Nonconvex and Nonsmooth Optimization for Robust Control and Power Systems

Ye Shi

Faculty of Engineering and Information Technology
University of Technology Sydney

This dissertation is submitted for the degree of
Doctor of Philosophy

September 2018

Declaration

I, Ye Shi, declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Production Note:

Signature: Signature removed prior to publication.

Date: 10th. Sep. 2018.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Tuan D. Hoang for his guidance, motivation and continuous support throughout my Ph.D study and related research. Prof. Tuan has brought me into the topics of optimization in smart grid and robust control system. His brilliant insights and guidance have helped me in all the time of research and writing of this thesis. I feel extremely fortunate to have a supervisor who helps me so lot in both Ph.D study and my life.

I am deeply appreciative of my co-supervisor Prof. Steven W. Su for his constant and patient help for my research and life. He has kindly offered help to resolve confusions in research and difficulties I meet in life. His insightful and generous advises inspire me to always keep positive and enthusiastic when I encounter the difficulties.

I would also like to thank Prof. Pierre Apkarian (the French Aerospace Lab, France), Prof. Andrey V. Savkin (University of New South Wales, Australia), Prof. H. Vincent Poor (Princeton University, USA), Prof. Trung Q. Duong (Queen's University Belfast, UK), and Professor Li Li (University of Technology Sydney, Australia) for their precious instruction for my research.

I owe my sincere gratitude my colleagues in FEIT of UTS, in particular, Enlong Che, Zhichao Sheng, Tao Zhang, Lin Ye and Wentian Zhang for their selfless support and help. I also enjoy the precious memory with my friends, particularly, Yi Xu and Tianqi Lu who gave me their help and time in listening to me and helping me work out my problems during the Ph.D studies.

Last my thanks would go to my beloved parents Biguang Shi and Caixia Zhang and my entire family for their loving considerations and great confidence in me through these years.

Abstract

There is considerable interest in structured $\mathcal{H}_\infty$ control in linear parameter varying (LPV) systems and Tagaki-Sugeno (T-S) fuzzy systems. The optimal power flow (OPF) problem, Electrical vehicles (EVs) charging problem and optimal placement of phasor measurement unit (PMU) have emerged as promising areas in power systems. This dissertation focuses on noncovex and nonsmooth optimization for the structured $\mathcal{H}_\infty$ control problem in LPV systems and T-S fuzzy systems and the OPF problem, the EV charging problem and the PMU placement problem in power systems.

We first consider reduced order LPV-LFT (linear parameter varying-linear fractional transformational) control synthesis. The reduced order control synthesis can be reformulated as a linear matrix inequality (LMI) optimization subject to a rank constraint on a matrix-valued affine function of the Lyapunov matrix variables. Finding a good reduced-order stabilizing controller is not an easy task because its computation is a NP-hard problem. A novel approach proposed in this thesis is to equivalently express the rank constraints on a positive semi-definite matrix-valued affine function by spectral nonlinear functions. We then show a simple but effective nonsmooth optimization technique leading to a path-following optimization procedure for these problems. An intensive simulation shows the clear advantage of the proposed method over the state-of-the-art nonlinear matrix inequality solvers.

In the second part of the dissertation, we investigate the $\mathcal{H}_\infty$ Proportional-integral-derivative (PID) control design in fuzzy systems. To gain the practicability and tractability of fuzzy systems, this thesis develops a parameterized bilinear matrix inequality characterization for the $\mathcal{H}_\infty$ fuzzy PID control design, which is then relaxed into a bilinear matrix inequality optimization problem of nonconvex optimization.

Several computational procedures are then developed for its solution. The merit of the developed algorithms is shown through the benchmark examples.

Thirdly, we consider the optimal power flow (OPF) problem over transmission networks. The OPF problem is to locate a steady state operating point such that the cost of electric power generation is minimized subject to operating constraints and meeting demand. Due to the highly nonlinear operation constraints, the OPF problem has been known as an NP-hard problem. The existing nonlinear solvers may fail in yielding a feasible point. Semi-definite relaxation (SDR) could provide the global solution only when the matrix solution of the relaxed semidefinite program (SDP) is of rank-one, which does not hold in general. We develop a nonsmooth optimization approach to address this difficult OPF problem, which is an iterative process to generate a sequence of improved points. We also develop an efficient decomposition for the large-scale OPF problem, which involves reduced numbers of the rank-one constraints on matrices of moderate size for expressing the network nonlinear constraints. Simulations for OPF problems and large-scale OPF problems demonstrate the efficiency of our approaches.

In the fourth section of this dissertation, we study the charging scheduling of plug-in electric vehicles (PEVs) and power control in smart grid. PEV charging scheduling aims at minimizing the potential impact of the massive integration of PEVs into smart grid to save service costs to customers while power control aims at minimizing the cost of power generation subject to operating constraints and meeting demand. A model predictive control (MPC)-based approach is proposed to address the joint PEV charging scheduling and power control to minimize both PEV charging cost and energy generation cost in meeting both residence and PEV power demands. Unlike in related works, no assumptions are made about the probability distribution of PEVs' arrivals, the known PEVs' future demand, or the unlimited charging capacity of PEVs. The proposed approach is shown to achieve a globally optimal solution. Numerical results for IEEE benchmark power grids serving Tesla Model S PEVs show the merit of this approach.

Finally, we consider the PMU placement problem for power grid state estimation under different degrees of observability. Observability degree is the depth of the buses'

reachability by the placed PMUs and thus constitutes an important characteristic for PMU placement. However, the sole observability as addressed in many works still does not guarantee a good estimate for the grid state. Some existing works also considered the PMU placement for minimizing the mean squared error or maximizing the mutual information between the measurement output and grid state. However, they ignore the obsvervability requirements for computational tractibility and thus potentially lead to artificial results such as acceptance of the estimate for an unobserved state component as its unconditional mean. In this dissertation, the PMU placement optimization problem is considered by minimizing the mean squared error or maximizing the mutual information between the measurement output and grid state, under grid observability constraints. The provided solution is free from the mentioned fundamental drawbacks in the existing PMU placement designs. The problems are posed as binary nonlinear optimization problems, for which this paper develops efficient algorithms for computational solutions. The performance of the proposed algorithms is analyzed in detail through numerical examples on large scale IEEE power networks.

Publications

The contents of this thesis are based on the following papers that have been published, accepted, or submitted to peer-reviewed journals and conferences.

Journal Papers:

1. Y. Shi, H. D. Tuan and P. Apkarian, "Nonconvex Spectral Optimization Algorithms for Reduced-Order H_∞ LPV-LFT controllers", International Journal of Robust and Nonlinear Control, vol. 27, pp. 4421-4442, 2017.

2. Y. Shi, H. D. Tuan, H. Tuy and S. W. Su, "Global Optimization for Optimal Power Flow over Transmission Networks", Journal of Global Optimization, vol. 69, pp. 745-760, 2017.

