

Copyright
by
Lin Xu
2009

The Dissertation Committee for Lin Xu
certifies that this is the approved version of the following dissertation:

**Analyzing Strategic Behaviors in Electricity Markets
Via Transmission-constrained Residual Demand**

Committee:

Ross Baldick, Supervisor

Aristotle Arapostathis

Constantine Caramanis

David Morton

Steven Puller

Jeffrey McDonald

**Analyzing Strategic Behaviors in Electricity Markets
Via Transmission-constrained Residual Demand**

by

Lin Xu, B.S., M.S.E.

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2009

Dedicated to my parents, who have not had the opportunities to receive the
highest education, but have been encouraging his son to do so.

Acknowledgments

I would like to thank the many people who made this dissertation possible.

It is difficult to overstate my gratitude to my adviser, Dr. Ross Baldick, for his guidance during my research and study at the University of Texas at Austin. I have been working with Dr. Baldick for 6 years, and it has been such a pleasant and valuable period of time in my life. I always feel lucky that I could have such a knowledgeable and nice adviser.

I am grateful to my manager at the California Independent System Operator, Dr. Jeffrey McDonald, who has been very supportive for me to pursue the PhD degree. He shared with me his own PhD experiences, helped me develop project plans, and gave me suggestions about research directions. Without his help, the process of accomplishing the PhD degree could not have been so smooth.

I wish to thank all other committee members: Dr. David Morton, Dr. Aristotle Arapostathis, Dr. Steven Puller, and Dr. Constantine Caramanis. I have been interacting with them in many different ways, such as taking courses, working on projects, discussing research topics, and receiving constructive advices. They have helped and inspired me so much during my doctoral study.

I am also indebted to many researchers and colleagues, who have shared their thoughts with me and provided insightful comments to my research. I am especially grateful to Dr. Julián Barquín at Pontificia Comillas University, Dr. Tong Wu at Pacific Gas and Electric Company, and Dr. Andrew Philpott at the University of Auckland.

Lastly, my deepest gratitude goes to my family. My parents have been always staying behind me, supporting me with their care, and pushing me forward with their love. I must apologize to my parents for having them waiting so long to see his son become “Dr. Xu”. I also owe a lot to my wife Guangjuan. I would write another dissertation to express my love for you. Your patience, love and encouragement have upheld me, particularly in those many days in which I spent more time on my dissertation than with you. Now it is your turn to get the PhD degree, and I promise I will support you the same way you have supported me.

Analyzing Strategic Behaviors in Electricity Markets Via Transmission-constrained Residual Demand

Publication No. _____

Lin Xu, Ph.D.

The University of Texas at Austin, 2009

Supervisor: Ross Baldick

This dissertation studies how to characterize strategic behaviors in electricity markets from a transmission-constrained residual demand perspective. This dissertation generalizes the residual demand concept, widely used by economists in general markets, to electricity markets, which are constrained by transmission networks. The transmission-constrained residual demand is characterized by a sensitivity analysis of the optimal power flow program, which is the electricity market clearing engine. Methods are proposed to optimize a generator or generation firm's profit utilizing the residual demand sensitivity information, which has several advantages over existing methods. The transmission-constrained residual demand concept and the methods are helpful for market participants to develop bidding strategies and for market monitors to analyze market power in electricity markets.

Table of Contents

Acknowledgments	v
Abstract	vii
List of Tables	xi
List of Figures	xii
Chapter 1. Introduction	1
1.1 Electric Power Industry Restructuring: From Utilities to Electricity Markets	2
1.2 Typical electricity market elements	4
1.3 Analyzing Strategic Behaviors in Electricity Markets via Residual Demand	10
Chapter 2. Characterizing Strategic Behaviors in Electricity Markets Without Transmission Constraints	19
2.1 Introduction	19
2.2 The SFE Model and the Equilibrium Continuum	24
2.2.1 SFE and Best response Mapping	24
2.2.2 SFE Continuum	28
2.3 Stability Analysis of SFE	31
2.3.1 Piecewise polynomial perturbations	31
2.3.2 Stability Analysis	34
2.4 Computational Example	36
2.4.1 Example Setup	37
2.4.2 Regression Model Setup	38
2.4.3 Stability Analysis Results	39
2.4.4 Practical Implication	42
2.5 Conclusion	44

Chapter 3. Characterizing Strategic Behaviors via Transmission-Constrained Residual Demand Derivative	46
3.1 Residual Demand and Its Derivative	47
3.1.1 Non-binding slack bus generation capacity constraint .	51
3.1.2 Binding slack bus generation capacity constraint . . .	53
3.1.3 Sensitivity analysis	54
3.1.4 Weighted least squares regression interpretation	58
3.1.5 Non-differentiable case	65
3.2 TCRDD Examples	66
3.2.1 Example 1: intuitive 2-bus case	66
3.2.2 Example 2: numerical 4-bus case	67
3.2.3 Example 3: numerical 3-bus case with perfectly elastic supply	70
3.3 Improving TCRDD Calculation	72
3.3.1 Practical issues with the TCRDD calculation to be resolved	73
3.3.2 Generalizing the TCRDD Calculation	74
3.4 Maximize A Generator's Profit Using the TCRDD	83
3.5 Computational Example	95
3.6 Conclusion	96
 Chapter 4. Characterizing Strategic Behaviors Via Transmission-constrained Residual Demand Jacobian Matrix	 99
4.1 Calculating the TCRDJ	100
4.2 Properties of the TCRDJ	104
4.3 Handling binding quantity offers and binding price offers . . .	106
4.4 Maximizing A Generation Firm's Profit	109
4.4.1 Bundle idea	110
4.4.2 Algorithm	113
4.5 Computational Example	116
4.5.1 Firm strategy vs single generator strategy	117
4.5.2 Different starting points	118
4.5.3 Performance Test	123
4.6 Conclusion	124

Chapter 5. Conclusion	126
5.1 Summary	126
5.2 Future research	129
Appendices	134
Appendix A. Ordinary Least Squares Problem and Weighted Least Squares Problem	135
Bibliography	139
Vita	147

List of Tables

1.1	Typical electricity market elements	5
2.1	Asymptotically stable SFEs under polynomial function (1-piece) deviations	40
2.2	Asymptotically stable SFEs under polynomial function (3-piece) deviations	40
2.3	Spectral radius in the base scenario	41
2.4	Asymptotically stable SFEs under inertia	44
3.1	Results comparison	68
3.2	Determine Maximizer Of $\hat{\Pi}(q; q_{lo}; q_{hi})$	91
4.1	Generator data	116
4.2	Optimization trajectory starting from (200,200)	118
4.3	Optimization solutions: firm vs single generator	118
4.4	Optimization trajectory starting from (300, 500)	120
4.5	Optimization trajectory starting from (450,250)	120
4.6	Optimization trajectory starting from (450,550)	120
4.7	Performance test	124

List of Figures

1.1	Offers and bids	7
1.2	Profit maximization given residual demand curve	14
2.1	A continuum of symmetric SFEs	30
2.2	Stable SFEs	42
3.1	4-bus system	67
3.2	3-bus example with perfectly elastic supply	71
3.3	Estimating the derivative for step offer function	75
3.4	Maximizing generation profit based on TCRDD	84
3.5	4-bus system	84
3.6	Profit function and residual demand curve for generator 1 . . .	86
3.7	Profit function and residual demand curve for generator 2 . . .	86
3.8	Determine Maximizer Of $\hat{\Pi}(q; q_{lo}; q_{hi})$	92
3.9	Profit function and residual demand curve for generator 5 . . .	97
3.10	Bisection loop iterations	97
4.1	Optimization trajectory starting from (200, 200)	121
4.2	Optimization trajectory starting from (300, 500)	121
4.3	Optimization trajectory starting from (450, 250)	122
4.4	Optimization trajectory starting from (450,550)	122

Chapter 1

Introduction

Restructuring has been a worldwide trend in the electric power industry in the last two decades. Starting in Chile and UK in the 1980s, electric industry restructuring has built electricity markets almost all over the world [39]. For example, there are electricity markets in:¹

- Europe: United Kingdom, Germany, Spain, Nordic countries (Denmark, Finland, Norway and Sweden)
- North America: ISO-NE (New England), PJM (originally Pennsylvania, New Jersey and Maryland, since expanded to other states), CAISO (California), ERCOT (Texas), Midwest ISO, Alberta (Canada)
- South America: Chile, Brazil, Argentina, Colombia
- Asia: Japan

In this chapter, we will briefly go over the backgrounds of this significant restructuring in the electric power industry, introduce key elements in major electricity markets in the US, and review electricity market analysis methods.

¹This is *not* a comprehensive list of the electricity markets in the world.

1.1 Electric Power Industry Restructuring: From Utilities to Electricity Markets

The content of this section mainly comes from [41]. Historically, the power system is managed by a traditional vertically integrated utility, which is in charge of electricity generation, transmission and distribution. Utilities were understood to be natural monopolies due to economies of scale, and must be regulated since monopolies generally do not have incentives to increase efficiency, reduce costs, or improve quality of service [41].

On the other hand, new developments in transmission technology enable power to be efficiently transferred over long distances. This means generators located at different physical locations, far or near from the load, could possibly compete to serve the load. In addition, more standardized reliability criteria are imposed to regulate utilities' operation. This also makes electricity more of a standardized product so that it can be priced through competitive markets instead of regulation.

All these new developments in technology and regulation eventually help to introduce competition in generation, because competitive markets can potentially increase efficiency, reduce costs, and improve quality of service [41].

Electric industry restructuring started with separating out generation from the traditional utilities, and introducing wholesale competition. Regulated utilities still own the transmission and distribution networks, but the transmission networks are operated by an Independent System Operator (ISO), which is under regulation to provide "open access" to all market participants.

Newly developed electricity markets have included more ingredients than generation competition, but generation competition is still the core of all electricity market design.

With more than two decades of experiences up to now, both successful and frustrating, there is a consensus that economic theories and engineering laws should be integrated into electricity market design, and a successful electricity market should achieve a good balance between economy and engineering characteristics.

An electricity market cannot operate independently of the power system. The quality of electricity energy, mainly in terms of reliability, is determined by the power system standards. Without the power system standards for electricity reliability, electricity energy would not even be considered as a commodity. The power system is the basis the electricity market is built upon. The established ways power engineers maintain system reliability greatly affect electricity market design. In this sense, an electricity market has the dual nature of being a market, and being an engineering system. In other words, it is a market operated within engineering boundaries. For example, if transmission capacity is not enough to accommodate the most economic generators to compete at certain locations in the system, electricity markets will work out a less economic way to serve the load within the transmission limitations.

Generally speaking, the relative importance of the two characteristics depend on the time frame. The closer to real time system operation, the more important the engineering nature; the longer time frame, the more prominent

the market nature. This is because close-to-real-time engineering boundaries must be strictly respected in order to avoid compromising power system reliability, while long-term power transactions are viewed more as market behaviors, and rely less on engineering practices.

1.2 Typical electricity market elements

Although the objectives for restructuring the electric power industry are similar, that are to increase efficiency, reduce costs, and improve quality of service [41], the restructuring has differed a lot in scope and details in their early stage. Nowadays, with more and more experiences accumulated, the various electricity market designs eventually show a trend toward uniformity. The Federal Energy Regulatory Commission (FERC) even promoted a “Standard Market Design” (SMD) for electricity markets in North America based on economic theories, engineering laws, and experiences, to offer a streamlined best practice. After that, electricity markets in the US have tended to be designed to be compatible with SMD. For example, the CAISO nodal market, and the proposed ERCOT nodal market are both compatible with SMD, and thus they share a lot of similarities in design.

In this dissertation, we will consider the common design elements of the major power markets in the US, referring to them collectively as a “typical” market compatible with FERC’s SMD. All the studies covered in the dissertation focus on a typical market, so the results could be applicable and useful for most of the existing electricity markets directly or with minor modifications.

Element	Time Frame	centralized
bilateral contract	years to weeks prior	no
congestion revenue right	a year to a month prior	partially yes
day-ahead market	a day prior	yes
real-time market	an hour to 10 minutes prior	yes

Table 1.1: Typical electricity market elements

This chapter briefly introduces the some of the most important typical electricity market elements that are energy related. There are other key elements in electricity markets, such as ancillary services, which are not directly related to energy, and they will not be covered here.

We will start the introduction with sorting the key electricity market elements in chronological order as listed in Tab. 1.1 from top to bottom. A market participant can arrange transactions going through each element from top to bottom, and finally the electric energy is physically sold to or purchased from the electric power grid depending on it is a generator or load.

First, we introduce the day-ahead market and real-time market, because they are where the major restructuring took place, and are also the most important subjects of this study.

The real-time market is an offer based centralized auction market operated by the ISO. As its name suggests, it operates very close to the physical trade time, typically one hour to 10 minutes prior. Generators can submit offers into the real-time market. A generator offer is a price function of quantity specifying at each output level what is the minimal acceptable price. A

typical generator offer is illustrated in Fig. 1.1. The ISO uses the offers as inputs to an optimal power flow (OPF) program clear the market. The OPF is an optimization problem minimizing total generation cost to meet the load forecast in the trade interval, subject to transmission constraints. The trade interval is typically 5-minute to 15-minute granularity. The cleared offers will get a Locational Marginal Price (LMP) that is the incremental cost to serve one extra unit of load at the same location. The detailed formulation and pricing mechanism will be covered in detail in chapter 3 and 4. The offers cleared in the real-time market are financially binding.

The day-ahead market is similar to the real-time market, except that it is a forward market that clears typically a day before the trade day. Not only generators but also loads can bid into the day-ahead market. A typical demand bid is also illustrated in Fig. 1.1. The day-ahead market clearing mechanism is very similar to the real-time market except that instead of meeting the load forecast, the ISO tries to clear as much as possible balanced generation and demand based on the their offers and bids. From optimization point of view, it is simply treating the demand bids as negative generation offers, and setting the load forecast at zero. Because day-ahead market and real-time market share the same clearing engine, the OPF, the studies in this dissertation will be applicable to both of them. The offers and bids that are cleared in the day-ahead market are also financially binding.

The real-time market can be viewed as a fine tuning market for the day-ahead market. It provides the opportunity for market participants to make

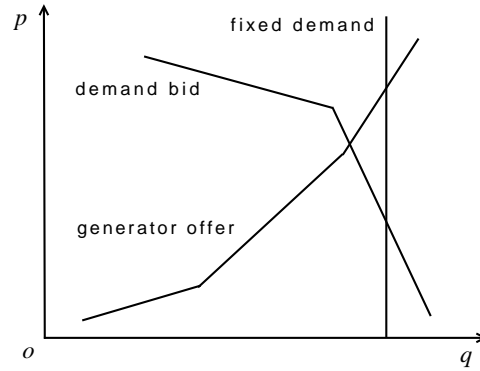


Figure 1.1: Offers and bids

changes to the day-ahead cleared offers. A generator cleared 100 MW in the day-ahead at \$30 /MWh, and an additional 10 MW in the real-time market at \$40 /MWh, will be paid $100 \cdot 30 + 10 \cdot 40 = 140$ \$/h. This is called the two-settlement system.

In a longer time frame, market participants can sign bilateral contracts with other participants several years before the physical delivery time. The bilateral contracts represent the counterparties' willingness to trade electric energy, and they are purely individual decisions of the counterparties without the ISO's involvement. The bilateral contracts are more of a market nature, but the transmission capabilities need to be considered to avoid infeasible transactions at physical delivery time. In other words, the transactions should not overload lines or line groups under normal operation condition and certain contingency conditions. Otherwise, the transactions will be at the risk of being curtailed to maintain power system security at delivery time. The

bilateral contracts will also need to go through the day-ahead market and/or real-time market to implicitly let the ISO know the transactions. It is implicit because the ISO only knows the offers and bids of the counterparties, but does not know there are the bilateral contracts between them when ISO clears the market. From the ISO's perspective, all bids are treated the same, and will be settled at the LMP. The counterparties need to do extra settlement outside the day-ahead market and real-time market to fulfill the contract price. For example, a generator and a load sign a contract of 100 MW at \$30 price, and the day-ahead market clears them at \$40 each. In this case, the generator receives \$10 above the contract price, and needs to refund $10 \cdot 100 = 1000$ \$/h to the load. The load pays \$40 /MWh to the ISO, and receives \$10 /MWh from the generator, so its net price is \$30 /MWh, which is the contract price.

However, there will be a problem if the generator and load are not at the same location, so they get different LMPs. For example, a generator and a load sign a contract of 100 MW at \$30 /MWh price, but due to transmission congestion, the day-ahead market settles generator at \$20 /MWh and the load at \$50 /MWh. The generator receives \$10 /MWh less than the contract price, and the load pays \$20 /MWh more than the contract price. The counterparties are short by $(20 + 10) \cdot 100 = 3000$ \$/h to be able to fulfill the contract price. The amount is sometimes referred as congestion cost. The congestion revenue right (CRR) is a financial instrument to hedge to congestion cost. Nowadays, all electricity markets, both the existing ones and those under development in the US, have CRRs built into the market design. For example, the proposed

ERCOT nodal market will have two kinds of CRRs, namely point to point rights to receive compensation according to LMP differences, and flowgate rights (FGRs), which are path based rights to receive compensation according to transmission shadow prices [17]. Take the point to point CRR as an example, the ISO will pay the CRR owner the difference in LMP between the source and sink times the CRR amount. If the counterparties own 100MW CRR with the source at the generator location and the sink at the load location, the ISO will pay them $(50 - 20) \cdot 100 = 3000$ \$/h, which is exactly the amount to cover the congestion cost. In this case, the congestion cost is perfectly hedged through CRR.

The ISO will auction CRR a month to a year prior to the trade day. Only part of the capacities in the transmission network will be auctioned. The rest of the capacities are either allocated to transmission owners as CRRs prior to the CRR auction, or held as margins.

A generator or generation firm can sell electric energy through these market elements in the following chronological order.

- Contact potential buyers to sign long-term bilateral contracts.
- Participate in CRR auction to procure CRR in order to hedge congestion cost risk.
- Make offers into day-ahead market trying to fulfill the contracted amount, and try to make more profit with the remaining capacity.

- Fine tune offers in real-time market if it is more profitable.
- Deliver cleared schedule into the transmission grid.

1.3 Analyzing Strategic Behaviors in Electricity Markets via Residual Demand

The electricity market is an oligopoly market from the supply side. The majority of total electric energy is supplied by a few big generation firms. Due to this reason, they are strategic players in the market. The main topic of this dissertation is to study how to analyze the strategic behaviors in centralized offer-based day-ahead market and real-time market. In the rest of the dissertation, unless stated otherwise, “electricity market” is specifically used to refer to a day-ahead market or a real-time market. In other words, “electricity market” is used through the rest of the dissertation to mean centralized offer-based electricity energy market that is cleared by an OPF program, and priced by LMP.

We assume the generators or generation firms pursue maximum profit in electricity markets. Their strategic behaviors are characterized by or at least bounded by the profit maximizing strategies. One major task of this dissertation is to find the profit maximizing strategy. The approach we are going to take is the residual demand approach, which has been widely used by economists in general markets, as well as electricity markets without considering the transmission constraints. For example, [26] and [40] analyze the bidding behaviors in ERCOT zonal balancing market using the residual demand

approach, [42] compares the Cournot model and supply function equilibrium model for German electricity market using the residual demand approach, and [5] optimizes a generators' expected profit based on historical residual demand curves.

Following Borenstein, Bushnell, and Stoft [11], let us consider a generator's profit maximization problem from the residual demand point of view in a transmission-constrained network. For simplicity, assume the generator under consideration does not have any forward bilateral contracts or CRRs. As discussed in [33], handling the contracts only involves shifting the offer curve by the contracted amount. CRRs can be handled in a similar fashion as handling the bilateral contracts.

Conceptually, each generator is facing a residual demand curve. The residual demand curve specifies at each price level the maximum market share left for the generator. Without transmission constraints, the residual demand is defined by the total system demand minus the total supply for other generators. The residual demand for generator i is

$$R_i(p) = D(p) - S_{-i}(p) \quad (1.1)$$

where

- p is the market price,
- $D(p)$ is the system total demand function, either price elastic or price inelastic,

- $S_{-i}(p)$ is the aggregated supply from other generators

$$S_{-i}(p) = \sum_{j \neq i} S_j(p),$$

where $S_j(p)$ is the supply from generator j .

For example, if at price \$50/MWh, the system total demand is 1000 MW, and the aggregated supply from other generators is 800 MW, then there is 200 MW market share left for the generator, so the residual demand at \$50/MWh is 200 MW. The price-demand pair (50, 200) is a point on the residual demand curve.

From generator i 's point of view, the market clears at the point where its supply meets its residual demand,

$$S_i(p_i) = R_i(p_i). \quad (1.2)$$

The generator i 's profit as a function of the output quantity is

$$\Pi_i(q_i) = P_i(q_i)q_i - C_i(q_i), \quad (1.3)$$

where $P_i(q_i) = R_i^{-1}(q_i)$. That is, $P_i(\bullet)$ is the inverse function of $R_i(\bullet)$, so that $(P_i(q_i), q_i)$ is a point on the residual demand curve.

Although this discussion has been in terms of output quantity as the strategic variable, the choice of the generator's strategy decision variable is not important. For example, the generator's strategy can be: price, as in the Bertrand model; quantity, as in the Cournot model; or a supply function, as in

the supply function model.² Without loss of generality, we choose the quantity to be the generator's strategic variable.

As illustrated in Fig. 1.2, at production q_i , generator i 's profit $\Pi_i(q_i)$ is the shaded area between quantities 0 and q_i , above the marginal cost function $C'_i(\bullet)$ and below the price $P_i(q_i)$. Intuitively, the problem of maximizing profit boils down to finding a point on the residual demand curve that maximizes the shaded area.

We assume piecewise quadratic cost functions and piecewise linear offer function such that the residual demand curve is piecewise linear, and the profit function is piecewise quadratic. The first order necessary condition (FONC) for maximizing generator i 's profit is

$$\begin{aligned}\frac{d\Pi_i}{dq_i}(q_i+) &= P_i(q_i) + P'_i(q_i)q_i - C'_i(q_i) \leq 0, \\ \frac{d\Pi_i}{dq_i}(q_i-) &= P_i(q_i) + P'_i(q_i)q_i - C'_i(q_i) \geq 0,\end{aligned}\tag{1.4}$$

with the definition

$$\begin{aligned}\frac{d\Pi_i}{dq_i}(q_i^{\max}+) &= 0, \\ \frac{d\Pi_i}{dq_i}(q_i^{\min}-) &= 0,\end{aligned}$$

where q_i^{\max} and q_i^{\min} are generator i 's output upper and lower limit respectively.

²Although the generator's strategy decision variable is not important in calculating the generator's maximum profit, it may make a huge difference for a Nash Equilibrium model. In other words, the Nash Equilibrium, if it exists, largely depends on the strategy space used in the model as has been observed in many references, such as [6]. Empirically characterizing the Nash Equilibrium of electricity markets is out of the scope of this dissertation. Latest researches in this direction include [40] and [42].

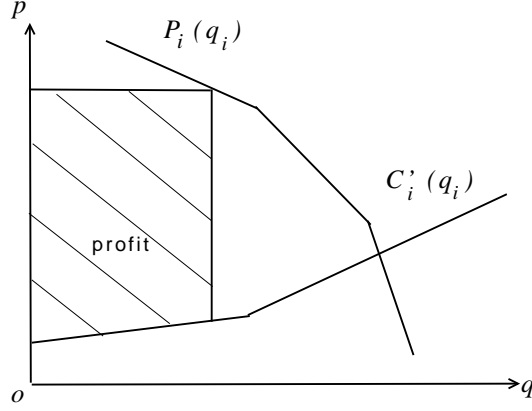


Figure 1.2: Profit maximization given residual demand curve

The point on the residual demand curve satisfying (1.4) maximizes the generator's profit. The derivative of the residual demand is especially useful because of its role in the FONC (1.4).

The transmission constraints in electricity markets are a challenge to characterizing the residual demand. The uniqueness of electricity markets comes largely from the transmission network [12]. The electric transmission network that connects the suppliers and consumers for trading electricity has to obey physical laws. In contrast to a transportation network, electricity cannot flow across the electric transmission network arbitrarily. The electric power flows are governed by Kirchoff's Current Law (KCL) and Kirchoff's Voltage Law (KVL).

Some studies are based on over simplified representation of the transmission network. For example, [11] considers only two nodes connected by

a single line, and [43] models the transmission network as a transportation network. These models are not able to capture the characteristics of a looped electric transmission network. There are other studies circumventing the challenge of transmission constraints, and make conjectured assumptions on the market price response or competitors' responses to the strategic behaviors, such as [15, 29]. Because these conjectured variation models are based on very arbitrary assumptions, they cannot capture the characteristics of a looped electric transmission network either. Yao and Oren propose an Equilibrium Programming with Equilibrium Constraints (EPEC) model to calculate the market Nash equilibrium in a transmission-constrained network [49]. They assume the residual demand functions are explicitly given and do not change as transmission congestion conditions change. In essence, it is the same as the conjectured variation model.

There have been many numerical case studies that look into the impacts of binding transmission constraints, such as [10, 12, 48]. For example, Cardell, Hitt and Hogan use a Cournot model in a three-bus looped network, and numerically demonstrate that strategic behaviors involving the transmission constraints could lead to a market outcome that is different from the usual analysis of imperfect competition [12]. Their work reinforces the need for more rigorous characterization of the effect of the transmission constraints on residual demand. Xu and Yu use a linear supply function model, and calculate the transmission-constrained supply function equilibrium (SFE) [48]. They consider the impact of the transmission, and calculate the best response

numerically. These studies have provide valuable insights, but they have not characterized the effects of transmission constraints systematically, and are far from practical applications.

To the other extreme, there exist many electricity market simulation tools, which can model the market clearing process for production level systems. These tools are used in industry to support business decisions. Different bidding strategies can be tested in simulations to find out the most profitable ones. Because of simulation software limitations, this often involves tremendous human judgment and intervention. It is true that these simulation tools can assist in decision making, but they are black boxes incapable of providing insights about how the market fundamentals drive the market outcomes. It is a big challenge and of great importance for market participants and market monitors to gain insights about how the transmission network acts as a main driver for the market outcome, and to know how to leverage these insights.

A perfect method to study the strategic behaviors should combine the power of two methods above:

- from the practical point of view, the method is able to handle large scale production level systems, and it is better if the existing advanced market simulation engines can be reused;
- from the theoretical point of view, the method is able to represent the transmission networks to the same details as in the true market clearing process, and systematically find the profit maximizing strategy.

There is another type of method that aims at these two goals: the mathematical program with complementary constraints (MPEC) method, proposed for application in electricity markets by Hobbs, Metzler, and Pang [23]. The idea is to formulate a generator or generation firm's profit maximization as a two-level optimization problem, where the lower level is the OPF problem, and the upper level is the generator's own profit maximization problem. One way to solve this problem is to add the KKT conditions for the OPF into the upper level profit maximization as constraints, and form an integrated optimization problem of MPEC [23]. As discussed in [23], the MPEC problem for generator i can be formulated in the following form:

$$\begin{aligned}
& \max_{\boldsymbol{\alpha}_i} \quad \text{profit}_i \\
& \text{s.t.} \quad \underline{\boldsymbol{\alpha}}_i \leq \boldsymbol{\alpha}_i \leq \overline{\boldsymbol{\alpha}}_i, \\
& \quad \text{OPF KKT conditions,}
\end{aligned} \tag{1.5}$$

where $\boldsymbol{\alpha}_i$ is generator i 's strategy variable vector, and $\underline{\boldsymbol{\alpha}}_i$ and $\overline{\boldsymbol{\alpha}}_i$ are its lower bound and upper bound respectively. Note that the OPF KKT conditions is nested in the generator's MPEC problem (1.5).

A production level security constrained OPF problem typically models hundreds of generators, thousands of buses and lines, tens to hundreds of contingencies [13], which makes the MPEC problem beyond the computational capability of existing MPEC solvers [30]. Moreover, although there exist advanced algorithms and software to solve the OPF, which utilize the property that only a few transmission constraints are likely to be simultaneously binding, it is difficult to reuse them in solving the MPEC due to the nested

structure. The MPEC approach essentially requires power system application developers to start from scratch in order to implement such an algorithm. Due to these reasons, there is rarely any implementation of the MPEC method to calculate a generator's optimal offer in practice. The MPEC method has difficulty in meeting the first goal satisfactorily due to this limitation.

In summary, there is no available methods that could achieve the two goals up to now. The major contributions of this dissertation are to propose methods to meet the two goals. The method is based on the residual demand concept. Through the rest of the dissertation we are going to:

- generalize the residual demand concept to transmission-constrained electricity markets, and
- systematically find the profit maximizing strategy based on transmission-constrained residual demand.

Chapter 2

Characterizing Strategic Behaviors in Electricity Markets Without Transmission Constraints

Before dealing with the transmission constraints, we first review how the residual demand can be used to characterize strategic behaviors and Nash equilibrium in the absence of transmission constraints. The characterization is based on the supply function Nash equilibrium (SFE) model, which is a residual demand method in essence.

2.1 Introduction

As introduced in chapter 1, in typical restructured electricity markets, market participants make offers into the market. The offer is a function specifying the minimum acceptable prices for different output levels. Offer functions may be fixed for an extended time horizon, and the market is cleared for every market clearing interval in the time horizon. An example is a day-ahead market, where the offers may be fixed for a day, and the market is cleared for each hour of the day. Because the system demand changes over time during the time horizon, the clearing price and quantities also change over time.

