIP	MIP-B	MIP	MIP-B	MISOCP	BnB	BnB-B
200	1.545	0.344	11.295	1.488	6.879	6.639	2.215
500	6.817	1.295	245.968	4.961	31.154	33.011	5.038
1000	40.185	5.242	730.632	13.026	886.495	809.142	20.467

Table 2.1: Computational time summary (in seconds) for various algorithms applied to GS, CVaRS, and HMCRS problems with scenario sets of K scenarios based on historical data, on a model with 7 demand nodes and 6 candidate locations.

The conducted computational study indicates that the proposed Benders decomposition allows for drastic reductions in running time for both GS model and CVaRS models as compared to the default CPLEX MIP solver, and the computational improvements tend to increase with the number of scenarios. With regard to the nonlinear HMCRS model, we observe that the "BnB" method that only exploits the structure of p-order cone constraints via polyhedral approximations and the corresponding cutting-plane algorithm, offers relatively mild improvements over "MISOCP", the default CPLEX Barrier MIP solver (which also employs polyhedral approximations). In contrast, the proposed Benders-based "BnB-B" algorithm significantly outperforms the other two approaches, especially as the number of scenarios increases.

	GS		CVaRS		HMCRS		
K	MIP	MIP-B	MIP	MIP-B	MISOCP	BnB	BnB-B
500	8.097	1.341	60.238	6.396	371.638	118.956	23.000
1000	60.855	5.413	938.591	13.650	906.556	774.880	69.564
2000	284.840	29.156	4061.700	27.144	10803.40	6501.98	217.885

Table 2.2: Computational time summary (in seconds) for various algorithms applied to GS, CVaRS, and HMCRS problems with scenario sets of K scenarios based on simulated data, on a model with 7 demand nodes and 6 candidate locations.

	GS		CVaRS		HMCRS		
K	MIP	MIP-B	MIP	MIP-B	MISOCP	BnB	BnB-B
100	5.039	3.495	32.749	25.147	55.413	39.564	42.960
200	17.277	2.356	111.346	59.576	409.750	182.297	19.641
500	88.779	7.784	1067.050	113.749	1216.690	922.894	289.238

Table 2.3: Computational time summary (in seconds) for various algorithms applied to GS, CVaRS, and HMCRS problems with scenario sets of K scenarios based on historical data, on a model with 14 demand nodes and 8 candidate locations.

2.4.3 Out-of-sample Solution Analysis

In this section, we conduct an out-of-sample analysis of the constructed wind farm location models. To this end, we consider optimal solutions of the risk-neutral and risk-averse problems (GS, CVaRS, and HMCRS, respectively) for given fixed sets of parameters and scenarios, and then compute appropriate quantities of interest (for example, power shortages) under the assumption that one of the relevant parameters of the model assumes values that are different from those used in the corresponding optimization problem (i.e., we "test" the obtained solutions on data samples that were not used during optimization). Importantly, our out-of-sample analysis focuses on the "extreme", or "worst-case" scenarios, in order to emphasize the effects of risk-aversion in the solutions of the proposed models.

2.4.3.1 Shortage Analysis

Specifically, we assume that the out-of-sample data is represented by power demands D_{ik} (obviously, all parameters in the respective mathematical programming problems can be replaced with out-of-sample values; but for simplicity and clarity of interpretation of the results, in this study the out-of-sample data includes only the power demands).

We analyzed out-of-sample shortages across the grid, $\sum_i (D_{ik} - \sum_j \zeta_{ij} K_{jk})_+$, induced by the optimal solutions of GS, CVaRS, and HMCRS problems in the case of a model with demand 7 nodes and 6 candidate locations and 2000 scenarios based on simulated data, as well as in the case of 14 demand nodes and 8 candidate locations,

both with value of parameter $\gamma = 0.24$. We randomly generated a dataset consisting of 2000 scenarios of consumer demand D_{ik} from the same mixed normal distribution, and selected the scenarios with shortages beyond 0.95 sample quantile in our out-of-sample scenario set as the "extreme", or "worst-case" scenarios (in other words, the out-of-sample scenario set contained a total of 100 scenarios that represent 5% of highest power shortage levels).

The results are presented in the shortage histograms and boxplots in Figures 2.1 and 2.2. In the case of the smaller model with 7 demand nodes and 6 candidate locations, the boxplots of shortages indicate that both risk-averse models, HMCRS and CVaRS, significantly outperform the risk-neutral GS model. The CVaRS model has smaller 0.75 quantile value of shortages than HMCRS model, but it has larger extreme points. This observation is in accord with γ sensitivity analysis presented in the next section. Analysis of shortage histograms shows that all shortages for HMCRS model are within 500 MW, and most of its shortages fall in the range of [0, 50) MW, while the fraction of zero shortages exceeds 30%. As regards the CVaRS model, shortage has exceeded 500 MW in one scenario, but most of its shortages do not exceed 250 MW, also with a high fraction of zero shortages. In contrast, "extreme" out-of-sample power shortages in GS model are always non-zero and fall mainly between 500 MW and 1000 MW, and can be as high as 1500 MW.

Similarly, in the case of models with 14 demand nodes and 8 candidate locations, boxplots in Figure 2.2 indicate that the HMCRS model has the lowest 0.25 quantile, median, 0.75 quantile, etc., and CVaRS model performs much better than

the GS model. The histograms of power shortages indicate that over 20% of out-of-sample shortages are within $[0, 250)$ MW for HMCRS model. However, no shortages fall into this range in the case of CVaRS and GS models. Also, 98% of "extreme" out-of-sample shortages are below 1000 MW for HMCRS model. Although over 70% of "extreme" out-of-sample shortages are under 1000 MW for CVaRS model, there is a substantial number of out-of-sample shortages within [1000,1500) MW, and even reaching 2000 MW in one scenario. As regards the GS model, most of its shortages are between 1000 MW and 2000 MW, and it has nearly 5% of "extreme" out-of-sample shortages beyond 2000 MW. The largest shortage that was observed in the GS model is close to 2500 MW.

2.4.3.2 Sensitivity with Respect to Shortage Penalty Parameter γ

Recall that the sensitivity to power shortages of the risk-averse CVaRS and HMCRS models is determined by the parameter γ, which represents the dollar cost of 1 MW of power short. The risk-neutral GS model is insensitive to (does not contain) the parameter γ, and, moreover, for the value of $\gamma = 0$, all three models yield the same solutions. In this section, we evaluate several aspects of the performance of the GS, CVaRS, and HMCRS models with respect to different levels of sensitivity to power shortages, corresponding to varying the value of the parameter γ from 0 to 0.95. Obviously, the solution of the GS model would not change with γ, and can be considered as the "reference" point in this comparison.

To evaluate the performance of three models, we consider four criteria: (1)

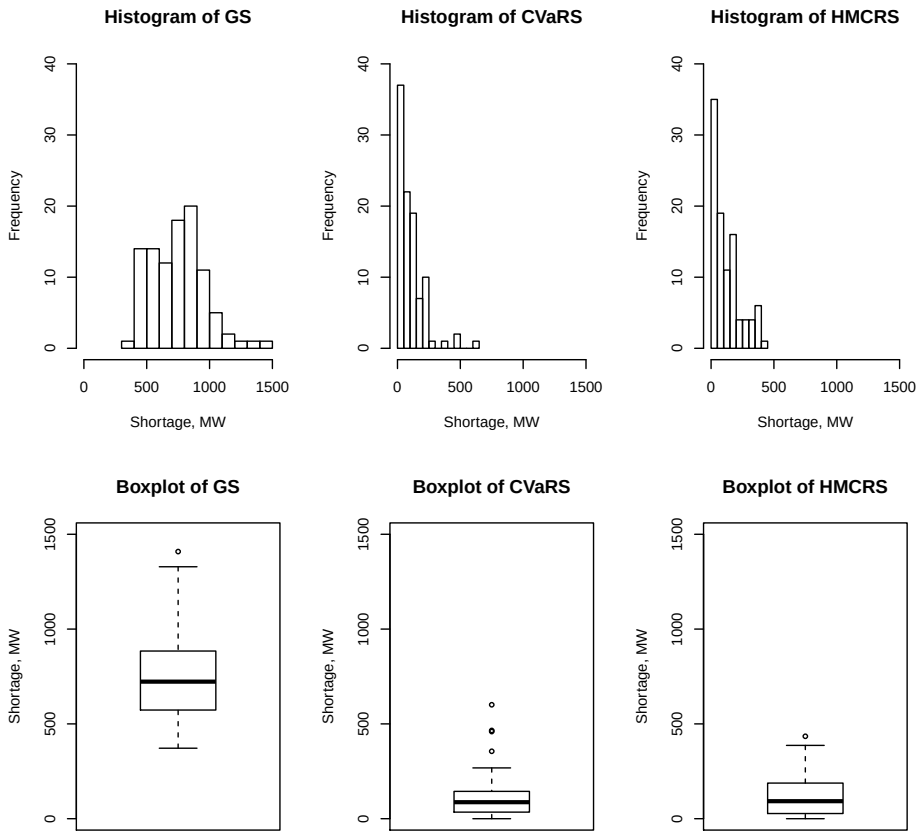

Figure 2.1: Shortages in out-of-sample extreme scenarios for model with 7 nodes and 6 locations

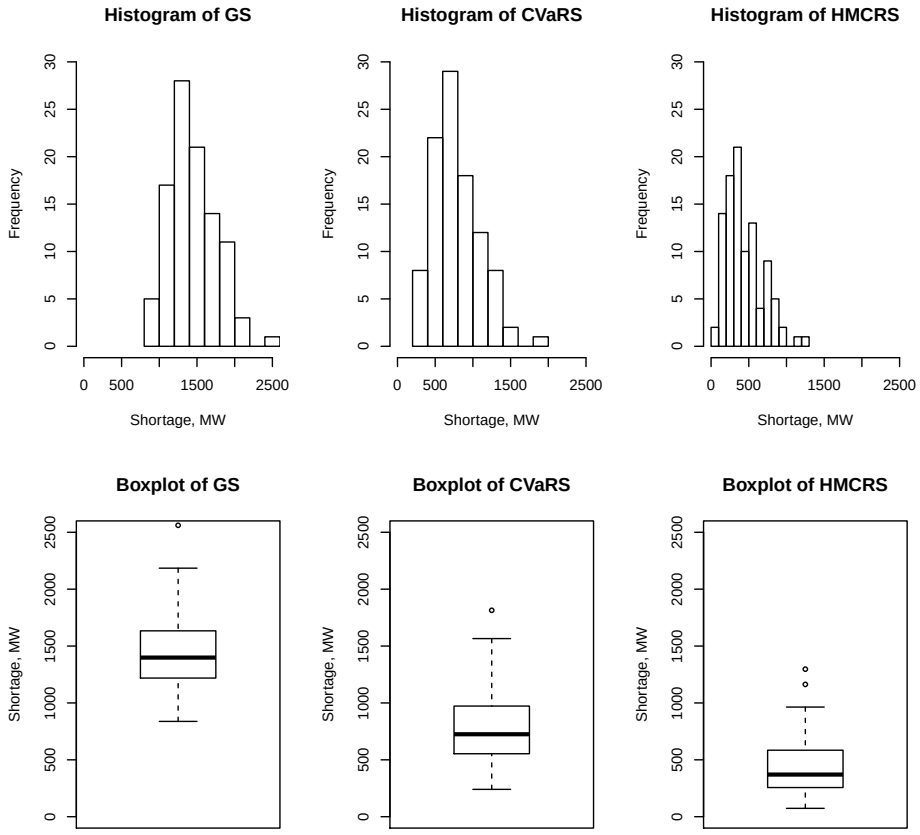

Figure 2.2: Shortages in out-of-sample extreme scenarios for model with 14 nodes and 8 locations

the amortized annual cost, (2) the mean cumulative shortage across the grid, (3) the number of shortage scenarios, i.e., the scenarios under which shortages occur, and (4) the mean number of demand nodes that experience shortages. The annual cost is computed as $\sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij}$; according to Section 2.4.3, the cumulative shortage at scenario k is defined as $\sum_i (D_{ik} - \sum_j \zeta_{ij} K_{jk})_+$, and thus mean shortage is $\mathsf{E}[\sum_i (D_{ik} - \sum_j \zeta_{ij} K_{jk})_+]$. As in the previous section, the out-of-sample analysis is conducted, in the sense that all the four criteria are evaluated on the set of 100 "extreme" out-of-sample scenarios determined as described above, e.g., the mean shortage and the mean number of demand nodes with shortages should be interpreted as a conditional expectations. The four criteria are thus computed for the case of 7 demand nodes and 6 candidate locations and 2000 scenarios, and values of γ varying from 0 to 0.95 with a step of 0.05. The results are presented in Figure 2.3, where the horizontal (constant) lines correspond to the GS model.

As expected, the annual costs of CVaRS and HMCRS models increase with γ. In contrast, mean shortage and mean number of shortage nodes in the CVaRS and HMCRS models decrease sharply with γ. Compared with the CVaRS model, the HMCRS model always performs better in terms of criteria (2)–(4), except for values γ around 0.5, but incurs higher annual costs. In conclusion, CVaRS and HMCRS models could be tuned to fit user's risk-averse preference so as to achieve better risk control of power shortages.