3. Y. Shi, H. D. Tuan, A. V. Savkin, S. W. Su, "Optimal Power Flow over Large-Scale Transmission Networks", Systems & Control Letters, vol. 118, pp. 16-21, 2018.

4. Y. Shi, H. D. Tuan, A. V. Savkin, T. Q. Duong and H. V. Poor, "Model Predictive Control for Smart Grids with Multiple Electric-Vehicle Charging Stations", accepted by IEEE Transaction on smart grid, 2017.

5. Y. Shi, H. D. Tuan, "Parameterized Bilinear Matrix Inequality Techniques in Fuzzy PID Control Design", under submission to IEEE Transactions on Fuzzy System, 2018.

6. Y. Shi, H. D. Tuan, A. V. Savkin, T. Q. Duong and H. V. Poor, "On-off Charging of Electrical Vehicles in Smart Grids", under submission to IEEE Transaction on smart grid, 2018.

7. Y. Shi, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "PMU Placement Optimization for Smart Grid Obervability and State Estimation", under submission to IEEE Transaction on smart grid, 2018.

Conference Papers:

1. Y. Shi, H. D. Tuan, S. W. Su and H. H. M. Tam, "Nonsmooth Optimization for Optimal Power Flow over Transmission Networks", the 3rd IEEE Global Conference on Signal and Information Processing, pp. 1141-1144, 2015, Orlando, America.

2. Y. Shi, H. D. Tuan, S. W. Su, and A. V. Savkin, "Multiple Matrix Rank Constrained Optimization for Optimal Power Flow over Large Scale Transmission Networks", proceedings of the 5th International Conference on Smart Cities and Green ICT Systems, vol. 1, pp. 384-389, 2016, Rome, Italy.

3. Y. Shi, H. D. Tuan, and S. W. Su, "Nonconvex Spectral Algorithm for Solving BMI on the Reduced Order H_∞ Control", the 6th IEEE International Conference on Control Systems, Computing and Engineering, 2016, Penang, Malaysia.

4. Y. Shi, H. D. Tuan, and A.V. Savkin, "Three-phase Optimal Power Flow for Smart Grids by Iterative Nonsmooth Optimization", the 6th International Conference on Smart Cities and Green ICT Systems, 2017, Porto, Portugal.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

This chapter starts with the motivation and scope of this thesis, then introduces some related research topics, and finally presents outlines of this thesis.

1.1 Motivation and Scope

Robust control is an approach of controller design to cope with system uncertainty [8]. Robust control method aims at achieving robust performance and stability with given modeling errors[9]. The theory of robust control started in the late 1970s and soon a few techniques handling system uncertainty were proposed[8]. One of the most important applications of a robust control technique is $\mathcal{H}_\infty$ control, which was originally developed by George Zames [10]. $\mathcal{H}_\infty$ control method is to minimize the sensitivity of a system, such that the closed-loop system is internally stable. Although the $\mathcal{H}_\infty$ control has been studied for a long time, there still have been considerable interest in the structured $\mathcal{H}_\infty$ control, such as reduced order controller design [11] and proportional-integral-derivative (PID) controller design[12–14] in the last few years. It is well known that the structured $\mathcal{H}_\infty$ control problem can be reformulated as an optimization problem with bilinear matrix inequality (BMI)[11, 15, 16]. Within the past few years, it has been realized that all BMI solvers [3, 4], which address the control synthesis for linear time varying (LTI) systems in state space using Lyapunov functions, could hardly compete with the nonsmooth optimization solver developed earlier in [17],

which addresses the problems directly in the frequency domain to bypass the Lyapunov variables of high dimension. Nowadays, the Matlab systune command [18], which is based on [17], is the most powerful tool for control synthesis of LTI systems and is widely used in industry. This means that structured $\mathcal{H}_\infty$ control problem should seek applications outside uncertain LTI systems such as linear parameter varying (LPV) systems [19] and Tagaki-Sugeno (T-S) fuzzy systems [20], where the Lyapunov function is irreplaceable.

The application of optimization techniques to power systems has been a promising area in recent years. The optimal power flow (OPF) is the key problem of the power system. This problem is complex economically, electrically and computationally[21]. However, we still lack a fast and robust solution for the OPF. Today's approximate method may result in unnecessary environmental harm and extra energy consumption. The optimal power flow (OPF) problem is to locate a steady state operating point in an AC power network such that the cost of electric power generation is minimized subject to operating constraints and meeting demand. Since its introduction by [22], OPF has received considerable interest (see e.g. [23–25] and references therein) but its solution remains largely open. As the basic quantities in power networks can be expressed in terms of the local bus voltages from Kirchhoff's voltage law, OPF can be represented by highly nonlinear optimization problems in voltage complex variables, whose NP-hard computational complexity has been particularly shown in [26]. The underlying difficulty of OPF lies on the multiple nonlinear constraints on the voltages variables due to the bus interconnections, hardware operating capacity and balance between power demand and supply. These nonlinear constraints are difficult so the state-of-the-art nonlinear optimization solvers may converge to just stationary points (see [27] and references therein).

1.2 Structured $\mathcal{H}_\infty$ control in LPV system and OPF in power system

This dissertation focuses on noncovex and nonsmooth optimization for the structured $\mathcal{H}_\infty$ control problem in LPV system and the OPF problem in power system.

1.2.1 Reduced order $\mathcal{H}_\infty$ control

The reduced order $\mathcal{H}_\infty$ control synthesis for linear time invariant (LTI) systems can be formulated by rank-constrained optimization involving rank constraints on matrix-valued functions of the decision variables[11]. It is as a linear matrix inequality (LMI) optimization subject to a rank constraint on a matrix-valued affine function of the Lyapunov matrix variables. Many other important and difficult problems in robust control are also reformulated in similar matrix-rank constrained optimizations [28]. The simplest approach is to relax or just to drop the rank constraints with hope that the optimal solution of the relaxed (convex) optimization would satisfy these matrix-rank constraints. For instance, matrix trace minimization and nuclear norm minimization were proposed to obtain low matrix rank of positive semi-definite matrix and rectangular matrices, respectively [29, 30]. These techniques are unable to address the matrix-rank constraints as they are. Indeed, just a trace of a matrix or its nuclear norm don't give any adequate indication on the matrix rank. Another attempt is to use a Newton-like method to find a projection of a positive semi-definite matrix to the manifold of fixed rank matrices [31, 32], which is an equally computationally difficult optimization due to complex geometry of this manifold [33], especially for lower fixed rank matrices of larger size. Realizing the challenge by these matrix-rank constraints on the Lyapunov matrix variables, most later developments in robust control preferred to avoid them in favor of alternative bilinear matrix inequality (BMI) [34, 28, 35–37, 3, 38, 4]. The state-of-the-art BMI solvers [3, 4] initialize from a reduced-order stabilizing controller and then move within a convex feasibility subset containing this initialized point. There are a few difficulties which arise from these kinds of feasibility algorithms. Firstly, finding a good reduced-order stabilizing controller is not an easy task because its computation is still a NP-hard problem [39]. Secondly, the feasibility set of reduced order stabilizing controllers is highly nonconvex, which is disconnected in general. This means moving within a convex neighborhood of such a reduced-order stabilizing controller may be trapped by local minima. Thirdly, usually the convergence of these kinds of algorithms are slow and dependent very much on the local geometry around such initial point [40], which may be unpredictable. Within the past few years, it has been realized that all BMI solvers [3, 4], which address the control synthesis for LTI systems in state space using Lyapunov functions, could hardly compete with the

nonsmooth optimization solver developed earlier in [17], which addresses the problems directly in the frequency domain to bypass the Lyapunov variables of high dimension. Nowadays, the Matlab systune command [18], which is based on [17], is the most powerful tool for control synthesis of LTI systems and is widely used in industry. This means that rank-constrained optimization and BMI should seek applications outside uncertain LTI systems such as linear parameter varying (LPV) systems [19], where Lyapunov function is irreplaceable.