Klemperer and Meyer introduced the supply function equilibrium (SFE) model in [27]. It is a Nash equilibrium model to cope with demand uncertainty. Following Green and Newbery [21], the SFE model fits the electricity market setting if we:

1. define the supply function to be the inverse of the offer function,
2. represent the load-time profile in an electric power system as being equivalent to load uncertainty in Klemperer and Meyer's formulation, and
3. perform a Nash equilibrium model analysis for the oligopolistic electricity market.

The SFE model treats the offer functions in electricity markets more realistically than other Nash equilibrium models such as the Cournot model and the Bertrand model.¹ Due to this advantage, the SFE model has been widely used to study strategic behavior and market power in electricity markets. Green and Newbery were among the first to use the SFE model in electricity market analysis. They applied the SFE model to the electricity market in England and Wales in [21] and [20]. Following them, SFE models gained more popularity in electricity markets [1–4, 7–9, 18, 24, 25, 36, 40, 48].

¹Von der Fehr and Harbord argued that the SFE model may not be appropriate for piecewise constant offers. In this case, the offers have to be approximated by continuous differentiable functions in order to apply the SFE model. The discussion in that direction is outside the scope of this chapter.

Klemperer & Meyer and Green & Newbery demonstrated that the SFE can be characterized as a system of differential equations [21, 27]. The end-point is specified by the price and quantity pair where the supply function intersects the highest demand curve. As the end-point condition for the differential equations changes, the solutions trace out a continuum of equilibria [25]. The continuum of SFE will be illustrated in section 2.2, and the multiplicity greatly reduces the predictive value of the SFE model. Various efforts have been carried out in order to reduce the range of SFEs or to set up criteria to select preferable or “focal” equilibria.

Klemperer and Meyer showed that if the support of realized prices is infinite and there are no capacity constraints, then there is a unique SFE. However, this is not a realistic assumption for electricity markets, because the load is always finite [25]. We need to be able to deal with supply function equilibria that have finite price and quantity support.

Green and Newbery chose the most profitable SFE in [21]. However, the most profitable SFE yields predicted prices that are substantially above actual prices [7], which greatly weakens the credibility of this kind of choice. Holmberg considered the effect of both capacity constraints and the price cap, and singled out a unique SFE for the special case of symmetric players with risk of power shortage [25]. The unique SFE selected in [25] involves the end-point of the differential equation with price equal to the price cap and quantity equal to the generation capacity. At least for a case with symmetric suppliers, this SFE is the most profitable one that can be achieved with the given generation

capacity and price cap. However, with high price caps, such as \$2500/MWh in actual electricity markets,² this method may not significantly improve over Green and Newbery’s method.

By observing that capacity constraints can possibly invalidate some SFEs [8, 21], Genc and Reynolds intend to eliminate part of the SFE set from the most competitive side by considering “pivotal” market participants [18], whose absence would result in load curtailment at the price cap. However, the eliminated SFE set are typically very small for realistic electricity markets.

Besides the efforts to eliminate some of the SFE by considering capacity constraints, there are other researchers pursuing the same goal by refining the equilibrium. Anderson & Xu performed a stability analysis for a two-player discrete SFE model in [2]. However, they found that it would be difficult to apply their analysis to a case with more players.

Baldick and Hogan applied a stability analysis to refine the equilibrium, and they found a unique equilibrium for the symmetric linear marginal cost case by ruling out equilibria that were unstable to a particular perturbation [8]. However, the perturbation considered in [8] to rule out equilibria is not “fair” enough, because one of the SFEs, the Linear SFE (LSFE), is not perturbed at all. It turns out that the only “not unstable” SFE under the perturbation is indeed the LSFE. If we instead choose a similar perturbation that *does* affect the LSFE, then the LSFE might be unstable as well. In a sense, their stability

²The CAISO nodal market currently imposes a \$2500/MWh energy price cap.

notion is too stringent to differentiate the continuum of SFEs into stable and unstable.

Baldick and Hogan proposed another stability analysis in [9]. They assumed that the supply functions are polynomials in order to be able to analyze stability in a finite dimensional space. However, the best response mapping is not a self map on the space of polynomial functions, so they assumed an approximation mechanism that maps an arbitrary function back to a polynomial function in order to study the SFE stability in the polynomial function space. Unfortunately, by making the approximation, they introduced another problem. That is, there is only one equilibrium (fixed point) for the approximated best response mapping, namely the LSFE. Although they argued that similar stability analysis could be performed for other SFEs as well, the analysis is only valid for the LSFE.

We propose a novel stability analysis method in this chapter aimed at overcoming the difficulties in [9]. Our method can analyze the stability of every SFE, not just the LSFE. This is a significant improvement over Baldick and Hogan's method in [9], because being able to scrutinize every SFE enables characterization of the stable SFE set. This could not be carried out with Baldick and Hogan's method.

A variation of the method in this chapter has been reported in [47]. In [47], we considered piecewise linear function perturbations based on Taylor expansion, but we found it very difficult to generalize the analysis to high order polynomial function perturbations. Therefore, we further developed the new

method of this chapter to be able to handle high order polynomial function perturbations. In this sense, [47] should be viewed as a preliminary result, and this chapter provides a more thorough and advanced analysis.

The organization of the rest of the chapter is as follows. Section 2.2 recalls the SFE model as well as the best response, and demonstrates the equilibrium continuum. Section 2.3 presents the SFE stability analysis method. Section 2.4 applies the method to a symmetric example. Section 2.5 concludes.

2.2 The SFE Model and the Equilibrium Continuum

We briefly review the SFE model in this section. We assume that each generation company's goal is to maximize its total profit over a time horizon, and its profit does not only depend on its own supply function, but also on its competitors' supply functions. Essentially, this is a game with the generation companies viewed as players, the supply functions as strategies, and profits as payoffs. A Nash Equilibrium for the game, if it exists, can be used to characterize the electricity market. This type of analysis has been applied to the England and Wales market of the 1990s [7, 20, 21], to the Electric Reliability Council of Texas (ERCOT) market [26, 40], and to the Germany market [42].

2.2.1 SFE and Best response Mapping

Consider a uniform price electricity market without transmission constraints. Our formulation mainly follows previous literatures [8, 9, 21]. Denote

the market price by p , the supply functions by $\mathbf{S} = (S_i)_{i=1,2,\dots,n}$, where n is the total number of players in the market, each $S_i = S_i(\bullet)$ is a function from price to quantity, and the cost functions by $\mathbf{C} = (C_i)_{i=1,2,\dots,n}$, where each $C_i = C_i(\bullet)$ is a function from quantity to cost per unit time. We assume a continuous load-time profile following Green and Newbery's approach in [21] to interpret the load-time profile as being equivalent to the uncertainty in Klemperer and Meyer's representation in [27]. That is, the demand function is in the following form:

$$D(p, t) = N(t) - \gamma p, \quad t \in [t_0, t_1],$$

where $N(t)$ is the load-time profile function. The range of N is $[N_{\min}, N_{\max}]$, where

$$N_{\max} = \max_{t \in [t_0, t_1]} \{N(t)\}, \quad N_{\min} = \min_{t \in [t_0, t_1]} \{N(t)\}.$$

Without loss of generality, we assume the supply functions are defined on a finite price support $[p_{\min}, p_{\max}]$ where p_{\min} and p_{\max} are, respectively, the minimum and maximum realizable prices. For example, they can be the price floor and price cap respectively if the price floor and price cap have been set for the electricity market. We assume the supply functions are in the space of second-order differentiable functions, i.e. $(S_i)_{i=1,2,\dots,n} \in \mathcal{C}^2([p_{\min}, p_{\max}])$.

Suppose the generation marginal cost has the following linear form:

$$C_i^{(1)}(q) = c_i q_i + e_i, \quad \forall q > 0, \forall i = 1, 2, \dots, n. \quad (2.1)$$

(Superscript (1) will be used throughout to represent first order differentiation.) Each player tries to maximize its total profit over all the market clearing

intervals. We consider a process where each generation company optimally (profit-maximizing) updates its supply function in response to observations of the residual demand function. The transmission-unconstrained best response in a supply function model is characterized in [21, 27, 36]. It is the same equation of (1.4) assuming the residual demand function is continuous differentiable, and using the supply function instead of the offer function. We rewrite the characterization as

$$S_i(p) = -(p - C'_i(S_i(p))) R'_i(p), \quad (2.2)$$

where $R'_i(p)$ can be calculated by taking derivative on both sides of (1.1).

We distinguish corresponding values before and after the best response update by adding superscripts “old” and “new” respectively, and rewrite (2.2) in a best response form as in [21, 27]:

$$\frac{S_i^{\text{new}}(p)}{p - c_i S_i^{\text{new}}(p) - e_i} = \gamma + S_{-i}^{\text{old}(1)}(p), \quad \forall p \in [p_{\min}, p_{\max}],$$

so:

$$S_i^{\text{new}}(p) = \frac{(\gamma + S_{-i}^{\text{old}(1)}(p))(p - e_i)}{1 + c_i(\gamma + S_{-i}^{\text{old}(1)}(p))}, \quad (2.3)$$

$$\forall p \in [p_{\min}, p_{\max}], \forall i = 1, 2, \dots, n,$$

where subscript $-i$ means all others except i , i.e.

$$S_{-i}^{\text{old}(1)} = \sum_{j \neq i} S_j^{\text{old}(1)}.$$

Denote the best response mapping in (2.3) by $\beta(\bullet)$ such that:

$$\mathbf{S}^{\text{new}} = \beta(\mathbf{S}^{\text{old}}). \quad (2.4)$$

The domain of $\beta(\bullet)$ is $(\mathcal{C}^2([p_{\min}, p_{\max}]))^n$. Every SFE \mathbf{S}^* is a fixed point of the best response mapping $\beta(\bullet)$, i.e.:

$$\mathbf{S}^* = \beta(\mathbf{S}^*),$$

or explicitly:

$$S_i^*(p) = \frac{\left(\gamma + S_{-i}^{*(1)}(p)\right)(p - e_i)}{1 + c_i \left(\gamma + S_{-i}^{*(1)}(p)\right)}, \quad (2.5)$$

$$\forall p \in [p_{\min}, p_{\max}], \forall i = 1, \dots, n.$$

To qualify for a SFE, besides (2.5), the following must be satisfied:

1. every supply function is non-decreasing, i.e.

$$S_i^{*(1)}(p) \geq 0, \quad \forall p \in [p_{\min}, p_{\max}], \forall i = 1, \dots, n,$$

2. every supply function satisfies the (for example, second-order sufficient) conditions for profit maximization given all the other supply functions.

Notice that each SFE as characterized above does not depend on any specific load-time profile $N(t)$. Following Anderson and Hu [2], we call it a “strong” SFE. By “strong” we mean the supply function optimality holds for every possible price realization within the finite load support $[N_{\min}, N_{\max}]$. We only consider “strong” SFEs in this chapter.

2.2.2 SFE Continuum

As described in [8], (2.5) can be transformed into the standard form of non-linear differential equations:

$$\mathbf{S}^{*(1)}(p) = \left[\frac{1}{n-1} \mathbf{1}\mathbf{1}^T - \mathbf{I} \right] \begin{bmatrix} \frac{S_1^*(p)}{p - C_1^{(1)}(S_1^*(p))} \\ \frac{S_2^*(p)}{p - C_2^{(1)}(S_2^*(p))} \\ \vdots \\ \frac{S_n^*(p)}{p - C_n^{(1)}(S_n^*(p))} \end{bmatrix} - \frac{\gamma}{n-1} \mathbf{1},$$

where:

$$\mathbf{1} = [1 \quad 1 \quad \cdots \quad 1]^T,$$

and \mathbf{I} is the identity matrix.

With different end point conditions, we can trace out a continuum of equilibria by solving (2.5) [25, 27]. For example, let us consider the simple symmetric SFE with quadratic cost, i.e. the symmetric SFE with linear marginal cost case. Suppose each player has the same cost function:

$$C(q) = \frac{1}{2}cq^2 + eq.$$

Then a symmetric solution to (2.5) is reduced to solving the following differential equation:

$$S^{(1)}(p) = \frac{1}{n-1} \left(\frac{S(p)}{p - cS(p) - e} - \gamma \right). \quad (2.6)$$

Klemperer and Meyer characterized the full range of the continuum of the equilibria using the following function:

$$f(p, S) \triangleq S^{(1)}(p) = \frac{1}{n-1} \left(\frac{S}{p - cS - e} - \gamma \right).$$

As analyzed by Klemperer and Meyer, there is a continuum of SFEs existing between the two loci $f(p, S) = 0$ and $f(p, S) = \infty$ in the p - S plane [27].

We demonstrate the continuum of SFEs in Fig. 2.1. The graph corresponds to the symmetric linear marginal cost computational example we will discuss in detail in section 2.4. In this case:

$$\begin{aligned} f(p, S) = 0 & \Leftrightarrow p = \left(c + \frac{1}{\gamma}\right)S + e, \\ f(p, S) = \infty & \Leftrightarrow p = cS + e, \end{aligned}$$

which define straight lines.

Every point on the line $D(p) = -\gamma p + N_{\max}$ that is between the loci $f(p, S) = \infty$ and $f(p, S) = 0$ can be an end point to solve the differential equation (2.6). Therefore, there is a continuum of solutions to (2.6). The “strong” SFEs are the portion of these solutions that are between the minimum demand $D(p) = -\gamma p + N_{\min}$ and the maximum demand $D(p) = -\gamma p + N_{\max}$. The boundaries of the SFE continuum are the “least competitive SFE,” whose end point is the intersection of $f(p, S) = 0$ and $D(p) = -\gamma p + N_{\max}$, and the “most competitive SFE,” whose end point is the intersection of $f(p, S) = \infty$ and $D(p) = -\gamma p + N_{\max}$.

Note that there is wide range between the “least competitive SFE” and the “most competitive SFE,” which greatly limits the predictive value of the SFE model. In Fig. 2.1, we plot 50 SFEs evenly distributed between the “least competitive SFE” and the “most competitive SFE,” to represent the SFE continuum.

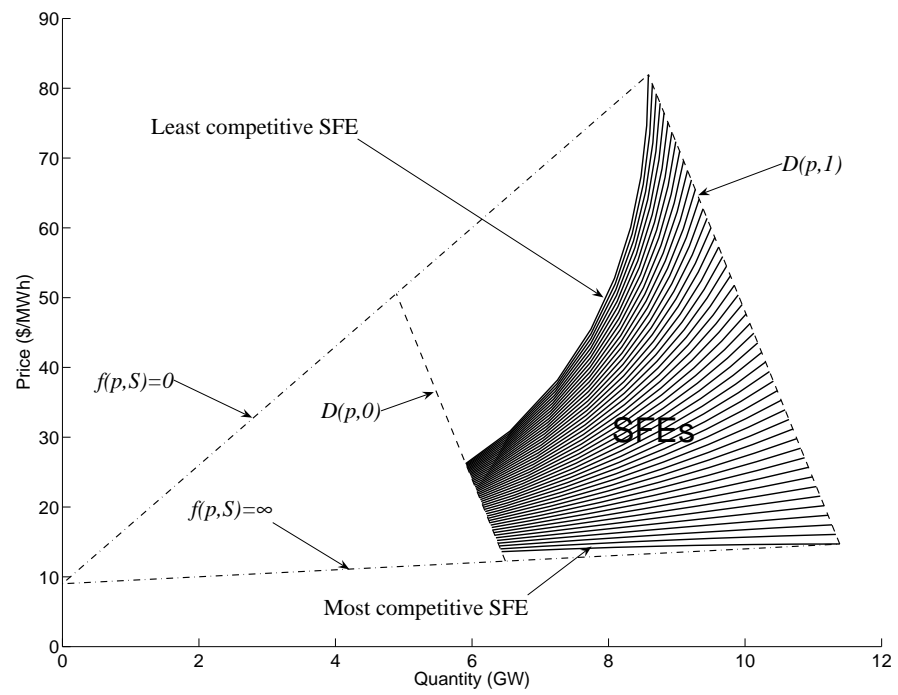


Figure 2.1: A continuum of symmetric SFEs

2.3 Stability Analysis of SFE

We have illustrated that there exists a continuum of strong SFEs. The multiplicity of the SFE greatly limits its predictive value. We want to refine the SFEs by a stability analysis. The major contribution of this chapter is to refine the equilibria for finite demand support through a stability analysis.

The idea of stability analysis is motivated by Baldick and Hogan's paper [9]. Baldick and Hogan did not study strong SFEs as we do in this chapter. Instead, they considered supply function equilibrium in the polynomial function space. Unfortunately, they were not able to find any other equilibrium in the polynomial function space other than the linear SFE. As a result, they studied the stability of the linear SFE in the polynomial function space, but were not able to apply the same analysis to other nonlinear SFEs. In contrast, we propose a stability analysis that is applicable to all strong SFEs.

2.3.1 Piecewise polynomial perturbations

We redefine the best response in terms of deviations. Consider deviations from \mathbf{S}^* ,

$$\Delta \mathbf{S} = \mathbf{S} - \mathbf{S}^*,$$

we denote the best response mapping in deviations by $\hat{\beta}$, which is characterized by:

$$\hat{\beta}(\Delta \mathbf{S}) = \beta(\Delta \mathbf{S} + \mathbf{S}^*) - \mathbf{S}^*, \quad \forall \Delta \mathbf{S} \in (\mathcal{C}^2([p_{\min}, p_{\max}]))^n. \quad (2.7)$$

Or, explicitly, $\hat{\beta}$ is derived by rewriting (2.3) in terms of polynomial deviation functions:

$$\begin{aligned}\Delta S_i^{\text{new}}(p) &= S_i^{\text{new}}(p) - S_i^*(p) \\ &= \frac{\left(\gamma + \Delta S_i^{\text{old}(1)}(p) + S_i^{*(1)}(p)\right)(p - e_i)}{1 + c_i \left(\gamma + \Delta S_i^{\text{old}(1)}(p) + S_i^{*(1)}(p)\right)} - S_i^*(p), \quad (2.8) \\ &\forall p \in [p_{\min}, p_{\max}], \forall i = 1, 2, \dots, n.\end{aligned}$$

Because \mathbf{S}^* is a fixed point of $\beta(\bullet)$, $\mathbf{S}^* - \mathbf{S}^* = \mathbf{0}$ is a fixed point of $\hat{\beta}(\bullet)$. Therefore, the stability problem of \mathbf{S}^* under $\beta(\bullet)$ is equivalent to the stability problem of $\mathbf{0}$ under $\hat{\beta}(\bullet)$.

In this chapter, we consider a specific class of perturbations to the strong SFEs, namely perturbations specified by polynomial functions. For each player i , the deviations ΔS_i can be any order- m polynomial function, where m is a predetermined positive integer. The polynomial deviations for all players are in the following form:

$$\Delta \mathbf{S} = \begin{bmatrix} \Delta S_1 \\ \Delta S_2 \\ \vdots \\ \Delta S_{n-1} \\ \Delta S_n \end{bmatrix} = \begin{bmatrix} \sum_{\ell=0}^m \alpha_{1\ell} \frac{(p-p_0)^\ell}{\ell!} \\ \sum_{\ell=0}^m \alpha_{2\ell} \frac{(p-p_0)^\ell}{\ell!} \\ \vdots \\ \sum_{\ell=0}^m \alpha_{(n-1)\ell} \frac{(p-p_0)^\ell}{\ell!} \\ \sum_{\ell=0}^m \alpha_{n\ell} \frac{(p-p_0)^\ell}{\ell!} \end{bmatrix}. \quad (2.9)$$

where p_0 is a price reference.

Before we can study the SFE stability in the polynomial function deviation space, there is one problem that needs to be resolved. As discussed in [9], the range of the best response map (2.2.1) is the space of all continuous differentiable functions, so generally the best response applied to polyno-

mial deviation functions will not produce polynomial deviation functions, i.e. $\hat{\beta}(\Delta \mathbf{S}^{\text{old}})$ may not be a polynomial function. In order to be able to analyze the stability, we need a self map on a Banach space. In order to make the best response map of deviations a self map defined on the polynomial function space, following [9], we regress the true best response map $\hat{\beta}(\Delta \mathbf{S}^{\text{old}})$ on the polynomial function space for observations of the true map at particular prices p_1, \dots, p_k , where $k \geq m$. The coefficients of the regression polynomial $\alpha_i^{\text{new}}, i = 1, \dots, m$ are specified by:

$$\alpha_{i.}^{\text{new}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_i$$

where:

$$\alpha_{i.}^{\text{new}} = \begin{bmatrix} \alpha_{i0}^{\text{new}} \\ \alpha_{i1}^{\text{new}} \\ \vdots \\ \alpha_{i(m-1)}^{\text{new}} \\ \alpha_{im}^{\text{new}} \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} 1 & (p_1 - p_0) & \frac{(p_1 - p_0)^2}{2!} & \dots & \frac{(p_1 - p_0)^{m-1}}{(m-1)!} & \frac{(p_1 - p_0)^m}{m!} \\ 1 & (p_2 - p_0) & \frac{(p_2 - p_0)^2}{2!} & \dots & \frac{(p_2 - p_0)^{m-1}}{(m-1)!} & \frac{(p_2 - p_0)^m}{m!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & (p_{k-1} - p_0) & \frac{(p_{k-1} - p_0)^2}{2!} & \dots & \frac{(p_{k-1} - p_0)^{m-1}}{(m-1)!} & \frac{(p_{k-1} - p_0)^m}{m!} \\ 1 & (p_k - p_0) & \frac{(p_k - p_0)^2}{2!} & \dots & \frac{(p_k - p_0)^{m-1}}{(m-1)!} & \frac{(p_k - p_0)^m}{m!} \end{bmatrix},$$

$$\mathbf{Y}_i = \begin{bmatrix} \Delta S_i^{\text{new}}(p_1) \\ \Delta S_i^{\text{new}}(p_2) \\ \vdots \\ \Delta S_i^{\text{new}}(p_{k-1}) \\ \Delta S_i^{\text{new}}(p_k) \end{bmatrix},$$

with $\Delta S_i^{\text{new}}(\bullet)$ defined in (2.8). This defines the best response map in polynomial deviation function space, or more precisely, in the space of the polynomial

coefficients $\boldsymbol{\alpha} \in \mathbb{R}^{n(m+1)}$. Denote the map by $\beta_p(\bullet) : \mathbb{R}^{n(m+1)} \rightarrow \mathbb{R}^{n(m+1)}$, such that:

$$\beta_p(\boldsymbol{\alpha}^{\text{old}}) = \begin{bmatrix} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_1 \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_2 \\ \vdots \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_{n-1} \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_n \end{bmatrix},$$

where

$$\boldsymbol{\alpha}^{\text{old}} = \begin{bmatrix} \alpha_{1\cdot}^{\text{old}} \\ \alpha_{2\cdot}^{\text{old}} \\ \vdots \\ \alpha_{(n-1)\cdot}^{\text{old}} \\ \alpha_{n\cdot}^{\text{old}} \end{bmatrix}.$$

By definition, $\mathbf{0}$ is a fixed point of $\beta_p(\bullet)$ in $\mathbb{R}^{n(m+1)}$ as shown below:

$$\begin{aligned} \beta_p(\mathbf{0}) &= \begin{bmatrix} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{0} \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{0} \\ \vdots \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{0} \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{0} \end{bmatrix} \\ &= \mathbf{0}, \end{aligned}$$

since $\mathbf{0}$ is a fixed point of $\hat{\beta}(\bullet)$.

The analysis can be easily generalized to piecewise polynomial function perturbations, if the regression is performed in a piecewise manner, and we repeatedly apply the same analysis on each of the polynomial segments. We will demonstrate this through an example in section 2.4.

2.3.2 Stability Analysis

We analyze the stability of $\beta_p(\bullet)$ at $\mathbf{0}$ in $\mathbb{R}^{n(m+1)}$.

Definition 1 (Lyapunov Stability). *Let $\|\bullet\|$ be a norm on vector space $\mathbb{R}^{n(m+1)}$. The n -player SFE \mathbf{S}^* is said to be Lyapunov stable to order- m polynomial deviation functions, which have the coefficients $\boldsymbol{\alpha} \in \mathbb{R}^{n(m+1)}$, if for any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that if $\|\boldsymbol{\alpha}\| < \delta(\epsilon)$,*

$$\|\beta_p^t(\boldsymbol{\alpha})\| < \epsilon, \forall t \in \mathbb{N}. \quad (2.10)$$

Definition 2 (Asymptotic Stability). *The n -player SFE \mathbf{S}^* is said to be asymptotically stable if it is Lyapunov stable and if there exists δ such that if $\|\boldsymbol{\alpha}\| < \delta$,*

$$\lim_{t \rightarrow \infty} \|\beta_p^t(\boldsymbol{\alpha})\| = \mathbf{0}. \quad (2.11)$$

A sufficient condition for asymptotic stability is the spectral radius of the Jacobian matrix of $\beta_p(\bullet)$ evaluated at $\mathbf{0}$ is less than 1, i.e.

$$\rho\left(\frac{\partial \boldsymbol{\alpha}^{\text{new}}}{\partial \boldsymbol{\alpha}^{\text{old}}}(\mathbf{0})\right) < 1, \quad (2.12)$$

where $\rho(\bullet)$ denotes the spectral radius. In the vicinity of $\mathbf{0}$, (2.12) implies (2.11), and (2.11) implies the boundedness of (2.10), so that the asymptotic stability definition is satisfied.

Similar to [9], after simplification, the partial derivatives are in the following form:

$$\begin{aligned} \frac{\partial \boldsymbol{\alpha}_{i.}^{\text{new}}}{\partial \boldsymbol{\alpha}_{i.}^{\text{old}}}(\mathbf{0}) &= \mathbf{0}, & \forall i = 1, \dots, n, \\ \frac{\partial \boldsymbol{\alpha}_{i.}^{\text{new}}}{\partial \boldsymbol{\alpha}_{j.}^{\text{old}}}(\mathbf{0}) &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Lambda}_i \mathbf{X} \boldsymbol{\Phi}_i, & \forall j \neq i, \end{aligned} \quad (2.13)$$

where

$$\mathbf{\Lambda}_i = \begin{bmatrix} \frac{1}{\left(1+c_i\left(\gamma+S_{-i}^{*(1)}(p_1)\right)\right)^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\left(1+c_i\left(\gamma+S_{-i}^{*(1)}(p_2)\right)\right)^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\left(1+c_i\left(\gamma+S_{-i}^{*(1)}(p_k)\right)\right)^2} \end{bmatrix}$$

$$\mathbf{\Phi}_i = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m \end{bmatrix}$$

Now we can calculate the Jacobian matrix and its spectral radius in order to characterize the asymptotic stability.

2.4 Computational Example

In this section, we apply the stability analysis discussed in section 2.3 to a computational example. First, we will demonstrate how the stability analysis can effectively refine the SFEs. Then, we will also study the robustness of the analysis against variations in regression models, which involve changes in:

- price reference,
- number of observations, and
- number of polynomial segments.

2.4.1 Example Setup

The example is derived from [8]. All the parameters remain the same as in [8] except the demand range $[N_{\min}, N_{\max}]$. The reason we change the demand range in the example is to represent the daily load variation for an electricity market more realistically.

The parameter values for the example are as follows. The demand curve is $D(p, t) = -\gamma p + N(t)$. The parameter values are $\gamma = 0.125$, $n = 3$, $c = 0.5$, $e = 9$, $N(t) = 21 + 15t$, $t \in [0, 1]$. The range of $N(t)$ is $[21, 36]$, which represents a typical summer day load variation.³

Because there are infinitely many SFEs, it is impossible to explicitly study the stability of every SFE. We discretize the SFE set into 50 representative SFEs by choosing 50 different end points on $D(p, 1)$, that are evenly distributed between $f(p, S) = 0$ and $f(p, S) = \infty$ as shown in Fig. 2.1.⁴ We sort them in ascending order of competitiveness, so the “most competitive SFE” is #1, and the “least competitive SFE” is #50. To gauge the effectiveness of the refinement, we will consider the number of equilibria among the samples that are asymptotically stable to perturbation under various model scenarios of the regression model setup.

³For example, the CAISO day-ahead market load varied from 22.2 GW to 36.4 GW for trade day July 10th, 2009.

⁴If the original SFE set had been discretized more finely than just 50 samples, we may be able to observe a more accurate stable SFE distribution in the continuum, but for the purpose of this study, a sample size of 50 seems to be large enough.

2.4.2 Regression Model Setup

There are many ways to set up a polynomial regression model, and it is a subjective choice. We will test the robustness of this analysis against variations in the regression model. We will apply the same stability analysis to several regression model scenarios that cover the following choices:

- price reference,
 1. $p_0 = 0$, 0 is a common choice for polynomial regression;
 2. $p_0 = e$, as a cost function parameter, e might be a “focal” choice;
- number of observations,
 1. $k = 24$, assuming the observations are hourly, i.e. there are 24 observations corresponding to 24 hours of the day;
 2. $k = 2(m + 1)$, assuming the number of observations adapt to the number of degrees of freedom in the regression but with a fixed redundancy;
- number of polynomial segments
 1. $seg = 1$, one piece regression;
 2. $seg = 3$, break the supply function into 3 pieces evenly, and do piecewise regression;

With each factor having 2 choices, we will explore a total of 8 scenarios for the regression model. The scenario with $p_0 = 0$, $k = 24$, and $seg = 1$ is referred as the base scenario.