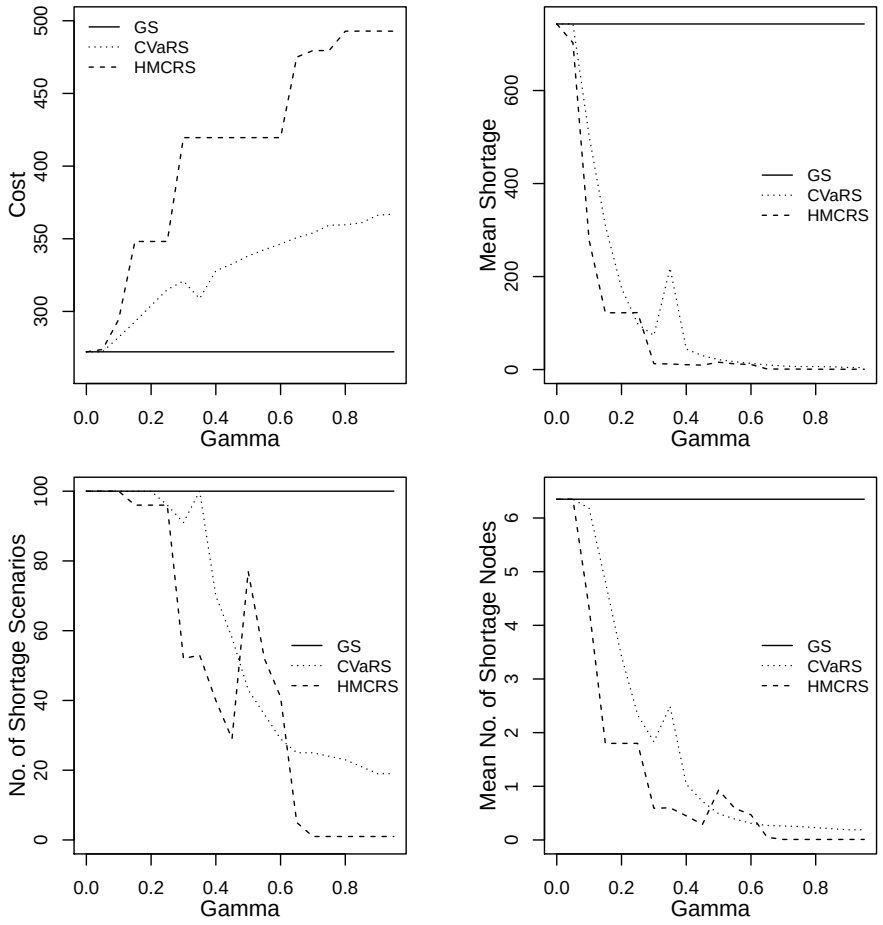

Figure 2.3: Out-of-sample performance of GS, CVaRS and HMCRS with regard to γ

2.5 Conclusions

In this chapter, we have considered three different stochastic optimization models for strategic wind farm location and operation: a risk-neutral model, a two models where risk preferences are represented by a linear risk measure (Conditional Value-at-Risk), and a nonlinear risk measure (Higher-Moment Coherent Risk measure). We proposed a branch-and-bound algorithm based on Benders decomposition technique to solve the resulting linear and p-order cone mixed-integer programming problems. The conducted numerical study demonstrates the efficiency of developed algorithms, and also indicates the risk-averse models allow for drastic reduction of wind power shortages, and can effectively be used in strategic location and planning problems.

CHAPTER 3
EFFECTS OF ENERGY STORAGE ON RISK-AVERSE STRATEGIC PLANNING OF RENEWABLE POWER GRIDS

3.1 Introduction

In this chapter, we also investigate how risk-averse planning of renewable energy grids can be effective in hedging of risks of power shortages due to stochastic variations in power generation and demand. We further develop the models of (2.4.1) by introducing two risk-averse stochastic models that take into consideration energy storage devices, which are supposed to compensate the possible future shortages to certain extent. The benefits of combining renewable sources with energy storage devices have been extensively studied. Paper [117] discusses the effects of wind power generation with compressed air energy storage on system operation, and [33] presents a dynamic programming algorithm to solve the problem of operating and scheduling a virtual power plant by coupling a wind farm and an energy storage facility. A stochastic dynamic programming model has been developed in [133], so as to co-optimize the energy storage which can be used for multiple applications. Similarly, a multi-stage stochastic program and a dynamic program are proposed in [79] to analyze different storage technologies and discuss their applications in power markets. An approximate dynamic programming method for energy storage problem with the supply of wind energy is presented in [55]. Approximate dynamic programming method has also been used in [100] to address the problem of optimizing an energy system with storage devices where the wind power supply, market electricity price, and electricity

demand are all stochastic.

Although time is an important factor in the formulation when taking energy storage capabilities into account, however, we here seek to mimic the storage and dispatch process by accumulating extra energy from "some" scenarios and consuming stored energy in "other" scenarios. In other words, we assume that the superfluous energy from some scenarios can be kept in storage devices (such as pumped hydro-electric storage and compressed air energy storage system), which would be utilized to supply certain amount of power demand in other scenarios to mitigate the shortage risk, under the condition that scenarios might represent a sequence of events.

The main aim of these extended models (models with energy storage) is to illustrate the impact of storage devices on operation of power grids and future power shortage possibility. Obviously, it is more reasonable to formulate the extended problems as stochastic dynamic programming problems to consider the storage and compensation in all time periods as a whole. The way we present the models here can be considered as one-period planning, which can be thought of as statically approximating the stochastic dynamic problem, and which serves our purpose to demonstrate the energy storage benefits.

The remainder of this chapter is organized as follows. In Section 3.2 we formulate two stochastic wind farm location models in presence of energy storage, with different degrees of risk aversion. Dataset generation, computational results, and corresponding solution analysis are presented in Section 3.3, while Section 3.4 summarizes and concludes this study.

3.2 Stochastic Wind Farm Location Models with Energy Storage

In this section, we introduce two stochastic models that integrate energy storage devices into renewable power grids and employ the Conditional Value-at-Risk (CVaR) measure and its nonlinear generalization Higher Moment Coherent Risk (HMCR), and therefore represent extensions of the risk-averse stochastic models described in Section 2.2.

First, it is worthwhile to comment on the difference between power and energy, which has an important influence on how we formulate the following models. Basically speaking, energy measures the ability to perform work, which can be stored and can also flow. The speed of energy flowing is defined as power, which is equivalent to the amount of energy consumed per unit of time, but cannot be conserved. In reality, we are concerned about the power shortage that would occur under the condition that the power supply is insufficient to meet the required power demand. In order to introduce power shortage risk measures and energy storage through constraints simultaneously, we suggest considering the power as the "instant" energy (an amount of energy consumed per unit of time) in the following models. As a result, extra "power" can be stored and re-delivered. From another perspective, even if the time multipliers are placed on both sides of energy storage constraints, they can always be simply cancelled out which would leave the constraints only in presence of power.

It is also assumed that the energy storage device is common to all wind farms, and is connected to these sites via HVDC lines. Using the property that HVDC transmission line allows for switching flow direction, the device can absorb surplus

energy from a given set of wind farms and dispatch the stored energy to final customers to compensate for power shortages, rerouting it via the transmission lines of the corresponding wind farms.

The notations of variables and parameters follow the definitions of Section 2.2, with additional variables and parameters defined as follows. Let κ_1 be the annualized cost (per MW) of installing energy storage device, and κ_2 be the storage efficiency of the device (i.e., if S MWh of the surplus energy are stored, then only $\kappa_2 S$ MWh would be recovered for energy supply). For instance, the energy efficiency of pumped hydroelectric storage system varies between 70% and 80% in practice, and compressed air energy storage system could achieve an energy efficiency between 60% and 90%.

Introducing additional variables s_{ijk} as the surplus power at candidate wind farm location j that generates power for demand node i under scenario k, ξ_{ik} as the amount of "stored" power used by demand node i under scenario k and $\overline{S}$ as the capacity of storage device, we can formulate the CVaR-based stochastic model with storage (CVaRSS) as:

$$\min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \kappa_1 \overline{S} + \gamma \left(\eta + \frac{1}{1-\alpha} \sum_k q_k U_k \right) \tag{3.1a}$$

$$\text{s. t.} \quad \sum_j x_j = h, \tag{3.1b}$$

$$y_{ij} \leq x_j, \qquad \forall i, j, \tag{3.1c}$$

$$\zeta_{ij} \leq M y_{ij}, \qquad \forall i, j, \tag{3.1d}$$

$$p_{ijk} + s_{ijk} \leq K_{jk} \zeta_{ij}, \qquad \forall i, j, k, \tag{3.1e}$$

$$\sum_k q_k \left(\sum_j p_{ijk} + \xi_{ik} \right) \geq \overline{D}_i, \qquad \forall i, \tag{3.1f}$$

$$\sum_i \xi_{ik} \leq \kappa_2 \sum_k q_k \sum_i \sum_j s_{ijk}, \qquad \forall k, \tag{3.1g}$$

$$\xi_{ik} \leq M \sum_j y_{ij}, \qquad \forall i, k, \tag{3.1h}$$

$$\sum_i \sum_j s_{ijk} \leq \overline{S}, \qquad \forall k, \tag{3.1i}$$

$$U_k \geq \sum_i \left(D_{ik} - \sum_j p_{ijk} - \xi_{ik} \right)_+ - \eta, \qquad \forall k, \tag{3.1j}$$

$$U_k \in \mathbb{R}_+, \tag{3.1k}$$

$$x_j, y_{ij} \in \{0, 1\}, \ \zeta_{ij} \in \mathbb{Z}_+, \ p_{ijk}, s_{ijk}, \xi_{ik}, \overline{S} \in \mathbb{R}_+. \tag{3.1l}$$

The objective function (3.1a) represents an utility function which consists of the cumulative annual cost and penalty for power shortage risk, to be minimized. Constraint (3.1b) stipulates that exactly h wind farms are to be located. Constraint (3.1c) states that a demand node i cannot be assigned to a wind farm j unless a wind farm

is located at site j. Constraint (3.1d) limits the number of wind turbines at site j that can be assigned for serving bus i. Constraint (3.1e) ensures that power supplied by candidate site j to demand node i under scenario k plus the corresponding possible extra power for "storage" is less than or equal to the total capacity of all wind turbines assigned at candidate site j to demand node i. Constraint (3.1f) ensures that the expected power supplied (direct supply from wind farms and indirect supply from storage devices) to demand node i from all sites does at least meet the expected demand at demand node i. Constraint (3.1g) requires that the total amount of "stored" power that is available under each scenario k should be less than or equal to the average amount of "stored" power over all the scenarios. This constraint also ensures that the total "consumed" power from storage device across all scenarios does not exceed the total "stored" power. Constraint (3.1h) states that if there were no connections between demand node i and any candidate wind farm site j, the "stored" power could not be be delivered to that demand node since we have assumed that the storage device can only "dispatch" power via the HVDC transmission lines between the wind farms and their respective customers. Constraint (3.1i) maintains that at any given scenario k, the total amount of stored power, as generated by all windfarms, cannot exceed the capacity of the storage device $\overline{S}$. Constraints (3.1j) relates to implementation of the Conditional Value at Risk measure. Lastly, constraints (3.1k)–(3.1l) prescribe the values that decision variables can take, where $\mathbb{Z}_+$ and $\mathbb{R}_+$ denote the sets of non-negative integer and real numbers, respectively. In what follows, the feasible set defined by constraints (3.1b)–(3.1i) and (3.1l) is denoted

by $\mathcal{P}'$.

Similarly, the HMCR-based stochastic model with storage (HMCRSS) is formulated as:

$$\min \quad \sum_j f_j x_j + \sum_i \sum_j c_j \zeta_{ij} + \sum_i \sum_j \lambda d_{ij} y_{ij} + \kappa_1 \overline{S} + \gamma\big(\eta + (1-\alpha)^{-1} U_0\big) \tag{3.2a}$$

$$\text{s.t.} \quad q_k^{-1/p} U_k \geq \sum_i \left(D_{ik} - \sum_j p_{ijk} - \xi_{ik} \right)_+ - \eta, \qquad \forall k, \tag{3.2b}$$

$$U_0 \geq \left(U_1^p + \ldots + U_K^p \right)^{1/p}, \tag{3.2c}$$

$$U_0, U_k \in \mathbb{R}+, \tag{3.2d}$$

$$x_j, y_{ij}, \zeta_{ij}, p_{ijk}, s_{ijk}, \xi_{ik}, \overline{S} \in \mathcal{P}'. \tag{3.2e}$$

The difference here is that we apply a higher moment risk measure to quantify the shortage risk in the formulation.

To reiterate, the proposed stochastic models CVaRSS and HMCRSS are "statically" approximating the dynamic storage and dispatch process, by ignoring time factor, i.e., the sequential nature of the "store-dispatch" operation. In a sense, such an approach to modeling the storage devices regards the operation of the grid in a "steady state", where scenarios may repeat over and over, with the prescribed probabilities. Constraints (3.1e)–(3.1g) determine how much energy can be stored in one scenario, the amount of energy can be extracted from storage device and how much energy can be used to supply in another scenario. From this point of view, if we consider different scenarios as a sequence of events, storing energy in one scenario and using it in another scenario amounts to conserving energy at some time point t and consuming it at another time point t'.

3.3 Computational Study

In this section, we analyze the out-of-sample performance of the proposed wind farm location models with storage (CVaRSS and HMCRSS), and draw a comparison with the models without storage (CVaRS and HMCRS) described in Section 2.2.