1.2.2 $\mathcal{H}_\infty$ fuzzy PID control

Tagaki-Sugeno (T-S) fuzzy model [20] has been proved as one of the most practical tools for representing complex nonlinear systems by gain-scheduling systems, which are easily implemented online. Treating T-S fuzzy models as gain-scheduling systems allows the application of advanced gain-scheduling control techniques in tackling state feedback and output feedback stabilization of nonlinear systems [41, 42]. Until now, most of the gain-scheduling controllers are assumed structure-free and full-rank to admit computationally tractable parameterized linear matrix inequality (PLMI) or linear matrix inequality (LMI) formulations [43, 41, 42].

Meanwhile, a proportional-integral-derivative (PID) structured controller is the indispensable component of industrial control so that PID control theory is still the subject of recent research [44–49], mainly concerning linear time-invariant systems in the frequency domain. PID controller for fuzzy systems has been considered in [13]. Reference [12] proposed an LMI based iterative algorithm for a proportional-integral (PI) controller in T-S systems under the specific structure of both system and controller. A recent work [14] transformed the fuzzy diagonal PID controller into a static output feedback problem with the dimension of controller dramatically increased. That is why all its testing examples are restricted to single input and single output systems with two states.

1.2.3 OPF problem in power system

The optimal power flow (OPF) problem is to locate a steady state operating point in an AC power network such that the cost of electric power generation is minimized subject to operating constraints and meeting demand. Since its introduction by [22], OPF has received considerable interest (see e.g. [23–25] and references therein) but its solution remains largely open. As the basic quantities in power networks can be expressed in terms of the local bus voltages from Kirchhoff's voltage law, OPF can be represented by highly nonlinear optimization problems in voltage complex variables, whose NP-hard computational complexity has been particularly shown in [26]. The underlying difficulty of OPF lies on the multiple nonlinear constraints on the voltages variables due to the bus interconnections, hardware operating capacity and balance between power demand and supply. These nonlinear constraints are difficult so the state-of-the-art nonlinear optimization solvers may converge to just stationary points (see [27] and references therein).

There has been recently renewed interest in the reformulation of the nonlinear constraints as convex constraints plus the nonconvex rank-one constraints on the matrix of outer product of voltage vector variables for the solution of OPF [50, 26]. For instance, by modifying numerical settings of some IEEE benchmark networks, [26] found that the matrix solution of the semi-definite relaxation (SDR) by dropping the rank-one matrix constraint, is of rank-one and hence the global solution of the OPF problem is found. Such SDR, dropping the rank-one matrix constraint, also provides the rank-one matrix solution in the so-called load over-satisfaction conditions, under which the quadratic equality constraints are essentially relaxed to loose one-sided inequality constraints. However, the point found from SDR is not necessarily feasible for the nonlinear equality constraints.

Indefinite quadratic programs find the most difficulty from the multiple quadratic equality constraints, which lead to nonzero duality gap and thus rank more than one of the matrix solution of SDR [51]. Another important contribution in SDR is provided in [7], where SDR has been shown to provide the rank-one matrix solution in the power distribution networks. When the networks are transmissions, [7] aims at a theoretical low-rank matrix solution of SDR, which could not lead to a feasible point for the

original nonconvex OPF problem. The common drawback of the SDR approaches in [26, 7, 52] is that once a rank-one matrix solution is not found there is no way to find even a feasible point of the original OPF problem. It should also be noted that in many examples, SDRs might have the rank-one matrix solution among their multiple matrix solutions. However, the SDP solvers can output a matrix solution of rank-more-than-one. The modification in [26] adds a small resistance for the lines of zero resistance, leading to the rank-one matrix solution of the SDR, which is implicitly related to the features of its used software.

Meanwhile, [5] adopts the high-order SDR method of [53] to find the global solution of the OPF problem over power transmission networks. Theoretically, such a method is able to generate a sequence of higher-order convex approximations, which converges to the original nonconvex problem in terms of the solution, regardless of whether the former is unique (as required in [53]) or not [54]. However, the dimension of these convex approximation problems grows dramatically in terms of decision variables, size and number of semi-definite constraints, making this approach suitable only for networks with a very small number of buses. Not surprisingly, [5] tests the performance of this method for networks with 2, 3 and 5 buses, where there are only 2, 3 and 5 bus voltages variables with the objective function linear in generation power. Another drawback of the high-order SDR method is that it works for real variables only so the dimension of the complex voltage variables become double for its utilization.

Power transmission networks in modern smart grids are often devised with a few thousand buses [55–57]. Under a such large number n of buses it is impossible to use the single matrix $W \in \mathbb{C}^{n\times n}$, which involves $n(n+1)/2 \approx O(10^7)$ complex variables. On the other hand, the number of the flow lines for bus connection is relatively moderate so only a small portion of the crossed nonlinear terms $V_k V_m^*$ appears in the nonlinear constraints. The common approach is to use the outer products of overlapped groups of the voltage variables to cover them [58, 59, 52]. All rank-one constraints on these outer products are then dropped for SDR. Obviously, the optimal solution of this SDR usually is not of rank-one and thus does not have any physical meaning. There is no technique to retrieve a feasible rank-one point from the rank-more-than-one solution of SDR. Multiple matrix rank constrained optimization has received a great attention due

to its potential application in robust control synthesis [31, 3] but to our best knowledge there is no effective computation so far.