2.4.3 Stability Analysis Results

The asymptotically stable SFEs for various orders m in the 8 different scenarios are summarized in Tab. 2.1 and Tab. 2.2, and are plotted in Fig. 2.2. We also list some of the calculated spectral radius data for the Jacobian in Tab. 2.3⁵. The data in Tab. 2.3 is for the base scenario only, which corresponds to the second column in Tab. 2.1, and we have restricted the list to SFE #1 to #10, as all the spectral radii are greater than one for SFE # greater than 10. With the data in Tab. 2.3, one can verify the asymptotic stability results in the second column of Tab. 2.1, where a less than one spectral radius in Tab. 2.3 indicates a asymptotically stable SFE in the second column of Tab. 2.1. The asymptotically stable SFEs under the other 7 scenarios are determined in exactly the same way, and the asymptotic stability results are listed in Tab. 2.1 and Tab. 2.2.

The first observation is that the stability analysis is very effective in refining the SFEs. As shown in Tab. 2.1 and Tab. 2.2, even with linear func-

⁵The spectral radius calculation may encounter numerical difficulties due to ill-conditioned Jacobian matrix, especially when m is larger than 5. This is caused by the fact that high order regressors do not add much variability to the regression, and are nearly linearly dependent on the low order regressors. Dealing with the numerical difficulty is out of the scope of this chapter. On the other hand, this fact suggests that it is unnecessary to pursue unrealistically high order polynomials in the regression, because it is not going to improve the goodness of fit.

m	$p_0 = 0$ $k = 24$ $seg = 1$	$p_0 = e$ $k = 24$ $seg = 1$	$p_0 = 0$ $k = 2(m + 1)$ $seg = 1$	$p_0 = e$ $k = 2(m + 1)$ $seg = 1$
1	#1 – #8	#1 – #8	#1 – #8	#1 – #8
2	#2 – #4	#2 – #4	#2 – #4	#2 – #4
3	#3, #4	#3, #4	#3, #4	#3, #4
≥ 4	none	none	none	none

Table 2.1: Asymptotically stable SFEs under polynomial function (1-piece) deviations

m	$p_0 = 0$ $k = 24$ $seg = 3$	$p_0 = e$ $k = 24$ $seg = 3$	$p_0 = 0$ $k = 2(m + 1)$ $seg = 3$	$p_0 = e$ $k = 2(m + 1)$ $seg = 3$
1	#2 – #7	#2 – #7	#2 – #7	#2 – #7
2	#3 – #4	#3 – #4	#3 – #4	#3 – #4
3	#4	#4	#4	#4
≥ 4	none	none	none	none

Table 2.2: Asymptotically stable SFEs under polynomial function (3-piece) deviations

tion perturbations, i.e. $m = 1$, only 6 or 8 SFEs out of the 50 samples are stable. Empirically, the stable SFE set shrinks as m increases, and if a SFE is stable under higher m , it is also stable under lower m . For cubic function perturbations, i.e. $m = 3$, only 1 or 2 SFEs out of the 50 samples are stable.

If m is greater or equal than 4, the stable SFE set shrinks to an empty set. This is consistent with Baldick and Hogan’s analysis of the stability of the linear SFE [9], and they have raised the concern whether a SFE could be sustained in practice. We will discuss more about this in the next section.

m	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
1	0.72	0.34	0.02	0.31	0.54	0.72	0.87	0.99	1.09	1.17
2	1.25	0.78	0.47	0.52	1.16	1.61	1.94	2.20	2.41	2.59
3	1.93	1.34	0.95	0.89	1.83	2.57	3.10	3.51	3.83	4.10
4	2.75	1.95	1.45	1.30	2.54	3.58	4.30	4.86	5.30	5.65
5	2.67	1.91	1.95	1.71	3.26	4.61	5.54	6.24	6.79	7.23

Table 2.3: Spectral radius in the base scenario

The second observation is that stable SFEs are located on the competitive side of the continuum. There are no stable SFEs among the samples with SFE # greater than 8. In other words, very uncompetitive SFEs are unstable. This is consistent with the stability analysis in [47] and the actual market results in [7].

The third observation is that the results are not very sensitive to the choice of p_0 and the number of observations. Let us compare different columns in Tab. 2.1. The asymptotically stable sample SFE set does not change across the selected different choices in price reference and number of observations. The same phenomenon is observed in Tab. 2.2.

Compared with the price reference and the number of observations, the results are slightly more sensitive to the number of regression segments. For example, the $m = 1$ asymptotically stable SFE lower boundary in 2.1 is #1, and the upper boundary is #8, while the $m = 1$ stable SFE lower boundary in 2.2 is #2, and the upper boundary is #7.

Overall, the analysis is robust against variations in the regression model,

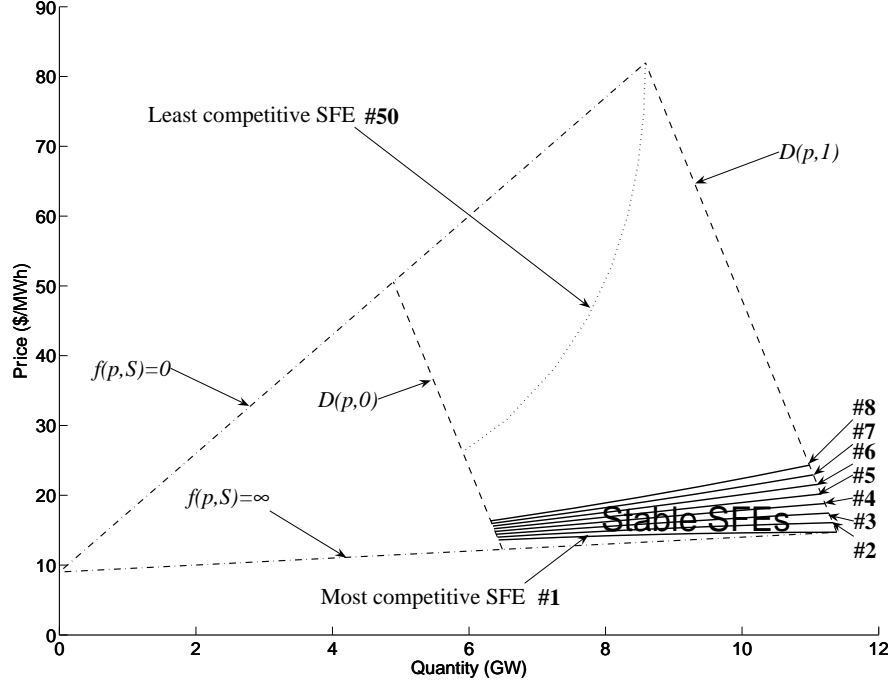


Figure 2.2: Stable SFEs

which adds to its predictive value.

2.4.4 Practical Implication

We have demonstrated the effectiveness of the stability analysis method in refining the SFEs in the previous section. However, with m greater than or equal than 4, the stable SFE set shrinks to an empty set, which raises the same concern as in [9] as to whether a SFE could exist in practice.

The concern boils down to the question of how the deviation functions

are likely to be modeled in practice. Since piecewise linear functions and piecewise quadratic functions are very common choices for approximating the deviation functions in practice, the concern may not be a real problem, because there exists stable SFEs that are able to withstand piecewise linear or piecewise quadratic function perturbations.

Another factor that will further relieve the concern is the market participants' inertia in changing their offers in the real world. Consider a simple model of inertia with a scalar as follows:

$$\mathbf{S}^{\text{new}} = \mathbf{S}^{\text{old}} + \kappa \left(\beta \left(\mathbf{S}^{\text{old}} \right) - \mathbf{S}^{\text{old}} \right),$$

where $\kappa \leq 1$, meaning that the supply function update can be slower than the best response. Note that $\kappa = 1$ corresponds to the best response (2.4), and smaller κ value represents larger inertia. In this case,

$$\boldsymbol{\alpha}^{\text{new}} = (1 - \kappa) \boldsymbol{\alpha}^{\text{old}} + \kappa \begin{bmatrix} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_1 \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_2 \\ \vdots \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_{n-1} \\ (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_n \end{bmatrix},$$

so similar to (2.13), the elements in the Jacobian will be in the following form:

$$\begin{aligned} \frac{\partial \boldsymbol{\alpha}_{i.}^{\text{new}}}{\partial \boldsymbol{\alpha}_{i.}^{\text{old}}}(\mathbf{0}) &= (1 - \kappa) \mathbf{I}, & \forall i = 1, \dots, n, \\ \frac{\partial \boldsymbol{\alpha}_{i.}^{\text{new}}}{\partial \boldsymbol{\alpha}_{j.}^{\text{old}}}(\mathbf{0}) &= \kappa (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Lambda}_i \mathbf{X} \boldsymbol{\Phi}_i, & \forall j \neq i. \end{aligned}$$

For the base scenario, the stable SFE set with $\kappa = 0.6$, $\kappa = 0.4$, $\kappa = 0.1$, and $\kappa = 0.01$ are listed in Tab. 2.4. With some inertia, more SFEs are stable.

m	$\kappa = 0.6$	$\kappa = 0.4$	$\kappa = 0.1$	$\kappa = 0.01$
1	#1 – #8	#1 – #8	#1 – #8	#1 – #8
2	#1 – #4	#1 – #4	#1 – #4	#1 – #4
3	#1 – #4	#1 – #4	#1 – #4	#1 – #4
4	none	#1 – #3	#1 – #3	#1 – #3
5	none	#1	#1 – #3	#1 – #3

Table 2.4: Asymptotically stable SFEs under inertia

A small inertia of $\kappa = 0.6$ makes SFE #1 and #2 stable for $m = 3$, which are unstable in Tab. 2.1. A large inertia of $\kappa = 0.1$ makes SFE #1, #2, and #3 to stand order 5 polynomial function perturbations ⁶, and the stable set does not change with a larger inertia of $\kappa = 0.01$.

Because deviation functions are unlikely to be represented by high order polynomials in practice, we can therefore expect SFEs to be exhibited through market interaction, particularly if market participants are conservative in changing their offers through actions such as inertia.

2.5 Conclusion

In this chapter, we demonstrated how to use residual demand to characterize the strategic behaviors and Nash equilibrium in the absence of transmission constraint in a SFE model. We proposed a stability analysis method to refine the SFEs. The system dynamics are defined by the best response

⁶Estimating the inertia in practice is out of the scope of this chapter. Hortaçsu and Puller did some empirical study about the bidding behaviors in the ERCOT market [26], which suggests that the inertia varies for different market participants.

mapping based on the residual demand function, and the stability analysis is performed in deviations space, i.e. we study the stability of the fixed point $\mathbf{0}$ under the best response mapping. We restrict our attention to polynomial functions (and piecewise polynomial functions) perturbations, which result from a regression of the true perturbation functions. We find that the stable SFE set can be much smaller than the original set. Moreover, the method is also very robust against variations in the regression model. This study also implies that for practical deviations in the real world, stable SFEs can be sustainable. The stability analysis method provides an effective way to refine the SFEs, and thus improves the value of the SFE model in characterizing the performance of electricity markets.

Chapter 3

Characterizing Strategic Behaviors via Transmission-Constrained Residual Demand Derivative

In this chapter, we will generalize the residual demand concept to a transmission-constrained network, and demonstrate how to use it to characterize the strategic behaviors in electricity markets.

The organization of the rest of the chapter is as follows. Section 3.1 presents the concept of transmission-constrained residual demand, and the analytic calculation of its derivative. Section 3.2 provides one intuitive example and two numerical examples taken from previous papers to verify our calculation. In addition, Appendix A provides background knowledge about ordinary least squares (OLS) problem and weighted least squares (WLS) problem, which are used in proving some of the transmission-constrained residual demand derivative (TCRDD) properties. Section 3.3 improves the TCRDD calculation efficiency and practical implementation. Section 3.4 proposes a decoupled method to maximize a generator's profit based on the TCRDD. Section 3.5 applies the proposed profit maximization approach to the IEEE 118 bus system. Section 3.6 concludes.

3.1 Residual Demand and Its Derivative

In a typical electricity market, let us assume that different generators are located at different buses, and index the generators by the bus number. Under nodal pricing, transmission constraints will generally lead to different nodal prices for different buses, so instead of having only one uniform market price p , we will have a vector of nodal equilibrium prices:

$$\mathbf{p} = [p_1 \ p_2 \ \dots \ p_n]^T$$

assuming there are n buses in the system.

We assume that the demand at each bus depends on only its local price, because in the short term, it is unlikely that a market participant could shift loads between buses according to real time nodal prices. Currently, the real-time prices are not published quickly enough to support this kind of load response in electricity markets.

Accordingly, a generator's residual demand will be a function of its local nodal price. Write out the energy balance condition:

$$\sum_{j=1}^n (D_j(p_j) - S_j(p_j)) = 0. \quad (3.1)$$

By keeping a specific $S_i(p_i) - D_i(p_i)$ on the left-hand side, and moving all other terms to the right-hand side, we obtain

$$S_i(p_i) - D_i(p_i) = \sum_{j=1, j \neq i}^n (D_j(p_j) - S_j(p_j)). \quad (3.2)$$

Before we continue, to simplify notation, we combine each generator's supply function with the demand curve at the same bus by treating demand as negative supply, and this process will result in only one supply function left at each bus so that (3.2) becomes:

$$S_i(p_i) = \sum_{j=1, j \neq i}^n -S_j(p_j). \quad (3.3)$$

The market clearing condition for the residual market at bus i implies that the residual demand equals supply, i.e.:

$$R_i(p_i) = S_i(p_i). \quad (3.4)$$

Without loss of generality, we are going to characterize the transmission-constrained residual demand, and its derive $R'_i(p_i)$ at a specific bus i . We choose this bus to be the slack bus, and reorder all the system buses to number this chosen bus to be bus n . That is, we want to characterize the residual demand, and calculated its derivative $R'_n(p_n)$ at the slack bus n . The corresponding offer cost function, whose derivative is the offer function or the inverse supply function $S_i(\bullet)$, is denoted by $O_i(\bullet)$. That is, we define $O_i(\bullet)$ by

$$p_i = O'_i(S_i(p_i)), \forall p_i. \quad (3.5)$$

Also we assume the functions $O_i(\bullet)$, $\forall i = 1, \dots, n$ are strictly convex and twice differentiable.

Following [38], we assume the market is cleared by solving the following

DC OPF problem:

$$\min_q \sum_{i=1}^n O_i(q_i), \quad (3.6)$$

$$\text{s.t. } \mathbf{H}\mathbf{q} \leq \mathbf{Z}, \quad (3.7)$$

$$\mathbf{q}_n^{\min} \leq \mathbf{q}_n \leq \mathbf{q}_n^{\max}, \quad (3.8)$$

$$\mathbf{1}^T \mathbf{q} = 0. \quad (3.9)$$

where

- bus n is the slack bus,
- $\mathbf{q} = [q_1 \ q_2 \ \dots \ q_n]^T$ is the nodal power injection quantity vector,
- (3.7) consists of the transmission constraints and the generation capacity constraints for non-slack buses (suppose there are totally m of them),
- \mathbf{H} is a $m \times n$ matrix consisting of the submatrix of power transfer distribution factors (PTDFs) corresponding to the transmission constraints and the submatrix representing the capacity constraints for non-slack buses,
- \mathbf{Z} consists of the transmission capacity limits and the generation capacity limits for non-slack buses,
- $\mathbf{1} = \underbrace{[1 \ 1 \ \dots \ 1]}_n^T$,
- (3.8) is the generation capacity constraint, that specifies the upper and lower limits of the domain of the offer cost function at the slack bus, and

- (3.9) is the energy balance constraint.

There are two widely used OPF formulations. One is to consider elastic demands, and the OPF objective is to maximize total social welfare, as is used in the day-ahead market; the other is to consider inelastic demand, and the OPF objective is to minimize total generation cost, as is used in the real-time market. We use the first OPF formulation in this chapter to derive the residual demand derivative. However, we stress that the methodology is applicable to both OPF formulations.

We intend to calculate the residual demand derivative evaluated at the current market operating point. The current market operating point is determined by the OPF solution. Therefore, the residual demand derivative calculation is a post-OPF analysis. Given an OPF solution, we know which constraints are binding in the OPF formulation. Given these binding OPF constraints at the solution, we will form the Lagrangian for the OPF problem (3.6)–(3.9) including only binding constraints. Let us denote the binding constraints by subscript “b”. The calculation needs to be separated into two cases:

- the slack bus generation capacity constraint (3.8) is *not* binding, and
- the slack bus generation capacity constraint (3.8) is binding,

but we will see that the two cases result in the same expression for the residual demand derivative.

3.1.1 Non-binding slack bus generation capacity constraint

The Lagrangian for (3.6)–(3.9) including only the binding constraints is as follows:

$$\mathcal{L} = \sum_{i=1}^n O_i(q_i) - \lambda \sum_{i=1}^n q_i + \boldsymbol{\mu}_b^T (\mathbf{H}_b \mathbf{q} - \mathbf{Z}_b). \quad (3.10)$$

The first-order necessary conditions (FONCs) of (3.10) are

$$O'_i(q_i) = \lambda + \boldsymbol{\mu}_b^T \bar{\mathbf{H}}_{bi}, \quad i = 1, \dots, n-1, \quad (3.11)$$

$$O'_n(q_n) = \lambda, \quad (3.12)$$

$$\bar{\mathbf{H}}_b \bar{\mathbf{q}} = \mathbf{Z}_b, \quad (3.13)$$

$$q_n = - \sum_{i=1}^{n-1} q_i. \quad (3.14)$$

where

- $\bar{\mathbf{H}}_b$ is a $m_b \times (n-1)$ matrix generated by eliminating the n -th column (all elements in this column are zeros) of \mathbf{H}_b ,
- $\bar{\mathbf{H}}_{bi}$ is the i -th column of $\bar{\mathbf{H}}_b$,
- $\bar{\mathbf{q}}$ is obtained from \mathbf{q} by eliminating the n -th entry,
- \mathbf{Z}_b consists of the binding transmission capacity limits and the binding generation capacity limits for non-slack buses, and
- by definition, λ is the LMP for the at the slack bus n , and the right hand side of (3.11) is the LMP for other buses.

There are $n + m_b + 1$ equations, and $n + m_b + 1$ variables in the FONCs (3.11)–(3.14), so the FONCs are uniquely solvable assuming regularity conditions. Denote the solution by

$$\begin{bmatrix} \hat{q}_1 & \dots & \hat{q}_n & \hat{\lambda} & \hat{\mu}_{b1} & \dots & \hat{\mu}_{bm_b} \end{bmatrix}^T.$$

The FONCs characterize a market clearing point. From the perspective of the generator located at the slack bus, the market clearing point $(\hat{\lambda}, \hat{q}_n)$ is the intersection of its own offer cost function and the residual demand function. If the generator located at the slack bus changes its offer function, the market will clear at a different point, which is the new intersection point of the changed offer function and the residual demand function. In other words, the market clearing points generated by changing the offer function of the generator located at the slack bus are all points on the residual demand function. Therefore, the residual demand function is characterized by the locus of the market clearing points (λ, q_n) obtained by changing the generator's offer function.

That is, if we remove the equations that contain the offer information of the generator located at the slack bus from the FONCs, the remaining equations characterize the residual demand function, because the residual demand function should not depend on a generator's own offer information. In the FONCs (3.11)–(3.14), (3.12) contains $O'_n(\bullet)$, which is based on the offer cost function of the generator located at the slack bus, so we should remove it from the FONCs to characterize the residual demand function. With (3.12)

removed, we have $n + m_b$ equations, namely (3.11), (3.13) and (3.14), and $n + m_b + 1$ variables left in the FONCs, so there is one degree of freedom. The one degree of freedom implicitly characterizes a locus of (λ, q_n) , i.e. the residual demand curve.

3.1.2 Binding slack bus generation capacity constraint

Consider that the upper generation capacity constraint is binding at the slack bus. The Lagrangian for (3.6)–(3.9) including only the binding constraints is as follows:

$$\mathcal{L} = \sum_{i=1}^n O_i(q_i) - \lambda \sum_{i=1}^n q_i + \boldsymbol{\mu}_b^T (\mathbf{H}_b \mathbf{q} - \mathbf{Z}_b) + \rho^{\max} (q_n - q_n^{\max}), \quad (3.15)$$

where without loss of generality, we have assumed that the maximum generation capacity constraint is binding. A similar analysis applies for the minimum generation capacity constraint. The FONCs of (3.15) are:

$$O'_i(q_i) = \lambda + \boldsymbol{\mu}_b^T \bar{\mathbf{H}}_{bi}, \quad i = 1, \dots, n-1, \quad (3.16)$$

$$O'_n(q_n) - \rho^{\max} = \lambda, \quad (3.17)$$

$$\bar{\mathbf{H}}_b \bar{\mathbf{q}} = \mathbf{Z}_b, \quad (3.18)$$

$$q_n = - \sum_{i=1}^{n-1} q_i, \quad (3.19)$$

$$q_n = q_n^{\max} \quad (3.20)$$

Denote the solution of (3.16)–(3.20) by

$$\begin{bmatrix} \hat{q}_1 & \dots & \hat{q}_n & \hat{\lambda} & \hat{\mu}_{b1} & \dots & \hat{\mu}_{bm_b} & \hat{\rho}^{\max} \end{bmatrix}^T.$$

We need to clarify that generally speaking, we could not choose a bus with binding generation capacity constraint as the slack bus to solve problems involving power flow. The reason why we can do this in this case is that the OPF is already solved, so in this post-OPF analysis, from the optimality conditions point of view, it does not matter which bus is the slack bus as long as the FONCs (3.16)–(3.20) are satisfied.

For the same reason as in section 3.1.1, we need to remove the equations that contain the offer information of the generator located at the slack bus from the FONCs (3.16)–(3.20). Again, (3.17) contains $O'_n(\bullet)$, and thus should be removed. In this case, in addition, another equation, (3.20), should also be removed, because it specifies the upper limit of the offer function domain, and thus $O_n(\bullet)$ and (3.20) together characterize the offer information.

Note that if (3.17) and (3.20) are removed, the remaining equations (3.16), (3.18) and (3.19) are exactly the same as (3.11), (3.13) and (3.14).

3.1.3 Sensitivity analysis

Now we are going to calculate the residual demand derivative at bus n by simultaneously solving (3.11), (3.13), and (3.14), which has one degree of freedom that characterizes a locus of (λ, q_n) .

Consider simultaneous equations (3.11), (3.13), and (3.14) parameterized by λ . By the implicit function theorem, if second-order sufficient condi-

tions hold, then a unique function

$$\begin{bmatrix} \tilde{q}_1 & \dots & \tilde{q}_n & \tilde{\mu}_{b1} & \dots & \tilde{\mu}_{bm_b} \end{bmatrix}^T (\lambda)$$

exists in a neighborhood of $\hat{\lambda}$ that solves (3.11), (3.13), and (3.14). Furthermore,

$$\begin{bmatrix} \tilde{q}_1 & \dots & \tilde{q}_n & \tilde{\mu}_{b1} & \dots & \tilde{\mu}_{bm_b} \end{bmatrix}^T (\hat{\lambda}) = \begin{bmatrix} \hat{q}_1 & \dots & \hat{q}_n & \hat{\mu}_{b1} & \dots & \hat{\mu}_{bm_b} \end{bmatrix}^T. \quad (3.21)$$

Because the equations, which contain the offer information at the slack bus, have been removed, the quantity q_n in (3.14) is actually the residual demand quantity R_n , so we replace q_n by R_n in the left-hand side of (3.14), and therefore (3.14) becomes:

$$R_n(\lambda) = - \sum_{i=1}^{n-1} \tilde{q}_i(\lambda). \quad (3.22)$$

We are interested in the residual demand derivative, i.e. the derivative of $R_n(\lambda)$ with respect to λ evaluated at $\hat{\lambda}$:

$$\frac{dR_n}{d\lambda}(\hat{\lambda}) = - \sum_{i=1}^{n-1} \frac{d\tilde{q}_i(\hat{\lambda})}{d\lambda}. \quad (3.23)$$

Sensitivity analysis enables us to calculate the derivative of

$$\begin{bmatrix} \tilde{q}_1 & \dots & \tilde{q}_n & \tilde{\mu}_{b1} & \dots & \tilde{\mu}_{bm_b} \end{bmatrix} (\lambda)$$

with respect to λ evaluated at $\hat{\lambda}$. It may be that the sensitivity is not defined due to non-differentiability, and we will briefly discuss this case in section 3.1.5 below.

From (3.11) and (3.13), we get

$$O'_i(\tilde{q}_i(\lambda)) = \lambda + \tilde{\boldsymbol{\mu}}_b^T(\lambda) \bar{\mathbf{H}}_{bi}, \quad i = 1, \dots, n-1, \quad (3.24)$$

$$\bar{\mathbf{H}}_b \tilde{\mathbf{q}}(\lambda) = \mathbf{Z}_b, \quad (3.25)$$

where

$$\tilde{\mathbf{q}}(\lambda) = [\tilde{q}_1 \quad \tilde{q}_2 \quad \dots \quad \tilde{q}_{n-1}]^T(\lambda).$$

The vector $\frac{d\tilde{\mathbf{q}}}{d\lambda}(\hat{\lambda})$ can be calculated by totally differentiating (3.24) and (3.25)

with respect to λ to obtain

$$\begin{bmatrix} O''_1(\hat{q}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & O''_{n-1}(\hat{q}_{n-1}) \end{bmatrix} \frac{d\tilde{\mathbf{q}}}{d\lambda}(\hat{\lambda}) - \bar{\mathbf{H}}_b^T \frac{d\tilde{\boldsymbol{\mu}}_b}{d\lambda}(\hat{\lambda}) = \bar{\mathbf{1}}, \quad (3.26)$$

$$\bar{\mathbf{H}}_b \frac{d\tilde{\mathbf{q}}}{d\lambda}(\hat{\lambda}) = \mathbf{0}, \quad (3.27)$$

where $\bar{\mathbf{1}} = \underbrace{[1 \quad 1 \quad \dots \quad 1]^T}_{n-1}$. Solving (3.26) and (3.27), we obtain

$$\frac{d\tilde{\boldsymbol{\mu}}_b}{d\lambda}(\hat{\lambda}) = -(\bar{\mathbf{H}}_b \boldsymbol{\Lambda} \bar{\mathbf{H}}_b^T)^{-1} \bar{\mathbf{H}}_b \boldsymbol{\Lambda} \bar{\mathbf{1}}, \quad (3.28)$$

$$\frac{d\tilde{\mathbf{q}}}{d\lambda}(\hat{\lambda}) = \boldsymbol{\Lambda} \left(\bar{\mathbf{1}} - \bar{\mathbf{H}}_b^T (\bar{\mathbf{H}}_b \boldsymbol{\Lambda} \bar{\mathbf{H}}_b^T)^{-1} \bar{\mathbf{H}}_b \boldsymbol{\Lambda} \right) \bar{\mathbf{1}}, \quad (3.29)$$

where

$$\begin{aligned} \boldsymbol{\Lambda} &= \begin{bmatrix} O''_1(\hat{q}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & O''_{n-1}(\hat{q}_{n-1}) \end{bmatrix}^{-1}, \\ &= \begin{bmatrix} S'_1(\hat{p}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & S'_{n-1}(\hat{p}_{n-1}) \end{bmatrix}, \end{aligned} \quad (3.30)$$

and $\bar{\mathbf{I}}$ is the $(n-1) \times (n-1)$ identity matrix. Because we assume the strict convexity of $O_i''(\bullet)$, $\forall i = 1, \dots, n$, the inverse in (3.30) exists, and $\mathbf{\Lambda}$ is positive definite. Therefore,

$$\begin{aligned} \frac{d\tilde{R}_n}{d\lambda}(\hat{\lambda}) &= -\bar{\mathbf{I}}^T \frac{d\tilde{\mathbf{Q}}}{d\lambda}(\hat{\lambda}), \\ &= -\bar{\mathbf{I}}^T \mathbf{\Lambda} \bar{\mathbf{I}} + \bar{\mathbf{I}}^T \mathbf{\Lambda} \bar{\mathbf{H}}_b^T (\bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{H}}_b^T)^{-1} \bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{I}}. \end{aligned} \quad (3.31)$$

We need to clarify that this residual demand derivative is the residual demand derivative for bus n as a whole, i.e. the local actual demand at bus n has been combined with the supply at the same bus. Therefore, from the point of view of the generator located at bus n , its residual demand derivative is actually (3.31) plus the local demand derivative at bus n , if there is any local demand.

For convenience in the above analysis, we calculate the residual demand derivative at the slack bus. Notice that (3.31) indicates that the residual demand derivative only depends on $\mathbf{\Lambda}$ and $\bar{\mathbf{H}}_b$, which are reduced matrices (with rows and/or columns corresponding to the slack bus deleted). For residual demand at an arbitrary bus k , all we need to do is to reconstruct $\mathbf{\Lambda}$ and $\bar{\mathbf{H}}_b$ assuming that bus k is chosen as the slack bus in order to make use of (3.31).

The approach in this chapter can also be viewed as a generalization of the methods in [32] and [28]. References [32] and [28] calculate the sensitivities of the generation dispatches to offer prices considering the transmission constraints. Our approach generalizes their methods in that we can handle any type of offer functions and not just fixed price offers. Our emphasis is on the

residual demand, whose derivative reveals the sensitivities of the generation dispatch to the incremental market price changes. We demonstrated that these sensitivities are independent of the offer functional forms but do depend on the binding generation capacity and transmission constraints. We can handle all these constraints without any difficulty, whereas the methods in [19] and [20] cannot directly handle the binding generation capacity constraints.