The "storage-based" and "storage-less" models are compared using out-of-sample data using the following procedure. First, optimal solutions for both "storage-based" and "storage-less" models are obtained for a given set of (in-sample) scenarios. The integer variables of these models represent long-term, or strategic planning decisions, which pertain to the locations of wind farms, number of wind turbines at each wind farm site, set of customers served by each wind farm. The continuous variables, such as power production, power dispatched to a given customer, power stored, etc., are scenario-dependent and represent "short-term" decisions that are made in response to or in anticipation of wind speed levels and power demands. Since the integer variables are scenario-independent, and continuous variables are scenario-dependent, they can also be regarded as first- and second-stage variables in the two-stage stochastic programming setting [17].

In view of this, we generate an out-of-sample scenario data for wind speed and power demand, and resolve the "storage-less" and "storage-based" models with the values of integer (first-stage) variables fixed to their optimal values obtained using the in-sample data, while the continuous (second-stage) variables are "optimized" in response to out-of-sample data. To formalize the outlined procedure, let $P(\mathbf{x}_1, \mathbf{x}_2, D_{\text{det}}, D_{\text{stoch}})$ denote a wind-farm location model, where $\mathbf{x}_1$, $\mathbf{x}_2$, D_{det}, and

D_{stoch} denote, respectively, the integer (first-stage) variables, continuous (second-stage) variables, deterministic data, and stochastic (wind speed, power demand) data. Then, the described procedure is implemented as follows.

1. Generate the in-sample data $D_{\text{det}}^{(\text{in})}, D_{\text{stoch}}^{(\text{in})}$.
2. Obtain an optimal solution $(\mathbf{x}_1^*, \mathbf{x}_2^*)$ of problem P with the in-sample data:

$$(\mathbf{x}_1^*, \mathbf{x}_2^*) \in \underset{\mathbf{x}_1, \mathbf{x}_2}{\arg\min}\, P\big(\mathbf{x}_1, \mathbf{x}_2, D_{\text{det}}^{(\text{in})}, D_{\text{stoch}}^{(\text{in})}\big).$$

3. Generate out-of-sample stochastic data $D_{\text{stoch}}^{(\text{out})}$.
4. Solve problem P with out-of-sample stochastic data and values of first-stage variables fixed to their optimal values $\mathbf{x}_1^*$ obtained in Step 1:

$$\mathbf{x}_2^{**} \in \underset{\mathbf{x}_1 = \mathbf{x}_1^*,\, \mathbf{x}_2}{\arg\min}\, P\big(\mathbf{x}_1, \mathbf{x}_2, D_{\text{det}}^{(\text{in})}, D_{\text{stoch}}^{(\text{out})}\big).$$

5. Evaluate the out-of-sample performance of the model P based on the solution obtained in Step 4.

Above, the "storage-less" models CVaRS and HMCRS and "storage-based" models CVaRSS and HMCRSS are used as model P.

We follow the data generation methods described in Section 2.4.1. In our numerical experiments, we use instances of wind farm location models with 7 demand nodes and 6 candidate locations.

3.3.1 Criteria for Out-of-sample Analysis

For models without storage (CVaRS and HMCRS), we adopt the same method to conduct the out-of-sample analysis which is elaborated in 2.4.3. In contrast, we

would have a set of possible shortage profiles due to the existence of storage device for models with storage (CVaRSS and HMCRSS). That means storage device could provide the option to reallocate "stored" power to compensate certain power shortage at some demand nodes under certain scenarios. In order to make their performance comparable to that of models without storage, we choose the expected shortage $\sum_k q_k \sum_i (D_{ik} - \sum_j p_{ijk} - \xi_{ik})_+$ as the objective function to be minimized over the feasible set.

The optimization model is defined as:

$$\min \quad \sum_k q_k U_k \tag{3.3a}$$

$$\text{s.t.} \quad p_{ijk} + s_{ijk} \leq K_{jk} \cdot \hat{\zeta}_{ij}, \qquad \forall i, j, k, \tag{3.3b}$$

$$\sum_i \xi_{ik} \leq \kappa_2 \sum_k q_k \sum_i \sum_j s_{ijk}, \qquad \forall k, \tag{3.3c}$$

$$\sum_i \sum_j s_{ijk} \leq \hat{\overline{S}}, \qquad \forall k, \tag{3.3d}$$

$$U_k \geq \sum_i \left(D_{ik} - \sum_j p_{ijk} - \xi_{ik} \right)_+, \qquad \forall k, \tag{3.3e}$$

$$p_{ijk}, s_{ijk}, \xi_{ik}, U_k \in \mathbb{R}_+. \tag{3.3f}$$

where $\hat{\zeta}_{ij}$ and $\hat{\overline{S}}$ are the values of variables ζ_{ij} and $\overline{S}$ obtained from optimal solutions of problems CVaRSS or HMCRSS, respectively. Constraint (3.3b) limits the magnitude of the "immediate" power supply p_{ijk} to customers and power "to be stored" s_{ijk} by optimal values of integer variable ζ_{ij} from the first stage solution. Constraint (3.3c) takes into account storage efficiency. Constraint (3.3d) ensures that the overall storage capacity is predetermined in the first stage.

The reason that model (3.3) omits constraints (3.1b)-(3.1d) is straightforward since these constraints only involve binary and integer variables and thus these constraints would definitely be satisfied with the corresponding values of discrete variables from the first stage. For the out-of-sample data, we cannot guarantee that the expected power supplied (partially determined by the values of binary and integer variables from the first stage) to each demand node i does at least meet the expected demand at the corresponding location, and consequently we leave out constraints (3.1f). Since we have constraints (3.1h) in models (3.1) and (3.2), which are necessary for the models to make sure that there is at least one connection between wind farm and demand node, we can deliver any available amount of "stored" power to any final customer via connection between wind farms and demand nodes. Therefore, we decide not to include (3.1h) in the optimization model (3.3).

3.3.2 Out-of-sample Solution Analysis

The out-of-sample data consists of one thousand randomly generated samples for parameters D_{ik} and K_{jk}, respectively, according to the methods described in Section 2.4.1.0.2. Accounting for different storage efficiency in practice, we set $\kappa_2 = 0.5$ and 0.75 for testing. In another respect, we also seek to investigate the influence of storage efficiency on renewable power grid planning. For the other parameter κ_1, we select it to assume values 0.04, 0.08, 0.12, 0.16, and 0.24, respectively, in order to investigate its impact on optimal solutions of CVaRRS and HMCRSS models.

We present the out-of-sample analysis results of stochastic wind farm location

models with storage (CVaRSS and HMCRSS) and without storage (CVaRS and HMCRS), in the following figures. Note that there are certain points with their X-axis values equal to Inf in the figures, which in fact represent the analysis results from models without storage (CVaRS and HMCRS). The reason for such an interpretation is that the value of $\kappa_1 = +\infty$ in models CVaRSS and HMCRSS corresponds to energy storage being extremely expensive for any practical use, so that the corresponding formulation is equivalent to the one without storage.

Figure 3.1 shows the out-of-sample test results of the CVaRS model ($\kappa_1 = +\infty$) and the CVaRSS model by setting $\kappa_1 = 0.04$, 0.08, 0.12, 0.16, and 0.24, and $\kappa_2 = 0.5$ and 0.75, respectively. Two series of data points representing results from setting $\kappa_2 = 0.5$ and 0.75 in fact demonstrate similar pattern, into which we would give certain insights first. From the top-left graph in Figure 3.1, we can observe that the fixed cost for configuring a power grid system increases with an increase in the annualized cost of installing storage device κ_1. The primary reason for this is that the required number of wind turbines would decrease if κ_1 goes down, which results from the fact that a smaller κ_1 always encourages a larger storage capacity (larger value of variable $\overline{S}$ as shown in middle-left picture of Figure 3.1), which in turn is able to provide more "capability" to compensate possible power shortage. These results indicate that one should take advantage of energy storage device when its annualized installation cost is low. The remaining three graphs demonstrate the performance of the CVaRSS model and CVaRS model on the out-of-sample dataset. The performance metrics include the mean power shortage, the number of shortage scenarios, and the mean number of

shortage nodes, which have been described at the beginning of this section. We can observe another trend: as the annualized installation cost κ_1 decreases, the obtained optimal configuration of renewable power grid performs better. That means if the annualized cost of installing storage device reduces, not only can one benefit from a lower fixed cost of configuring a renewable power grid, but also from bearing less risk exposure to power shortages.

Comparison between two series of data points reveals the influence of storage efficiency parameter κ_2 on the configuration of renewable power grid. In particular, when the storage efficiency is higher ($\kappa_2 = 0.75$), the optimal storage capacity offered by the CVaRSS model is higher, accompanied by a lower fixed cost and a smaller number of installed wind turbines. Additionally, higher storage efficiency would make the model less vulnerable to power shortage risk, according to the adopted performance metrics of mean power shortage, number of shortage scenarios, and mean number of shortage nodes. Another interesting point is that as κ_1 increases, the performance of CVaRSS solution approaches that of CVaRS solution. More specifically, when κ_1 is equal to 0.24, the CVaRSS model with $\kappa_2 = 0.75$ would have the same solution as the CVaRS model. Similarly, when κ_1 is greater than 0.16 and $\kappa_2 = 0.5$, the CVaRSS solution always suggests against incorporating storage devices into power grid.

Next, we present a comparison of the out-of-sample performance of the HMCRS model and the HMCRSS model by setting $\kappa_1 = 0.04$, 0.08, 0.12, 0.16, and 0.24 and $\kappa_2 = 0.5$ and 0.75 in Figure 3.2. Generally, the main tendencies are analogous to those in Figure 3.1, although the values of mean shortage oscillate as κ_1 increases.

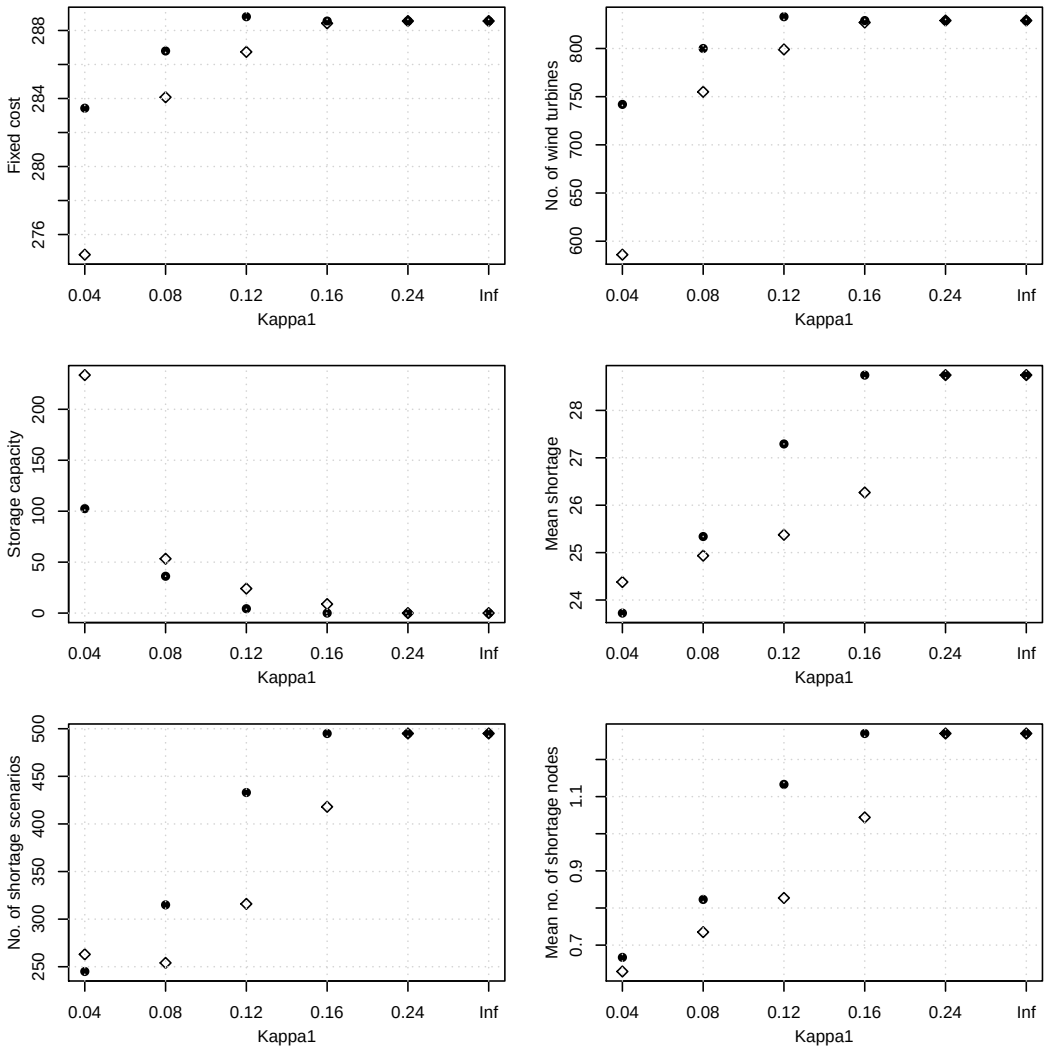

Figure 3.1: Comparison of models CVaRSS ($\kappa_1 = 0.04$, 0.08, 0.12, 0.16 and 0.24) and CVaRS ($\kappa_1 = \text{Inf}$), where black dots represent storage efficiency $\kappa_2 = 0.5$ and diamonds represent storage efficiency $\kappa_2 = 0.75$.