1.2.4 Joint OPF-PEV charging problem in smart grid

Electrical vehicles (EVs) have emerged as a promising solution to resolve both the economic and environmental concerns in the transportation industry [60]. Using a smart power grid in concurrently serving residences and charging EVs constitutes one of the most important applications of the smart grid technology. However, the massive integration of plug-in EVs (PEVs) into the grid causes many potential impacts such as voltage deviation, increased load variations and power loss of the grid [61], which requires different strategies for load shifting and energy trading and storage in the grid [62–65]. The main difficulty in scheduling of PEV charging to manage the cost and impact of PEV integration is that individual PEVs randomly arrive for charging with their individual demands on charging load and deadlines, which cannot be known beforehand. In other words, the future charging demand of PEVs cannot be known a priori. Many existing works consider a simple smart grid with a single charging station (CS) to exclusively serve PEVs. For instance, [66] sets no charging deadlines for PEVs, whose arrival process follows a probability distribution, while [67] assumes that the future load demand is perfectly known a priori. The future load demand is also assumed to be known in [68] as all PEVs are assumed to arrive at the same time with no charging deadline. It is assumed in [69] that only statistics of demand are known but the PEVs can be fully charged in a single time slot [69, (30)]. It should be realized that serving PEVs is typically considered during a 12-hour time period (for instance from 8:00 pm to 8:00 am), where the integration of a massive number of PEVs has a sizable effect on the power grid, and as such, the length of a time slot is rationally set by 30 minutes or one hour. In other words, the charging scheduling should be considered over a finite horizon of 12-24 time slots, but not over an infinite horizon as considered in [70]. Due to their physical limitations, PEVs are rarely able to be fully charged just during a single time slot.

1.2.5 Optimal PMU Placement in smart grid

Phasor measurement unit (PMU) is an advanced digital meter, which is used in smart power grids for real-time monitoring of grid operations [71]. By installing it at a buse, the state-of-the-art PMU can measure not only the phasor of the bus voltage but also the current phasors of incident power branches with high accuracy [72]. These measurements are explored by the modern energy management systems (EMSs) for critical applications such as optimal power flow, contingency analysis, and cyber security, etc. [73–75].

As phasor measurement units (PMUs) are costly, there is a vast amount of literature on PMU placement optimization to target the minimal number of PMUs. Under different degrees of observability, the mission is accomplished by binary linear programming (BLP) [76, 77]. Here, the complete observability means that there is no bus left unobserved by the placed PMUs, while depth-of-n unobservability means that there are at most n connecting buses left unobserved by the placed PMUs [78], making as many states as possible observed by restricted number of PMUs. An exhaustive binary search was proposed in [79] to deal with this objective under the complete observability condition and additional operating conditions such as the single branch outage and the presence of zero power injections. A binary particle swarm optimization algorithm was proposed in [80] to deal with it while maintaining the complete observability conditions under the contingencies of PMU loss or branch outage. Binary quadratic programming and BLP were respectively used in [81] and [82] to study the effect of conventional measurements and zero bus injections to the complete observability.

Apparently, observability alone does not necessarily lead to a meaningful state estimate or an informative PMU configuration. In fact, PMU configurations, which use the same number of PMUs to make the grid completely observable, can result in quite different estimation accuracies [83]. Intuitively, a better estimator can be obtained by appropriately employing more PMUs. PMU placement optimization to minimize the mean squared error (of grid state estimation) or to maximize the mutual information between the measurement output and grid state under a fixed allowable number of PMUs was considered in [2] and [84], respectively. Obviously, these placement tasks are mathematically modelled by optimization of nonlinear objective functions of binary

variables subject to a simple linear constraint for fixing the number of PMUs. A convex relaxation with the binary constraint $\{0, 1\}$ for binary variables relaxed to the box constraint $[0, 1]$ was proposed in [2], which not only fails to provide even a local optimal solution in general but also is not scalable in the grid dimension as it involves an additional large-size semi-definite matrix variable. A greedy algorithm proposed in [84] does not provide a local optimal solution either. More importantly, both [2] and [84] ignored observability constraints for computational tractability. It was argued in [84] that its proposed mutual information criterion includes the grid complete observability, which is obviously not right simply because as shown later in the paper, the latter differentiates the state estimate from its unconditional mean, which is the trivial estimate, while the former does not.

1.3 Dissertation Outline

The outline of the dissertation is as follows:

Chapter 1

This chapter presents the motivation and scope, the research topics and the outline of the dissertation.

Chapter 2

A brief review of power system is presented in this chapter. Then an overview of a robust control system is introduced. In the last, an overview of optimization theory including nonsmooth optimization and d.c. (difference of two convex functions) programming are provided.

Chapter 3

In this chapter, we consider the reduced order LPV-LFT (linear parameter varying-linear fractional transformational) control synthesis. The reduced order control synthesis can be reformulated as a linear matrix inequality (LMI) optimization subject to a rank constraint on a matrix-valued affine function of the Lyapunov matrix variables. k-order robust control synthesis for a LPV plant of order n leads to rank-$(n + k)$ constraint on the positive semi-definite matrix-valued affine function of size $(2n) \times (2n)$. A novel approach proposed in this thesis is to equivalently express these rank-$(n + k)$ constraints on positive semi-definite matrix-valued affine function by rank-k constraint

on the matrix-valued nonlinear function of size n, which are then exactly expressed by spectral nonlinear functions. We then show a simple but effective nonsmooth optimization technique leading to a path-following optimization procedure for these problems. An intensive simulation shows the clear advantage of the proposed method over the state-of-the-art nonlinear matrix inequality solvers.

The work in this chapter has been published in:

- Y. Shi, H. D. Tuan and P. Apkarian, "Nonconvex Spectral Optimization Algorithms for Reduced-Order H_∞ LPV-LFT controllers", International Journal of Robust and Nonlinear Control, vol. 27, pp. 4421-4442, 2017.

Chapter 4

In this chapter, we consider the $\mathcal{H}_\infty$ Proportional-integral-derivative (PID) control design in fuzzy systems. To gain the practicability and tractability of fuzzy systems, we develop a parameterized bilinear matrix inequality characterization for the $\mathcal{H}_\infty$ fuzzy PID control design, which is then relaxed into a bilinear matrix inequality optimization problem of nonconvex optimization. Several computational procedures are then developed for its solution. The merit of the developed algorithms is shown through the benchmark examples.

The work in this chapter has been published in or submitted to:

- Y. Shi, H. D. Tuan, "Parameterized Bilinear Matrix Inequality Techniques in Fuzzy PID Control Design", under submission to IEEE Transactions on Fuzzy System, 2017.
- Y. Shi, H. D. Tuan, and S. W. Su, "Nonconvex Spectral Algorithm for Solving BMI on the Reduced Order H_∞ Control", the 6th IEEE International Conference on Control Systems, Computing and Engineering, 2016, Penang, Malaysia.

Chapter 5

In this chapter, we propose a nonsmooth optimization algorithm to obtain a near global solution for the OPF problem, which is aiming at locating a steady state operating point such that cost of electric power generation is minimized subject to operating constraints and meeting demand. As existing SDR may not locate a feasible

solution when the relaxed matrix variable is not rank-one, we propose a nonsmooth optimization algorithm to address the matrix rank constraint, which is an iterative process to generate a sequence of improved points that converge to a global or nearly global solution in most cases. We also develop an efficient decomposition for the large-scale OPF problem, which involves reduced numbers of the rank-one constraints on matrices of moderate size for expressing the network nonlinear constraints. Simulations for OPF problems and large-scale OPF problems demonstrate the efficiency of our approaches.