3.1.4 Weighted least squares regression interpretation

The formula in (3.31) is the negative summed square error (SSE) of the following linear weighted least squares (WLS) regression problem: regress 1 on each column of $\bar{\mathbf{H}}_b$, and use $\mathbf{\Lambda}$ as the weight matrix. Suppose we have $n - 1$ observations (Y_i, \mathbf{X}_i) , $\forall i = 1, \dots, n - 1$, where

$$\begin{bmatrix} Y_1 & Y_2 & \dots & Y_{n-1} \end{bmatrix}^T = \bar{\mathbf{1}} \equiv \mathbf{Y}, \quad (3.32)$$

$$\begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \dots & \mathbf{X}_{n-1} \end{bmatrix} = \bar{\mathbf{H}}_b \equiv \mathbf{X}^T. \quad (3.33)$$

The linear WLS regression problem is to find an optimal $m_b \times 1$ vector $\boldsymbol{\beta}$ that minimizes the weighted sum of squared errors (see Appendix A for details):

$$\min_{\boldsymbol{\beta}} \quad SSE^{\text{WLS}}(\boldsymbol{\beta}) = \sum_{i=1}^{n-1} w_i (Y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2, \quad (3.34)$$

where w_i , $i = 1, \dots, n$, satisfy

$$\begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{n-1} \end{bmatrix} = \mathbf{\Lambda}. \quad (3.35)$$

The solution to this WLS problem is

$$\mathbf{b}^{\text{WLS}} = (\bar{\mathbf{H}}_{\text{b}} \boldsymbol{\Lambda} \bar{\mathbf{H}}_{\text{b}}^{\text{T}})^{-1} \bar{\mathbf{H}}_{\text{b}} \boldsymbol{\Lambda} \bar{\mathbf{1}} = \left| \frac{d\tilde{\boldsymbol{\mu}}_{\text{b}}}{d\lambda} (\hat{\lambda}) \right|, \quad (3.36)$$

$$\begin{aligned} SSE^{\text{WLS}}(\mathbf{b}^{\text{WLS}}) &= \bar{\mathbf{1}}^{\text{T}} \boldsymbol{\Lambda} \bar{\mathbf{1}} - \bar{\mathbf{1}}^{\text{T}} \boldsymbol{\Lambda} \bar{\mathbf{H}}_{\text{b}}^{\text{T}} (\bar{\mathbf{H}}_{\text{b}} \boldsymbol{\Lambda} \bar{\mathbf{H}}_{\text{b}}^{\text{T}})^{-1} \bar{\mathbf{H}}_{\text{b}} \boldsymbol{\Lambda} \bar{\mathbf{1}}, \\ &= \left| \frac{d\tilde{R}_n}{d\lambda} (\hat{\lambda}) \right|. \end{aligned} \quad (3.37)$$

Such a least squares interpretation helps us gain insight into the original problem. (It is interesting that least squares interpretations are widely observed. Another example in the context of nodal prices can be found in [35], although the specific topic is somewhat different.)

From WLS theory (see Appendix A), we know that the WLS problem (3.34) could be transformed into an equivalent ordinary least squares (OLS) problem. Define

$$\mathbf{Y}^* = \boldsymbol{\Lambda}^{1/2} \mathbf{Y}, \quad (3.38)$$

$$\mathbf{X}^* = \boldsymbol{\Lambda}^{1/2} \mathbf{X}, \quad (3.39)$$

where

$$\boldsymbol{\Lambda}^{1/2} = \begin{bmatrix} w_1^{1/2} & 0 & \dots & 0 \\ 0 & w_2^{1/2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{n-1}^{1/2} \end{bmatrix}. \quad (3.40)$$

The equivalent OLS to (3.34) is

$$\min_{\boldsymbol{\beta}} SSE^{\text{OLS}}(\boldsymbol{\beta}) = \sum_{i=1}^{n-1} (Y_i^* - \mathbf{X}_i^{*\text{T}} \boldsymbol{\beta})^2. \quad (3.41)$$

The solution to this OLS (3.41) problem is exactly the same as the solution of the WLS problem (3.34):

$$\mathbf{b}^{\text{OLS}} = (\bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{H}}_b^T)^{-1} \bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{1}} = \left| \frac{d\tilde{\boldsymbol{\mu}}_b}{d\lambda}(\hat{\lambda}) \right|, \quad (3.42)$$

$$\begin{aligned} SSE^{\text{OLS}}(\mathbf{b}^{\text{OLS}}) &= \bar{\mathbf{1}}^T \mathbf{\Lambda} \bar{\mathbf{1}} - \bar{\mathbf{1}}^T \mathbf{\Lambda} \bar{\mathbf{H}}_b^T (\bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{H}}_b^T)^{-1} \bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{1}}, \\ &= \left| \frac{d\tilde{R}_n}{d\lambda}(\hat{\lambda}) \right|. \end{aligned} \quad (3.43)$$

The residual of the OLS problem (3.41) with \mathbf{b}^{OLS} specified in (3.42) is

$$\mathbf{e}^* = \mathbf{Y}^* - \mathbf{X}^* \mathbf{b}^{\text{OLS}} = \mathbf{\Lambda}^{1/2} \left(\bar{\mathbf{1}} + \bar{\mathbf{H}}_b^T \frac{d\tilde{\boldsymbol{\mu}}_b}{d\lambda}(\hat{\lambda}) \right). \quad (3.44)$$

From OLS theory (see Appendix A), we know \mathbf{X}^* is orthogonal to \mathbf{e}^* ,

i.e.

$$\mathbf{X}^{*T} \mathbf{e}^* = 0, \quad (3.45)$$

i.e.

$$\bar{\mathbf{H}}_b \mathbf{\Lambda} \left(\bar{\mathbf{1}} + \bar{\mathbf{H}}_b^T \frac{d\tilde{\boldsymbol{\mu}}_b}{d\lambda}(\hat{\lambda}) \right) = 0, \quad (3.46)$$

or

$$\bar{\mathbf{H}}_b \mathbf{\Lambda} \left(\bar{\mathbf{1}} - \bar{\mathbf{H}}_b^T (\bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{H}}_b^T)^{-1} \bar{\mathbf{H}}_b \mathbf{\Lambda} \bar{\mathbf{1}} \right) = 0, \quad (3.47)$$

by substituting $\frac{d\tilde{\boldsymbol{\mu}}_b}{d\lambda}(\hat{\lambda})$ from (3.28). Note that (3.47) is the same as (3.27).

The orthogonality condition (3.45) produces the same equation as (3.27), so we would like to call equation (3.27) the orthogonality equation of binding constraints. The meaning of (3.27) is that when binding constraints do not change, the Jacobian of the constraints ($\bar{\mathbf{H}}_b$ in this case) is orthogonal to the

direction of change $\frac{d\tilde{\mathbf{q}}}{d\lambda}(\hat{\lambda})$. In other words, the direction of change $\frac{d\tilde{\mathbf{q}}}{d\lambda}(\hat{\lambda})$ will not change these binding constraints: no active constraints will be violated, and no active constraints will become non-binding. It is the orthogonality equation of binding constraints that makes possible the WLS interpretation.

In addition, from WLS theory (see Appendix A), we can deduce the following properties about the residual demand derivative.

Proposition 3. *If $\mathbf{\Lambda}$ is positive definite, then the residual demand derivative is less than or equal to zero.*

Proposition 4. *Enforcing a new linearly independent binding constraint in the OPF will reduce the residual demand derivative in absolute value if*

$$\mathbf{M}_{\mathbf{X}^*} (\bar{\mathbf{H}}_{\mathbf{b}}^{\text{Added}})^{\text{T}} \neq 0, \quad (3.48)$$

and

$$\mathbf{M}_{\mathbf{X}^*} \bar{\mathbf{1}}^{\text{T}} \neq 0, \quad (3.49)$$

where $\mathbf{X}^* = \mathbf{\Lambda}^{1/2} \bar{\mathbf{H}}_{\mathbf{b}}^{\text{T}}$ as defined in (3.39),

$$\mathbf{M}_{\mathbf{X}^*} = \bar{\mathbf{I}} - \mathbf{X}^* (\mathbf{X}^{*\text{T}} \mathbf{X}^*)^{-1} \mathbf{X}^{*\text{T}}, \quad (3.50)$$

and $\bar{\mathbf{H}}_{\mathbf{b}}^{\text{Added}}$ is an added row to $\bar{\mathbf{H}}_{\mathbf{b}}$.

In particular, if there is no transmission congestion, the residual demand derivative is $-\bar{\mathbf{1}}^{\text{T}} \mathbf{\Lambda} \bar{\mathbf{1}}$. When transmission is congested, the residual demand derivative decreases in absolute value, because:

$$\left| \frac{dR_n}{d\lambda}(\hat{\lambda}) \right| = SSE_{\text{C}}^* \leq \bar{\mathbf{1}}^{\text{T}} \mathbf{\Lambda} \bar{\mathbf{1}} = SSE_{\text{NC}}^*, \quad (3.51)$$

where the subscript “C” denotes “transmission congested,” and the subscript “NC” denotes “transmission uncongested.” This implies that when transmission constraints bind, the players have more incentive to exert market power because of the decrease in magnitude of the residual demand derivative.

We have assumed that there is no perfectly elastic supply at any bus in the derivation of (3.31). If there is perfectly elastic supply at some bus then (3.31) is not valid, because we cannot invert a singular matrix to get $\mathbf{\Lambda}$ in (3.30). To consider perfectly elastic supply, we will analyze the limit of the residual demand derivative as some diagonal elements of $\mathbf{\Lambda}$ go to infinity. We are especially interested in the conditions under which the residual demand derivative goes to infinity, i.e. the residual demand is perfectly elastic.

Theorem 5. *Suppose there are l buses (i_1, i_2, \dots, i_l) each with its supply derivative going to infinity, and denote the set composed of i_1, i_2, \dots, i_l as \mathbf{N} . If the following equation (3.52) has solution then the residual demand derivative at the slack bus is bounded; otherwise, the residual demand derivative at the slack bus goes to infinity.*

$$\begin{bmatrix} \mathbf{X}_{i_1}^T \\ \mathbf{X}_{i_2}^T \\ \vdots \\ \mathbf{X}_{i_l}^T \end{bmatrix} \boldsymbol{\beta} = \mathbf{1}_l, \quad (3.52)$$

where $\mathbf{1}_l = \underbrace{\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T}_l$.

Proof. 1) Suppose (3.52) has a solution $\hat{\beta}$. From (3.34), we have

$$\min_{\beta} SSE^{\text{WLS}}(\beta) \leq SSE^{\text{WLS}}(\hat{\beta}) = \sum_{i=1, i \notin \mathbf{N}}^{n-1} w_i \left(1 - \mathbf{X}_i^{\text{T}} \hat{\beta}\right)^2. \quad (3.53)$$

Therefore, if (3.52) has solution, the residual demand derivative at the slack bus is bounded.

2) Suppose (3.52) does not have a solution. Choose any $\tilde{\beta}$ that satisfies

$$\tilde{\beta} \in \arg \min_{\beta} \sum_{i \in \mathbf{N}} \left(1 - \mathbf{X}_i^{\text{T}} \beta\right)^2.$$

By assumption,

$$\sum_{i \in \mathbf{N}} \left(1 - \mathbf{X}_i^{\text{T}} \tilde{\beta}\right)^2 \neq 0. \quad (3.54)$$

From (3.34), we have

$$\begin{aligned} \min_{\beta} SSE^{\text{WLS}}(\beta) &= \min_{\beta} \sum_{i=1}^{n-1} w_i \left(Y_i - \mathbf{X}_i^{\text{T}} \beta\right)^2 \\ &\geq \min_{\beta} \sum_{i \in \mathbf{N}} w_i \left(1 - \mathbf{X}_i^{\text{T}} \beta\right)^2 \\ &\geq \min_{\beta} \sum_{i \in \mathbf{N}} \left(\min_{k \in \mathbf{N}} w_k\right) \left(1 - \mathbf{X}_i^{\text{T}} \beta\right)^2 \\ &= \left(\min_{k \in \mathbf{N}} w_k\right) \left(\min_{\beta} \sum_{i \in \mathbf{N}} \left(1 - \mathbf{X}_i^{\text{T}} \beta\right)^2\right). \end{aligned}$$

By the definition of $\tilde{\beta}$,

$$\min_{\beta} SSE^{\text{WLS}}(\beta) \geq \left(\min_{k \in \mathbf{N}} w_k\right) \sum_{i \in \mathbf{N}} \left(1 - \mathbf{X}_i^{\text{T}} \tilde{\beta}\right)^2. \quad (3.55)$$

Therefore,

$$\begin{aligned} &\lim_{w_i \rightarrow \infty, \forall i \in \mathbf{N}} \left(\min_{\beta} SSE^{\text{WLS}}(\beta)\right) \\ &\geq \lim_{w_i \rightarrow \infty, \forall i \in \mathbf{N}} \left(\left(\min_{k \in \mathbf{N}} w_k\right) \sum_{i \in \mathbf{N}} \left(1 - \mathbf{X}_i^{\text{T}} \tilde{\beta}\right)^2\right) = \infty. \end{aligned} \quad (3.56)$$

Therefore, if (3.52) does not have solution, the residual demand derivative at the slack bus is unbounded. \square

Generally speaking, if the number of buses with perfectly elastic supply is greater than the number of binding constraints, then the residual demand derivative at the slack bus is unbounded, because there are more equations than variables in (3.52), unless enough number of equations in (3.52) are redundant. When the residual demand derivative is bounded, it could be calculated from the WLS problem (3.34). Essentially, all buses that have perfectly elastic supply must have zero residual:

$$e_i^* = 0, \quad \forall i \in \mathbf{N}, \quad (3.57)$$

in order to zero out the arbitrarily large w_i , $\forall i \in \mathbf{N}$. We will show an example for this case in section 3.2. The WLS interpretation and the fact (3.57) have an important implication. Increasing the quantity-price response makes the electricity market more competitive. As shown in (3.35), larger quantity-price response will have a larger weight in the WLS problem. However, increased quantity-price response is not effective if binding transmission constraints prevent the large quantity-price responses from “spreading out” to the whole system. Our analysis allows the determination of the quantity-price response in the presence of transmission constraints.

3.1.5 Non-differentiable case

The OPF solution might be at a point of non-differentiability. In other words, the sensitivity with respect to the slack bus price is not defined. This occurs when there are just binding constraints, and/or the current OPF solution is at a point of non-differentiability of a supply function. If we encounter the case where the left side and right side residual demand derivative exist, but they are not equal, we could calculate the left side and right side residual demand derivative, respectively. If we can determine which constraints are binding as the price at bus n increases or decreases then the left side and right side residual demand derivatives are specified by (3.31) or can be calculated by the equivalent WLS problem for perfectly elastic supply/demand case with the corresponding binding constraints. Determining the binding constraints in general involves enumerating each possible combination of binding constraints and, for each, checking if the solution implied by sensitivity analysis will indeed induce the same the set of binding constraints. Similar analysis is also necessary in the case that the OPF solution is at a kink point of a supply function; that is, the left side derivative does not equal right side derivative.

3.2 TCRDD Examples

In the section, we verify the TCRDD calculation in 3 examples.

3.2.1 Example 1: intuitive 2-bus case

Consider the simplest case of a two-bus system connected by a single transmission line, and the line is congested.

Many researchers have adopted this example to illustrate the transmission effect on the equilibrium. (For example, see Borenstein, Bushnell, and Stoft [11].) It is observed that the market equilibrium just resembles the combination of two decoupled single bus system equilibria. We will demonstrate that the analysis in section 3.1 is consistent with this observation. We compute the residual demand at bus 2, and choose it as the slack bus. Because $m_b = 1$ and $n = 2$, $\bar{\mathbf{H}}_b$ is a 1×1 matrix, and $\bar{\mathbf{H}}_b = 1$. Substitute $\bar{\mathbf{H}}_b = 1$ into (3.28) and (3.29), we get $\frac{d\mu_b}{d\lambda}(\hat{\lambda}) = -1$ and $\frac{dq_1}{d\lambda}(\hat{\lambda}) = 0$, which implies that the supply at bus 1 is not affected by the price at bus 2. Also Substitute $\bar{\mathbf{H}}_b = 1$ into (3.31), and we get $\frac{dR_2}{d\lambda}(\hat{\lambda}) = 0$. Since we have combined the supply and demand at bus 2, so this implies that the residual demand at bus 2 is just the derivative of the actual local demand derivative at bus 2, and the market at bus 1 does not affect the residual demand derivative at bus 2. Similar results hold for bus 1. These results verify the intuitive result that the market is decoupled in this case.

3.2.2 Example 2: numerical 4-bus case

This example is illustrated in Fig. 3.1 We consider a two-loop system from [48]. Each branch admittance equals 0.1. There are two generators located at bus 1 and 2, and two loads located at bus 3 and 4.

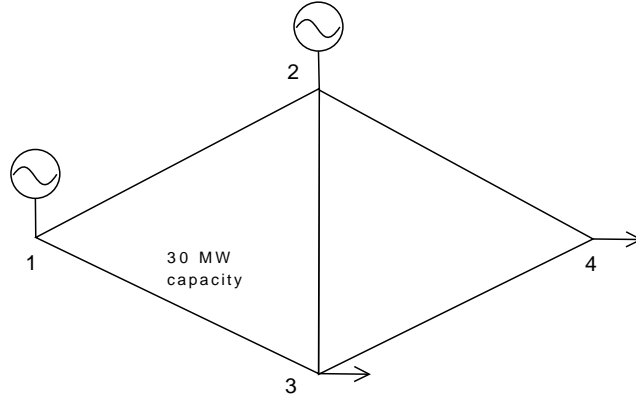


Figure 3.1: 4-bus system

The marginal generation cost functions and demand functions in bus order are:

Generation	Load
$p_{G1} = 10 + 0.35q_1$	$D_{L3} = 200 - 1.92p_3$
$p_{G2} = 10 + 0.45q_2$	$D_{L4} = 150 - 1.54p_4$

Branch 1–3 has a transmission capacity limit of 30 MW, all the other branches have transmission capacity limits of 400 MW. Following Xu and Yu, we consider the existence of a transmission-constrained linear SFE for this case. The transmission-constrained linear SFE

$$S_i(p_i) = \beta_i(p - \alpha_i), \quad i = 1, 2, \dots, n$$

is characterized by [48]

$$\frac{\beta_i^*}{1 - c_i \beta_i^*} = -R_i^*, \quad (3.58)$$

$$\alpha_i^* = e_i. \quad (3.59)$$

where generator i 's linear marginal cost function is $C'_i(q_i) = c_i q_i + e_i$.

$tol = 0.01$	Bus 1	Bus 2
$\hat{\beta}_i$	1/7.803	1/1.256
\hat{R}'_i	-0.1342	-1.2407
R'^*_i	-0.1290	-1.1824
RE_i	4%	5%
$\hat{\beta}_i$	1/8.343	1/1.329
\hat{R}'_i	-0.1251	-1.1377
R'^*_i	-0.1246	-1.1325
RE_i	0.4%	0.5%
$\hat{\beta}_i$	1/8.399	1/1.337
\hat{R}'_i	-0.1242	-1.1274
R'^*_i	-0.1241	-1.1276
RE_i	0.08%	0.02%

Table 3.1: Results comparison

Now we demonstrate that our analysis confirms to the numerical results in [48]. The way we verify it is as follows.

A linear supply function best response could be calculated directly from a generator's profit maximization problem as Xu and Yu did in [48]. Because the supply function best response and the residual demand derivative satisfy (2.2), we could solve for the residual demand derivative with a given supply function best response. The solution is (3.58). We used Xu and Yu's numerical supply function best response solutions as input, and solved for numerical residual demand derivatives using (3.58). Recall that we have derived the analytical solution of the residual demand derivative in (3.31). Therefore, we could compare our analytic solutions derived in this chapter with the numerical solutions recovered from Xu and Yu's results using (3.58).

In particular, denote the output of the algorithm in [48] by $\hat{\beta}$, which is the input to recover R_i^* using (3.31), and denote the so recovered R_i^* by \hat{R}_i' . Then we calculate analytically the R_i^* values evaluated at $\hat{\beta}$, and compare R_i^* calculated analytically to \hat{R}_i' from the output of the algorithm in [48]. The results are summarized in Table I.

Define the relative numerical error by

$$RE_i = \frac{|R_i^* - \hat{R}_i'|}{|R_i^*|} \times 100\%. \quad (3.60)$$

The smaller the RE_i , the more closely our results conform with Xu and Yu's results.

Table I shows that the relative error is in the range of 5% to 0.02%. Moreover, as the numerical accuracy of Xu and Yu's results increases as specified by a tolerance parameter in the stopping criterion of Xu and Yu's algorithm, RE_i decreases which indicates a better confirmation. The parameter Tol is the tolerance control parameter in [48], which controls the output accuracy of the bidding parameter $\hat{\beta}$. Note that smaller Tol results in smaller RE_i .

From the results we conclude that the characterization in this chapter is consistent with the numerical results in [48].

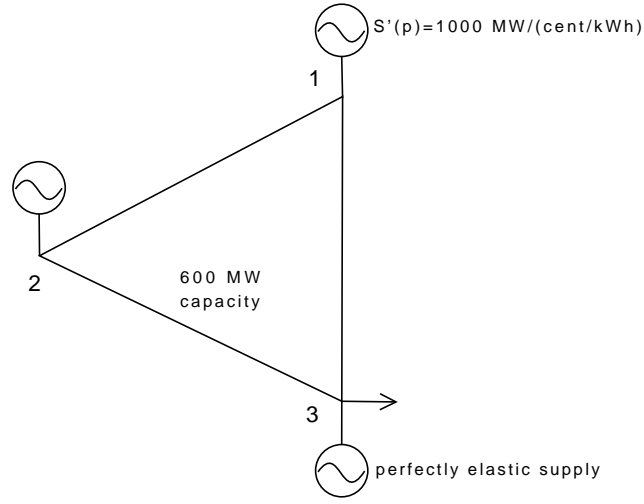


Figure 3.2: 3-bus example with perfectly elastic supply

3.2.3 Example 3: numerical 3-bus case with perfectly elastic supply

This example is taken from [12], and we simplify it by only considering information that is relevant to the residual demand derivative calculation. As illustrated in Fig. 3.2, it is a three-bus system with line 2–3 having a capacity 600 MW. All other lines have very large capacity, and cannot be congested. All three lines have the same impedance. There are three generators G1, G2, and G3 located at the corresponding buses. G1 bids a linear supply function with slope:

$$S'_1(p) = 1000 \text{ MW} / (\text{cent} / \text{kWh}),$$

and G3 has perfectly elastic supply. There is only one load, which is located at bus 3.

We want to calculate the residual demand derivative at bus 2, so we designate bus 2 as the slack bus. Since there is a perfectly elastic supply at bus 3, we can use the WLS interpretation technique to calculate the residual demand derivative as the limit as a weight approaches infinity. We have for this case

$$\bar{\mathbf{H}}_{\mathbf{b}} = \begin{bmatrix} -\frac{1}{3} & -\frac{2}{3} \end{bmatrix},$$

and where W_2 denotes an arbitrarily large number. Form the WLS problem as follows

$$\min_{\beta} \quad SSE^{\text{WLS}}(\beta) = 1000 \left(1 + \frac{1}{3}\beta\right)^2 + M \left(1 + \frac{2}{3}\beta\right)^2.$$

The solution of this problem is

$$b^{\text{WLS}} = \frac{-2000 - 2W_2}{\frac{2000}{3} + \frac{4W_2}{3}}.$$

Take the limit of b^{WLS} as $W_2 \rightarrow \infty$, we have:

$$\lim_{W_2 \rightarrow \infty} b^{\text{WLS}} = \lim_{W_2 \rightarrow \infty} \frac{-2000 - 2W_2}{\frac{2000}{3} + \frac{4W_2}{3}} = -\frac{3}{2}.$$

Therefore,

$$\begin{aligned} \lim_{W_2 \rightarrow \infty} \left(\min_{\beta} \quad SSE^{\text{WLS}}(\beta) \right) &= \lim_{W_2 \rightarrow \infty} (SSE^{\text{WLS}}(b^{\text{WLS}})), \\ &= 1000 \left(1 + \frac{1}{3} \left(-\frac{3}{2}\right)\right)^2 = 250. \end{aligned}$$

Therefore, the residual demand derivative at bus 2 is -250 .

Notice that the optimal b^{WLS} makes

$$1 + \frac{2}{3}b^{\text{WLS}} = 0.$$

Generally, the optimal WLS coefficients should make the residual, corresponding to a bus with perfectly elastic supply, equal to zero. Otherwise, the WLS problem will be unbounded.

In [12], Cardell, Hitt, and Hogan characterized the inverse of the residual demand derivative at bus 2 in equation (2). Their calculation is based on their intuitive price relationship:

$$p_2 = 2p_1 - p_3.$$

They obtained the solution -0.004 for the inverse residual demand derivative as indicated in equation (3.1) in [12]. Therefore, our calculation is consistent with this solution, because

$$\frac{1}{-0.004} = -250.$$

3.3 Improving TCRDD Calculation

In this section, we will discuss how to improve the TCRDD calculation to make it practical for real applications.

3.3.1 Practical issues with the TCRDD calculation to be resolved

For simplicity, the TCRDD formula (3.31), is derived under several assumptions. From the practical perspective, it has several limitations.

First, we assumed that each bus has at most one generator in deriving (3.31). This is not realistic for power system. In addition, offer functions and

marginal cost functions are typically modeled by piecewise linear functions in practice, because piecewise linear functions balance the accuracy and computational efforts very well. Existing advanced OPF solvers can handle piecewise linear offer functions. They typically model a piecewise linear offer function by multiple linear segments, and each segment has an associated variable. Each variable will have its capacity constraints to make sure it cannot go out of the segment. In this case, we may have multiple offer segments on every bus that may belong to different generators. Without loss of generality, we can model each generator offer segment as a “generator,” so we may have multiple “generators” at each bus. The first generalization is to handle multiple generators at a single bus.

Second, some electricity markets require step offer functions, such as the California nodal market. If the offer functions are step functions, then “generator” g will have $O_g''(\hat{q}_g) = 0$, which may make $\mathbf{\Lambda}$ in (3.30) undefined because of matrix singularity. Generally speaking, this is a modeling issue, because $O_g''(\hat{q}_g)$ is only used for calculating the TCRDD, and it is not a required information in an OPF solver. This issue can be handled by estimating the $O_g''(\hat{q}_g)$ information in the TCRDD calculation. Some estimating techniques are discussed in [44]. As illustrated in Fig. 3.3, a step offer function segment can be approximated by a straight line connecting middle points of adjacent segments in the TCRDD calculation, so that the dashed line will have a non-zero $O_g''(\hat{q}_g)$. Note that the $O_g''(\hat{q}_g)$ estimation is only needed for post-OPF TCRDD calculation, and does not affect the OPF solver. The TCRDD cal-

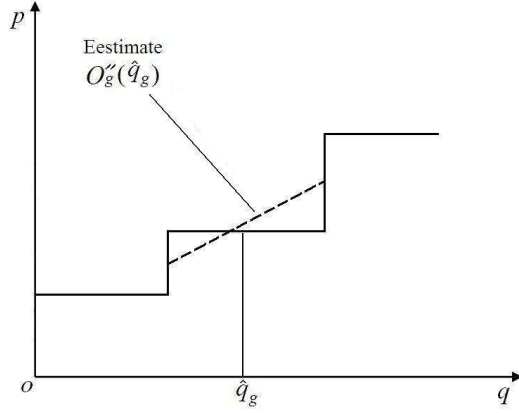


Figure 3.3: Estimating the derivative for step offer function

ulation as an OPF post-processing step is flexible as to the choice of OPF solver and offer function approximation algorithm.

Even with slope estimation, it is still possible to have a zero $O_g''(\hat{q}_g)$. An example is a generator bidding all its capacity at a constant price, and it is setting the LMP at its bus. In this case, we need to explicitly consider zero $O_g''(\hat{q}_g)$ in the TCRDD derivation in order to get an applicable formula.

Third, another problem arises when there are some generators' output binding at their capacity limit or a constant output level, in which case $O_g''(\hat{q}_g)$ in [44] does not exist. Treating piecewise segments as “generators” in order to handle piecewise offer functions makes this this much more likely to happen. For a generator with multiple offer segments, either all segments or all but one segment will be binding at the corresponding segment size limit. We need to explicitly consider this case in the TCRDD derivation as well.

3.3.2 Generalizing the TCRDD Calculation

To overcome the three limitations discussed in the previous section, we are going to derive a more general formula in this section. We consider the following simplified DC OPF model:

$$\min_{\mathbf{q}_G} \sum_{g \in G} O_g(q_g), \quad (3.61)$$

$$\text{s.t.} \quad \mathbf{H}_G \mathbf{q}_G - \mathbf{H}_L \mathbf{q}_L \leq \mathbf{Z}, \quad (3.62)$$

$$\mathbf{q}_G^{\min} \leq \mathbf{q}_G \leq \mathbf{q}_G^{\max}, \quad (3.63)$$

$$\sum_{l \in L} q_l - \sum_{g \in G} q_g = 0, \quad (3.64)$$

where

- G is the set of generators, with each segment of an offer represented by a distinct element,
- L is the set of loads,
- \mathbf{q}_G is the generator output variable vector,
- \mathbf{q}_L is the load vector,
- $O_g(q_g)$ is generator g 's total offer cost function, whose derivative, $O'_g(q_g)$, is generator g 's offer function,
- \mathbf{H}_G is the generator shift factor matrix corresponding to the transmission constraints,

- \mathbf{H}_L is the load shift factor matrix corresponding to the transmission constraints,
- \mathbf{q}_G^{\min} consists of generator output capacity lower limits,
- \mathbf{q}_G^{\max} consists of generator output capacity upper limits,
- \mathbf{Z} consists of the transmission capacity limits,
- (3.62) consists of the transmission constraints,
- (3.63) is the generator capacity constraint, and
- (3.64) is the energy balance constraint.