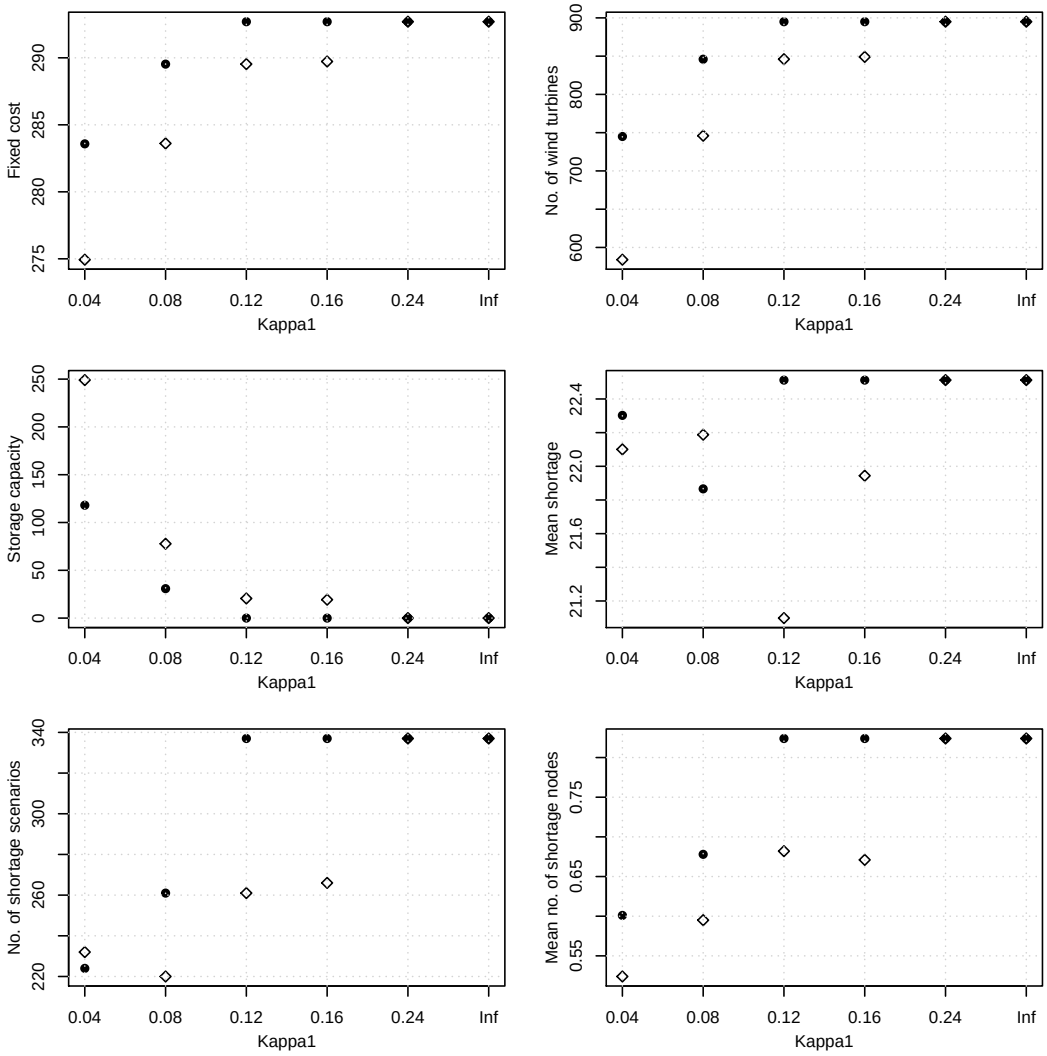

Figure 3.2: Comparison of models HMCRSS ($\kappa_1 = 0.04$, 0.08, 0.12, 0.16 and 0.24) and HMCRS ($\kappa_1 = \text{Inf}$), where black dots represent storage efficiency $\kappa_2 = 0.5$ and diamonds represent storage efficiency $\kappa_2 = 0.75$.

The fixed cost rises but the corresponding robustness of power grid decreases as the value of κ_1 goes up. With a lower storage efficiency ($\kappa_2 = 0.5$), the HMCRSS model suggests a larger number of wind turbines should be installed and a smaller storage capacity, when compared to the model with a higher storage efficiency ($\kappa_2 = 0.75$). In addition, higher storage efficiency can reduce the HMCRSS model's exposure to power shortage risk. If $\kappa_1 = 0.24$ and $\kappa_2 = 0.75$, the optimal solution from HMCRSS indicates that no storage devices should be installed. In contrast, the HMCRSS solution always suggests not incorporating any storage devices into power grid if $\kappa_2 = 0.5$ and κ_1 is greater than 0.12.

When we compare the results of HMCRSS model to that of CVaRSS model with the same κ_1 and κ_2 settings, we may observe that the HMCRSS model recommends a larger storage capacity and a higher number of installed wind turbines. As a result, the HMCRSS model provides less exposure to the risk of power shortage as compared with the CVaRSS model in terms of the employed performance metrics (mean power shortage, number of shortage scenarios, and mean number of shortage nodes).

In conclusion, we observe that the stochastic wind farm location models with storage (CVaRSS and HMCRSS) outperform the models without storage (CVaRS and HMCRS) in the out-of-sample test when the annualized cost of installing storage device (κ_1) is relatively low. Additionally, the CVaRSS and HMCRSS models tend to perform better as annualized cost of installing storage device κ_1 decreases and storage efficiency κ_2 increases. Lastly, the HMCRSS model generally suggests larger storage

capacity and larger number of installed wind turbines compared with the CVaRSS model, in exchange for less exposure to power shortage risk.

3.4 Conclusions

In this chapter, we have considered two different stochastic optimization models with energy storage for strategic wind farm location and operation: a model where risk preference is represented by a linear risk measure (Conditional Value-at-Risk), and the other adopts a nonlinear risk measure (Higher-Moment Coherent Risk measure). The storage and dispatch is mimicked by conserving extra energy from "some" scenarios and consuming stored energy in "other" scenarios. Compared with risk-averse stochastic models (CVaRS and HMCRS), the proposed models (CVaRSS and HMCRSS) demonstrate that employing energy storing devices in a renewable energy power grids could lead to significant benefits in reduction of power shortage risk, provided that the annualized installation cost energy storage devices is relatively low.

CHAPTER 4
CONVEX RELAXATIONS OF OPTIMAL POWER FLOW PROBLEM OVER RADIAL NETWORKS

In this chapter, we consider the optimal power flow (OPF) problem over networks with a tree topology, or radial networks. The reason that we are interested in radial networks is two-fold. In practice, it becomes increasingly crucial to solve OPF problem for distribution network because of the penetration of distributed generation and controllable demand. From the computational perspective, distribution networks generally have a tree topology, and as a result it would simplify the OPF problem as compared with general non-tree, or meshed networks.

The challenges associated with OPF models are due to the non-convexity of the corresponding optimization formulations, which introduce significant computational difficulties in determining an optimal solution of OPF problems even in small-scale instances. Thus, considerable attention has been given in literature to development of *convex relaxations* of the non-convex nonlinear OPF models. The most common methods of constructing such convex relaxations rely on reformulating the OPF model as a (non-convex) semidefinite programming (SDP) problem, and then relaxing some of the non-convex constraints in the resulting SDP formulations in order to obtain a convex SDP problem.

The remainder of this chapter is organized as follows. In Section 4.1, we first review the recent advances in the convex relaxation of the OPF problem mainly based on the work [74, 75, 40, 19, 136, 106, 106, 71, 44, 43, 136, 69], and then

show the exactness of convex relaxation on OPF problem over radial networks under certain assumptions. Proofs of exactness of a specific conic relaxation of OPF problem ([58, 59]) under over-satisfaction condition and that of its "extended" formulation are presented in Section 4.2.

4.1 Convex Relaxation of OPF Problem over Radial Networks

We can model a power network with an undirected graph $G(\mathcal{N}, \mathcal{E})$ consisting of a set of nodes (buses) and a set of edges (lines), where $\mathcal{N} := \{1, 2, \ldots, N\}$, and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. Since graph $G(\mathcal{N}, \mathcal{E})$ is undirected, then $(j, k) \in \mathcal{E}$ if and only if $(k, j) \in \mathcal{E}$. In what follows, we denote an undirected edge (j, k) by "$j \sim k$" and a cycle $c := \{n_1, \ldots, n_K\}$ is an ordered set of nodes such that $n_k \in \mathcal{N}$, $(n_K, n_1) \in \mathcal{E}$ and $(n_k, n_{k+1}) \in \mathcal{E}$ for $k = 1, \ldots, K-1$.

4.1.1 Optimal Power Flow Problem

The power flow equations for each node j can be described as (see Section 1.3.1.1 for details):

$$S_j = V_j \sum_{k:j\sim k} Y_{jk}^H V_k^H, \qquad \forall j \in \mathcal{N}, \tag{4.1}$$

where V_j and S_j are the complex voltage and the complex power respectively, and Y_{jk} represents elements in admittance matrix Y, which is defined by

$$Y_{jk} = \begin{cases} \sum_{i:i\sim j} y_{ji}, & \text{if } j = k, \\ -y_{jk}, & \text{if } j \neq k \text{ and } j \sim k, \\ 0, & \text{otherwise.} \end{cases}$$

In addition to constraints (4.1), the basic optimal power flow problem also

requires all voltage magnitudes to be box-constrained as:

$$\underline{V}_j \leq |V_j| \leq \overline{V}_j, \qquad \forall j \in \mathcal{N}, \tag{4.2}$$

where $\underline{V}_j$ and $\overline{V}_j$ represent bounds on voltage magnitudes at each bus j. Further, the power injections at each node j must satisfy:

$$\underline{S}_j \leq S_j \leq \overline{S}_j, \qquad \forall j \in \mathcal{N}, \tag{4.3}$$

where $\underline{S}_j$ and $\overline{S}_j$ are given lower and upper bounds.

If we substitute expressions (4.1) for variables S_j in constraints (4.3), we can formulate the OPF problem only in terms of the complex voltage variables, which can be presented as

$$\min \quad f(V) \tag{4.4a}$$

$$\text{s.t.} \quad \underline{V}_j \leq |V_j| \leq \overline{V}_j, \qquad \forall j \in \mathcal{N}, \tag{4.4b}$$

$$\underline{S}_j \leq V_j \sum_{k:j\sim k} Y_{jk}^H V_k^H \leq \overline{S}_j, \qquad \forall j \in \mathcal{N}, \tag{4.4c}$$

where V is the complex voltage vector and $f(V)$ is the cost function, which is most commonly specified as function (1.9) representing real power generating cost at all generator buses.

4.1.2 Convex Relaxations of the OPF problem

In the what follows, we will first present three different relaxations of the OPF problem based on semidefinite programming (SDP), chordal and second-order cone programming (SOCP) due to [74, 75], and then discuss their properties and exactness.

Before introducing the convex relaxations mentioned above, the definitions of partial matrix and its completion are given as follows. Suppose we have an undirected graph U with N vertices and M edges, then a partial matrix W_U is defined as:

$$W_U := \big\{[W_U]_{jj}, [W_U]_{jk}, [W_U]_{kj} \;\big|\; \text{nodes } j \text{ and edges } (j,k),(k,j) \text{ of } U\big\}.$$

with cardinality equal to $2M + N$. It is a matrix with corresponding elements equal to these complex numbers. Any fully specified $n \times n$ matrix W such that

$$[W]_{jj} = [W_U]_{jj}, \quad [W]_{jk} = [W_U]_{jk} \qquad \forall j, (j,k) \in U,$$

is a completion of partial matrix W_U. For partial matrices, the Hermitian is defined as $[W_U]_{jk} = [W_U]^H_{kj}$, for all $(j,k) \in U$. The positive semidefiniteness (PSD) property of W_U, denoted as $W_U \succeq 0$, is given that W_U is Hermitian and for any clique q of graph U the corresponding principal submatrix $W_U(q)$ is PSD. If the principal submatrices $W_U(q)$ all have rank-1 for any clique q on graph U, the rank of matrix W_U is equal to one.

A chordal graph is a simple graph in which all cycles of length four and greater have a chord, which is an edge connecting two vertices of the cycle but not belonging to the cycle. Thus, the chordal extension $c(U)$ of graph U can be interpreted as a chordal graph containing U, with the same vertices but having a superset of graph U's edge set. Correspondingly, we denote $W_{c(U)}$ as a chordal extension of the partial matrix W_U.

If we specify the undirected graph U as the power network $G = (\mathcal{N}, \mathcal{E})$, a

partial matrix W_G is then defined as

$$[W_G]_{jj} := |V_j|^2, \qquad \forall j \in \mathcal{N}, \tag{4.5a}$$

$$[W_G]_{jk} := V_j V_k^H =: [W_G]_{kj}^H, \qquad \forall (j,k) \in \mathcal{E}. \tag{4.5b}$$

Then, the constraints (4.4b) and (4.4c) can be reformulated as

$$\underline{V}_j^2 \leq [W_G]_{jj} \leq \overline{V}_j^2, \qquad \forall j \in \mathcal{N}, \tag{4.6a}$$

$$\underline{S}_j \leq \sum_{k:j\sim k} Y_{jk}^H [W_G]_{jk} \leq \overline{S}_j, \qquad \forall j \in \mathcal{N}. \tag{4.6b}$$

However, it is not always possible to recover a solution of model (4.4) from this unless the completion W of W_G is PSD with rank-1 since thus W would not only satisfy constraints (4.6), but also can be uniquely decomposed as $W = VV^H$. Therefore, additional conditions need to be added to ensure the completion of W_G is PSD with rank-1.