The work in this chapter has been published in:

- Y. Shi, H. D. Tuan, H. Tuy and S. W. Su, "Global Optimization for Optimal Power Flow over Transmission Networks", Journal of Global Optimization, vol. 69, pp. 745-760, 2017.
- Y. Shi, H. D. Tuan, A. V. Savkin, S. W. Su, "Optimal Power Flow over Large-Scale Transmission Networks", Systems & Control Letters, vol. 118, pp. 16-21, 2018.
- Y. Shi, H. D. Tuan, S. W. Su and H. H. M. Tam, "Nonsmooth Optimization for Optimal Power Flow over Transmission Networks", the 3rd IEEE Global Conference on Signal and Information Processing, pp. 1141-1144, 2015, Orlando, America.
- Y. Shi, H. D. Tuan, S. W. Su, and A. V. Savkin, "Multiple Matrix Rank Constrained Optimization for Optimal Power Flow over Large Scale Transmission Networks", proceedings of the 5th International Conference on Smart Cities and Green ICT Systems, vol. 1, pp. 384-389, 2016, Rome, Italy.
- Y. Shi, H. D. Tuan, and A.V. Savkin, "Three-phase Optimal Power Flow for Smart Grids by Iterative Nonsmooth Optimization", the 6th International Conference on Smart Cities and Green ICT Systems, 2017, Porto, Portugal.

Chapter 6

In this chapter, we consider the charging scheduling of plug-in electric vehicles (PEVs) and power control in smart grid. PEV charging scheduling aims at minimizing the

potential impact of the massive integration of PEVs into smart grid to save service costs to customers while power control aims at minimizing the cost of power generation subject to operating constraints and meeting demand. A model predictive control (MPC)-based approach is proposed to address the joint PEV charging scheduling and power control to minimize both PEV charging cost and energy generation cost in meeting both residence and PEV power demands. Unlike in related works, no assumptions are made about the probability distribution of PEVs' arrivals, the known PEVs' future demand, or the unlimited charging capacity of PEVs. The proposed approach is shown to achieve a globally optimal solution. Numerical results for IEEE benchmark power grids serving Tesla Model S PEVs show the merit of this approach.

The work in this chapter has been published in or submitted to:

- Y. Shi, H. D. Tuan, A. V. Savkin, T. Q. Duong and H. V. Poor, "Model Predictive Control for Smart Grids with Multiple Electric-Vehicle Charging Stations", accepted by IEEE Transaction on smart grid, 2017.
- Y. Shi, H. D. Tuan, A. V. Savkin, T. Q. Duong and H. V. Poor, "On-off Charging of Electrical Vehicles in Smart Grids", under submission to IEEE Transaction on smart grid, 2018.

Chapter 7

In this chapter, we consider the PMU placement problem for power grid state estimation under different degrees of observability. Observability degree is the depth of the buses' reachability by the placed PMUs and thus constitutes an important characteristic for PMU placement. However, the sole observability as addressed in many works still does not guarantee a good estimate for the grid state. In this chapter, the PMU placement optimization problem is considered by minimizing the mean squared error or maximizing the mutual information between the measurement output and grid state, under grid observability constraints. The provided solution is free from the mentioned fundamental drawbacks in the existing PMU placement designs. The problems are posed as binary nonlinear optimization problems, for which this paper develops efficient algorithms for computational solutions. The performance of the proposed algorithms is analyzed in detail through numerical examples on large scale IEEE power networks.

The work in this chapter has been submitted to:

- Y. Shi, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "PMU Placement Optimization for Smart Grid Obvervability and State Estimation", under submission to IEEE Transaction on smart grid, 2018.

Chapter 8

This chapter summarizes works of this dissertation and proposes future directions of related researches.

Chapter 2

Background

In this chapter, we first briefly present the foundation of H_∞ control for linear time varying system and linear parameter varying system, respectively. As a wide variety of system and control problems, such as optimal controller design, robust stability, pole placement can be reformulated as an optimization problem with linear matrix inequality (LMI) constraints and bilinear matrix inequality (BMI) [85, 15, 16]. Then we introduce the definition and property of the LMI and the BMI. At last, we introduce the optimization theory, including convex optimization and d.c. programming.

2.1 H_∞ Control for Linear time varying system

Consider a continuous time system

$$\begin{aligned} \dot{x} &= Ax + B_\infty w + Bu \\ z &= C_\infty x + D_\infty w + D_{\infty u} u \\ y &= Cx + D_{y\infty} w \end{aligned} \tag{2.1}$$

where $x \in R_x^n$ is the system state, $u \in R^{n_u}$ is the system control, $y \in R^{n_y}$ is the measured output, $w \to z$ is the H_∞ performance channel of the same dimension n_∞.

We seek a structured output feedback control

$$\begin{pmatrix} \dot{x}_K \\ u \end{pmatrix} = \mathbf{K} \begin{pmatrix} x_K \\ y \end{pmatrix}, \mathbf{K} \subset R^{(n_K+n_u)\times(n_K+n_y)} \tag{2.2}$$

where n_K is the controller order, $\mathbf{K}$ is a convex structure such that

- The closed-loop system is internally stable
- The H_∞-performance $T_{w\to z}$ is minimized

$$T_{w\to z} \to \min \tag{2.3}$$

2.2 H_∞ Control for Linear parameter varying system

Consider a continuous LPV system in linear fractional transformation (LFT) [86–88]

$$\begin{aligned} \begin{bmatrix} \dot{x}(t) \\ z_\Delta(t) \\ z(t) \\ y(t) \end{bmatrix} &= \begin{bmatrix} A & B_\Delta & B_1 & B_2 \\ C_\Delta & D_{\Delta\Delta} & D_{\Delta 1} & D_{\Delta 2} \\ C_1 & D_{1\Delta} & D_{11} & D_{12} \\ C_2 & D_{2\Delta} & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ w_\Delta(t) \\ w(t) \\ u(t) \end{bmatrix} \\ w_\Delta(t) &= \Delta(\alpha(t)) z_\Delta(t) \end{aligned} \tag{2.4}$$

where

$$\Delta(\alpha(t)) = \sum_{i=1}^{L} \alpha_i(t)\Delta_i, \alpha_i(t) \geq 0, \sum_{i=1}^{L} \alpha_i(t) = 1. \tag{2.5}$$

Here, $x(t) \in R^n$, $y(t) \in R^{n_y}$, $z(t) \in R^{n_z}$, $w(t) \in R^{n_z}$, $z_\Delta(t) \in R^{n_\Delta}$, $w_\Delta(t) \in R^{n_\Delta}$. Note that we assume without loss of generality that $z(t)$ and $w(t)$ ($z_\Delta(t)$ and $w_\Delta(t)$, resp.) have the same dimension. The pair (w_Δ, z_Δ) is regarded as the gain-scheduling channel. All matrices in (3.1)-(3.2) are given with appropriate size. Parameters $\alpha_i(t)$ are measured online and exploited by the controller.