Although this OPF model seems very simple, it is general enough to capture the advanced features in production level OPF programs.¹ Currently, most production level OPF programs in existing and proposed nodal electricity markets, are either DC, or solve AC by successive linearization, such as the CAISO nodal market and the proposed ERCOT nodal market. For DC OPF, our model and results are directly applicable. For AC OPF solved by successive linearization, the typical scheme is discussed in [13], where the AC OPF is decoupled into two subproblems, namely the optimization problem and the network assessment problem. The network assessment problem is to solve

¹This is a single interval OPF model, which is applicable to most the real-time electricity markets in the US, such as PJM, ISO-NE, and ERCOT. Very few electricity markets, such as the California nodal market, solve a multi-interval OPF with looking-ahead capability in the real-time market, where they include the inter-temporal constraints, such as the ramp constraints. Handling these inter-temporal constraints is out of the scope of this dissertation.

the AC power flow and contingency analysis, and generate a list of overloaded (and/or nearly overloaded) lines to be passed to the optimization problem as constraints. It also calculates certain sensitivities, such as the shift factors, so that the constraints can be modeled linearly in the optimization problem. The optimization problem solves the OPF program using the linearized constraints passed from the network assessment problem. The optimization problem and network assessment problem will be solved iteratively, so that if the process converges, it converges to a solution to the original OPF problem. The advantage of this decoupled method is that the optimization problem with only linearized constraints is much easier to solve than the original problem, so this method is able to solve large scale problem. The linearized OPF optimization problem in production level OPF program is essentially the same as the DC OPF problem we consider in this chapter. The security constraints and stability constraints can be modeled in (3.62) with the \mathbf{H}_G and \mathbf{H}_L being the outage compensated shift factors ([45] chapter 11 Appendix), and the Nomo-gram shift factors. Therefore, the DC OPF model and results are also directly applicable to an AC OPF solved by successive linearization.

For pure AC OPF solved by nonlinear programming, it requires extra work to be able to use the method in this chapter. Basically, one needs to calculate the shift factors and other sensitivities in order to formulate the problem as (3.61) – (3.64).

Therefore, generally speaking, working with the DC OPF model (3.61) – (3.64) is not limiting the value of this work. The purpose of this work is not

to compete with the most advanced existing OPF programs, but to make the best use of them to develop bidding strategies rather than start from scratch.

This OPF model is more general than the one used in section 3.1, because multiple generators at the same bus can be explicitly modeled using the generator shift factors instead of node shift factors. Also, we separate the transmission constraints and generation capacity constraints here, whereas section 3.1 considers them together.

The subsequent derivation largely follows section 3.1, and we will only cover the steps that are different from section 3.1 in order to overcome the three limitations discussed in the previous section.

As discussed in section 3.1, the TCRDD is a post-OPF calculation. An OPF variable with a hat represents the OPF solution, and the binding transmission constraints are denoted by a subscript “b”. We assume the market is cleared by LMPs determined from the OPF solution. Denote the LMP at the slack bus by λ . The OPF solves at the slack bus price $\lambda = \hat{\lambda}$. Following section 3.1, without loss of generality, we calculate the TCRDD for a generator s located at the slack bus n . Because the TCRDD is a post-OPF calculation, any bus can be designated as the “slack” bus for TCRDD calculation purpose, and it is not necessary to use the same slack bus that is used in the OPF.

Partition the generators other than generator s into three subsets: the generator segments with binding output quantities, denoted by \mathbf{f} ; the generator segments with binding price offers, denoted by \mathbf{z} ; and the generator segments

with offers having non-zero slopes, denoted by \mathbf{v} , so that

$$\mathbf{q}_G = \begin{bmatrix} \mathbf{q}_v \\ \mathbf{q}_f \\ \mathbf{q}_z \\ q_s \end{bmatrix}.$$

Accordingly, partition \mathbf{H}_{Gb} into:

$$\mathbf{H}_{Gb} = \begin{bmatrix} \mathbf{H}_{vb} & \mathbf{H}_{fb} & \mathbf{H}_{zb} & 0 \end{bmatrix}.$$

Similarly to section 3.1, we consider the OPF solution to be parameterized by the price at the slack bus λ , with the TCRDD is defined by

$$\begin{aligned} \frac{dR_G}{d\lambda}(\hat{\lambda}) &= \frac{d\left(\sum_{l \in \mathbf{L}} q_l - \sum_{g \in G, g \neq s} q_g\right)}{d\lambda} \\ &= - \sum_{g \in G, g \neq s} \frac{dq_g}{d\lambda}(\hat{\lambda}) \\ &= - \sum_{g \in \mathbf{v}} \frac{dq_g}{d\lambda}(\hat{\lambda}) - \sum_{g \in \mathbf{f}} \frac{dq_g}{d\lambda}(\hat{\lambda}) - \sum_{g \in \mathbf{z}} \frac{dq_g}{d\lambda}(\hat{\lambda}) \\ &= -\mathbf{1}_v^T \frac{d\mathbf{q}_v}{d\lambda}(\hat{\lambda}) - \mathbf{1}_f^T \frac{d\mathbf{q}_f}{d\lambda}(\hat{\lambda}) - \mathbf{1}_z^T \frac{d\mathbf{q}_z}{d\lambda}(\hat{\lambda}), \end{aligned} \tag{3.65}$$

where $\mathbf{1}_v$, $\mathbf{1}_f$, and $\mathbf{1}_z$ are column vectors of 1s whose dimensions equal the number of generators in set \mathbf{v} , \mathbf{f} , and set \mathbf{z} respectively.

Because by definition,

$$\frac{d\mathbf{q}_f}{d\lambda}(\hat{\lambda}) = \mathbf{0}_f, \tag{3.66}$$

we only need to calculate $\frac{d\mathbf{q}_v}{d\lambda}(\hat{\lambda})$ and $\frac{d\mathbf{q}_z}{d\lambda}(\hat{\lambda})$ in order to compute the TCRDD.

Construct the Lagrangian of the OPF with only the binding constraints:

$$\begin{aligned}\mathcal{L} = & \sum_{g \in \mathbf{G}} O_g(q_g) + \lambda \left(\sum_{l \in \mathbf{L}} q_l - \sum_{g \in \mathbf{G}} q_g \right) \\ & + \boldsymbol{\mu}_b^T (\mathbf{H}_{vb} \mathbf{q}_v + \mathbf{H}_{fb} \mathbf{q}_f + \mathbf{H}_{zb} \mathbf{q}_z - \mathbf{H}_{lb} \mathbf{q}_L - \mathbf{Z}_b) \\ & + \boldsymbol{\rho}_{\max}^T (\mathbf{q}_f - \mathbf{q}_f^{\max}) + \boldsymbol{\rho}_{\min}^T (\mathbf{q}_f - \mathbf{q}_f^{\min}).\end{aligned}$$

Similarly to section 3.1, we calculate $\frac{d\mathbf{q}_v}{d\lambda}(\hat{\lambda})$ and $\frac{d\mathbf{q}_z}{d\lambda}(\hat{\lambda})$ as well as $\frac{d\boldsymbol{\mu}_b}{d\lambda}(\hat{\lambda})$ from a sensitivity analysis of the following OPF first-order necessary conditions:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \mathbf{q}_v} &= \mathbf{0}_v, \\ \frac{\partial \mathcal{L}}{\partial \mathbf{q}_z} &= \mathbf{0}_z, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}_b} &= \mathbf{0}_b,\end{aligned}$$

i.e.

$$\begin{aligned}\mathbf{O}'_v(\mathbf{q}_v) - \lambda \mathbf{1}_v - \mathbf{H}_{vb}^T \boldsymbol{\mu}_b &= \mathbf{0}_v, \\ \mathbf{O}'_z(\mathbf{q}_z) - \lambda \mathbf{1}_z - \mathbf{H}_{zb}^T \boldsymbol{\mu}_b &= \mathbf{0}_z,\end{aligned}\tag{3.67}$$

$$\mathbf{H}_{vb} \mathbf{q}_v + \mathbf{H}_{fb} \mathbf{q}_f + \mathbf{H}_{zb} \mathbf{q}_z - \mathbf{H}_{lb} \mathbf{q}_L = \mathbf{Z}_b,$$

where

- $\mathbf{O}'_v(\mathbf{q}_v) = \nabla_{\mathbf{q}_v} \left(\sum_{g \in \mathbf{G}} O_g(q_g) \right)$,
- $\mathbf{O}'_z(\mathbf{q}_z) = \nabla_{\mathbf{q}_z} \left(\sum_{g \in \mathbf{G}} O_g(q_g) \right) = \mathbf{p}_z$, i.e. binding offer prices,
- $\mathbf{0}_v$, $\mathbf{0}_f$, and $\mathbf{0}_z$ are column vectors of 0s whose dimensions equal the number of generators in set v , f , and z respectively.

Differentiate both sides of (3.67) with respect to λ , we get

$$\begin{aligned} \mathbf{O}_v'' \frac{d\mathbf{q}_v}{d\lambda} - \mathbf{1}_v - \mathbf{H}_{vb}^T \frac{d\boldsymbol{\mu}_b}{d\lambda} &= \mathbf{0}_v, \\ -\mathbf{1}_z - \mathbf{H}_{zb}^T \frac{d\boldsymbol{\mu}_b}{d\lambda} &= \mathbf{0}_z, \\ \mathbf{H}_{vb} \frac{d\mathbf{q}_v}{d\lambda} + \mathbf{H}_{zb} \frac{d\mathbf{q}_z}{d\lambda} &= \mathbf{0}_b, \end{aligned} \tag{3.68}$$

where

$$\mathbf{O}_v'' = \nabla_{\mathbf{q}_v \mathbf{q}_v}^2 \left(\sum_{g \in G} O_g(q_g) \right).$$

If the number of binding price offers in set z is less than the number of binding transmission constraints, and the number of binding transmission constraints is less than the number of non-binding offers in set v , then (3.68) has a unique solution:

$$\begin{aligned} \frac{d\mathbf{q}_v}{d\lambda}(\hat{\lambda}) &= (\mathbf{O}_v'')^{-1} \left(\mathbf{1}_v + \mathbf{H}_{vb}^T \frac{d\boldsymbol{\mu}_b}{d\lambda}(\hat{\lambda}) \right) \\ \frac{d\mathbf{q}_z}{d\lambda}(\hat{\lambda}) &= \mathbf{Q}^{-1} \mathbf{1}_z - \mathbf{Q}^{-1} \mathbf{H}_{zb}^T \mathbf{M}^{-1} \mathbf{H}_{vb} (\mathbf{O}_v'')^{-1} \mathbf{1}_v, \\ \frac{d\boldsymbol{\mu}_b}{d\lambda}(\hat{\lambda}) &= -\mathbf{M}^{-1} \left(\mathbf{H}_{zb} \frac{d\mathbf{q}_z}{d\lambda}(\hat{\lambda}) + \mathbf{H}_{vb} (\mathbf{O}_v'')^{-1} \mathbf{1}_v \right) \end{aligned} \tag{3.69}$$

where

$$\begin{aligned} \mathbf{M} &= \mathbf{H}_{vb} (\mathbf{O}_v'')^{-1} \mathbf{H}_{vb}^T, \\ \mathbf{Q} &= \mathbf{H}_{zb}^T \mathbf{M}^{-1} \mathbf{H}_{zb}. \end{aligned}$$

The TCRDD can be calculated by substituting (3.66) and (3.69) into (3.65).

The whole derivation process follows section 3.1 with necessary changes to overcome its limitations. One change is in the modeling. All generators

except s are explicitly modeled here, and are divided into three sets, namely the generators with binding capacity constraints, the generators with binding price offers, and the generators with varying prices and output quantities. The three sets of generators are treated differently. Transmission constraints and generator capacity constraints are also treated differently. The new derivation overcomes the limitations in section 3.1 discussed in the previous section, and it is more widely applicable.

Another change is in the result. The new TCRDD formula only depends on binding transmission constraints and the offers of the generators having non-zero slopes, while the formula in section 3.1 depends on both transmission constraints and generation capacity constraints. If the offers are piecewise linear or piecewise continuous functions, there will be a lot of binding generation capacity segments, which makes the matrix bigger and the calculation inefficient. Comparing with the formula in section 3.1, this new formula improves the computational performance.

What remains unchanged from section 3.1 is the concept of the TCRDD. Similar to section 3.1, note that the LMP of the slack bus is a dual variable of the OPF, so this is a variable to variable sensitivity analysis. Generator s 's offer function does not factor into this sensitivity analysis (it does factor into the OPF solution), so the sensitivities here evaluated at the OPF solution are independent of the slack bus generator's offer function parameters. This unique property makes this kind of sensitivity different from the ordinary variable to parameter sensitivities, such as in [14] and [32].

3.4 Maximize A Generator's Profit Using the TCRDD

As discussed in section 1.3, the profit maximization can be calculated based on the given residual demand curve, which suggests that the generator's profit maximization problem can be decoupled into two subproblems: the upper subproblem of maximizing the generator's profit based on TCRDD, and the lower subproblem of calculating the TCRDD. This gives rise to the new generator profit maximization approach in this chapter, and we will refer it as the TCRDD approach hereafter. The TCRDD approach is illustrated in Fig. 3.4. Note that in the TCRDD approach, the upper problem and lower problem are not in nested structure, and can be solved separately. The whole problem can be solved by iteratively solving the two subproblems. This is a significant advantage over the MPEC approach because solving the two smaller problems is much easier than solving the problem as a whole. In addition, the lower subproblem is a standard OPF problem with a lightweight post-processing step of the TCRDD calculation, so that existing advanced OPF algorithms and solvers can be reused. This is another significant advantage over the MPEC approach.

In this section, we propose an algorithm to maximize a generator's profit utilizing the TCRDD information that can be calculated in the way discussed in section 3.3.

To make the idea easy to understand, we will illustrate the algorithm in a very simply 4-bus system example. A larger scale 118-bus system example will be solved and discussed in section 3.5. The 4-bus example is very similar

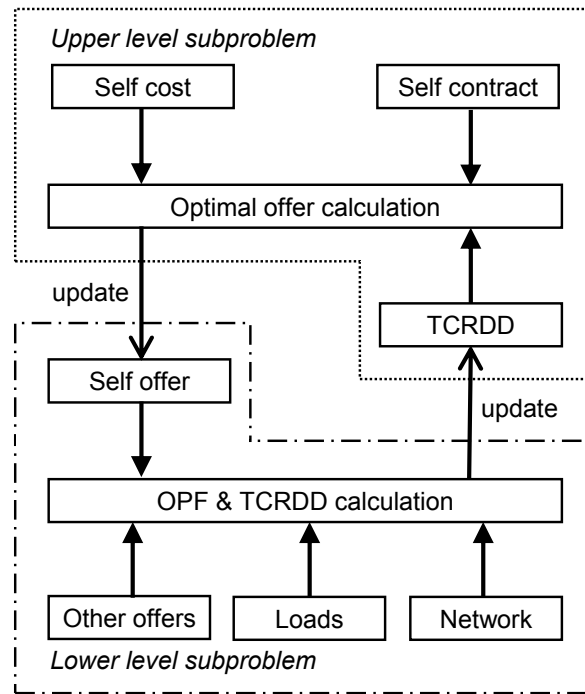


Figure 3.4: Maximizing generation profit based on TCRDD

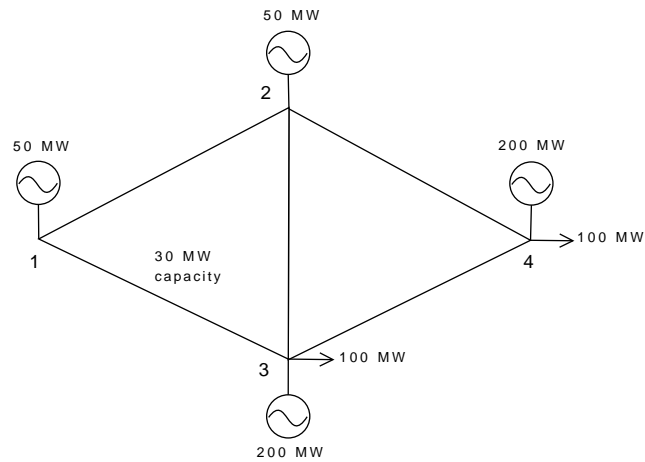


Figure 3.5: 4-bus system

to the one in section 3.2.2. As illustrated in Fig. 3.5, there are four generators in the system, and each one is located at different bus. Generator 1 and 2 are 50 MW unit each. Generator 3 and 4 are 200 MW each. There are two loads, 100 MW each, located at bus 3 and bus 4 respectively. Branch 1-3 has a capacity limit of 30 MW. All other branches have capacities of 200 MW. All branches have the same impedances. The generators' cost functions are

$$C_1(q_1) = 0.175q_1^2 + 10q_1,$$

$$C_2(q_2) = 0.497q_2^2 + 10q_2,$$

$$C_3(q_3) = 0.260q_3^2 + 20q_3,$$

$$C_4(q_4) = 0.325q_4^2 + 20q_4.$$

Assume all generators bid in their true cost. We plot generator 1's and generator 2's profit functions and residual demand curves in Fig. 3.6 and Fig. 3.7 respectively.

From now on, to simplify notation, we drop the subscript i in all the variables and functions, p , q , RDD , Π and P , meaning that all the variables and functions belong to the generator under consideration by default.

Generator 2 has a linear residual demand curve. Suppose the OPF solves at q_0 and $p_0 = P(q_0)$ for generator 2, and the TCRDD evaluated at q_0 is $RDD(q_0)$. We represent the inverse function of the residual demand curve, $P(q)$, by its tangent $P(q; q_0)$ at q_0 :

$$P(q; q_0) = \frac{1}{RDD(q_0)}(q - q_0) + p_0, \quad (3.70)$$

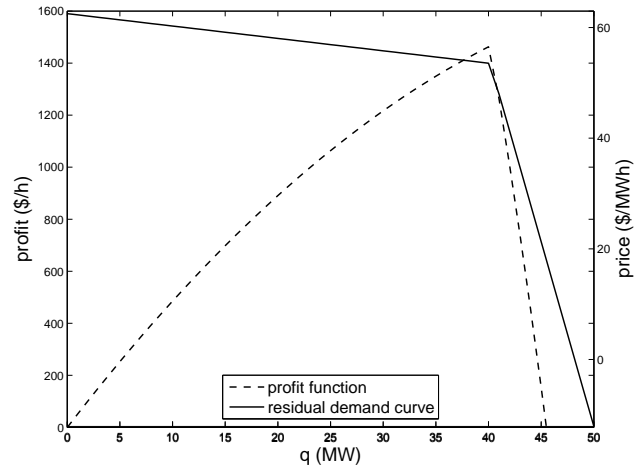


Figure 3.6: Profit function and residual demand curve for generator 1

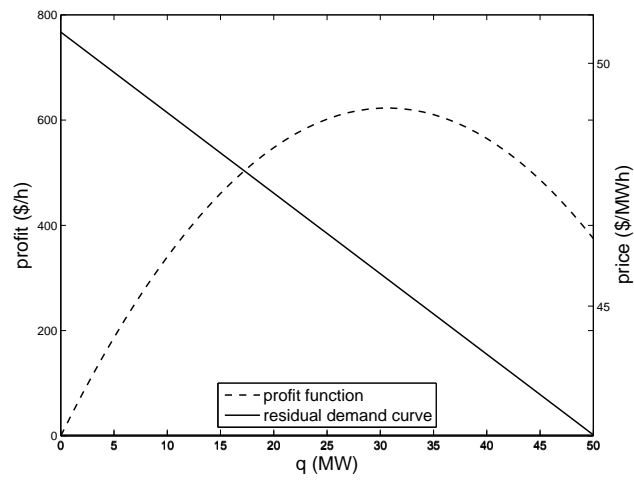


Figure 3.7: Profit function and residual demand curve for generator 2

which is the solid line in Fig. 3.7. In this case, $P(q; q_0)$ matches $P(q)$ exactly.

We also represent generator 2's profit $\Pi(q)$ by $\bar{\Pi}(q; q_0)$ using $P(q; q_0)$ in (1.3):

$$\bar{\Pi}(q; q_0) = \left(\frac{1}{RDD(q_0)}(q - q_0) + p_0 \right) q - C(q). \quad (3.71)$$

In this case, $\bar{\Pi}(q; q_0)$ also matches $\Pi(q)$ exactly, which is the dashed curve shown in Fig. 3.7. Because generator 2's cost function is quadratic, the profit function is also quadratic.

The situation is more complicated for generator 1, because the residual demand curve may have kinks when the set of binding transmission constraints changes. This is demonstrated in Fig. 3.6, if generator 1's output is less than 43 MW, branch 1-3 is not congested; while if generator 1's output is greater than 43 MW, branch 1-3 is congested (i.e. a binding constraint). The kink in the residual demand curve will also result in a kink in the profit function. In Fig. 3.6, the profit function is composed of two quadratic pieces with the curvature change right at the kink, q_x , of the residual demand curve.

Suppose we have two solved OPFs: one OPF solves at (q_{lo}, p_{lo}) , with $q_{lo} \leq q_x$, and the TCRDD evaluated at q_{lo} is $RDD(q_{lo})$; the other OPF solves at (q_{hi}, p_{hi}) , with $q_{hi} \geq q_x$, and the TCRDD evaluated at q_{hi} is $RDD(q_{hi})$. We represent the inverse function of generator 1's residual demand curve, $P(q)$, by two tangent lines at q_{lo} and q_{hi} respectively:

$$\begin{aligned} P(q; q_{lo}) &= \frac{1}{RDD(q_{lo})}(q - q_{lo}) + p_{lo}, \quad \forall q \leq q_x, \\ P(q; q_{hi}) &= \frac{1}{RDD(q_{hi})}(q - q_{hi}) + p_{hi}, \quad \forall q > q_x. \end{aligned} \quad (3.72)$$

Accordingly, we represent the profit function $\Pi(q)$ by

$$\begin{aligned}\bar{\Pi}(q; q_{lo}) &= \left(\frac{1}{RDD(q_{lo})} (q - q_{lo}) + p_{lo} \right) q - C(q), \forall q \leq q_x, \\ \bar{\Pi}(q; q_{hi}) &= \left(\frac{1}{RDD(q_{hi})} (q - q_{hi}) + p_{hi} \right) q - C(q), \forall q > q_x,\end{aligned}\tag{3.73}$$

and the two quadratic segments intersect at q_x . In this case, our representations of the residual demand curve and the profit function are both exact.

Note that in this case the profit function is still concave, but in other cases it may end up with two local optima with each of the residual demand segment containing one local optimum, which also makes the profit function not concave. It is generally difficult to find the global optimizer, and the algorithm in this section is aiming at finding a local optimizer as most other methods do, such as the MPEC method [23]. However, the approach in this chapter can be provided with different initial points based on knowledge of the kink to explore a broader region in order to get closer to the global optimizer.

If the whole residual demand curve is known, the task to find a local profit optimizer is not difficult. However, constructing the whole residual demand curve is computationally intense when there are changing binding constraints, because one solved OPF only produces one point on the residual demand curve. Conceptually, one can continuously solve OPFs to trace out the residual demand curve for a generator as we do in Fig. 3.6 and Fig. 3.7 for the 4-bus system example, but for the purpose of profit maximization, tracing out the whole residual demand curve is inefficient and unnecessary.

Because the decision variable for a generator is its output level, the

problem to find a local optimum is basically a line search. There are various existing standard *inexact* line search algorithms, such as the Wolfe condition and the Armijo-Goldstein condition ([34] chapter 3), that are not designed for finding an exact solution. An exact line search typically requires a large number of function evaluations ([34] chapter 3), which may not be efficient for this specific problem of a generator's profit, because evaluating a generator's profit involves solving the OPF, which is computationally intense for large scale systems.

In order to improve the performance, we developed a special algorithm aimed at requiring less profit function evaluations. As will be demonstrated, the algorithm will be able to find the profit optimum for generator 1 or generator 2 in the 4-bus system example within one iteration. Its performance will be further tested in the IEEE standard 118-bus system example in section 3.5.

The algorithm we propose is a special bisection search scheme based on approximating the residual demand curve by a two-piece linear function as the generator 1's residual demand curve in the 4-bus system example.

Suppose we have two output levels q_{lo} and q_{hi} with $q_{lo} < q_{hi}$. The residual demand curve can have many segments in $[q_{lo}, q_{hi}]$ depending on the set of binding constraints while changing its output level. Without the exact knowledge of residual demand curve in $[q_{lo}, q_{hi}]$, we estimate the residual demand curve in a similar way as in (3.72). Denote the estimated two-piece quadratic

profit function by

$$\begin{aligned}\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}}) &= \bar{\Pi}(q; q_{\text{lo}}), \quad q \in [q_{\text{lo}}, \max\{q_{\text{lo}}, q_x\}], \\ \hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}}) &= \bar{\Pi}(q; q_{\text{hi}}), \quad q \in [\min\{q_{\text{hi}}, q_x\}, q_{\text{hi}}],\end{aligned}\tag{3.74}$$

where q_x denotes an intersection of $P(q; q_{\text{lo}})$ and $P(q; q_{\text{hi}})$ if they intersect, otherwise, let $q_x = q_{\text{hi}}$.

Function $\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}})$ exactly match the profit function in the vicinity of $\Pi(q)$ at q_{lo} and q_{hi} , but may differ from the profit function if the residual demand curve differ from the estimated two-piece residual demand function evaluated at q_{lo} and q_{hi} respectively.

Denote the maximizer of $\bar{\Pi}(q; \bullet)$ by ²

$$\bar{q}(\bullet) = \underset{q}{\operatorname{argmax}}\{\bar{\Pi}(q; \bullet)\}.$$

We know there exists a local maximum in $[q_{\text{lo}}, q_{\text{hi}}]$ if

$$\begin{aligned}q_{\text{lo}} &\leq \bar{q}(q_{\text{lo}}), \\ q_{\text{hi}} &\geq \bar{q}(q_{\text{hi}}),\end{aligned}\tag{3.75}$$

because the profit function is continuous. We also know if (3.75) is satisfied,

$$\operatorname{argmax}\{\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}})\} \subseteq \{\bar{q}(q_{\text{lo}}), \bar{q}(q_{\text{hi}}), q_x\},\tag{3.76}$$

because $\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}})$ is a two-piece quadratic function. If $q_{\text{lo}} \leq q_x \leq q_{\text{hi}}$, $\operatorname{argmax}\{\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}})\}$ can be determined as specified in Tab. 3.2 and illustrated in Fig. 3.8.