First, the cycle condition of a partial matrix W_G can be described as:

$$\sum_{(j,k)\in c} \angle [W_G]_{jk} = 0, \tag{4.7}$$

where c is any cycle in graph G and $\angle [W_G]_{jk}$ is the voltage phase differences of line (j,k). Additionally, we take the following conditions into account

$$W \succeq 0, \qquad \text{rank}(W) = 1, \tag{4.8}$$

$$W_{c(G)} \succeq 0, \qquad \text{rank}(W_{c(G)}) = 1, \tag{4.9}$$

$$W_G(j,k) \succeq 0, \qquad \text{rank}(W_G(j,k)) = 1, \qquad (j,k) \in \mathcal{E}. \tag{4.10}$$

where W is a $(N+1) \times (N+1)$ matrix, W_G and $W_{c(G)}$ are partial matrices.

Theorem 1 (Low, 2014, Theorem 2). *Fix a graph G on $N+1$ nodes and any chordal extension $c(G)$ of G. Assuming $W_{jj} > 0$, $[W_{c(G)}]_{jj} > 0$ and $[W_G]_{jj} > 0$, $j \in \mathcal{N}$, we have:*

1. *Given an $(N+1) \times (N+1)$ matrix W that satisfies (4.8), its submatrix $W_{c(G)}$ satisfies (4.9).*
2. *Given a partial matrix $W_{c(G)}$ that satisfies (4.9), its submatrix W_G satisfies (4.10) and cycle condition (4.7).*
3. *Given a partial matrix W_G that satisfies (4.10) and the cycle condition (4.7), there is a completion W of W_G that satisfies (4.8).*

Loosely speaking, Theorem 1 indicates the completion W of partial matrix W_G is PSD with rank-1 (condition (4.8)) if and only if it has a PSD rank-1 chordal extension $W_{c(G)}$ (condition (4.9)), if and only if the partial matrix W_G satisfies the cycle condition (4.7) and is 2×2 PSD rank-1 (condition (4.10)).

Based on the constraints (4.6) and conditions (4.7)-(4.10), the following sets are defined:

$$\mathcal{W} := \{W \mid W \text{ satisfies } (4.6), (4.8)\}, \tag{4.11}$$

$$\mathcal{W}_{c(G)} := \{W_{c(G)} \mid W_{c(G)} \text{ satisfies } (4.6), (4.9)\}, \tag{4.12}$$

$$\mathcal{W}_G := \{W_G \mid W_G \text{ satisfies } (4.6), (4.7), (4.10)\}. \tag{4.13}$$

Corollary 1 (Low, 2014, Corollary 3). *Given a partial matrix $W_{c(G)} \in \mathcal{W}_{c(G)}$ or $W_G \in \mathcal{W}_G$ there is a unique PSD rank-1 completion $W \in \mathcal{W}$.*

Theorem 2 (Low, 2014, Theorem 4). $\mathcal{V} \equiv \mathcal{W} \equiv \mathcal{W}_{c(G)} \equiv \mathcal{W}_G$.

According to Theorem 2, solving the OPF model (4.4) is equivalent to solving the following problem:

$$\begin{aligned} \min_{W} \quad & f(W_G) \\ \text{s.t.} \quad & W \text{ satisfies } \mathcal{W} \text{ or } \mathcal{W}_{c(G)} \text{ or } \mathcal{W}_G \end{aligned} \tag{4.14}$$

which is still difficult to solve because rank-1 conditions and the cycle condition (4.7) make the feasible regions $\mathcal{W}$, $\mathcal{W}_{c(G)}$ and $\mathcal{W}_G$ nonconvex. However, we can obtain SDP, chordal, and SOCP relaxations of OPF problem, respectively, by removing the latter restrictions, which leads to:

OPF-sdp:

$$\begin{aligned} \min_{W} \quad & f(W_G) \\ \text{s.t.} \quad & W \text{ satisfies (4.6)} \\ & W \succeq 0 \end{aligned} \tag{4.15}$$

OPF-ch:

$$\begin{aligned} \min_{W_{c(G)}} \quad & f(W_G) \\ \text{s.t.} \quad & W_{c(G)} \text{ satisfies (4.6)} \\ & W_{c(G)} \succeq 0 \end{aligned} \tag{4.16}$$

OPF-socp:

$$\begin{aligned} \min_{W_G} \quad & f(W_G) \\ \text{s.t.} \quad & W_G \text{ satisfies (4.6)} \\ & W_G(j,k) \succeq 0, (j,k) \in \mathcal{E}. \end{aligned} \tag{4.17}$$

Note that the above constraint $W_G(j,k) \succeq 0$ is in fact equivalent to $[W_G]_{jk} = [W_G]_{kj}^H$ and

$$[W_G]_{jj} > 0, \quad [W_G]_{kk} > 0, \quad [W_G]_{jj}[W_G]_{kk} \geq [W_G]_{jk}[W_G]_{jk},$$

which is a rotated second-order cone making the convex relaxation **OPF-socp** (4.17) a second-order cone programming (SOCP) problem. Relaxation **OPF-sdp** (4.15) is the most computationally intensive, especially for large sparse networks. In contrast, relaxation **OPF-socp** (4.17) is the simplest to solve, and relaxation **OPF-ch** (4.16) usually needs less computation compared to **OPF-sdp** (4.15). The relaxations **OPF-sdp** (4.15) and **OPF-ch** (4.16) are equally tight but strictly tighter than **OPF-socp** (4.17) for mesh networks. However, they are equally tight for radial networks.

To guarantee the exactness of relaxations **OPF-sdp** (4.15), **OPF-ch** (4.16) and **OPF-socp** (4.17) over radial networks, three types of sufficient conditions which have implications on power injections, voltage magnitudes and voltage angles respectively are given and discussed in [75].

4.2 A Specific Conic Relaxation of OPF Problem for Radial Networks

The second-order cone programming (SOCP) for radial networks was first proposed in [58]. Although the reformulation has been illustrated numerically, however, whether or when it would turn out to be exact has not been studied. In this section, we will first present the SOCP reformulation by Jabr [58], and then discuss its exactness given the over-satisfaction condition. Lastly, we propose an "extended" OPF by considering storage variables and prove the exactness of it over radial networks.

4.2.1 SOCP Relaxation and Exactness

Based on Section 1.3.1.1, the OPF problem in polar form can be expressed as follows, with a quadratic objective function minimizing the cost of real power

generation (the same as function (1.9) described in Section 1.3.3):

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) \tag{4.18a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} = \sum_{k=1}^{N} |V_j||V_k| \left[G_{jk} \cos(\theta_j - \theta_k) + B_{jk} \sin(\theta_j - \theta_k) \right], \tag{4.18b}$$

$$Q_{G_j} - Q_{L_j} = \sum_{k=1}^{N} |V_j||V_k| \left[G_{jk} \sin(\theta_j - \theta_k) - B_{jk} \cos(\theta_j - \theta_k) \right], \tag{4.18c}$$

$$\underline{P}_{G_j} \le P_{G_j} \le \overline{P}_{G_j}, \tag{4.18d}$$

$$\underline{Q}_{G_j} \le Q_{G_j} \le \overline{Q}_{G_j}, \tag{4.18e}$$

$$\underline{V}_j \le |V_j| \le \overline{V}_j. \tag{4.18f}$$

According to [58, 59], by introducing additional variables $u_j := |V_j|^2$, $R_{jk} := |V_j||V_k| \cos(\theta_j - \theta_k)$, $I_{jk} := |V_j||V_k| \sin(\theta_j - \theta_k)$, we can reformulate the OPF problem in polar form (4.18) equivalently as:

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) \tag{4.19a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} = u_j G_{jj} + \sum_{k=1, k\neq j}^{N} \left[G_{jk} R_{jk} + B_{jk} I_{jk} \right], \tag{4.19b}$$

$$Q_{G_j} - Q_{L_j} = -u_j B_{jj} + \sum_{k=1, k\neq j}^{N} \left[G_{jk} I_{jk} - B_{jk} R_{jk} \right], \tag{4.19c}$$

$$u_j u_k = R_{jk}^2 + I_{jk}^2, \tag{4.19d}$$

$$\theta_j - \theta_k = \tan^{-1} \frac{I_{jk}}{R_{jk}}, \tag{4.19e}$$

$$\underline{P}_{G_j} \le P_{G_j} \le \overline{P}_{G_j}, \tag{4.19f}$$

$$\underline{Q}_{G_j} \le Q_{G_j} \le \overline{Q}_{G_j}, \tag{4.19g}$$

$$\underline{V}_j^2 \le u_j \le \overline{V}_j^2. \tag{4.19h}$$

where (4.19e) corresponds to the cycle condition (4.7). Since we focus on studying OPF problem over radial networks, the cycle condition is vacuous and can be removed. Thus the exact conic reformulation for OPF problem with a tree topology can be expressed as:

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) \tag{4.20a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} = u_j G_{jj} + \sum_{k=1, k\neq j}^{N} \left[G_{jk} R_{jk} + B_{jk} I_{jk} \right], \tag{4.20b}$$

$$Q_{G_j} - Q_{L_j} = -u_j B_{jj} + \sum_{k=1, k\neq j}^{N} \left[G_{jk} I_{jk} - B_{jk} R_{jk} \right], \tag{4.20c}$$

$$u_j u_k = R_{jk}^2 + I_{jk}^2, \tag{4.20d}$$

$$\underline{P}_{G_j} \leq P_{G_j} \leq \overline{P}_{G_j}, \tag{4.20e}$$

$$\underline{Q}_{G_j} \leq Q_{G_j} \leq \overline{Q}_{G_j}, \tag{4.20f}$$

$$\underline{V}_j^2 \leq u_j \leq \overline{V}_j^2. \tag{4.20g}$$

However, the simplified problem (4.20) is still nonconvex due to constraints (4.20d). These constraints correspond to the requirements of $W_G(j,k) \succeq 0$ and $\text{rank}(W_G(j,k)) = 1$ in conditions (4.10). Relaxing constraints (4.20d) to $u_j u_k \geq R_{jk}^2 + I_{jk}^2$ is equivalent to removing the rank-1 requirement. By relaxing the equality in constraints (4.20d), the conic relaxation, which can be considered as one specific form of **OPF-socp** (4.17)

is as follows:

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) \tag{4.21a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} = u_j G_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} R_{jk} + B_{jk} I_{jk} \right], \tag{4.21b}$$

$$Q_{G_j} - Q_{L_j} = -u_j B_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} I_{jk} - B_{jk} R_{jk} \right], \tag{4.21c}$$

$$u_j u_k \geq R_{jk}^2 + I_{jk}^2, \tag{4.21d}$$

$$\underline{P}_{G_j} \leq P_{G_j} \leq \overline{P}_{G_j}, \tag{4.21e}$$

$$\underline{Q}_{G_j} \leq Q_{G_j} \leq \overline{Q}_{G_j}, \tag{4.21f}$$

$$\underline{V}_j^2 \leq u_j \leq \overline{V}_j^2. \tag{4.21g}$$

If the optimal solution of conic relaxation problem (4.21) makes constraint (4.21d) active (achieving equality), it will also be the optimal solution for the original OPF problem over radial networks. In [58], it seeks to incorporate a function of R_{jk} into objective in order to enforce constraints (4.21d) to be active but this has not been proved. That is, we can not exclude the possibility the optimal solution would make some of constraints (4.21d) not active. According to our computational study results, usually not all of constraints (4.21d) can achieve equality if we use the utility function consisting of generation cost and function of variables R_{jk} as the objective to be minimized. Note that the inactiveness of constraints (4.21d) corresponds to the condition that the optimal solution of **OPF-socp** (4.17) can not satisfy the rank-1 requirement. The load over-satisfaction condition in [19, 106, 105] provides a possibility to ensure the exactness of this conic relaxation, which is in fact a special form

of power injections conditions described in [75]. It suggests replacing "equality" with "greater than" in constraints (4.21b)-(4.21c) representing that over-delivery of power is permitted at nodes. According to [106, 105], a practical power network would still be in a normal condition even if load over-satisfaction is allowed since receiving extra power for free is unreasonable. More specifically, power may loss through transmission lines and the generation cost would increase as amount of power increases.