The standard $\mathcal{H}_\infty$ LPV control design is to find k-order controller in LFT

$$\begin{aligned}\begin{bmatrix}\dot{x}_K(t)\\ u(t)\\ z_K(t)\end{bmatrix} &= \begin{bmatrix}A_K & B_{K1} & B_{K\Delta}\\ C_{K1} & D_{K11} & D_{K1\Delta}\\ C_{K\Delta} & D_{K\Delta 1} & D_{K\Delta\Delta}\end{bmatrix}\begin{bmatrix}x_K(t)\\ y(t)\\ w_K(t)\end{bmatrix}\\ w_K(t) &= \Delta_K(\alpha(t))z_K(t)\end{aligned} \tag{2.6}$$

with

$$\Delta_K(\alpha(t)) = \sum_{i=1}^{L}\alpha_i\Delta_{Ki} \tag{2.7}$$

such that the closed-loop system is internally stable and satisfies

$$\int_0^T ||z(t)||^2dt \le \gamma^2\int_0^T ||w(t)||^2dt \quad \forall w(.)\in L_2, T<+\infty, \tag{2.8}$$

initialized from $x(0)=0$. Here $x_K(t)\in R^k$, $z_K(t)\in R^{n_\Delta}$ and $w_K(t)\in R^{n_\Delta}$. k is called the control order and the pair (w_K, z_K) is regarded as the control's gain-scheduling channel.

Note that (3.1) and (3.3) are the following LPV LFTs

$$\begin{aligned}\begin{bmatrix}\dot{x}(t)\\ z(t)\\ y(t)\end{bmatrix} = \Bigg(&\begin{bmatrix}A & B_1 & B_2\\ C_1 & D_{11} & D_{12}\\ C_2 & D_{21} & 0\end{bmatrix} + \begin{bmatrix}B_\Delta\\ D_{1\Delta}\\ D_{2\Delta}\end{bmatrix}(I-\Delta(\alpha(t))D_{\Delta\Delta})^{-1}\Delta(\alpha(t))\\ &\times\begin{bmatrix}C_\Delta & D_{\Delta 1} & D_{\Delta 2}\end{bmatrix}\Bigg)\begin{bmatrix}x(t)\\ w(t)\\ u(t)\end{bmatrix}\end{aligned} \tag{2.9}$$

and

$$\begin{aligned}\begin{bmatrix}\dot{x}_K(t)\\ u(t)\end{bmatrix} = \Bigg(&\begin{bmatrix}A_K & B_{K1}\\ C_{K1} & D_{K11}\end{bmatrix} + \begin{bmatrix}B_{K\Delta}\\ D_{K1\Delta}\end{bmatrix}(I-\Delta_K(\alpha(t))D_{K\Delta\Delta})^{-1}\Delta_K(\alpha(t))\\ &\times\begin{bmatrix}C_{K\Delta} & D_{K\Delta 1}\end{bmatrix}\Bigg)\begin{bmatrix}x_K(t)\\ y(t)\end{bmatrix},\end{aligned} \tag{2.10}$$

respectively.

2.3 Linear matrix inequality and bilinear matrix inequality

Many control applications including structure control, robust control, and optimal design of experiments can be formulated as LMIs and BMIs[15]. In this section, a brief introduction of LMIs and BMIs will be presented.

2.3.1 Linear matrix inequality

A linear matrix inequality (LMI) has the form:

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0, \tag{2.11}$$

where $x \in \mathcal{R}^m$, $F_i \in \mathcal{R}^{n\times n}$, it should be noted that $F(x)$ is a positive definite matrix and it is also an affine function of the variable x given the fixed symmetric matrix F_i.

A set $\mathcal{D}$ is said to be convex if $\forall\ x, y \in \mathcal{D}$ and $\theta \in [0,1]$, it is true that $\theta x + (1-\theta)y \in \mathcal{D}$. To see the convexity of LMIs, let x and y be two vectors such that $F(x) > 0$ and $F(y) > 0$, and let $\theta \in [0,1]$. Then,

$$\begin{aligned} F(\theta x + (1-\theta)y) &= F_0 + \sum_{i=1}^{m}(\theta x_i + (1-\theta)y_i)F_i \\ &= \theta F_0 + \theta \sum_{i=1}^{m} x_i F_i + (1-\theta)F_0 + (1-\theta)\sum_{i=1}^{m} y_i F_i \\ &= \theta F(x) + (1-\theta)F(y) \\ &> 0. \end{aligned} \tag{2.12}$$

Thus, LMIs have the good property of convexity. Another important property is that multiple LMIs can be expressed as a single LMI. Consider a set of s LMIs:

$$F^1(x) > 0, F^2(x) > 0, \cdots, F^s(x) > 0, \tag{2.13}$$

Then an equivalent single LMI is:

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i = \text{diag}\{F^1(x), \cdots, F^s(x)\}, \tag{2.14}$$

where $\text{diag}\{F^1(x), \cdots, F^s(x)\}$ is a block diagonal matrix with blocks $F^1(x), \cdots, F^s(x)$.

2.3.2 Bilinear matrix inequality

A bilinear matrix inequality (BMI) is as follow [15]:

$$F(x, y) = F_0 + \sum_{i=1}^{m} x_i F_i + \sum_{j=1}^{n} y_i G_i + \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j H_{ij} > 0, \tag{2.15}$$

where F_i, G_j and H_{ij} are symmetric matrices of dimension $\mathcal{R}^{k \times k}$, and $y \in \mathcal{R}^n$. BMIs are not necessarily convex and thus they can describe much more constraints than LMIs. And BMIs can be applied to more types of optimization and control problems. But they are much more difficult to handle comparing with LMIs.

2.4 Optimization Theory

In this section, we first review convex optimization and then describe sequential convex programming.

2.4.1 Convex Optimization

Fundamental definitions in convex optimization [89] are given as follows,

Definition 2.1 *A set $\mathcal{D}$ is said to be convex if $\forall\ x, y \in \mathcal{D}$ and $\theta \in [0, 1]$, it is true that $\theta x + (1 - \theta) y \in \mathcal{D}$.*

Definition 2.2 *A function $f(x)$ is convex, if for all $x, y \in \mathcal{D}$ and $0 \leq \theta \leq 1$, it is true that*

$$f(\theta x + (1 - \theta) y) \leq \theta f(x) + (1 - \theta) f(y). \tag{2.16}$$

Definition 2.3 *To the problem* $\min\{f(x)|x \in \mathcal{D}\}$*, a point* $x^* \in \mathcal{D}$ *such that*

$$f(x^*) \leq f(x),\ \forall x \in \mathcal{D}, \tag{2.17}$$

is called a global minimizer. If there exists a neighborhood $\mathcal{N}$ *of* x' *satisfying*

$$f(x') \leq f(x),\ \forall x \in D \cap \mathcal{N}, \tag{2.18}$$

$x' \in \mathcal{N}$ *is called a local minimizer.*

A convex optimization problem is represented as

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0,\ i = 1, \ldots, m, \end{aligned} \tag{2.19}$$

where $f(x) and g_i(\cdot), i = 1, \ldots, m$ are convex functions. The most important characteristics of a convex optimization problem are

- Any local minimizer of (2.19) is also its global minimizer;
- (2.19) is computationally tractable and can be solved efficiently by off-the-shelf software within polynomial time.