²Strictly speaking, $\operatorname{argmax}\{\bullet\}$ represents the set of maximizers, but because it is a singleton for a quadratic objective function, we use the notation the maximizer “=” $\operatorname{argmax}\{\bullet\}$.

condition	characteristic	maximizer of $\hat{\Pi}(q; q_{lo}; q_{hi})$
$\bar{q}(q_{lo}) \notin [q_{lo}, q_x], \bar{q}(q_{hi}) \notin [q_x, q_{hi}]$	no hump	q_x
$\bar{q}(q_{lo}) \in [q_{lo}, q_x], \bar{q}(q_{hi}) \notin [q_x, q_{hi}]$	left hump	$\bar{q}(q_{lo})$
$\bar{q}(q_{lo}) \notin [q_{lo}, q_x], \bar{q}(q_{hi}) \in [q_x, q_{hi}]$	right hump	$\bar{q}(q_{hi})$
$\bar{q}(q_{lo}) \in [q_{lo}, q_x], \bar{q}(q_{hi}) \in [q_x, q_{hi}]$	double hump	$\bar{q}(q_{lo})$ or $\bar{q}(q_{hi})$

Table 3.2: Determine Maximizer Of $\hat{\Pi}(q; q_{lo}; q_{hi})$

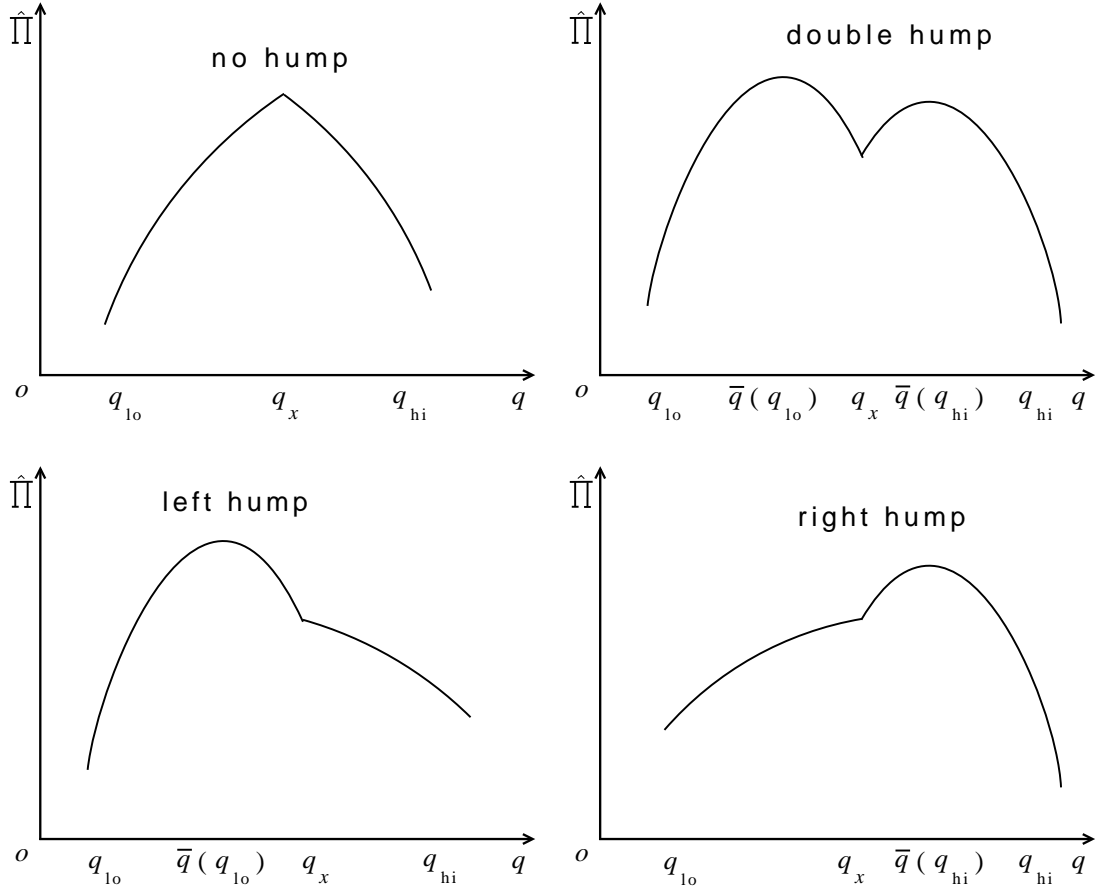


Figure 3.8: Determine Maximizer Of $\hat{\Pi}(q; q_{lo}; q_{hi})$

Based on the profit function approximation $\arg\max\{\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}})\}$, we design a special bisection search to find the optimizer. It has two consecutive parts. The first part is to search for an interval $[q_{\text{lo}}, q_{\text{hi}}]$ that satisfies (3.75), and we call this part the local optimum searching loop. The second part is to find a point in $[q_{\text{lo}}, q_{\text{hi}}]$ that satisfies (1.4) through a special bisection procedure, and we call it the bisection loop.

The bisection procedure is special in that the bisection point is chosen from the candidate maximizers of $\hat{\Pi}(q; q_{\text{lo}}; q_{\text{hi}})$, instead of the standard choice of $0.5(q_{\text{lo}} + q_{\text{hi}})$, unless the candidate choices $\{\bar{q}(q_{\text{lo}}), \bar{q}(q_{\text{hi}}), q_x\} \subseteq \{q_{\text{lo}}, q_{\text{hi}}\}$.

The whole algorithm is as follows.

Local Optimum Searching Loop:

1. Start with an initial point $q = q^0$.
2. Solve OPF with $q = q^0$, and calculate $RDD(q_0)$ and $\bar{q}(q_0)$.
3. If $\bar{q}(q^0) = q^0$, optimal solution found with $q^* = q^0$, stop.
 Or if $q^0 = q^{\max}$ and $\bar{q}(q_0) > q^{\max}$, optimal solution found with $q^* = q^{\max}$, stop.
 Or if $q^0 = q^{\min}$ and $\bar{q}(q_0) < q^{\min}$, optimal solution found with $q^* = q^{\min}$, stop.
4. If $\bar{q}(q_0) > q^0$, let $q^1 = \min\{\bar{q}(q^0), q^{\max}\}$, else let $q^1 = \max\{\bar{q}(q^0), q^{\min}\}$.
 Solve OPF with $q = q^1$, and calculate $RDD(q^1)$ and $\bar{q}(q^1)$.

5. If $\bar{q}(q^1) = q^1$, optimal solution found with $q^* = q^1$, stop.
 Or if $q^1 = q^{\max}$ and $\bar{q}(q_1) > q^{\max}$, optimal solution found with $q^* = q^{\max}$, stop.
 Or if $q^1 = q^{\min}$ and $\bar{q}(q_1) < q^{\min}$, optimal solution found with $q^* = q^{\min}$, stop.
6. If $(\bar{q}(q^1) - q^1)(\bar{q}(q^0) - q^0) < 0$, local optimum exists in

$$[\min\{q^0, q^1\}, \max\{q^0, q^1\}].$$

Let $q_{\text{lo}} = \min\{q^0, q^1\}$, and $q_{\text{hi}} = \max\{q^0, q^1\}$, stop.

Otherwise, $q^0 = q^1$, and continue with step 4.

Bisection Loop:

1. If $q_{\text{hi}} - q_{\text{lo}} < \epsilon$, where $\epsilon > 0$ is the tolerance threshold, optimal solution found with $q^* = 0.5(q_{\text{lo}} + q_{\text{hi}})$. Stop.
2. Calculate the bisection point q_{mid} as follows.
 - If $P(q; q_{\text{lo}})$ and $P(q; q_{\text{hi}})$ specify the same function, $q_{\text{mid}} = \bar{q}(q_{\text{lo}})$.
 - If $P(q; q_{\text{lo}})$ and $P(q; q_{\text{hi}})$ do not intersect, or they intersect at (q_x, p_x) with $q_x \notin [q_{\text{lo}}, q_{\text{hi}}]$, then $q_{\text{mid}} = 0.5(q_{\text{lo}} + q_{\text{hi}})$;
 - Otherwise, determine the bisection point q_{mid} as specified in Tab. 3.2.

For the double hump case, let

$$q_{\text{mid}} = \underset{q \in \{\bar{q}(q_{\text{lo}}), \bar{q}(q_{\text{hi}})\}}{\operatorname{argmin}} |q - 0.5(q_{\text{lo}} + q_{\text{hi}})|.$$

3. Solve OPF with $q = q_{\text{mid}}$, and calculate $RDD(q_{\text{mid}})$ and $\bar{q}(q_{\text{mid}})$.
4. If $\bar{q}(q_{\text{mid}}) = q_{\text{mid}}$, optimal solution reached with $q^* = q_{\text{mid}}$. Stop.
5. If $q_{\text{mid}} = q_x$ and $p_{\text{mid}} = p_x$. Do an incremental test as follows.
 - If $RDD(q_{\text{mid}}) = RDD(q_{\text{lo}})$, run OPF with $q = q_x + \epsilon$, and calculate $RDD(q_x + \epsilon)$, $\bar{q}(q_x + \epsilon)$. If $\bar{q}(q_x + \epsilon) \leq q_x + \epsilon$, then optimal solution reached with $q^* = q_x$ because q_x satisfies (1.4). Stop.
 - If $RDD(q_{\text{mid}}) = RDD(q_{\text{hi}})$, run OPF with $q = q_x - \epsilon$, and calculate $RDD(q_x - \epsilon)$, $\bar{q}(q_x - \epsilon)$. If $\bar{q}(q_x - \epsilon) \geq q_x - \epsilon$, then optimal solution reached with $q^* = q_x$ because q_x satisfies (1.4). Stop.
6. If $\bar{q}(q_{\text{mid}}) > q_{\text{mid}}$, $q_{\text{lo}} = q_{\text{mid}}$, else $q_{\text{hi}} = q_{\text{mid}}$. Continue with step 1.

After the local optimum searching loop, we either have found a local optimum, or we end up with two output levels q_{lo} and q_{hi} that satisfies (3.75) so that we can enter the bisection loop. In the 4-bus system example, for generator 2, the local optimum searching loop will find the maximizer in step 5 after one iteration; for generator 1, the local optimum searching loop will find two output levels q_{lo} and q_{hi} that satisfies (3.75) after one iteration.

The bisection loop has a quadratic rate of convergence if the residual demand curves that contain the local optimizer have been correctly identified. In the 4-bus system example, for either generator 1, the bisection loop will find the maximizer in step 5 after one iteration.

3.5 Computational Example

In this section, we apply the algorithm to the IEEE 118 bus test system. There are 186 branches and 54 generators in the system. The total load in the system is 4,242 MW. We optimize the profit for generator 5 located at bus 10 with 550 MW capacity. Branches 30-17, 26-30, 38-37 have capacities 200 MW respectively, so they are likely to be the binding transmission constraints. All other branches have capacities large enough such that the flows will be within their limits. We plotted the residual demand curve and the profit function in Fig. 3.9.

From the profit curve in Fig. 3.9, one can tell the maximizer is between 344 MW and 345 MW. We start with 40 MW for the optimization, which is far away from the optimizer. The local optimum searching loop terminates after one iteration with $q_{lo} = 40$ and $q_{hi} = 439.4$. The bisection terminated after two iterations with the optimal solution $q^* = 344.76$. The first bisection iteration is a right hump case, and the second bisection iteration is a left hump case, as specified in Tab. 3.2. The bisection trajectory is illustrated in Fig. 3.10.

3.6 Conclusion

We characterize the residual demand at the slack bus based on the FONCs of the OPF problem. The residual demand curve is implicitly characterized by eliminating the equations in the FONCs that contain the offer information of the generator located at the slack bus. After doing that, there is one degree of freedom left in the FONCs that defines a locus of (λ, q_n) , i.e.

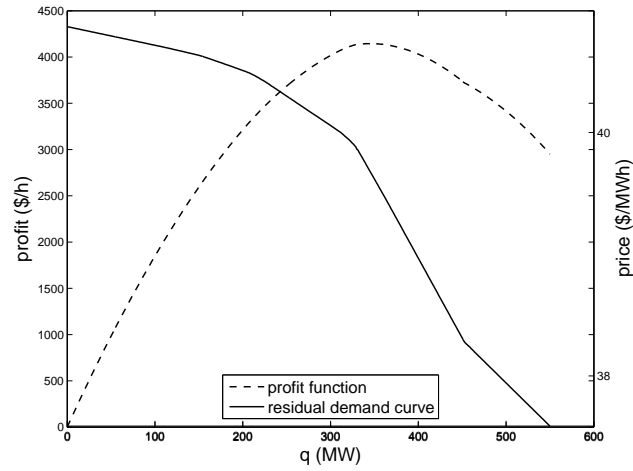


Figure 3.9: Profit function and residual demand curve for generator 5

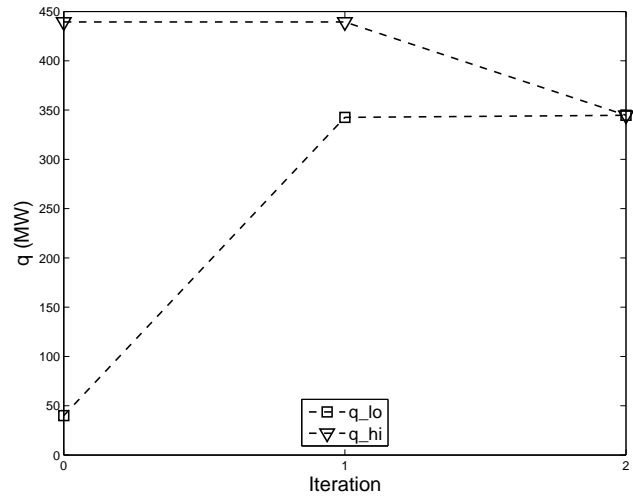


Figure 3.10: Bisection loop iterations

the residual demand curve. We obtain the residual demand derivative formula (3.31) by sensitivity analysis viewing the price at the slack bus as a parameter and assuming there is no perfectly elastic supply in the system. The solution has a suggestive WLS interpretation. Several useful properties of the residual demand derivative are implied by WLS theory. If there is perfectly elastic supply at some bus in the system, it is convenient to use the WLS interpretation to analyze the limit of the residual demand derivative as the quantity-price response of the bus goes to infinity. We establish the condition under which the residual demand derivative at another bus will be bounded or unbounded in this case.

The correctness of the residual demand derivative analytic formulation is verified using an intuitive 2-bus system, a numerical 4-bus 2-loop system from [48], and a 3-bus 1-loop system from [12].

We also improved the TCRDD calculation to make more appeals for practical applications. We use the residual demand derivative formulation to construct optimal bidding strategies in transmission-constrained networks. We proposed a decoupled approach, which iteratively solves the subproblem of calculating the TCRDD, and the subproblem of calculating the optimal offer. Due to the decoupled structure of the TCRDD approach, it is suitable for solving large scale problems. We designed a special bisection search method to find the profit maximization strategy, which requires less function evaluations than standard line search algorithms. We demonstrated the effectiveness of this method in the IEEE 118-bus system.

Chapter 4

Characterizing Strategic Behaviors Via Transmission-constrained Residual Demand Jacobian Matrix

A generation firm in electricity markets may own multiple generators located at multiple locations. To characterize a generation firm's residual demand, this chapter generalizes the concept of transmission-constrained residual demand derivative from a single generator's perspective to the concept of the transmission-constrained residual demand Jacobian (TCRDJ) matrix from a generation firm's perspective. We will derive the formula to calculate the Jacobian matrix based on a multi-parameter sensitivity analysis of the optimal power flow, and also demonstrate how to use the matrix to optimize a generation firm's profit.

The organization of the rest of the chapter is as follows. Section 4.1 derives the TCRDJ in a transmission-constrained network. Section 4.2 proves the symmetry and negative semidefinite properties of the TCRDJ. Section 4.3 deals with binding quality offers and binding price offers. Section 4.4 proposes a bundle trust region method to maximize a generation firm's profit using the TCRDJ. Section 4.5 demonstrate the effectiveness of proposed method in the IEEE 118-bus system. Section 4.6 concludes.

4.1 Calculating the TCRDJ

In this section, we will derive the TCRDJ formula. Following section 3.1 and section 3.3.2, we consider an offer-based electricity market. The market is cleared by an OPF program minimizing the total generation offer cost. In this section, We adopt the more general DC OPF model (3.61) – (3.64) used in section 3.3.2.

Partition the generators into two sets \mathbf{A} and $\bar{\mathbf{A}}$ such that $\mathbf{G} = \mathbf{A} \cup \bar{\mathbf{A}}, \mathbf{A} \cap \bar{\mathbf{A}} = \emptyset$, where set \mathbf{A} consists of all the generators owned by a particular generation firm A , and set $\bar{\mathbf{A}}$ consists of all the generators that do not belong to generation firm A .

Partition $\mathbf{q}_{\mathbf{G}}$ into sub-vectors $\mathbf{q}_{\mathbf{A}}$ and $\mathbf{q}_{\bar{\mathbf{A}}}$ corresponding to the sets of generators \mathbf{A} and $\bar{\mathbf{A}}$. Consider the following optimization problem parameterized by $\mathbf{q}_{\mathbf{A}}$:

$$\min_{\mathbf{q}_{\bar{\mathbf{A}}}} \sum_{\bar{a} \in \bar{\mathbf{A}}} O_{\bar{a}}(q_{\bar{a}}), \quad (4.1)$$

$$\text{s.t.} \quad \mathbf{H}_{\bar{\mathbf{A}}} \mathbf{q}_{\bar{\mathbf{A}}} \leq \mathbf{Z} + \mathbf{H}_{\mathbf{L}} \mathbf{q}_{\mathbf{L}} - \mathbf{H}_{\mathbf{A}} \mathbf{q}_{\mathbf{A}}, \quad (4.2)$$

$$\mathbf{q}_{\bar{\mathbf{A}}}^{\min} \leq \mathbf{q}_{\bar{\mathbf{A}}} \leq \mathbf{q}_{\bar{\mathbf{A}}}^{\max}, \quad (4.3)$$

$$\sum_{\bar{a} \in \bar{\mathbf{A}}} q_{\bar{a}} = \sum_{l \in \mathbf{L}} q_l - \sum_{a \in \mathbf{A}} q_a. \quad (4.4)$$

Denote the Lagrange multipliers corresponding to the constraints (4.2), (4.3), and (4.4), by $\boldsymbol{\mu}(\mathbf{q}_{\mathbf{A}})$, $\boldsymbol{\rho}(\mathbf{q}_{\mathbf{A}})$, and $\lambda(\mathbf{q}_{\mathbf{A}})$, respectively. The optimization problem (4.1)–(4.4) implicitly defines a vector function representing the LMPs (as

derived in (3.11) and (3.67)) for the generators in the set \mathbf{A} :

$$\mathbf{P}_{\mathbf{A}}(\mathbf{q}_{\mathbf{A}}) = \mathbf{1}_{\mathbf{A}}\lambda(\mathbf{q}_{\mathbf{A}}) + \mathbf{H}_{\mathbf{A}}^T\boldsymbol{\mu}(\mathbf{q}_{\mathbf{A}}), \quad (4.5)$$

where $\mathbf{1}_{\mathbf{A}}$ is a column vector of 1s whose dimension equals the number of generators in set \mathbf{A} . By the inverse function theorem, if the Jacobian of $\mathbf{P}_{\mathbf{A}}(\bullet)$ is nonsingular, there exists an inverse function for $\mathbf{P}_{\mathbf{A}}(\bullet)$ locally, which is the vector residual demand function $\mathbf{R}_{\mathbf{A}}(\bullet)$ faced by the generation firm, i.e. $\mathbf{P}_{\mathbf{A}}^{-1}(\bullet) = \mathbf{R}_{\mathbf{A}}(\bullet)$, and the Jacobian of $\mathbf{R}_{\mathbf{A}}(\bullet)$ equals the inverse of the Jacobian of $\mathbf{P}_{\mathbf{A}}(\bullet)$.

Now we calculate the Jacobian of $\mathbf{P}_{\mathbf{A}}(\bullet)$, evaluated at an OPF solution. For derivation simplicity, we assume there are no generators with binding offer quantities or binding offer prices. Dealing with those binding offer quantities and binding offer prices is the topic of section 4.3.

Construct the Lagrangian of the optimization problem (4.1)–(4.4) with only the binding transmission constraints:

$$\begin{aligned} \mathcal{L} = & \sum_{\bar{a} \in \bar{\mathbf{A}}} O_{\bar{a}}(q_{\bar{a}}) + \lambda \left(\sum_{l \in \mathbf{L}} q_l - \sum_{\bar{a} \in \bar{\mathbf{A}}} q_{\bar{a}} - \sum_{a \in \mathbf{A}} q_a \right) \\ & + \boldsymbol{\mu}_{\mathbf{b}}^T (\mathbf{Z}_{\mathbf{b}} + \mathbf{H}_{\mathbf{Lb}}\mathbf{q}_{\mathbf{L}} - \mathbf{H}_{\mathbf{Ab}}\mathbf{q}_{\mathbf{A}} - \mathbf{H}_{\bar{\mathbf{A}}\mathbf{b}}\mathbf{q}_{\bar{\mathbf{A}}}), \end{aligned}$$

where the subscript “b” means the rows or entries corresponding to binding constraints. We assume that the second-order sufficient conditions hold.

For all $a \in \mathbf{A}$ we calculate $\frac{\partial \lambda}{\partial q_a}$ and $\frac{\partial \boldsymbol{\mu}_{\mathbf{b}}}{\partial q_a}$ from a sensitivity analysis of the

following first-order necessary conditions (FONCs):

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \mathbf{q}_{\bar{A}}} &= \mathbf{0}_{\bar{A}}, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}_{\mathbf{b}}} &= \mathbf{0}_{\mathbf{b}}, \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 0.\end{aligned}$$

Write the FONCs explicitly:

$$\begin{aligned}\mathbf{O}'_{\bar{A}}(\mathbf{q}_{\bar{A}}) - \mathbf{1}_{\bar{A}}\lambda - \mathbf{H}_{\bar{A}\mathbf{b}}^T \boldsymbol{\mu}_{\mathbf{b}} &= \mathbf{0}_{\bar{A}}, \\ \mathbf{H}_{\bar{A}\mathbf{b}} \mathbf{q}_{\bar{A}} &= \mathbf{Z}_{\mathbf{b}} + \mathbf{H}_{\mathbf{L}\mathbf{b}} \mathbf{q}_{\mathbf{L}} - \mathbf{H}_{\mathbf{A}\mathbf{b}} \mathbf{q}_{\mathbf{A}}, \\ \mathbf{1}_{\bar{A}}^T \mathbf{q}_{\bar{A}} &= \mathbf{1}_{\mathbf{L}}^T \mathbf{q}_{\mathbf{L}} - \mathbf{1}_{\mathbf{A}}^T \mathbf{q}_{\mathbf{A}},\end{aligned}\tag{4.6}$$

where we view $\mathbf{q}_{\bar{A}}$, $\boldsymbol{\mu}_{\mathbf{b}}$, and λ as implicit functions of $\mathbf{q}_{\mathbf{A}}$ and

- $\mathbf{O}'_{\bar{A}}(\mathbf{q}_{\bar{A}}) = \nabla_{\mathbf{q}_{\bar{A}}} (\sum_{\bar{a} \in \bar{A}} O_{\bar{a}}(q_{\bar{a}}))$,
- $\mathbf{0}_{\bar{A}}$ is a column vector of 0s whose dimension is equal to the number of generators in the set \bar{A} ,
- $\mathbf{1}_{\mathbf{A}}$, $\mathbf{1}_{\bar{A}}$ and $\mathbf{1}_{\mathbf{L}}$ are column vectors of 1s whose dimensions equal the number of generators in set \mathbf{A} , the number of generators in set \bar{A} , and the number of loads in \mathbf{L} respectively.

Differentiate both sides of (4.6) with respect to $q_a, \forall a \in \mathbf{A}$, we obtain:

$$\mathbf{O}''_{\bar{A}} \frac{\partial \mathbf{q}_{\bar{A}}}{\partial q_a} - \mathbf{1}_{\bar{A}} \frac{\partial \lambda}{\partial q_a} - \mathbf{H}_{\bar{A}\mathbf{b}}^T \frac{\partial \boldsymbol{\mu}_{\mathbf{b}}}{\partial q_a} = \mathbf{0}_{\bar{A}},\tag{4.7}$$

$$\mathbf{H}_{\bar{A}\mathbf{b}} \frac{\partial \mathbf{q}_{\bar{A}}}{\partial q_a} = -\mathbf{H}_{a\mathbf{b}},\tag{4.8}$$

$$\mathbf{1}_{\bar{A}}^T \frac{\partial \mathbf{q}_{\bar{A}}}{\partial q_a} = -1,\tag{4.9}$$

where

$$\mathbf{O}_{\bar{A}}'' = \nabla_{\mathbf{q}_{\bar{A}} \mathbf{q}_{\bar{A}}}^2 \left(\sum_{\bar{a} \in \bar{A}} O_{\bar{a}}(q_{\bar{a}}) \right).$$

Assume $\mathbf{O}_{\bar{A}}''$ is positive definite (P.D.), so that it is invertible. This is always true if there are no offers in \bar{A} with binding quantities or prices. We will deal with binding quantity and binding price offers in section 4.3. Now define:

$$\begin{aligned} \mathbf{M} &= \mathbf{H}_{\bar{A}b} (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{H}_{\bar{A}b}^T, \\ N &= \mathbf{1}_{\bar{A}}^T (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{1}_{\bar{A}}. \end{aligned}$$

Assume that the rows of $\mathbf{H}_{\bar{A}b}$ are linearly independent and that the number of rows is less than the number of generators in the set \bar{A} so that \mathbf{M} is invertible.

Multiply both sides of (4.7) on the left by $\mathbf{H}_{\bar{A}b} (\mathbf{O}_{\bar{A}}'')^{-1}$, substitute in (4.8), and then multiply both sides on the left by \mathbf{M}^{-1} , we get

$$\frac{\partial \boldsymbol{\mu}_b}{\partial q_a} = -\mathbf{M}^{-1} \mathbf{H}_{ab} - \mathbf{M}^{-1} \mathbf{H}_{\bar{A}b} (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{1}_{\bar{A}} \frac{\partial \lambda}{\partial q_a}. \quad (4.10)$$

Multiply both sides of (4.7) on the left by $\mathbf{1}_{\bar{A}}^T (\mathbf{O}_{\bar{A}}'')^{-1}$, substitute in (4.9), we get

$$-1 - N \frac{\partial \lambda}{\partial q_a} - \mathbf{1}_{\bar{A}}^T (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{H}_{\bar{A}b}^T \frac{\partial \boldsymbol{\mu}_b}{\partial q_a} = 0. \quad (4.11)$$

Solve (4.10) and (4.11), we get

$$\frac{\partial \lambda}{\partial q_a} = V^{-1} \left(1 - \mathbf{1}_{\bar{A}}^T (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{H}_{\bar{A}b}^T \mathbf{M}^{-1} \mathbf{H}_{ab} \right), \quad (4.12)$$

where

$$V = -N + \mathbf{1}_{\bar{A}}^T (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{H}_{\bar{A}b}^T \mathbf{M}^{-1} \mathbf{H}_{\bar{A}b} (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{1}_{\bar{A}}. \quad (4.13)$$

The derivative $\frac{\partial \boldsymbol{\mu}_b}{\partial q_a}$ can be calculated by substituting (4.12) into (4.10).

Now we can calculate the Jacobian matrix $\frac{\partial \mathbf{p}_A}{\partial \mathbf{q}_A}$, with the (g, a) element of the matrix defined by differentiating both sides of the g -th row of (4.5) with respect to q_a :

$$\frac{\partial p_g}{\partial q_a} = \frac{\partial \lambda}{\partial q_a} + \mathbf{H}_{ab}^T \frac{\partial \boldsymbol{\mu}_b}{\partial q_a}, \quad \forall g, a \in A.$$

After simplification,

$$\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A} = -\mathbf{H}_{Ab}^T \mathbf{M}^{-1} \mathbf{H}_{Ab} + \mathbf{U}^T \mathbf{U} V^{-1}, \quad (4.14)$$

where

$$\mathbf{U} = \mathbf{1}_A - \mathbf{H}_{Ab}^T \mathbf{M}^{-1} \mathbf{H}_{\bar{A}b} (\mathbf{O}_{\bar{A}}'')^{-1} \mathbf{1}_{\bar{A}}.$$

By the inverse function theorem, the residual demand Jacobian for the generation firm is:

$$\frac{\partial \mathbf{q}_A}{\partial \mathbf{p}_A} = \left(\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A} \right)^{-1}. \quad (4.15)$$

If the generation firm has only one generator, and the generator is located at the slack bus s , then $\mathbf{H}_s = \mathbf{0}$, where \mathbf{H}_s is the column of \mathbf{H} corresponding to the slack bus. In this case, $\mathbf{H}_{sb} = \mathbf{0}$ in (4.12), so $\frac{\partial \lambda}{\partial q_a} = V^{-1}$, where V is the TCRDD formula (3.31). This verifies the correctness of our calculation for this special case.

4.2 Properties of the TCRDJ

As proved in [19], the TCRDJ is symmetric and negative semidefinite (N.S.D.) if the DC OPF only models the branch flow constraints. The TCRDJ

derivation part of the paper can be viewed as a generalization of [19], because as discussed in section 3.3.2, our OPF formulation can include all types of linear constraints, such as contingency constraints and Nomograms. We are going to prove the symmetry and N.S.D. properties of the TCRDJ holds with all constraints being linear in the OPF.

Proposition 6. *The TCRDJ $\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A}$ is symmetric.*

Proof. Directly from the TCRDJ $\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A}$ formula (4.14),

$$\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A} = \left(\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A} \right)^T.$$

□

Proposition 7. *The TCRDJ $\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A}$ is negative semidefinite.*

Proof. Because $(\mathbf{O}_{\bar{A}}'')^{-1}$ is positive definite (P.D.), \mathbf{M} is also P.D. by definition under the assumption that the rows of $\mathbf{H}_{\bar{A}b}$ are linearly independent and that the number of rows is less than the number of generators in the set \bar{A} . Therefore, we only need to prove $V \leq 0$ to prove $\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A}$ is N.S.D. by its formula (4.14).

Similar to section 3.1.4 where we proved the TCRDD is less than or equal to zero, we prove $V \leq 0$ by a Weighted Least Squares (WLS) formulation. Following appendix A, consider the WLS problem specified by

$$\mathbf{X} = \mathbf{H}_{\bar{A}b}^T,$$

$$\mathbf{Y} = \mathbf{1}_{\bar{A}},$$

$$\mathbf{W} = (\mathbf{O}_{\bar{A}}'')^{-1}.$$

The Least Sum of Squares Error (SSE) is

$$\begin{aligned}
SSE^{\text{WLS}} &= \mathbf{Y}^T \mathbf{W} \mathbf{Y} - \mathbf{Y}^T \mathbf{W} \mathbf{X} (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y} \\
&= \mathbf{1}_{\bar{\mathbf{A}}}^T (\mathbf{O}_{\bar{\mathbf{A}}}'')^{-1} \mathbf{1}_{\bar{\mathbf{A}}} - \mathbf{1}_{\bar{\mathbf{A}}}^T (\mathbf{O}_{\bar{\mathbf{A}}}'')^{-1} \mathbf{H}_{\bar{\mathbf{A}}\mathbf{b}}^T \mathbf{M}^{-1} \mathbf{H}_{\bar{\mathbf{A}}\mathbf{b}} (\mathbf{O}_{\bar{\mathbf{A}}}'')^{-1} \mathbf{1}_{\bar{\mathbf{A}}} \\
&= -V \\
&\geq 0.
\end{aligned}$$

Therefore, $V \leq 0$, and the TCRDJ $\frac{\partial \mathbf{P}_{\mathbf{A}}}{\partial \mathbf{q}_{\mathbf{A}}}$ is negative semidefinite.