The conic relaxation of modified OPF problem with load over-satisfaction is formulated as:

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) \tag{4.22a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} \geq u_j G_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} R_{jk} + B_{jk} I_{jk} \right], \tag{4.22b}$$

$$Q_{G_j} - Q_{L_j} \geq -u_j B_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} I_{jk} - B_{jk} R_{jk} \right], \tag{4.22c}$$

$$u_j u_k \geq R_{jk}^2 + I_{jk}^2, \tag{4.22d}$$

$$\underline{P}_{G_j} \leq P_{G_j} \leq \overline{P}_{G_j}, \tag{4.22e}$$

$$\underline{Q}_{G_j} \leq Q_{G_j} \leq \overline{Q}_{G_j}, \tag{4.22f}$$

$$\underline{V}_j^2 \leq u_j \leq \overline{V}_j^2. \tag{4.22g}$$

Proposition 5. *The conic relaxation of modified OPF problem* (4.22) *has an optimal solution at which every inequality in constraints* (4.22d) *becomes an equality.*

Before the proof, some basic knowledge of admittance y_{jk} and admittance matrix entry Y_{jk} is given. First, the admittance $y_{jk} = g_{jk} + ib_{jk}$, where $g_{jk} = \frac{r_{jk}}{r_{jk}^2 + x_{jk}^2}$, $b_{jk} = -\frac{x_{jk}}{r_{jk}^2 + x_{jk}^2}$ and r_{jk}, x_{jk}, g_{jk}, b_{jk} are resistance, reactance, conductance and

susceptance respectively. According to $Y_{jk} = G_{jk} + iB_{jk} = -y_{jk}$, we can have $G_{jk} = -\frac{r_{jk}}{r_{jk}^2+x_{jk}^2}$ and $B_{jk} = \frac{x_{jk}}{r_{jk}^2+x_{jk}^2}$. Due to the nonnegativity of resistance r_{jk} and x_{jk} ([106, 105, 76, 7]), G_{jk} is nonpositive and B_{jk} is nonnegative in power flow equations.

Proof. Consider an arbitrary solution $(R^{opt}, I^{opt}, u^{opt}, P_G^{opt}, Q_G^{opt})$ of modified conic relaxation (4.22). Define $\hat{R}_{jk}^{opt} = \sqrt{u_j^{opt} u_k^{opt} - (I_{jk}^{opt})^2}$. Constraints (4.22d) become equality for $(\hat{R}^{opt}, I^{opt}, u^{opt})$. Furthermore, given a branch $(j,k) \in E$, one can write

$$G_{jk}\hat{R}_{jk}^{opt} + B_{jk}I_{jk}^{opt} \leq G_{jk}R_{jk}^{opt} + B_{jk}I_{jk}^{opt}, \tag{4.23a}$$

$$G_{jk}I_{jk}^{opt} - B_{jk}\hat{R}_{jk}^{opt} \leq G_{jk}I_{jk}^{opt} - B_{jk}R_{jk}^{opt}. \tag{4.23b}$$

Note that the above inequality is inferred from the fact that G_{jk} is nonpositive and B_{jk} is nonnegative. Thus,

$$u_j^{opt}G_{jj} + \sum_{k=1,k\neq j}^{N}\left[G_{jk}\hat{R}_{jk}^{opt} + B_{jk}I_{jk}^{opt}\right] \leq u_j^{opt}G_{jj} + \sum_{k=1,k\neq j}^{N}\left[G_{jk}R_{jk}^{opt} + B_{jk}I_{jk}^{opt}\right] \tag{4.24a}$$

$$\leq P_{G_j}^{opt} - P_{L_j},$$

$$-u_j^{opt}B_{jj} + \sum_{k=1,k\neq j}^{N}\left[G_{jk}I_{jk}^{opt} - B_{jk}\hat{R}_{jk}^{opt}\right] \leq -u_j^{opt}B_{jj} + \sum_{k=1,k\neq j}^{N}\left[G_{jk}I_{jk}^{opt} - B_{jk}R_{jk}^{opt}\right] \tag{4.24b}$$

$$\leq Q_{G_j}^{opt} - Q_{L_j}.$$

Therefore, $(\hat{R}^{opt}, I^{opt}, u^{opt}, P_G^{opt}, Q_G^{opt})$ is another solution of modified conic relaxation (4.22) at which every constraint (4.22d) becomes an equality.

Although we can guarantee that the constraints (4.22d) when solving relaxed model (4.22) are always active, however, the equality of constraints (4.22b) and (4.22c) are

not ensured. Therefore, the conic relaxation of modified OPF problem with load over-satisfaction cannot always be an exact relaxation of problem (4.18) although which is advocated intuitively in [106, 105] because receiving extra power for free is unreasonable. The following numerical example in Section 4.3 also verifies this.

4.2.2 SOCP Relaxation of "Extended" OPF and Exactness

By further defining storage variables P_{S_j} and Q_{S_j}, we can formulate an "extended" OPF problem which includes storage variables both in the objective function and power flow constraints. Since stored energy should be beneficial, the original objective function can be augmented by adding functions of P_{S_j}. The appearance of storage variables in power flow constraints can be thought of storing extra power, which in some extent correpsonds to the load over-satisfaction mentioned in Section 4.2.1. Thus, the "extended" OPF problem is formulated as:

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) - \sum_j f(P_{S_j}) \tag{4.25a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} - P_{S_j} = \sum_{k=1}^{N} |V_j||V_k| \left[G_{jk} \cos(\theta_j - \theta_k) + B_{jk} \sin(\theta_j - \theta_k) \right], \tag{4.25b}$$

$$Q_{G_j} - Q_{L_j} - Q_{S_j} = \sum_{k=1}^{N} |V_j||V_k| \left[G_{jk} \sin(\theta_j - \theta_k) - B_{jk} \cos(\theta_j - \theta_k) \right], \tag{4.25c}$$

$$\underline{P}_{G_j} \le P_{G_j} \le \overline{P}_{G_j}, \tag{4.25d}$$

$$\underline{Q}_{G_j} \le Q_{G_j} \le \overline{Q}_{G_j}, \tag{4.25e}$$

$$\underline{V}_j \le |V_j| \le \overline{V}_j. \tag{4.25f}$$

where $f(P_{S_j})$ is a concave function of variables P_{S_j}. As a result, the objective function is still convex.

The corresponding conic relaxation is expressed as:

$$\min \quad \sum_j (a_j + b_j P_{G_j} + c_j {P_{G_j}}^2) - \sum_j f(P_{S_j}) \tag{4.26a}$$

$$\text{s.t.} \quad P_{G_j} - P_{L_j} - P_{S_j} = u_j G_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} R_{jk} + B_{jk} I_{jk}\right], \tag{4.26b}$$

$$Q_{G_j} - Q_{L_j} - Q_{S_j} = -u_j B_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} I_{jk} - B_{jk} R_{jk}\right], \tag{4.26c}$$

$$u_j u_k \geq R_{jk}^2 + I_{jk}^2, \tag{4.26d}$$

$$\underline{P}_{G_j} \leq P_{G_j} \leq \overline{P}_{G_j}, \tag{4.26e}$$

$$\underline{Q}_{G_j} \leq Q_{G_j} \leq \overline{Q}_{G_j}, \tag{4.26f}$$

$$\underline{V}_j^2 \leq u_j \leq \overline{V}_j^2. \tag{4.26g}$$

Proposition 6. *The conic relaxation of "extended" OPF* (4.26) *always ends with an optimal solution at which inequality in constraints* (4.26d) *becomes an equality.*

Proof. Consider an arbitrary solution $(R^{opt}, I^{opt}, u^{opt}, P_G^{opt}, Q_G^{opt}, P_S^{opt}, Q_S^{opt})$ of conic relaxation of "extended" OPF (4.26). Define $\hat{R}_{jk}^{opt} = \sqrt{u_j^{opt} u_k^{opt} - (I_{jk}^{opt})^2}$, then $\hat{R}_{jk}^{opt} \geq R_{jk}^{opt}$. Constraints (4.26d) become equality for $(\hat{R}^{opt}, I^{opt}, u^{opt})$. Furthermore, we define

$$\hat{P}_{S_j}^{opt} = P_{G_j}^{opt} - P_{L_j}^{opt} - \left(u_j^{opt} G_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} \hat{R}_{jk}^{opt} + B_{jk} I_{jk}^{opt}\right]\right), \tag{4.27a}$$

$$\hat{Q}_{S_j}^{opt} = Q_{G_j}^{opt} - Q_{L_j}^{opt} - \left(-u_j^{opt} B_{jj} + \sum_{k=1,k\neq j}^{N} \left[G_{jk} I_{jk}^{opt} - B_{jk} \hat{R}_{jk}^{opt}\right]\right). \tag{4.27b}$$

We may find that $(\hat{R}^{opt}, I^{opt}, u^{opt}, P_G^{opt}, Q_G^{opt}, \hat{P}_S^{opt}, \hat{Q}_S^{opt})$ is a feasible solution of (4.26). If any constraint (4.26d) is inactive, there must exist some $\hat{R}_{jk}^{opt} > R_{jk}^{opt}$, and correspondingly $G_{jk}\hat{R}_{jk}^{opt} + B_{jk} I_{jk}^{opt} \leq G_{jk} R_{jk}^{opt} + B_{jk} I_{jk}^{opt}$. Thus, $\hat{P}_{S_j}^{opt} \geq P_{S_j}^{opt}$, which would

result in a smaller objective value since the values of P_G^{opt} do not change. That means, if some of constraints (4.26d) can not become active for $(R^{opt}, I^{opt}, u^{opt})$, then we can obtain a better solution $(\hat{R}^{opt}, I^{opt}, u^{opt}, P_G^{opt}, Q_G^{opt}, \hat{P}_S^{opt}, \hat{Q}_S^{opt})$ which contradicts the optimal solution assumption. Therefore, conic relaxation (4.26) always ends with an optimal solution at which inequality in (4.26d) becomes an equality.

Thus, we can always obtain the optimal solution of problem (4.25) by solving its conic relaxation (4.26).

4.3 Numerical Example

This section empirically demonstrates if conic relaxations (4.22) and (4.26) are exact based on a 6-bus electrical power network dataset from [87], as shown in Figure 4.1.

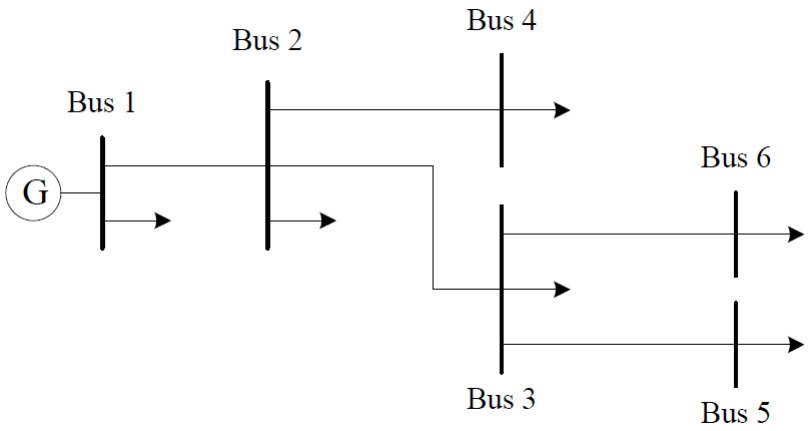

Figure 4.1: 6-bus radial network

The base voltage rating is 10 kV and the base power rating is 100 MVA. A detailed specification of parameters is given in Table 4.1 and Table 4.2. The upper and lower bounds of real and reactive power at generator bus 1 are specified as -0.08 and 0.08. The range of voltage at any bus is prescribed by 0.95 and 1.05. For simplicity, we specify parameters $a_j, c_j = 0$ and $b_j = 1$. Besides, we simplify the concave function $f(P_{S_j})$ as $g \cdot P_{S_j}$ where g represents a fixed number and here is chosen as 0.1.

Branch No.	Bus-i	Bus-j	r	x
1	2	1	0.0250	0.3182
2	3	2	0.2839	0.1931
3	4	2	0.2677	0.1765
4	5	3	0.0611	0.0597
5	6	3	0.2074	0.1440

Table 4.1: Branch data of 6-bus radial network

First, the conic relaxation of modified OPF problem with load over-satisfaction (4.22) has been solved with C++ and CPLEX 12.5 solver. We find that there are some constraints (4.22c) not active. In other words, we cannot claim the optimal solution of problem (4.22) will also be the optimal solution of problem (4.20), which in fact is an exact reformulation of problem (4.18) over distribution network. This result contradicts the intuitive expectation in [106, 105], which assert the modified relaxation problem allowing for load over-satisfaction condition would always achieve

Bus No.	MW	MVAR
1	0.0031	0.0023
2	0.0071	0.0062
3	0.0088	0.0030
4	0.0082	0.0061
5	0.0055	0.0029
6	0.0057	0.0060

Table 4.2: Load data of 6-bus radial network

equality for power flow constraints (like constraints (4.22b)–(4.22c)) because receiving extra power for free is unreasonable. Additionally, we further conduct some tests by augmenting the objective function of problem (4.22), with penalty on extra reactive power in the power network, which could be expressed as

$$\sum_j (a_j + b_j \cdot P_{G_j} + c_j \cdot {P_{G_j}}^2 + g' \cdot Q_{G_j}), \tag{4.28}$$

where g' is a fixed positive number representing penalty. Although the augmented formulation is obviously not any more a relaxation problem of original problem (4.18), we want to demonstrate here that by penalizing extra reactive power, the relaxed power flow constraints (allowing for load over-satisfaction condition) might become active. In this experiment, different values of parameter g', 1, 0.1, 0.01, 1e-03, 1e-04, 1e-05 and 1e-06 are used, and we find that if we solve the problem with $g' = 1$, 0.1 or 0.01, constraints (4.22b)–(4.22d) would become active. That means, certain penalty weights on reactive power would help the OPF problem under load over-satisfaction condition achieve equality for power flow constraints. However, this is

just an empirical observation without any rigorous proof.