2.4.2 D.C. optimization

Convexity is a nice property of functions but in the most cases of control system and power system, this property can not be preserved. To address the non-convexity challenge of those optimization problems, we may exploit the d.c. (difference of convex functions) structure, which is the common underlying structure of virtually all non-convex optimization problem [89].

Definition 2.4 *[89] Let* $\mathcal{B}$ *be a convex set in* $\mathcal{R}^n$*. We say that a function is d.c. on* $\mathcal{B}$ *if it can be expressed as the difference of two convex functions on* $\mathcal{B}$*, i.e. if* $f(x) = f_1(x) - f_2(x)$*, where* f_1, f_2 *are convex function on* $\mathcal{B}$*.*

Definition 2.5 *[89] An optimization problem is called a d.c. programming problem if it has the form*

$$\min_x \quad f(x), \quad s.t. \quad x \in \mathcal{B}, \tag{2.20}$$

where f is a d.c. function and $\mathcal{B}$ is a d.c. set.

It should be noted that, d.c. programming problems consists of a wide class of nonconvex programming problems. Some of the d.c. programming problems are extremely hard to solve, we are interested in the following d.c. programming, which minimize a d.c. function over a convex set.

$$\min_x \quad F(x), \quad \text{s.t.} \quad x \in \mathcal{D}, \tag{2.21}$$

where f is a d.c. function and $\mathcal{D}$ is a convex set.

Construct a convex functions $F_u(x)$ such that $F_u(x)$ is a upper bound of $F(x)$, i.e. $F(x) \leq F_\mu(x) \quad \forall x \in \mathcal{D}$. Initialized by a feasible point $x^{(0)}$, at the κ-th iteration, solve the following convex optimization,

$$\min_x \quad F_u^{(\kappa)}(x), \quad \text{s.t.} \quad x \in \mathcal{D}, \tag{2.22}$$

and generate the next feasible point $x^{(\kappa+1)}$.

Note that any feasible point for the convex optimization problem (2.22) is also feasible for the nonconvex optimization problem (2.19). As $x^{(\kappa)}$ and $x^{(\kappa+1)}$ are respectively feasible point and the optimal solution of (2.22), it is true that

$$F_u^{(\kappa)}(x^{(\kappa+1)}) < F_u^{(\kappa)}(x^{(\kappa)}). \tag{2.23}$$

Then the point $x^{(\kappa+1)}$ is better feasible point for the d.c. programming problem (2.22) than $x^{(\kappa)}$. Thus, $\{x^{(\kappa)}\}$ is a sequence of improved feasible points of the d.c. programming (2.22), which converges to a point satisfying first-order necessary optimality conditions.

Chapter 3

Nonconvex Spectral Optimization Algorithms for Reduced-Order $\mathcal{H}_\infty$ LPV-LFT controllers

3.1 Introduction

Rank-constrained optimization is referred to optimization problems involving rank constraints on matrix-valued functions of the decision variables. Initialized by the pioneering work [11], which reformulates the reduced order $\mathcal{H}_\infty$ control synthesis for linear time invariant (LTI) systems as linear matrix inequality (LMI) optimization subject to a rank constraint on a matrix-valued affine function of the Lyapunov matrix variables, many other important and difficult problems in robust control are also reformulated in similar matrix-rank constrained optimizations [28]. The simplest approach is to relax or just to drop that rank constraints with hope that the optimal solution of the relaxed (convex) optimization would satisfy these matrix-rank constraints. For instance, matrix trace minimization and nuclear norm minimization were proposed to obtain low matrix rank of positive semi-definite matrix and rectangular matrices, respectively [29, 30]. These techniques are unable to address the matrix-rank constraints as they are. Indeed, just a trace of a matrix or its nuclear norm don't give any adequate indication on the matrix rank. Another attempt is to use a Newton-like method

to find a projection of a positive semi-definite matrix to the manifold of fixed rank matrices [28, 32], which is equally computationally difficult optimization due to complex geometry of this manifold [33], especially for lower fixed rank matrices of larger size. Realizing the challenge by these matrix-rank constraints on the Lyapunov matrix variables, most later developments in robust control preferred to avoid them in favor of alternative bilinear matrix inequality (BMI) [34, 28, 35–37, 3, 38, 4]. The state-of-the-art BMI solvers [3, 4] initialize from a reduced-order stabilizing controller and then move within a convex feasibility subset containing this initialized point. There are a few difficulties arisen with this kind of feasibility algorithms. Firstly, finding a good reduced-order stabilizing controller is not an easy task because its computation is still a NP-hard problem [39]. Secondly, the feasibility set of reduced order stabilizing controllers is highly nonconvex, which is disconnected in general. This means moving within a convex neighborhood of such reduced-order stabilizing controller may be trapped by local minima. Thirdly, usually the convergence of this kind of algorithms is slow and is dependent very much on the local geometry around such initial point [40], which may be unpredictable. Within the past few years, it has been realized that all BMI solvers [3, 4], which address the control synthesis for LTI systems in state space using Lyapunov functions, could hardly compete with the nonsmooth optimization solver developed earlier in [17], which addresses the problems directly in the frequency domain to bypass the Lyapunov variables of high dimension. Nowadays, the Matlab `systune` command [18], which is based on [17], is the most powerful tool for control synthesis of LTI systems and is widely used in industry. This means that rank-constrained optimization and BMI should seek applications outside uncertain LTI systems such as linear parameter varying (LPV) systems [19], where Lyapunov function is irreplaceable.

Meanwhile, for solution of indefinite quadratic optimization in signal processing applications, [90–92] developed an approach for optimization on the rank-one constrained positive semi-definite outer product of decision vector variable. Intensive simulations even for large scale indefinite quadratic optimization [92] show that the rank-one matrices can be quickly located, which are turned out to be global optimal solutions of the considered indefinite quadratic problems in most cases. Reduced order robust LPV controller synthesis is more difficult than indefinite quadratic programming

and has not been appropriately considered in literature. The matrix-rank constraints in the former are much more challenging than the rank-one constraint in the latter. Indeed, they are lower fixed rank constraints on matrix-valued affine functions of larger size with very complex geometry. For instance, k-order robust control synthesis for a LPV plant of order n leads to rank-$(n+k)$ constraint on the positive semi-definite matrix-valued affine function of size $(2n) \times (2n)$ [86, 88]. A novel approach proposed in the present chapter is to equivalently express these rank-$(n+k)$ constraints on the positive semi-definite matrix-valued affine function by rank-k constraint on the matrix-valued nonlinear function of size n, which are then exactly expressed by spectral nonlinear functions. We then show a simple but effective optimization technique leading to a path-following optimization procedure for these problems. To the author's best knowledge, spectral nonlinear function optimization was not quite considered in the literature.