□

4.3 Handling binding quantity offers and binding price offers

Some offer functions may be binding at certain constant output levels, such as the capacity bounds, and some other offers may be binding at constant offer prices if they bid constant prices for some output stacks. As discussed in section 3.4, these special cases need to be handled separately, because (4.12) will be invalid under those circumstances. If an offer is binding at the output limit for the market clearing conditions, then the corresponding entry in $\mathbf{O}_{\bar{\mathbf{A}}}''$ is undefined. If an offer is binding at a constant price, $\mathbf{O}_{\bar{\mathbf{A}}}''$ is not invertible.

Following section 3.4, to handle these binding offer quantities and binding offer prices, partition the generators in set $\bar{\mathbf{A}}$ into three subsets: the generators with binding offer quantities, denoted by \mathbf{f} ; the generators with binding offer prices, denoted by \mathbf{z} ; and the generators with offers having non-zero

slopes, denoted by \mathbf{v} , so that

$$\mathbf{q}_{\bar{\mathbf{A}}} = \begin{bmatrix} \mathbf{q}_{\mathbf{v}} \\ \mathbf{q}_{\mathbf{f}} \\ \mathbf{q}_{\mathbf{z}} \end{bmatrix}.$$

Accordingly, partition $\mathbf{H}_{\bar{\mathbf{A}}\mathbf{b}}$ into:

$$\mathbf{H}_{\bar{\mathbf{A}}\mathbf{b}} = \begin{bmatrix} \mathbf{H}_{\mathbf{vb}} & \mathbf{H}_{\mathbf{fb}} & \mathbf{H}_{\mathbf{zb}} \end{bmatrix}.$$

Rewrite the Lagrangian of the OPF with only the binding constraints:

$$\begin{aligned} \mathcal{L} = & \sum_{g \in \mathbf{v}} O_g(q_g) + \sum_{g \in \mathbf{f}} O_g(q_g) + \sum_{g \in \mathbf{z}} O_g(q_g) \\ & + \lambda \left(\sum_{l \in \mathbf{L}} q_l - \sum_{g \in \mathbf{v}} q_g - \sum_{g \in \mathbf{f}} q_g - \sum_{g \in \mathbf{z}} q_g - \sum_{g \in \mathbf{A}} q_g \right) \\ & + \boldsymbol{\mu}_{\mathbf{b}}^T (\mathbf{H}_{\mathbf{vb}} \mathbf{q}_{\mathbf{v}} + \mathbf{H}_{\mathbf{fb}} \mathbf{q}_{\mathbf{f}} + \mathbf{H}_{\mathbf{zb}} \mathbf{q}_{\mathbf{z}} + \mathbf{H}_{\mathbf{Ab}} \mathbf{q}_{\mathbf{A}} - \mathbf{H}_{\mathbf{Lb}} \mathbf{q}_{\mathbf{L}} - \mathbf{Z}_{\mathbf{b}}) \\ & + \boldsymbol{\rho}_{\max}^T (\mathbf{q}_{\mathbf{f}} - \mathbf{q}_{\mathbf{f}}^{\max}) + \boldsymbol{\rho}_{\min}^T (\mathbf{q}_{\mathbf{f}} - \mathbf{q}_{\mathbf{f}}^{\min}). \end{aligned}$$

By definition, at the OPF solution,

$$\forall a \in \mathbf{A}, \frac{\partial \mathbf{q}_{\mathbf{f}}}{\partial q_a} = \mathbf{0}_{\mathbf{f}}. \quad (4.16)$$

Similar to section 3.1, we calculate $\frac{\partial \lambda}{\partial q_a}$ and $\frac{\partial \boldsymbol{\mu}_{\mathbf{b}}}{\partial q_a}$ from a sensitivity analysis of the following first-order necessary conditions:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{q}_{\mathbf{v}}} &= \mathbf{0}_{\mathbf{v}}, \\ \frac{\partial \mathcal{L}}{\partial \mathbf{q}_{\mathbf{z}}} &= \mathbf{0}_{\mathbf{z}}, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}_{\mathbf{b}}} &= \mathbf{0}_{\mathbf{b}}, \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 0, \end{aligned}$$

i.e.

$$\begin{aligned}
\mathbf{O}'_{\mathbf{v}}(\mathbf{q}_{\mathbf{v}}) - \mathbf{1}_{\mathbf{v}}\lambda - \mathbf{H}_{\mathbf{vb}}^{\mathbf{T}}\boldsymbol{\mu}_{\mathbf{b}} &= \mathbf{0}_{\mathbf{v}}, \\
\mathbf{O}'_{\mathbf{z}}(\mathbf{q}_{\mathbf{z}}) - \lambda\mathbf{1}_{\mathbf{z}} - \mathbf{H}_{\mathbf{zb}}^{\mathbf{T}}\boldsymbol{\mu}_{\mathbf{b}} &= \mathbf{0}_{\mathbf{z}}, \\
\mathbf{H}_{\mathbf{vb}}\mathbf{q}_{\mathbf{v}} + \mathbf{H}_{\mathbf{zb}}\mathbf{q}_{\mathbf{z}} &= \mathbf{Z}_{\mathbf{b}} - \mathbf{H}_{\mathbf{Lb}}\mathbf{q}_{\mathbf{L}} - \mathbf{H}_{\mathbf{Ab}}\mathbf{q}_{\mathbf{A}}, \\
\mathbf{1}_{\mathbf{A}}^{\mathbf{T}}\mathbf{q}_{\mathbf{A}} &= \mathbf{1}_{\mathbf{L}}^{\mathbf{T}}\mathbf{q}_{\mathbf{L}} - \mathbf{1}_{\mathbf{A}}^{\mathbf{T}}\mathbf{q}_{\mathbf{A}},
\end{aligned} \tag{4.17}$$

where we again view $\mathbf{q}_{\mathbf{v}}$, $\mathbf{q}_{\mathbf{z}}$, $\boldsymbol{\mu}_{\mathbf{b}}$, and λ as implicit function of $\mathbf{q}_{\mathbf{A}}$, and $\mathbf{O}'_{\mathbf{z}}(\mathbf{q}_{\mathbf{z}})$, $\mathbf{0}_{\mathbf{v}}$, $\mathbf{0}_{\mathbf{f}}$, and $\mathbf{0}_{\mathbf{z}}$ are similarly defined as in (3.67).

Differentiate both sides of (4.17) with respect to $q_a, \forall a \in \mathbf{A}$, we get

$$\begin{aligned}
\mathbf{O}''_{\mathbf{v}}\frac{\partial\mathbf{q}_{\mathbf{v}}}{\partial q_a} - \mathbf{1}_{\mathbf{v}}\frac{\partial\lambda}{\partial q_a} - \mathbf{H}_{\mathbf{vb}}^{\mathbf{T}}\frac{\partial\boldsymbol{\mu}_{\mathbf{b}}}{\partial q_a} &= \mathbf{0}_{\mathbf{v}}, \\
-\mathbf{1}_{\mathbf{z}}\frac{\partial\lambda}{\partial q_a} - \mathbf{H}_{\mathbf{zb}}^{\mathbf{T}}\frac{\partial\boldsymbol{\mu}_{\mathbf{b}}}{\partial q_a} &= \mathbf{0}_{\mathbf{z}}, \\
\mathbf{H}_{\mathbf{vb}}\frac{\partial\mathbf{q}_{\mathbf{v}}}{\partial q_a} + \mathbf{H}_{\mathbf{zb}}\frac{\partial\mathbf{q}_{\mathbf{z}}}{\partial q_a} &= -\mathbf{H}_{\mathbf{ab}}, \\
\mathbf{1}_{\mathbf{v}}^{\mathbf{T}}\frac{\partial\mathbf{q}_{\mathbf{v}}}{\partial q_a} + \mathbf{1}_{\mathbf{z}}^{\mathbf{T}}\frac{\partial\mathbf{q}_{\mathbf{z}}}{\partial q_a} &= -1,
\end{aligned} \tag{4.18}$$

where $\mathbf{O}''_{\mathbf{v}}$ is similarly defined as in (3.68). If the number of binding price offers in the set \mathbf{z} is less than the number of binding transmission constraints, and the number of binding transmission constraints is less than the number of offers in set \mathbf{v} , then (4.18) has a unique solution. The TCRDJ can be calculated by solving (4.18).

The symmetry and N.S.D. properties of the TCRDJ still hold with binding quantity and binding price offers as a result of the implicit function theorem. Because (4.7)–(4.9) are continuously partially differentiable in the

bidding slopes, under the assumption that (4.18) has a unique solution,

$$\begin{bmatrix} \frac{\partial \mathbf{q}_v}{\partial q_a} \\ \frac{\partial \mathbf{q}_z}{\partial q_a} \\ \frac{\partial \boldsymbol{\mu}_b}{\partial q_a} \\ \frac{\partial \lambda}{\partial q_a} \end{bmatrix}$$

can be viewed as a continuous differentiable implicit function of the bidding slopes of the generators in set \mathbf{z} in the vicinity of $\mathbf{0}_z$. If we define a sequence of positive bidding slopes of the generators in set \mathbf{z} , that monotonically approach $\mathbf{0}_z$, then the corresponding TCRDJs, calculated by solving (4.7)–(4.9), in the sequence are symmetric and N.S.D. as proved in section 4.2. At the limit of the sequence, (4.7)–(4.9) converge to (4.18). Therefore, the limit of the TCRDJ in the sequence is the TCRDJ calculated by solving (4.18), and it is symmetric and N.S.D. as well.

4.4 Maximizing A Generation Firm's Profit

Similar to maximizing a generator's profit using TCRDD, we can maximize a generation firm's profit using TCRDJ. In a nodal electricity market, a generation firm may own multiple generators located at different locations. Compared with a single generator, a generation firm has more resources to leverage, and thus may have more profitable strategies.

Similar to optimizing a single generator's profit based on TCRDD, one major task is to deal with the non-differentiability in the residual demand function. In section 3.4, we proposed a special bisection search algorithm to

find the optimizer for a single generator. However, the algorithm does not apply to higher dimensional residual demand function. In this chapter, we propose a bundle method to optimize a generation firm's bidding strategy in the residual demand function.

4.4.1 Bundle idea

The bundle concept is widely applied to non-differentiable function optimization. To be consistent with the convention of most bundled method literatures, such as [37], we model the problem as a minimization problem

$$\begin{aligned} \min_{\mathbf{x}} f(\mathbf{x}) \\ \text{s.t. } \underline{\mathbf{x}} \leq \mathbf{x} \leq \bar{\mathbf{x}}. \end{aligned} \tag{4.19}$$

where $f : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ is a Lipschitz continuous non-smooth convex function, where the decision variables are the output levels of the self-owned generators, i.e. $\mathbf{x} = \mathbf{q}_A$, and the function $f(\bullet)$ is the negative profit function,

$$f(\mathbf{q}_A) = - \sum_{a \in A} (P_a(q_a)q_a - C_a(q_a)).$$

The inverse residual demand function $\mathbf{P}_A(\mathbf{q}_A)$ may have kinks, which may make the objective function f non-convex. Let us assume the convexity of objective function for the current moment to introduce the bundle idea conveniently. We will cover how to handle non-convex objective function in section 4.4.2 when we go into the details of the algorithm.

The $\frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A}(\mathbf{q}_A)$ calculation in section 4.1 provides a subgradient of $\mathbf{p}_A(\bullet)$.

Therefore, one subgradient of $f(\bullet)$ is

$$\mathbf{g}(\mathbf{q}_A) = - \left(\mathbf{P}_A(\mathbf{q}_A) + \frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A}(\mathbf{q}_A) \mathbf{q}_A - \mathbf{C}'_A(\mathbf{q}_A) \right). \quad (4.20)$$

The subgradient information (4.20) is crucial for a bundle method.

The bundle concept has two features [37]:

1. Make use at iteration k , the bundle information

$$(f(\mathbf{x}^k), \mathbf{g}(\mathbf{x}^k)), (f(\mathbf{x}^{k-1}), \mathbf{g}(\mathbf{x}^{k-1})), \dots,$$

collected so far to build a model of the objective function f .

2. If, due to the kinked structure of f , this model does not characterize f accurate enough, then mobilize more subgradient information.

Feature 1 leads to the cutting plane approximation of f at \mathbf{x}^k from below.

$$\begin{aligned} f_{\text{CP}}(\mathbf{x}) &= \max_{1 \leq j \leq k} \{ \mathbf{g}(\mathbf{x}^j)^T (\mathbf{x} - \mathbf{x}^j) + f(\mathbf{x}^j) \} \\ &= \max_{1 \leq j \leq k} \{ \mathbf{g}(\mathbf{x}^j)^T (\mathbf{x} - \mathbf{x}^k) + \mathbf{g}(\mathbf{x}^j)^T (\mathbf{x}^k - \mathbf{x}^j) + f(\mathbf{x}^j) \} \\ &= \max_{1 \leq j \leq k} \{ \mathbf{g}(\mathbf{x}^j)^T (\mathbf{x} - \mathbf{x}^k) - \alpha_{k,j} + f(\mathbf{x}^k) \} \end{aligned} \quad (4.21)$$

where

$$\alpha_{k,j} = f(\mathbf{x}^k) - (f(\mathbf{x}^j) + \mathbf{g}(\mathbf{x}^j)^T (\mathbf{x}^k - \mathbf{x}^j)) \quad (4.22)$$

as defined in [37]. If $f(\bullet)$ is convex, $f_{\text{CP}}(\bullet)$ is an approximation of $f(\bullet)$ from below. It is typically a good approximation in the vicinity of \mathbf{x}^j , because it

coincides with f at all x^j . If enough bundle information has been collected, f_{CP} may also be a good approximation for f even for \mathbf{x} that is far away from \mathbf{x}^j , but without enough bundle information, f_{CP} is likely to be a poor approximation for f , in which case, more bundle information needs to be mobilized to improve the approximation accuracy.

On the other hand, following section 3.4, define

$$\bar{\mathbf{P}}_{\mathbf{A}}(\mathbf{q}_{\mathbf{A}}; \mathbf{q}_{\mathbf{A}}^k) = \frac{\partial \mathbf{P}_{\mathbf{A}}}{\partial \mathbf{q}_{\mathbf{A}}}(\mathbf{q}_{\mathbf{A}}^k)(\mathbf{q}_{\mathbf{A}} - \mathbf{q}_{\mathbf{A}}^k) + \mathbf{p}_{\mathbf{A}}^k.$$

$\bar{\mathbf{P}}_{\mathbf{A}}(\bullet; \mathbf{q}_{\mathbf{A}}^k)$ is an approximation of $\mathbf{P}_{\mathbf{A}}(\bullet)$ at iteration k , and it coincides with $\mathbf{P}_{\mathbf{A}}(\bullet)$ at $\mathbf{q}_{\mathbf{A}}^k$. Thus,

$$\bar{f}(\mathbf{q}_{\mathbf{A}}) = - \left(\bar{\mathbf{P}}_{\mathbf{A}}^{\text{T}}(\mathbf{q}_{\mathbf{A}}; \mathbf{q}_{\mathbf{A}}^k) \mathbf{q}_{\mathbf{A}} - \mathbf{1}_{\mathbf{A}}^{\text{T}} \mathbf{C}_{\mathbf{A}}(\mathbf{q}_{\mathbf{A}}) \right),$$

is another approximation of $f(\bullet)$. In the vicinity of $\mathbf{q}_{\mathbf{A}}$, \bar{f} is a better approximation than f_{CP} , because \bar{f} matches both the value and curvature of f .

As \bar{f} is a good local approximation function, and f_{CP} may be a good overall approximation given enough bundle information, it will be better to combine the power of the two. Consider the following optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{t} \bar{f}(\mathbf{x}) + f_{\text{CP}}(\mathbf{x}) \\ \text{s.t.} \quad & \underline{\mathbf{x}} \leq \mathbf{x} \leq \bar{\mathbf{x}}. \end{aligned} \tag{4.23}$$

where $t \in (0, 1]$ is a parameter to implicitly control the step size. The idea is to adaptively adjust t in the optimization progress, which resembles the trust region concept. Although we adjust t instead of the trust region, the effect is the same. If t is small, the “trust region” is small, and vice versa. If t is chosen

properly and enough bundle information has been collected, the minimizer of (4.23) also solves (4.19).

Problem (4.23) is equivalent to the following quadratic program

$$\begin{aligned} \min \quad & \frac{1}{t} \bar{f}(\mathbf{x}) + v \\ \text{s.t.} \quad & v \geq \mathbf{g}(\mathbf{x}^j)^T (\mathbf{x} - \mathbf{x}^k) - \alpha_{k,j} + f(\mathbf{x}^k), \quad \forall j \leq k, \\ & \underline{\mathbf{x}} \leq \mathbf{x} \leq \bar{\mathbf{x}}. \end{aligned} \tag{4.24}$$

We are going to solve (4.24) for \mathbf{x}^{k+1} , and depending on whether \mathbf{x}^{k+1} improves f , we will adjust t , add bundle information, and make progress.

Applying the trust region to the bundle concept leads to the following scheme [37]. If $f(\mathbf{x}^{k+1})$ is “sufficiently smaller” than $f(\mathbf{x}^k)$, that means (4.24) is a good approximation to the original problem, so we can make a Serious Step from the incumbent \mathbf{x}^k to \mathbf{x}^{k+1} , and in the mean time, t could be increased to enlarge the trust region. If $f(\mathbf{x}^{k+1})$ is not “sufficiently smaller” than $f(\mathbf{x}^k)$, that means (4.24) is not a good approximation to the original problem, so we need to make the approximation more accurate instead of proceed to the next step. The process is called a Null Step, which does the following: first add $\mathbf{g}^{k+1} = \mathbf{g}(\mathbf{x}^{k+1})$ to the bundle information, then decrease t to shrink the trust region, and stay with the incumbent \mathbf{x}^k instead of proceeding to \mathbf{x}^{k+1} .

4.4.2 Algorithm

To turn the bundle trust region idea into an algorithm, we need to qualify the “sufficiently smaller” criteria and an appropriate stopping criteria.

These two criteria can be implemented based on the ϵ -optimal criteria:
if \mathbf{x}^* is the optimizer of $f(\bullet)$, then \mathbf{x} is ϵ -optimal if

$$f(\mathbf{x}) \leq f(\mathbf{x}^*) + \epsilon \|\mathbf{x} - \mathbf{x}^*\| + \epsilon. \quad (4.25)$$

Based on (4.25), we say $f(\mathbf{x}^{k+1})$ is “sufficiently smaller” than $f(\mathbf{x}^k)$ if:

$$f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k) < -\epsilon \|\mathbf{x}^{k+1} - \mathbf{x}^k\| - \epsilon \text{ and } \|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq 1,$$

or

$$f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k) < -\epsilon \text{ and } \|\mathbf{x}^{k+1} - \mathbf{x}^k\| > 1.$$

The second condition above means when \mathbf{x}^{k+1} is sufficiently away from \mathbf{x}^k , it is likely the bundle information will get improved from the current model if we proceed to \mathbf{x}^{k+1} , so we would like to make a Serious Step even though the improvement in the objective function is small.

The stopping criteria needs more derivation. Lemma 2.2 in [37] is a sufficient condition for ϵ -optimal. Following [37], let us work out the stopping criteria, that satisfies Lemma 2.2 in [37], from the KKT conditions of (4.24):

$$\lambda^j(-v + \mathbf{g}(\mathbf{x}^j)^\top(\mathbf{x} - \mathbf{x}^k) - \alpha_{k,j} + f(\mathbf{x}^k)) = 0, \quad \forall j \leq k, \quad (4.26)$$

$$\frac{1}{t} \nabla \bar{f}(\mathbf{x}) + \sum_j \lambda^j \mathbf{g}(\mathbf{x}^j) = 0, \quad (4.27)$$

$$1 - \sum_j \lambda^j = 0. \quad (4.28)$$

By (4.27),

$$\sum_j \lambda^j \mathbf{g}(\mathbf{x}^j) = -\frac{1}{t} \nabla \bar{f}(\mathbf{x}). \quad (4.29)$$

Sum (4.26) over j , apply (4.27) and (4.28), and evaluate at $\mathbf{x} = \mathbf{x}^{k+1}$,

$$\sum_j \lambda^j \alpha_{k+1,j} = -v^{k+1} - \left(\frac{1}{t} \nabla \bar{f}(\mathbf{x}^{k+1}) \right)^T (\mathbf{x}^{k+1} - \mathbf{x}^k) + f(\mathbf{x}^k) \quad (4.30)$$

Substitute (4.29) and (4.30) into Lemma 2.2 in [37], we get the ϵ -optimal stopping criteria

$$\begin{aligned} \left\| \frac{1}{t} \nabla \bar{f}(\mathbf{x}^{k+1}) \right\| &\leq \epsilon, \\ -v^{k+1} - \left(\frac{1}{t} \nabla \bar{f}(\mathbf{x}^{k+1}) \right)^T (\mathbf{x}^{k+1} - \mathbf{x}^k) + f(\mathbf{x}^k) &\leq \epsilon. \end{aligned} \quad (4.31)$$

The algorithm is as follows.

1. Let $k = 0$, $\mathbf{x}^k = \mathbf{x}^0$, and $t^k = 1$.
2. Let $k = k + 1$, and solve (4.24) for \mathbf{x}^{k+1} .
3. If stopping criteria (4.31) is satisfied, optimal solution found with $\mathbf{x}^* = \mathbf{x}^k$, stop.
4. If $f(\mathbf{x}^{k+1})$ is “sufficiently smaller” than $f(\mathbf{x}^k)$, then make a Serious Step: $t^k = \min\{1, 2t^k\}$, compute $\mathbf{g}^{k+1} = \mathbf{g}(\mathbf{x}^{k+1})$, and continue with step 2. Otherwise, $t^k = 0.5t^k$, and make a Null Step: compute $\mathbf{g}^{k+1} = \mathbf{g}(\mathbf{x}^{k+1})$, let $\mathbf{x}^{k+1} = \mathbf{x}^k$, and continue with step 2.

With minor modifications, the algorithm can handle non-convex function as well. If the objective function is non-convex, (4.21) may not be an

generator	bus	q^{\min}	q^{\max}	marginal cost
1	1	0	100	$0.020q + 40$
2	4	0	100	$0.020q + 40$
3	6	0	100	$0.020q + 40$
4	8	0	100	$0.020q + 40$
5	10	0	550	$0.044q + 20$
30	69	0	805.2	$0.039q + 20$

Table 4.1: Generator data

approximation of $f(\bullet)$ from below. In this case, $\alpha_{k,j}$ in (4.22) may be negative. As proved in section 3 of [37], as long as we replace $\alpha_{k,j}$ by

$$\beta_{k,j} = \max\{\alpha_{k,j}, \epsilon_0 \|\mathbf{x}^k - \mathbf{x}^j\|\} \geq 0,$$

where ϵ_0 is a very small positive number, the algorithm will be able to handle non-convex objective function. Note that, in this case, the stopping criteria does not imply ϵ -optimality, it merely means \mathbf{x}^{k+1} is “almost” stationary as pointed by in [37].

4.5 Computational Example

We apply the algorithm to the IEEE 118-bus test system, which is also used in section 3.4. We optimize the profit for a fictitious generation firm A, who owns two generators: generator 5 located at bus 10 with 550 MW capacity, and generator 30 located at bus 69 with 805.2 MW capacity. Part of the generator data is listed in Tab. 4.1. Assume all generators other than generators 5 and 30 offer at their true marginal cost.

4.5.1 Firm strategy vs single generator strategy

We plotted the contour of the profit function and the optimization trajectory in Fig. 4.1, and the data is listed in Tab. 4.2. The “step” column in Tab. 4.2 indicates a Serious Step by 'S', a Null Step by 'N', and optimal solution by 'O'. The same convention also applies to other similar tables through the rest of the chapter. After 5 iterations, the algorithm found the optimizer $q_5^* = 356.58$ and $q_{30}^* = 434.17$, with a total profit of 9192.3. At the optimum, generator 5 is making a profit of 4299.9, and generator 30 is making a profit of 4892.3, as listed in column 2 of Tab. 4.3. The power flow involves the capacity constraints of branches 36-30, 38-26, and 51-38 binding at 200 MW each.

We also consider the case that generators 5 and generator 30 belong two different generation firms, respectively, which do not own any other units in the system. As listed in column 3 of Tab. 4.3, the optimal strategy for generator 5 is $q_5^* = 344.76$ with profit $\pi_5^* = 4188.88$, assuming all other generators, including generator 30, offer at their marginal costs. As listed in column 4 of Tab. 4.3, the optimal strategy for generator 30 is $q_{30}^* = 436.44$ with profit $\pi_{30}^* = 4650.6$, assuming all other generators, including generator 5, offer at their marginal costs. The sum of π_5^* and π_{30}^* is 8839.4, which is less than the generation firm A's profit 9192.3. The increased profit is achieved by decreasing q_{30} by 2.27 MW, which reduces generator 30's profit, but frees up some transmission capacity so that generator 5's output can increase by a larger amount, 11.82 MW, without overloading the transmission lines. This demonstrates that a generation firm's profit maximizing strategy may be sub-

iter.	q_5	q_{30}	profit	step
0	200.00	200.00	6509.6	S
1	435.27	499.92	8260.2	S
2	313.33	412.58	9069.1	S
3	374.58	445.66	9126.5	S
4	359.47	431.98	9191.6	S
5	356.59	434.17	9192.3	O

Table 4.2: Optimization trajectory starting from (200,200)

	Firm A		Single Generator	
	g_5	g_{30}	g_5	g_{30}
q^*	356.58	434.17	344.76	436.44
p^*	39.98	39.68	39.68	39.11
π^*	4299.9	4892.3	4188.88	4650.6

Table 4.3: Optimization solutions: firm vs single generator

optimal from each single generator's perspective, but is more profitable than unilaterally using each generator's optimal strategy. This makes monitoring and analyzing a generation firm's strategic behavior more challenging.

4.5.2 Different starting points

In addition to the starting point $q_5 = 200$ and $q_{30} = 200$, we initialized the algorithm with other starting points to test its robustness. Including the starting point $q_5 = 200$ and $q_{30} = 200$, we will test four different starting points:

Starting point	q_5	q_{30}
1	200.00	200.00
2	300.00	500.00
3	450.00	250.00
4	450.00	550.00

The optimization trajectories are listed in Tab. 4.2, Tab. 4.4, Tab. 4.5, and Tab. 4.6 respectively. The algorithm reliably finds the optimizer for each of the starting point within 4 to 6 iterations with run time less than 2 seconds each.

The robustness and performance of this approach is superior to the MPEC approach. As illustrated in [23], the MPEC approach suffers from the non-convexity due to the complementary constraints. The MPEC cannot reliably find the optimizer for certain starting points. As a result, one need to run the MPEC program multiple times in order to safely conclude an optimal solution. Our algorithm, in contrast, works with the primal variables, and thus has less difficulties with the non-convexity. Although theoretically it is possible that the optimization problem is non-convex, as a practical matter, the problem is typically convex or close to convex, which helps our algorithm to work robustly.