The conic relaxed model (4.26) is also solved with C++ and CPLEX 12.5 solver, and the optimal objective value is 0.0385565. According to its optimal solution, we can recover the optimal values of certain variables in "extended" OPF (4.25) with

$$|V_j| = \sqrt{u_j}, \tag{4.29a}$$

$$\theta_j - \theta_k = \arccos\left(\frac{R_{jk}}{\sqrt{u_j}\sqrt{u_k}}\right), \tag{4.29b}$$

where $\theta_1 = 0$. The exactness of conic relaxation (4.26) is shown in Table 4.3.

Constraint (4.26b)	Constraint (4.26c)	Constraint (4.26d)
5.94072e-009	9.71445e-017	4.48394e-008
1.39161e-007	-4.12864e-016	-9.39032e-009
3.50254e-007	1.09678e-015	-8.33856e-009
7.53595e-008	6.07153e-018	-1.78298e-008
2.14065e-007	-3.46945e-016	-7.77525e-009
9.09494e-008	2.1684e-016	

Table 4.3: Tightness of constraints in the conic relaxation of "extended" OPF

In addition, we modify the constraints (4.26d) by changing "greater or equal to" sign to "equal to" sign, to obtain a nonconvex variant of problem (4.26). Note that this nonconvex variant would be an exact reformulation of (4.25) over radial networks. Then, we solve the nonconvex variant in AMPL using nonlinear solvers CONOPT, KNITRO, MINOS and SNOPT respectively. All the nonlinear solvers obtain the same

optimal objective value 0.0385571 (CONOPT, MINOS and SNOPT give exactly the same solutions), which is slightly larger than 0.0385565 by solving problem (4.26) with C++ and CPLEX 12.5 solver. This difference might be due to numerical errors. Moreover, we also solve the problem (4.26) with KNITRO nonlinear solver and find its optimal solution is almost the same as solving the nonconvex variant, with maximum relative difference of 1e-05.

4.4 Conclusions

In this chapter, the progresses in convex relaxation of OPF problem over radial network are first discussed, and a specific SOCP relaxation originally presented in [58] is elaborated. Then, we propose an "extended" OPF problem by considering storage variables and its corresponding SOCP relaxation. This conic relaxation of "extended" OPF problem over distribution network has been proved to be exact theoretically and empirically.

The purpose that we focus on SOCP relaxation of OPF problem or "extended" OPF problem rather than other types of relaxations such as SDP, is that if a mathematical optimization problem not only includes power flow equations but also involves integer variables, a SOCP relaxation of power flow constraints would result in a mixed-integer second-order cone programming (MISOCP) relaxed problem which will be much easier to handle, compared with a mixed-integer semidefinite programming relaxed problem with SDP relaxation of power flow constraints for example, since at each branch and bound node a SOCP problem will be solved more efficiently than

SDP problem.

In summary, the contribution of this chapter is three-fold. First, we augment the OPF problem over distribution networks by including storage variables to build a meaningful "extended" OPF problem, and prove its SOCP relaxation to be exact. That means, we can efficiently solve a nonlinear nonconvex problem via solving its SOCP relaxation and guarantee the global optimality of the solution. Second, this "extended" OPF problem can be considered as a "subproblem" when formulating certain mathematical optimization problems in a broader view which would take power flow equations and storage devices into considerations. A good example might be a mathematical optimization model which integrates renewable energy (such as wind, hydro and solar) into current power grid. Last but not least, we find the assertion in [106, 105] the modified relaxation problem (4.22) allowing for load over-satisfaction condition would always achieve equality for power flow constraints (4.22b)–(4.22c)) because receiving extra power for free is unreasonable, is not correct according to the empirical study of the 6-bus radial network above.

REFERENCES

[1] Chad Abbey and Géza Joós. A stochastic optimization approach to rating of energy storage systems in wind-diesel isolated grids. *Power Systems, IEEE Transactions on*, 24(1):418–426, 2009.

[2] MA Abido. Optimal power flow using particle swarm optimization. *International Journal of Electrical Power & Energy Systems*, 24(7):563–571, 2002.

[3] Hafzullah Aksoy, Z Fuat Toprak, Ali Aytek, and N Erdem Ünal. Stochastic generation of hourly mean wind speed data. *Renewable energy*, 29(14):2111–2131, 2004.

[4] Sacha Alberici, Sil Boeve, and Pieter Breevoort. Subsidies and costs of EU energy. Technical report, European Commission, 2014.

[5] O Alsac, J Bright, M Prais, and B Stott. Further developments in lp-based optimal power flow. *Power Systems, IEEE Transactions on*, 5(3):697–711, 1990.

[6] O Alsac and B Stott. Optimal load flow with steady-state security. *Power Apparatus and Systems, IEEE Transactions on*, (3):745–751, 1974.

[7] Göran Andersson. Modelling and analysis of electric power systems. *EEH-Power Systems Laboratory, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland*, 2004.

[8] Cristina L Archer and Mark Z Jacobson. Evaluation of global wind power. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D12), 2005.

[9] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[10] X Bai and H Wei. Semi-definite programming-based method for security-constrained unit commitment with operational and optimal power flow constraints. *Generation, Transmission & Distribution, IET*, 3(2):182–197, 2009.

[11] Xiaoqing Bai and Hua Wei. A semidefinite programming method with graph partitioning technique for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 33(7):1309–1314, 2011.

[12] Xiaoqing Bai, Hua Wei, Katsuki Fujisawa, and Yong Wang. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30(6):383–392, 2008.

[13] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2 of *MPS/SIAM Series on Optimization*. SIAM, 2001.

[14] Aharon Ben-Tal and Arkadi Nemirovski. On polyhedral approximations of the second-order cone. *Mathematics of Operations Research*, 26(2):193–205, 2001.

[15] Jacques F Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik*, 4(1):238–252, 1962.

[16] Arthur R Bergen and Vijay Vittal. *Power systems analysis*. Prentice Hall, second edition, 2000.

[17] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer, New York, 1997.

[18] Steven W Blume. *Electric power system basics for the nonelectrical professional*, volume 32. John Wiley & Sons, 2008.

[19] Subhonmesh Bose, Dennice F Gayme, Steven Low, and K Mani Chandy. Optimal power flow over tree networks. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1342–1348. IEEE, 2011.

[20] Rodney J Brown and William F Tinney. Digital solutions for large power networks. *Power apparatus and systems, part iii. transactions of the american institute of electrical engineers*, 76(3):347–351, 1957.

[21] Waqquas A Bukhsh, Andreas Grothey, Ken IM McKinnon, and Paul A Trodden. Local solutions of the optimal power flow problem. *Power Systems, IEEE Transactions on*, 28(4):4780–4788, 2013.

[22] RC Burchett, HH Happ, and DR Vierath. Quadratically convergent optimal power flow. *Power Apparatus and Systems, IEEE Transactions on*, (11):3267–3275, 1984.

[23] Daniel J Burke and MJ O'Malley. Optimal wind power location on transmission systems-a probabilistic load flow approach. In *Probabilistic Methods Applied to Power Systems, 2008. PMAPS'08. Proceedings of the 10th International Conference on*, pages 1–8. IEEE, 2008.

[24] Tony Burton, Nick Jenkins, David Sharpe, and Ervin Bossanyi. *Wind energy handbook*. John Wiley & Sons, 2011.

[25] Florin Capitanescu, Mevludin Glavic, Damien Ernst, and Louis Wehenkel. Interior-point based algorithms for the solution of optimal power flow problems. *Electric Power systems research*, 77(5):508–517, 2007.

[26] J Carpentier. Contribution to the economic dispatch problem. *Bulletin de la Societe Francoise des Electriciens*, 3(8):431–447, 1962.

[27] Gang Chen, Mark S Daskin, Zuo-Jun Max Shen, and Stanislav Uryasev. The α-reliable mean-excess regret model for stochastic facility location modeling. *Naval Research Logistics (NRL)*, 53(7):617–626, 2006.

[28] Po-Hung Chen and Hong-Chan Chang. Large-scale economic dispatch by genetic algorithm. *Power Systems, IEEE Transactions on*, 10(4):1919–1926, 1995.

[29] Shin-Der Chen and Jiann-Fuh Chen. A new algorithm based on the newton-raphson approach for real-time emission dispatch. *Electric power systems research*, 40(2):137–141, 1997.

[30] TS Chung and YZ Li. A hybrid ga approach for opf with consideration of facts devices. *Power Engineering Review, IEEE*, 20(8):54–57, 2000.

[31] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[32] Conserve Energy Future. Cost of wind energy. `http://www.conserve-energy-future.com/WindEnergyCost.php`, 2013. [Online; accessed 12-Nov-2013].

[33] Luís M Costa, Franck Bourry, Jérémie Juban, and George Kariniotakis. Management of energy storage coordinated with wind power under electricity market conditions. In *Probabilistic Methods Applied to Power Systems, 2008. PMAPS'08. Proceedings of the 10th International Conference on*, pages 1–8. IEEE, 2008.

[34] Mark S Daskin, Susan M Hesse, and Charles S Revelle. α-reliable p-minimax regret: A new model for strategic facility location modeling. *Location Science*, 5(4):227–246, 1997.

[35] Hermann W Dommel and William F Tinney. Optimal power flow solutions. *power apparatus and systems, IEEE transactions on*, (10):1866–1876, 1968.

[36] HW Dommel, WF Tinney, and WL Powell. Further developments in newton's method for power system applications. In *IEEE Winter Power Meeting, Conference Paper*, number 70, 1970.

[37] Orhan Ekren and Banu Y Ekren. Size optimization of a pv/wind hybrid energy conversion system with battery storage using simulated annealing. *Applied Energy*, 87(2):592–598, 2010.

[38] Mohamed E El-Hawary. *Electrical energy systems*. CRC Press, 2007.

[39] Albert M Erisman, KW Neves, MH Dwarakanath, et al. *Electric Power Problems: The Mathematical Challenge: Proceedings of a Conference, Seattle, Washington, March 18-20, 1980*, volume 10. SIAM, 1980.

[40] Masoud Farivar, Christopher R Clarke, Steven H Low, and K Mani Chandy. Inverter var control for distribution systems with renewables. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, pages 457–462. IEEE, 2011.

[41] Stephen Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal power flow: a bibliographic survey i. *Energy Systems*, 3(3):221–258, 2012.

[42] Stephen Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal power flow: a bibliographic survey ii. *Energy Systems*, 3(3):259–289, 2012.

[43] Lingwen Gan, Na Li, Ufuk Topcu, and Steven H Low. Exact convex relaxation of optimal power flow in radial networks. *arXiv preprint arXiv:1311.7170*, 2013.

[44] Lingwen Gan, Na Li, Ufuk Topcu, and Steven H Low. Optimal power flow in distribution networks. In *Proc. 52nd IEEE Conference on Decision and Control*, 2013.

[45] A Garcia, JL Torres, E Prieto, and A De Francisco. Fitting wind speed distributions: a case study. *Solar Energy*, 62(2):139–144, 1998.

[46] Robert Glass, Walt Beyeler, Kevin Stamber, Laura Glass, Randall LaViolette, Stephen Conrad, Nancy Brodsky, Theresa Brown, Andy Scholand, and Mark Ehlen. Simulation and analysis of cascading failure in critical infrastructure. Technical report, Sandia National Laboratories, 2005.

[47] AF Glimn and GW Stagg. Automatic calculation of load flows. *Power apparatus and systems, part iii. transactions of the american institute of electrical engineers*, 76(3):817–825, 1957.

[48] Global Wind Energy Council. Global wind report 2013. Technical report, Global Wind Energy Council, 2014.

[49] Ajit Gopalakrishnan, Arvind U Raghunathan, Daniel Nikovski, and Lorenz T Biegler. Global optimization of optimal power flow using a branch & bound algorithm. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pages 609–616. IEEE, 2012.

[50] Sergio Granville. Optimal reactive dispatch through interior point methods. *Power Systems, IEEE Transactions on*, 9(1):136–146, 1994.

[51] Leonard L Grigsby. *Power systems*. CRC press, 2007.

[52] Leonard L Grigsby. *Electric Power Generation, Transmission, and Distribution*. CRC press, 2012.

[53] Leonard L Grigsby. *Power Systems*. CRC Press, third edition, 2012.

[54] N Grudinin. Reactive power optimization using successive quadratic programming method. *Power Systems, IEEE Transactions on*, 13(4):1219–1225, 1998.

[55] Lauren Hannah and David B Dunson. Approximate dynamic programming for storage problems. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 337–344, 2011.

[56] M Huneault, FD Galiana, and Que St Bruno. A survey of the optimal power flow literature. *IEEE Transactions on Power Systems*, 6(2), 1991.

[57] RA Jabr. Exploiting sparsity in sdp relaxations of the opf problem. *Power Systems, IEEE Transactions on*, 27(2):1138–1139, 2012.

[58] Rabih A Jabr. Radial distribution load flow using conic programming. *Power Systems, IEEE Transactions on*, 21(3):1458–1459, 2006.

[59] Rabih A Jabr. A conic quadratic format for the load flow equations of meshed networks. *Power Systems, IEEE Transactions on*, 22(4):2285–2286, 2007.

[60] Rabih A Jabr. Optimal power flow using an extended conic quadratic formulation. *Power Systems, IEEE Transactions on*, 23(3):1000–1008, 2008.

[61] YA Katsigiannis and PS Georgilakis. Optimal sizing of small isolated hybrid power systems using tabu search. *Journal of Optoelectronics and Advanced Materials*, 10(5):1241, 2008.

[62] Panos Kouvelis and Gang Yu. *Robust discrete optimization and its applications*. Kluwer Academic Publishers, Dordrecht, 1997.

[63] P. Krokhmal, M. Zabarankin, and S. Uryasev. Modeling and optimization of risk. *Surveys in Operations Research and Management Science*, 16(2):49–66, 2011.

[64] PAVLO A Krokhmal. Higher moment coherent risk measures. *Quantitative Finance*, 7(4):373–387, 2007.

[65] Ludwig Kuznia, Bo Zeng, Grisselle Centeno, and Zhixin Miao. Stochastic optimization for power system configuration with renewable energy in remote areas. *Annals of Operations Research*, pages 1–22, 2011.

[66] Javad Lavaei. Zero duality gap for classical opf problem convexifies fundamental nonlinear power problems. In *American Control Conference (ACC), 2011*, pages 4566–4573. IEEE, 2011.

[67] Javad Lavaei and Steven H Low. Convexification of optimal power flow problem. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 223–232. IEEE, 2010.

[68] Javad Lavaei and Steven H Low. Zero duality gap in optimal power flow problem. *Power Systems, IEEE Transactions on*, 27(1):92–107, 2012.

[69] Javad Lavaei, David Tse, and Baosen Zhang. Geometry of power flows and optimization in distribution networks. *Power Systems, IEEE Transactions on*, 29(2):572–583, 2013.

[70] Bernard C Lesieutre, Daniel K Molzahn, Alexander R Borden, and Christopher L DeMarco. Examining the limits of the application of semidefinite programming to power flow problems. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1492–1499. IEEE, 2011.

[71] Na Li, Lijun Chen, and Steven H Low. Exact convex relaxation of opf for radial networks using branch flow model. In *SmartGridComm*, pages 7–12. Citeseer, 2012.

[72] Wenyuan Li et al. *Reliability assessment of electric power systems using Monte Carlo methods*. Springer Science & Business Media, 2013.

[73] Anthony Lopez, Billy Roberts, Donna Heimiller, Nate Blair, and Gian Porro. U.S. renewable energy technical potentials: a gis-based analysis. Technical report, National Renewable Energy Laboratory, 2012.

[74] Steven H Low. Convex relaxation of optimal power flow, part i: Formulations and equivalence. *arXiv preprint arXiv:1405.0766*, 2014.

[75] Steven H Low. Convex relaxation of optimal power flow, part ii: Exactness. *arXiv preprint arXiv:1405.0814*, 2014.

[76] James D McCalley. the power flow equations. Lecture Notes, 2006.

[77] Shengwei Mei, Xuemin Zhang, and Ming Cao. *Power grid complexity*. Springer, 2011.

[78] Wang Min and Liu Shengsong. A trust region interior point algorithm for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 27(4):293–300, 2005.

[79] Pedram Mokrian and Moff Stephen. A stochastic programming framework for the valuation of electricity storage. In *26th USAEE/IAEE North American Conference*, pages 24–27, 2006.

[80] James A Momoh. A generalized quadratic-based model for optimal power flow. In *Systems, Man and Cybernetics, 1989. Conference Proceedings., IEEE International Conference on*, pages 261–271. IEEE, 1989.

[81] James A Momoh. *Electric power system applications of optimization*. CRC Press, 2000.

[82] James A Momoh, ME El-Hawary, and Ramababu Adapa. A review of selected optimal power flow literature to 1993. part i: Nonlinear and quadratic programming approaches. *IEEE transactions on power systems*, 14(1):96–104, 1999.

[83] James A Momoh, ME El-Hawary, and Ramababu Adapa. A review of selected optimal power flow literature to 1993. part ii: Newton, linear programming and interior point methods. *IEEE Transactions on Power Systems*, 14(1):105–111, 1999.

[84] Alcir Monticelli, Ariovaldo Garcia, and OR Saavedra. Fast decoupled load flow: Hypothesis, derivations, and testing. *Power Systems, IEEE Transactions on*, 5(4):1425–1431, 1990.

[85] Y. Morenko, A. Vinel, Z. Yu, and P. Krokhmal. On p-cone linear discrimination. *European Journal of Operational Research*, 231:784–789, 2013.

[86] Yurii Nesterov, Arkadii Nemirovskii, and Yinyu Ye. *Interior-point polynomial algorithms in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

[87] Branimir Novoselnik. Convex relaxations of optimal power flow.

[88] KS Pandya and SK Joshi. A survey of optimal power flow methods. *Journal of Theoretical & Applied Information Technology*, 4(5), 2008.

[89] Jong-Bae Park, Ki-Song Lee, Joong-Rin Shin, and Kwang Y Lee. A particle swarm optimization for economic dispatch with nonsmooth cost functions. *Power Systems, IEEE Transactions on*, 20(1):34–42, 2005.

[90] Norris M Peterson and W Scott Meyer. Automatic adjustment of transformer and phase-shifter taps in the newton power flow. *Power Apparatus and Systems, IEEE Transactions on*, 90(1):103–108, 1971.

[91] Dzung T Phan. Lagrangian duality and branch-and-bound algorithms for optimal power flow. *Operations Research*, 60(2):275–285, 2012.

[92] Lynn Powell. *Power System Load Flow Analysis (Professional Engineering)*. McGraw-Hill Professional, 2004.

[93] Konrad Purchala, Leonardo Meeus, Daniel Van Dommelen, and Ronnie Belmans. Usefulness of dc power flow for active power flow analysis. In *Power Engineering Society General Meeting, 2005. IEEE*, pages 454–459. IEEE, 2005.

[94] Zhifeng Qiu, Geert Deconinck, and Ronnie Belmans. A literature survey of optimal power flow problems in the electricity market context. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–6. IEEE, 2009.

[95] R Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

[96] R Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, 2002.

[97] Fabien Roques, Céline Hiroux, and Marcelo Saguan. Optimal wind power deployment in europea portfolio approach. *Energy Policy*, 38(7):3245–3256, 2010.

[98] AM Sasson, F Viloria, and F Aboytes. Optimal load flow solution using the hessian matrix. *Power Apparatus and Systems, IEEE Transactions on*, (1):31–41, 1973.

[99] Marc Schwartz, Donna Heimiller, Steve Haymes, and Musial Walt. Assessment of offshore wind energy resources for the United States. Technical report, National Renewable Energy Laboratory, 2010.

[100] W Scott and Warren B Powell. Approximate dynamic programming for energy storage with new results on instrumental variables and projected bellman errors. *Submitted to Operations Research (Under Review)*, 2012.

[101] Warren R Scott and Warren B Powell. Approximate dynamic programming for energy storage with new results on instrumental variables and projected bellman errors. *Submitted to Operations Research (Under Review)*, 2012.

[102] Tomonobu Senjyu, Daisuke Hayashi, Atsushi Yona, Naomitsu Urasaki, and Toshihisa Funabashi. Optimal configuration of power generating systems in isolated island with renewable energy. *Renewable Energy*, 32(11):1917–1933, 2007.

[103] Anssi Seppälä. *Load research and load estimation in electricity distribution*. Technical Research Centre of Finland, 1996.

[104] Nidul Sinha, R Chakrabarti, and PK Chattopadhyay. Evolutionary programming techniques for economic load dispatch. *Evolutionary Computation, IEEE Transactions on*, 7(1):83–94, 2003.

[105] Somayeh Sojoudi and Javad Lavaei. Network topologies guaranteeing zero duality gap for optimal power flow problem. *submitted for publication*, 2011.

[106] Somayeh Sojoudi and Javad Lavaei. Physics of power networks makes hard optimization problems easy to solve. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8. IEEE, 2012.

[107] AndrACOa A Sousa, Geraldo L Torres, and Claudio A Canizares. Robust optimal power flow solution using trust region and interior-point methods. *Power Systems, IEEE Transactions on*, 26(2):487–499, 2011.

[108] Glenn W Stagg, Ahmed H El-Abiad, and AH El-Abiad. *Computer methods in power system analysis*, volume 9. McGraw-Hill New York, 1968.

[109] B Stott. Effective starting process for newton-raphson load flows. In *Proceedings of the Institution of Electrical Engineers*, volume 118, pages 983–987. IET, 1971.

[110] B Stott. Decoupled newton load flow. *Power Apparatus and Systems, IEEE Transactions on*, 91(5):1955–1959, 1972.

[111] B Stott and Of Alsacc. Fast decoupled load flow. *power apparatus and systems, ieee transactions on*, 93(3):859–869, 1974.

[112] B Marinho Stott. Linear programming for power-system network security applications. *Power*, 1979.

[113] Brian Stott. Review of load-flow calculation methods. *Proceedings of the IEEE*, 62(7):916–929, 1974.

[114] Brian Stott and Eric Hobson. Power system security control calculations using linear programming, part i. *Power Apparatus and Systems, IEEE Transactions on*, (5):1713–1720, 1978.

[115] Brian Stott, Jorge Jardim, and Ongun Alsacc. Dc power flow revisited. *Power Systems, IEEE Transactions on*, 24(3):1290–1300, 2009.

[116] David I Sun, Bruce Ashley, Brian Brewer, Art Hughes, and William F Tinney. Optimal power flow by newton approach. *power apparatus and systems, ieee transactions on*, (10):2864–2880, 1984.

[117] Derk J Swider. Compressed air energy storage in an electricity system with significant wind power generation. *Energy Conversion, IEEE Transactions on*, 22(1):95–102, 2007.

[118] William F Tinney and Clifford E Hart. Power flow solution by newton's method. *Power Apparatus and Systems, IEEE Transactions on*, 86(11):1449–1460, 1967.

[119] Xiaojiao Tong and Mugang Lin. Semismooth newton-type algorithms for solving optimal power flow problems. In *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, pages 1–7. IEEE, 2005.

[120] Geraldo Leite Torres and Victor Hugo Quintana. An interior-point method for nonlinear optimal power flow using voltage rectangular coordinates. *Power Systems, IEEE Transactions on*, 13(4):1211–1218, 1998.

[121] U.S. Department of Interior. `http://www.usbr.gov/pn/agrimet/webaghrread.html`, 2014. Accessed: 17-Dec-2014.

[122] U.S. Energy Information Administration. Electricity storage can take advantage of daily price variations. `http://www.eia.gov/todayinenergy/detail.cfm?id=6350`, 2012. Accessed: 22-Sep-2014.

[123] Robert AM Van Amerongen. A general-purpose version of the fast decoupled load flow. *Power Systems, IEEE Transactions on*, 4(2):760–770, 1989.

[124] Juan Pablo Vielma, Shabbir Ahmed, and George L Nemhauser. A lifted linear programming branch-and-bound algorithm for mixed-integer conic quadratic programs. *INFORMS Journal on Computing*, 20(3):438–450, 2008.

[125] A. Vinel and P. Krokhmal. Polyhedral approximations in p-order cone programming. *Optimization Methods & Software*, 29:1210–1237, 2014.

[126] Alexander Vinel and P Krokhmal. On polyhedral approximations in p-order cone programming. *Optimization Methods and Software*, 2014.

[127] Hua Wei, Hiroshi Sasaki, Junji Kubokawa, and R Yokoyama. An interior point nonlinear programming for optimal power flow problems with a novel data structure. *Power Systems, IEEE Transactions on*, 13(3):870–877, 1998.

[128] DW Wells. Method for economic secure loading of a power system. In *Proceedings of the Institution of Electrical Engineers*, volume 115, pages 1190–1194. IET, 1968.

[129] Jerry C Whitaker. *AC power systems handbook*. CRC Press, third edition, 2007.

[130] Allen J Wood and Bruce F Wollenberg. *Power generation, operation, and control*. John Wiley & Sons, 2012.

[131] World Wind Energy Association. 2014 half-year report. Technical report, World Wind Energy Association, 2014.

[132] Yu-Chi Wu, Atif S Debs, and Roy E Marsten. A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows. *Power Systems, IEEE Transactions on*, 9(2):876–883, 1994.

[133] Xiaomin Xi, Ramteen Sioshansi, and Vincenzo Marano. A stochastic dynamic programming model for co-optimization of distributed energy storage. *Energy Systems*, 5(3):475–505, 2014.

[134] Hirotaka Yoshida, Kenichi Kawata, Yoshikazu Fukuyama, Shinichi Takayama, and Yosuke Nakanishi. A particle swarm optimization for reactive power and voltage control considering voltage security assessment. *Power Systems, IEEE Transactions on*, 15(4):1232–1239, 2000.

[135] Jason Yuryevich and Kit Po Wong. Evolutionary programming based optimal power flow algorithm. *Power Systems, IEEE Transactions on*, 14(4):1245–1250, 1999.

[136] Baosen Zhang and David Tse. Geometry of injection regions of power networks. *Power Systems, IEEE Transactions on*, 28(2):788–797, 2013.

[137] Jizhong Zhu. *Optimization of power system operation*, volume 49. John Wiley & Sons, 2009.

[138] Jizhong Zhu and James A Momoh. Multi-area power systems economic dispatch using nonlinear convex network flow programming. *Electric Power Systems Research*, 59(1):13–20, 2001.