Another observation is that because the bundle method depends on history, even if different optimization path intersect at a certain point, their subsequent trajectories may not be the same. For example, iteration 1 in Tab. 4.4 coincides with iteration 1 in Tab. 4.5, but after that, the two trajectories are very different. Nevertheless, they approach the same optimizer.

iter.	q_5	q_{30}	profit	step
0	300.00	500.00	9015.0	S
1	374.59	445.67	9126.4	S
2	355.56	393.19	9131.3	S
3	362.78	434.00	9191.3	S
4	356.60	434.18	9192.3	O

Table 4.4: Optimization trajectory starting from (300, 500)

iter.	q_5	q_{30}	profit	step
0	450.00	250.00	8293.7	S
1	374.59	445.67	9126.4	N
2	374.59	445.67	9126.4	N
3	374.59	445.67	9126.4	S
4	356.64	433.78	9191.9	S
5	356.61	433.98	9192.2	O

Table 4.5: Optimization trajectory starting from (450,250)

iter.	q_5	q_{30}	profit	step
0	450.00	550.00	7694.0	S
1	326.63	399.15	9084.1	S
2	374.59	445.67	9126.4	S
3	355.20	440.33	9191.5	S
4	356.60	434.18	9192.3	O

Table 4.6: Optimization trajectory starting from (450,550)

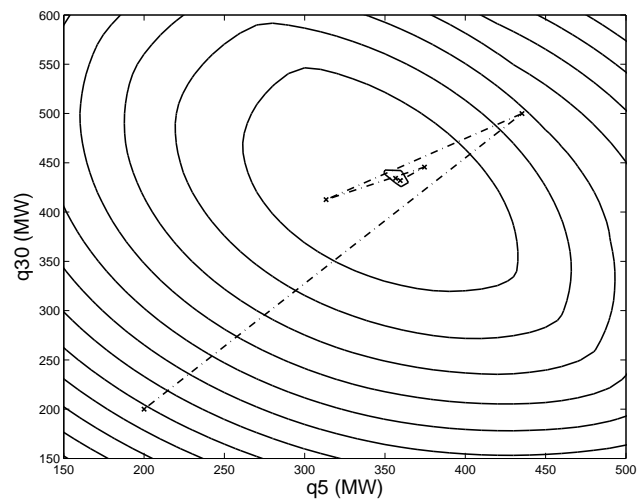


Figure 4.1: Optimization trajectory starting from (200, 200)

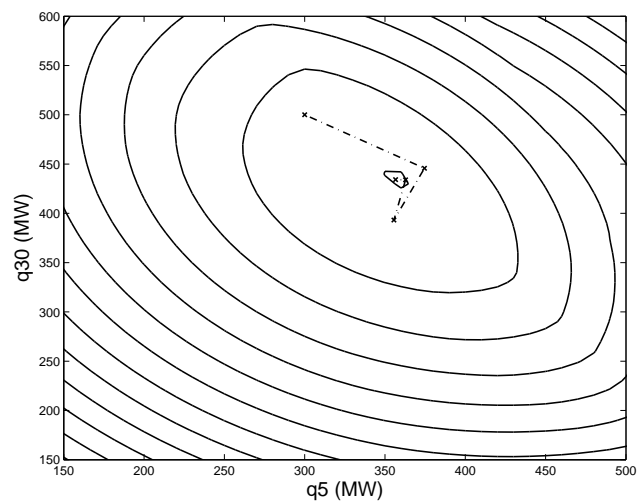


Figure 4.2: Optimization trajectory starting from (300, 500)

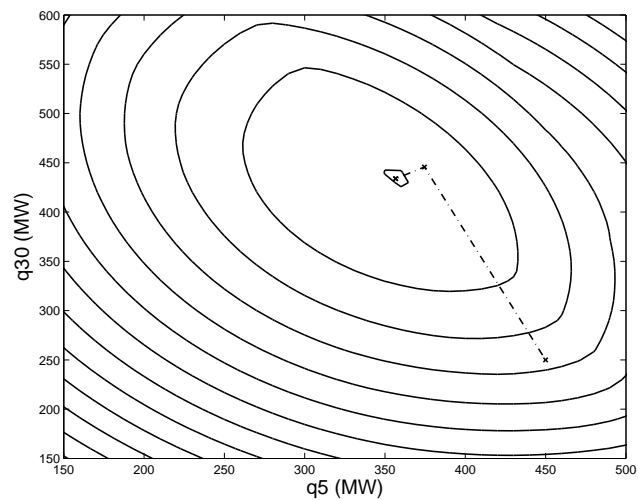


Figure 4.3: Optimization trajectory starting from $(450, 250)$

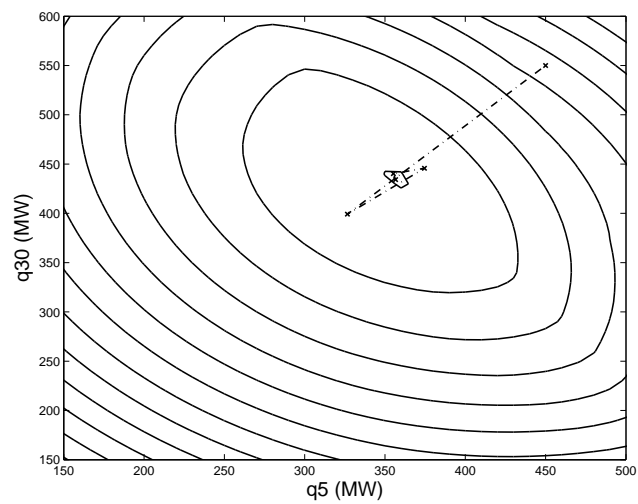


Figure 4.4: Optimization trajectory starting from $(450, 550)$

4.5.3 Performance Test

We test the performance of the algorithm by increasing the number of generators owned by generation firm A from 5 generators to 30 generators. All tests start from competitive output levels. The results are summarized in Tab. 4.7. All these scenarios finish within 15 seconds on a 2.2GHz Intel Core 2 Duo PC, with less than 35 total steps including both Serious Steps and Null Steps. The number of steps also implies the number of OPFs solved, because there is exactly one OPF solved in each step in order to evaluate the profit function, and calculate the TCRDJ. For large scale problem, iteratively solving the OPF is the most computationally intense part for this algorithm, so improving the OPF solver performance will directly improve the performance of this algorithm.

The 30 generator scenario has an infeasible OPF due to significant withholding from generation firm A. In this case, the generation firm A is pivotal meaning that without generators from generation firm A, the rest of the generators in the system are not enough to meet the total system demand. In this case, if there is no price cap in the market, generation firm A will behave pivotally to withhold till the system is short of supply to drive the price arbitrarily high. We can observe this phenomena in the 30 generator scenario in Tab. 4.7. After 20 iterations, the profit has gone above 10^{10} \$/h.

The performance of the algorithm does not directly depend on the number of generators owned by the generation firm. For example, the 20 generators scenario requires less iterations than the 15 generators scenario.

generators	Serious Steps	Null Steps	run time	profit
1–5	3	9	1.7 s	4.1448×10^3
1–10	3	0	0.6 s	5.0231×10^3
1–15	5	7	1.8 s	1.0109×10^4
1–20	3	0	0.8 s	1.0454×10^4
1–25	24	10	11.4 s	5.3263×10^5
1–30	9	11	4.4 s	$3.8379 \times 10^{10}{}^a$

^aEnergy balance has been violated.

Table 4.7: Performance test

This is because sometimes when the generation firm has more generators to leverage, some of the generator output levels can be more easily determined to stay at the capacity bounds, which leaves fewer effective decision variables.

Another observation is that some of the generators have larger impact on the profit than others. For example, the profit only changes about 1000 \$/h from the 5-generator scenario to the 10-generator scenario, while the profit changes about 5000 \$/h from the 10-generator scenario to the 15-generator scenario, which implies generator 6 – 10 are not as effective as generator 11 – 15. Similarly, generator 21 - 25 seem to have very large impact on the profit. The performance of the algorithm is affected more by the number of these highly effective generators. Generally speaking, more iterations are needed when highly effective generators are added to the portfolio.

4.6 Conclusion

In electricity markets, especially nodal electricity markets, a generation firm may own multiple generators located at multiple buses and exposed to

different LMPs. In this context, we generalize the residual demand concept from a single generator's perspective to a generation firm's perspective. The TCRDD for a single generator corresponds to a TCRDJ for a generation firm. We derived the TCRDJ based on a multi-parameter sensitivity analysis of the OPF. Then we proposed a bundle trust region algorithm to optimize a generation firm's profit based on the TCRDJ. The algorithm is applied to the IEEE 118-bus system to demonstrate the effectiveness of the method. The TCRDJ provides useful insights about a generation firm's strategic behavior. The algorithm provides an effective and promising approach for generation firms to bid into electricity markets.

Chapter 5

Conclusion

In the chapter, we summarize the dissertation and discuss future research topics.

5.1 Summary

Through the previous chapters, we have analyzed strategic behavior in electricity markets from a residual demand perspective. We started with a SFE stability analysis in chapter 2, which is an application of the residual demand method in the absence of transmission constraints, aiming at refining the electricity market supply function Nash equilibria.

Then we devoted major effort to deal with transmission constraints in chapters 3 and 4. We characterized the residual demand in a transmission-constrained network from a single generator's perspective, and then generalized the concept to a generation firm's perspective, which may own multiple generators located at different locations. The core of the residual demand characterization is the transmission-constrained residual demand derivative (TCRDD) and the transmission-constrained residual demand Jacobian (TCRDJ), which are very useful in characterizing the profit maximizing strategy. We not only

defined these concepts, but also improved their calculation efficiency and practical implementation. With those improvements, the TCRDD and TCRDJ calculation can be implemented as a light-weight OPF post-processing step. This enables us to decouple the generator or generation firm's profit maximization problem into two subproblems, the lower problem of post-OPF TCRDD or TCRDJ calculation, and the upper problem of profit maximization based on TCRDD or TCRDJ. The decoupled structure has the advantage of being able to reuse advanced OPF solvers, so this method is able to solve large scale problems.

If the whole residual demand function is given, as discussed in section 1.3, the upper problem is very easy to solve. Basically we need to find a point on the explicitly given residual demand function that maximizes the profit.

However, we do not have an explicit analytical representation of the residual demand function, and it is computationally expensive to numerically evaluate the function using the OPF. Therefore, we proposed methods to use the TCRDD or TCRDJ to find the optimizer that only require a few evaluations of OPF. In essence, the methods are based on Newton methods, as we calculate the profit function Hessian matrix using TCRDD or TCRDJ.

Generally, the residual demand function is not continuously differentiable, which results in “kinks” in the profit function. Thus another task is to deal with these non-differentiable points. We customized the bundle trust region method by Schramm and Zowe [37] to optimize the kinked profit function

based on TCRDJ. The idea is to build up a cutting plane approximation to the objective function from gathered information from previous iterations, that is, the “bundle.” We rely on the cutting plane approximation to determine the locations of the objective function kinks.

For a single generator’s profit maximization, because there is only one decision variable, profit maximization only involves a line search. We developed a special bisection search algorithm to find the optimizer. Considering the algorithm in the context of the bundle idea, one can find that the algorithm is building a two-piece quadratic function approximation in the interval $[q_{lo}, q_{hi}]$, so it is in essence a “quadratic bundle” method. For a scalar problem, the quadratic bundle problem can be solved very easily because there are at most three candidates for the optimizer. Therefore, we can use the more accurate quadratic bundle. In contrast, for a generation firm’s profit maximization, if we also use a quadratic bundle instead of the (linear) cutting plane bundle, it will introduce quadratic constraints into the optimization problem we solve in each iteration, and make it more difficult to solve in practice. This is why we use the cutting plane bundle for a generation firm’s profit maximization problem.

The contributions of the dissertation can be evaluated by the goals we discussed in chapter 1:

- from the practical point of view, the method is able to handle large scale production level systems, and it is better if the existing advanced market

simulation engines can be reused;

- from the theoretical point of view, the method is able to represent the transmission network to the same details as in the true market clearing process, and systematically find the profit maximizing strategy.

To achieve the second goal, we have represented the transmission constraints in the same way that the market clearing engine models them in deriving the TCRDD and TCRDJ. To achieve the first goal, we decouple the profit maximization problem into two easily solvable subproblems. The decoupled structure makes the computational capability of the method close to the computational capability of existing advanced OPF programs, which can solve very large scale problems. We have tested the performance of these methods on the IEEE 118-bus system. Our testing has been limited by the capability of our OPF solver and available test cases. In the future, we will do further tests in larger test systems (with thousands of buses, which is close to size of actual power systems, such as CAISO and ERCOT.) In summary, the residual demand approach provides a very promising methodology in achieving these goals.

5.2 Future research

The theory and methods proposed in the dissertation have broken ground for further research topics. This dissertation directly studies how to find the profit maximizing strategies for a generator or generation firm. This

can help small generation firms to participant in the market, but may also help oligopolistic generation firms understand how to leverage transmission constraints to exercise market power. One natural future research topic is from a market monitor’s perspective how to monitor potential market power in electricity markets.

The studies in this field have been very *ad hoc* historically. There are two major reasons for this. One reason is that the classic market power indices, such as the Herfindahl-Hirschman Index (HHI), do not have a solid connection with price markups or any other tangible measures unless in a Cournot model. The other reason is that the transmission constraints have been represented unrealistically.

In industry, several electricity markets have used capacity-based HHI or similar measures to monitor market power, which do not have any theoretical justification [41]. Some examples of such measures include the “Element Competitiveness Index” (ECI) [16] in the proposed ERCOT nodal market and the “three pivotal supplier test” [31] in the PJM market. The ECI considers each transmission line in the system separately to determine if there is enough competition to resolve congestion from the import side and the export side defined by positive and negative shift factors. Then the HHI is calculated for the import side and the export side. The representation of the transmission network model is partial in this measure in that the line-by-line examination neglects interaction of multiple binding constraints in the system. For example, a generator with 0.5 shift factor that is considered to be effective in resolving

congestion in the ECI model may not actually be able to resolve the congestion if it would then overload another line by doing so. The “three pivotal supplier test” in the PJM market is very similar to the ECI test except it is doing a pivotal supplier test instead of calculating the HHI. The effectiveness of these type of methods is very questionable.

The transmission-constrained residual demand concept can be used to derive market power indices with solid theoretical foundation. We will give two examples with the hope of spurring ideas.

Price Markup The price markup is defined as the difference between the strategic price and the competitive price. By (2.2), at the profit maximizing point,

$$p - C'_i(q_i) = -P'_i(q_i)q_i. \quad (5.1)$$

Note that the left hand side of (5.1) is the difference between the profit maximizing price and the corresponding marginal cost. This property can be used to measure potential price markup for a generator assuming the competitive price is higher than the marginal cost. The potential price markup characterizes how large the incentive is for a generator to drive up the price above the competitive level in the process of pursuing maximum profit. Similarly for a generation firm A:

$$p_A - C'_A(q_A) = -\frac{\partial P_A}{\partial q_A}q_A, \quad (5.2)$$

so the same logic could be applied to a generation firm.

Profit Markup The profit markup is defined as the difference between

the profit by bidding strategically and the competitive profit. At the profit maximizing point,

$$q_i (p - C'_i(q_i)) = -P'_i(q_i)q_i^2. \quad (5.3)$$

Note that the left hand side of (5.3) is the output quantity multiplied by the price markup above the marginal cost, which can be used to measure the potential profit markup in the process of pursuing maximum profit. Similarly for a generation firm A:

$$\mathbf{q}_A^T (\mathbf{p}_A - \mathbf{C}'_A(\mathbf{q}_A)) = \mathbf{q}_A^T \frac{\partial \mathbf{P}_A}{\partial \mathbf{q}_A} \mathbf{q}_A, \quad (5.4)$$

so the same logic could be applied to a generation firm as well.

The two examples above are both sensitivity based measures. They can be calculated with one run of the market clearing process. Similar measures can be calculated using simulation based methods. With simulation, one can clear the market with true cost, then clear the market with the offers and bids, and compare the differences in prices and profits. The appropriateness of sensitivity based measures versus simulation based measures depends on the application. The simulation based methods are more suitable for offline studies to determine overall market performance over time, say a month. The sensitivity based indices above are more suitable for online operations to assist market power identification and mitigation. We will continue working on using the transmission-constrained residual demand concept to define market power indices in assessing electricity market competitiveness, and design online market power mitigation mechanisms.

It is my hope to see real applications of the residual demand based methods in electricity markets in the future.

Appendices

Appendix A

Ordinary Least Squares Problem and Weighted Least Squares Problem

The formulation and results in this section are from [22]. An Ordinary least squares (OLS) problem is formulated as follows. Suppose there are n observations (Y_i, \mathbf{X}_i) , $\forall i = 1, \dots, n$. The objective is to find an optimal vector $\boldsymbol{\beta}$ that minimizes the Sum of Squares Error (SSE):

$$\min_{\boldsymbol{\beta}} \quad SSE^{\text{OLS}}(\boldsymbol{\beta}) = \sum_{i=1}^n (Y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2$$

The solution to this OLS problem is

$$\mathbf{b}^{\text{OLS}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y},$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \vdots \\ \mathbf{X}_n^T \end{bmatrix},$$
$$\mathbf{Y} = [Y_1 \ Y_2 \ \dots \ Y_n]^T,$$

assuming there is no multicollinearity, i.e. \mathbf{X} has linearly independent columns.

The minimal SSE is:

$$SSE^{\text{OLS}}(\mathbf{b}^{\text{OLS}}) = \mathbf{Y}^T \left(\mathbf{I} - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \mathbf{Y}.$$

Define a projection matrix $\mathbf{P}_{\mathbf{X}}$ by:

$$\mathbf{P}_{\mathbf{X}} = \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T.$$

Define another projection matrix \mathbf{M} by:

$$\mathbf{M}_{\mathbf{X}} = \mathbf{I} - \mathbf{P}_{\mathbf{X}}.$$

Both matrices $\mathbf{P}_{\mathbf{X}}$ and $\mathbf{M}_{\mathbf{X}}$ are *idempotent*, namely:

$$\mathbf{P}_{\mathbf{X}}^2 = \mathbf{P}_{\mathbf{X}},$$

$$\mathbf{M}_{\mathbf{X}}^2 = \mathbf{M}_{\mathbf{X}}.$$

In addition, both matrices \mathbf{P} and \mathbf{M} are positive semi-definite so that:

$$SSE^{\text{OLS}}(\mathbf{b}^{\text{OLS}}) = \mathbf{Y}^T \mathbf{M}_{\mathbf{X}} \mathbf{Y} \geq 0.$$

Define the residual:

$$\mathbf{e} = \mathbf{M}_{\mathbf{X}} \mathbf{Y},$$

then:

$$\mathbf{X}^T \mathbf{e} = \mathbf{0}.$$

Suppose we want to add one regressor to the problem. Now there are n observations (Y_i, \mathbf{X}_i, z_i) , $\forall i = 1, \dots, n$, with z_i added. Again assume there is no multicollinearity with z_i added.

The objective is to find an optimal vector $\boldsymbol{\beta}$ and γ that minimizes the SSE:

$$\min_{\boldsymbol{\beta}, \gamma} SSE^{\text{OLS}}(\boldsymbol{\beta}, \gamma) = \sum_{i=1}^n (Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - z_i \gamma)^2.$$

We have:

$$\left\{ \min_{\boldsymbol{\beta}, \gamma} SSE^{\text{OLS}}(\boldsymbol{\beta}, \gamma) \right\} = \left\{ \min_{\boldsymbol{\beta}} SSE^{\text{OLS}}(\boldsymbol{\beta}) \right\} - c^2 (z^T \mathbf{M}_{\mathbf{X}} z),$$

where $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_n]^T$, and:

$$c = (\mathbf{z}^T \mathbf{M}_{\mathbf{X}} \mathbf{z})^{-1} \mathbf{z}^T \mathbf{M}_{\mathbf{X}} \mathbf{Y}.$$

Note that if:

$$\mathbf{M}_{\mathbf{X}} \mathbf{z} \neq 0,$$

and:

$$\mathbf{M}_{\mathbf{X}} \mathbf{Y} \neq 0,$$

then:

$$\min_{\boldsymbol{\beta}, \gamma} SSE^{\text{OLS}}(\boldsymbol{\beta}, \gamma) < \min_{\boldsymbol{\beta}} SSE^{\text{OLS}}(\boldsymbol{\beta}).$$

If we put weights on different observations, then it is a weighted least squares (WLS) problem with the following objective:

$$\min_{\boldsymbol{\beta}} SSE = \sum_{i=1}^n w_i (Y_i - \mathbf{X}_i^T \boldsymbol{\beta})^2.$$

The solution to this WLS problem is:

$$\mathbf{b}^{\text{WLS}} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y},$$

where the weight matrix is defined by:

$$\mathbf{W} = \begin{bmatrix} W_1 & 0 & \dots & 0 \\ 0 & W_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & W_n \end{bmatrix}.$$

The minimal SSE is:

$$SSE^{\text{WLS}} \mathbf{b}^{\text{WLS}} = \mathbf{Y}^T \mathbf{W} \mathbf{Y} - \mathbf{Y}^T \mathbf{W} \mathbf{X} (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y}.$$

Let us assume \mathbf{W} is positive semi-definite so that $\mathbf{W}^{1/2}$ exists:

$$\mathbf{W}^{1/2} \mathbf{W}^{1/2} = \mathbf{W}.$$

The WLS could be transformed to an equivalent OLS problem by defining:

$$\mathbf{Y}^* = \mathbf{W}^{1/2} \mathbf{Y},$$

$$\mathbf{X}^* = \mathbf{W}^{1/2} \mathbf{X},$$

where:

$$\mathbf{W}^{1/2} = \begin{bmatrix} w_1^{1/2} & 0 & \dots & 0 \\ 0 & w_2^{1/2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{n-1}^{1/2} \end{bmatrix}.$$

The equivalent OLS to (3.34) is:

$$\min_{\boldsymbol{\beta}} SSE^{\text{OLS}}(\boldsymbol{\beta}) = \sum_{i=1}^{n-1} (Y_i^* - \mathbf{X}_i^{*T} \boldsymbol{\beta})^2.$$

Bibliography

- [1] E. J. Anderson and A. B. Philpott. Using supply functions for offering generation into an electricity market. *Operations Research*, 50(3):477–489, 2002.
- [2] E. J. Anderson and H. Xu. Nash equilibria in electricity markets with discrete prices. *mathematical methods of operations research*, 60(2):215–238, 2004.
- [3] E. J. Anderson and H. Xu. Supply function equilibrium in electricity spot markets with contracts and prices caps. *Journal of Optimization Theory and Applications*, 60(2):215–283, February 2005.
- [4] Edward J. Anderson and Xinmin Hu. Finding supply function equilibria with asymmetric firms. *Operations Research*, 56(3):697–711, 2008.
- [5] Alvaro Baillo, Mariano Ventosa, Michel Rivier, and Andres Ramos. Optimal offering strategies for generation companies operating in electricity spot markets. *IEEE Transactions on Power Systems*, 19(2):745–753, May 2004.
- [6] Ross Baldick. Electricity market equilibrium models: The effect of parametrization. *IEEE Transactions on Power Systems*, 17(4):1170–1176, November 2002.

- [7] Ross Baldick, Ryan Grant, and Edward Kahn. Theory and application of linear supply function equilibrium in electricity markets. *Journal of Regulatory Economics*, 25(2):143–167, 03 2004.
- [8] Ross Baldick and William Hogan. Capacity constrained supply function equilibrium models of electricity markets: Stability, nondecreasing constraints, and function space iterations. POWER Working Papers 089, University of California Energy Institute, December 2001.
- [9] Ross Baldick and William W. Hogan. Polynomial approximations and supply function equilibrium stability (aug-04). Working Paper Series rwp05-028, Harvard University, John F. Kennedy School of Government, Mar 2005.
- [10] Carolyn A. Berry, Benjamin F. Hobbs, William A. Meroney, Richard P. O’Neill, and William R. Stewart Jr. Understanding how market power can arise in network competition: a game theoretic approach. *Utilities Policy*, 8(3):139–158, September 1999.
- [11] Severin Borenstein, James. Bushnell, and Steven Stoft. The competitive effects of transmission capacity in a deregulated electricity industry. *RAND Journal of Economics*, 31(2):294–325, Summer 2000.
- [12] Judith B. Cardell, Carrie Cullen Hitt, and William W. Hogan. Market power and strategic interaction in electricity networks. *Resource and Energy Economics*, 19(1-2):109–137, March 1997.

- [13] A.I. Cohen, V. Brandwahjn, and S.-K. Chang. Security constrained unit commitment for open markets. In *Proceedings of the 21st 1999 IEEE International Conference*, pages 39–44, May 1999.
- [14] Antonio J. Conejo, Enrique Castillo, Roberto Minguez, and Federico Milano. Locational marginal price sensitivities. *IEEE Transactions on Power Systems*, 20(4):2026–2033, November 2005.
- [15] Christopher J. Day, Benjamin F. Hobbs, and Jong-Shi Pang. Oligopolistic competition in power networks: A conjectured supply function approach. *IEEE Transactions on Power Systems*, 17:597–607, 2002.
- [16] ERCOT. ERCOT nodal protocols section 3.19: Constraint competitiveness tests. <http://nodal.ercot.com/protocols/index.html>.
- [17] ERCOT. ERCOT nodal protocols section 7: Congestion revenue rights. <http://nodal.ercot.com/protocols/index.html>.
- [18] T. Genc and S. Reynolds. Supply function equilibria with pivotal electricity suppliers. In *Eller College Working Paper 1001-04*, July 2004.
- [19] Julian Barquin Gil. Symmetry properties of conjectural price responses. In *IEEE General Meeting 2008*. IEEE Power Engineering Society, 2008.
- [20] Richard J Green. Increasing competition in the British electricity spot market. *Journal of Industrial Economics*, 44(2):205–16, June 1996.

- [21] Richard J Green and David M Newbery. Competition in the British electricity spot market. *Journal of Political Economy*, 100(5):929–53, October 1992.
- [22] William H. Greene. *Econometric Analysis*. Prentice Hall, Upper Saddle River, New Jersey, 2000.
- [23] Benjamin F. Hobbs, Carolyn B. Metzler, and Jong-Shi Pang. Strategic gaming analysis for electric power systems: An MPEC approach. *IEEE Transactions on Power Systems*, 15(2):638–645, 2000.
- [24] Pär Holmberg. Asymmetric supply function equilibrium with constant marginal costs. Working Paper Series 2005:16, Uppsala University, Department of Economics, April 2005.
- [25] Pär Holmberg. Unique supply function equilibrium with capacity constraints. *Energy Economics*, 30(1):148–172, January 2008.
- [26] Ali Hortaçsu and Steven L. Puller. Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market. *RAND Journal of Economics*, 39(1):86–114, 2008.
- [27] Paul D Klemperer and Margaret A Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):1243–77, November 1989.
- [28] B. C. Lesieutre, H. S. Oh, R. J. Thomas, and V. Donde. Identification of market power in large-scale electric energy markets. In *HICSS '06: Pro-*

ceedings of the 39th Hawaii International Conference on System Sciences, Hawaii, USA, 2006.

- [29] S. Lopez de Haro, P. Sanchez Martin, J.E. de la Hoz Ardiz, and J. Fernandez Caro. Estimating conjectural variations for electricity market models. *European Journal of Operational Research*, 181(3):1322–1338, September 2007.
- [30] Hans Mittelmann. MPEC benchmark (5-2-2008)). Technical report, Arizona State University, May 2008.
- [31] PJM Market Monitor. 2006 state of market report volume ii: Detailed analysis. <http://www2.pjm.com/markets/market-monitor/downloads/mmu-reports/2006-som-volume-ii.pdf>.
- [32] C. E. Murillo-Sanchez, S. M. Ede, T. D. Mount, R. J. Thomas, and R. D. Zimmerman. An engineering approach to monitoring market power in restructured markets for electricity. In *Proceedings of the 24th Annual International Conference*, Houston, Texas, April 2001.
- [33] Hui Niu, Ross Baldick, and Guidong Zhu. Supply function equilibrium bidding strategies with fixed forward contracts. *IEEE Transactions on Power Systems*, 20(4):1859–1867, November 2005.
- [34] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, 1999.

- [35] T. Orfanogianni and G. Gross. A general formulation for LMP evaluation. *IEEE Transactions on Power Systems*, 22(3):1163–1173, 2007.
- [36] Aleksandr Rudkevich. Supply function equilibrium in power markets: learning all the way. Technical report, Tabors Caramanis and Associates, Cambridge, MA, September 1999.
- [37] Helga Schramm and Jochem Zowe. A version of the bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results. *SIAM J. Optimization*, 2(1):121–152, February 1992.
- [38] H. Singh, S. Hao, and A. Papalexopoulos. Transmission congestion management in competitive electricity markets. *IEEE Transactions on Power Systems*, 13(2):672–680, 1998.
- [39] Fereidoon P. Sioshansi and Wolfgang Pfaffenberger. *Electricity market reform: an international perspective*. Elsevier, Amsterdam, 2007.
- [40] Ramteen Sioshansi and Shmuel Oren. How good are supply function equilibrium models: an empirical analysis of the ERCOT balancing market. *Journal of Regulatory Economics*, 31(1):1–35, February 2007.
- [41] Steven Stoft. *Power System Economics: Designing Markets for Electricity*. Wiley-IEEE Press, 2002.
- [42] B. Willems, I. Rumiantseva, and H. Weigt. Cournot versus supply functions: What does the data tell us? *Energy Economics*, 31(1):38–47, January 2009.

- [43] Robert Wilson. Supply function equilibrium in a constrained transmission system. *Operations Research*, 56(2):369–382, March-April 2008.
- [44] Frank A. Wolak. Identification and estimation of cost functions using observed bid data: An application to electricity markets. NBER Working Papers 8191, National Bureau of Economic Research, Inc, Mar 2001.
- [45] Allen J. Wood and Bruce F. Wollenberg. *Power Generation, Operation, and Control 2nd ed.* John Wiley & Sons, Inc., New York, NY, 1996.
- [46] Lin Xu and Ross Baldick. Transmission-constrained residual demand derivative in electricity markets. *IEEE Transactions on Power Systems*, 22(4):1563–1573, November 2007.
- [47] Lin Xu and Ross Baldick. Stability of supply function equilibrium in electricity markets under piecewise polynomial function perturbations. In *Proceedings of the Allerton Conference on Communication, Control, and Computing*, September 2008.
- [48] Lin Xu and Yixin Yu. Transmission constrained linear supply function equilibrium in power markets: Method and example. In *PowerCon 2002: International Conference on Power System Technology Proceedings*, volume 3, pages 1349–1354. IEEE Power Engineering Society, 2002.
- [49] Jian Yao, Shmuel S. Oren, and Ilan Adler. Computing cournot equilibria in two settlement electricity markets with transmission constraints. In *HICSS '04: Proceedings of the Proceedings of the 37th Annual Hawaii*

International Conference on System Sciences (HICSS'04) - Track 2, page
20051.2, Washington, DC, USA, 2004. IEEE Computer Society.

Vita

Lin Xu received the Bachelor of Science degree in Engineering and the Master of Science degree in Engineering from Tianjin University in 2000 and 2003 respectively. He started his graduate studies in electrical engineering at the University of Texas at Austin in Fall 2003. In 2006, he worked as an internship at the Lower Colorado River Authority. From 2007 to 2008, he worked as a software developer and power engineering at the Electric Reliability Council of Texas. Currently, he is an engineering specialist in the Department of Market Monitoring, California Independent System Operator.

Permanent address: 3517 N Hills Dr. Apt. U101, Austin, TX 78731

This manuscript was typed by the author.