This chapter is organized as follows. After the Introduction, section 3.2 is devoted to algorithmic solutions for reduced-order LPV $\mathcal{H}_\infty$ controllers while section 3.3 is devoted to static output feedback LPV controllers. An intensive simulation is provided in Section 3.4 to support the algorithmic development of the previous sections. Section 3.5 concludes the chapter.

Notation. Notation used in this chapter is standard. Particularly, $X \succeq 0$, $X \succ 0$, $X \preceq 0$ and $X \prec 0$ mean that a symmetric matrix X is positive semi-definite, positive definite, negative semi-definite and negative definite, respectively, while $\langle X, Y \rangle$ is the dot product of the matrices X and Y. For simplicity, we also denote $\mathsf{tr}(X)$ as the trace of X. I is the identity matrix but when needed we also use I_n to emphasize the size $n \times n$ of I. In symmetric block matrices or long matrix expressions, we use $*$ as an ellipsis for terms that are induced by symmetry, e.g.,

$$K \begin{bmatrix} S + S^T & M^T \\ M & Q \end{bmatrix} K^T = K \begin{bmatrix} S + (*) & * \\ M & Q \end{bmatrix} *$$

The matrix variables are typed boldfaced in the chapter.

3.2 Dynamic reduced order $\mathcal{H}_\infty$ LPV control synthesis

Consider a continuous LPV system in linear fractional transformation (LFT) [86–88]

$$\begin{bmatrix} \dot{x}(t) \\ z_\Delta(t) \\ z(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B_\Delta & B_1 & B_2 \\ C_\Delta & D_{\Delta\Delta} & D_{\Delta 1} & D_{\Delta 2} \\ C_1 & D_{1\Delta} & D_{11} & D_{12} \\ C_2 & D_{2\Delta} & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ w_\Delta(t) \\ w(t) \\ u(t) \end{bmatrix} \tag{3.1}$$
$$w_\Delta(t) = \Delta(\alpha(t)) z_\Delta(t)$$

where

$$\Delta(\alpha(t)) = \sum_{i=1}^{L} \alpha_i(t)\Delta_i, \alpha_i(t) \geq 0, \sum_{i=1}^{L} \alpha_i(t) = 1. \tag{3.2}$$

Here, $x(t) \in R^n$, $y(t) \in R^{n_y}$, $z(t) \in R^{n_z}$, $w(t) \in R^{n_z}$, $z_\Delta(t) \in R^{n_\Delta}$, $w_\Delta(t) \in R^{n_\Delta}$. Note that we assume without loss of generality that $z(t)$ and $w(t)$ ($z_\Delta(t)$ and $w_\Delta(t)$, resp.) have the same dimension. The pair (w_Δ, z_Δ) is regarded as the gain-scheduling channel. All matrices in (3.1)-(3.2) are given with appropriate size. Parameters $\alpha_i(t)$ are measured online and exploited by the controller.

The standard $\mathcal{H}_\infty$ LPV control design is to find k-order controller in LFT

$$\begin{bmatrix} \dot{x}_K(t) \\ u(t) \\ z_K(t) \end{bmatrix} = \begin{bmatrix} A_K & B_{K1} & B_{K\Delta} \\ C_{K1} & D_{K11} & D_{K1\Delta} \\ C_{K\Delta} & D_{K\Delta 1} & D_{K\Delta\Delta} \end{bmatrix} \begin{bmatrix} x_K(t) \\ y(t) \\ w_K(t) \end{bmatrix} \tag{3.3}$$
$$w_K(t) = \Delta_K(\alpha(t)) z_K(t)$$

with

$$\Delta_K(\alpha(t)) = \sum_{i=1}^{L} \alpha_i \Delta_{Ki} \tag{3.4}$$

such that the closed-loop system is internally stable and satisfies

$$\int_0^T ||z(t)||^2 dt \leq \gamma^2 \int_0^T ||w(t)||^2 dt \quad \forall w(.) \in L_2, T < +\infty, \tag{3.5}$$

initialized from $x(0) = 0$. Here $x_K(t) \in R^k$, $z_K(t) \in R^{n_\Delta}$ and $w_K(t) \in R^{n_\Delta}$. k is called the control order and the pair (w_K, z_K) is regarded as the control's gain-scheduling channel.

Note that (3.1) and (3.3) are the following LPV LFTs

$$\begin{bmatrix} \dot{x}(t) \\ z(t) \\ y(t) \end{bmatrix} = \left(\begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{bmatrix} + \begin{bmatrix} B_\Delta \\ D_{1\Delta} \\ D_{2\Delta} \end{bmatrix} (I - \Delta(\alpha(t))D_{\Delta\Delta})^{-1}\Delta(\alpha(t)) \times \begin{bmatrix} C_\Delta & D_{\Delta 1} & D_{\Delta 2} \end{bmatrix} \right) \begin{bmatrix} x(t) \\ w(t) \\ u(t) \end{bmatrix} \tag{3.6}$$

and

$$\begin{bmatrix} \dot{x}_K(t) \\ u(t) \end{bmatrix} = \left(\begin{bmatrix} A_K & B_{K1} \\ C_{K1} & D_{K11} \end{bmatrix} + \begin{bmatrix} B_{K\Delta} \\ D_{K1\Delta} \end{bmatrix} (I - \Delta_K(\alpha(t))D_{K\Delta\Delta})^{-1}\Delta_K(\alpha(t)) \times \begin{bmatrix} C_{K\Delta} & D_{K\Delta 1} \end{bmatrix} \right) \begin{bmatrix} x_K(t) \\ y(t) \end{bmatrix}, \tag{3.7}$$

respectively. Figure 3.1 provides a block-diagram for a such system.

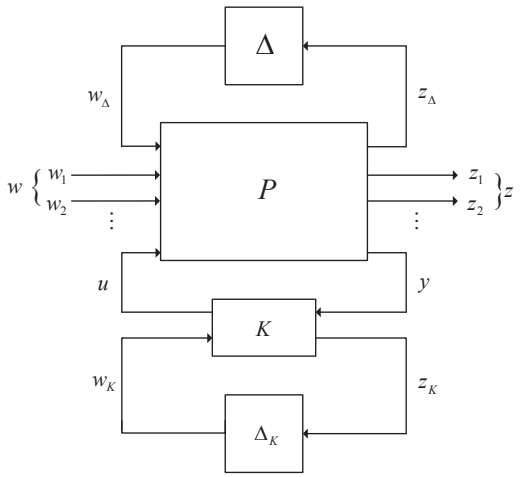

Fig. 3.1 Closed-loop LPV-LFT system

Let's state the following result adapted from [88]: the feasibility of the following matrix inequality in $\mathbf{X} \in R^{n\times n}$, $\mathbf{Y} \in R^{n\times n}$, $\mathbf{R}, \mathbf{H}, \mathbf{Q}, \mathbf{E}, \mathbf{\Delta}_{Ki}$ and

$$\hat{\mathbf{K}} := \begin{bmatrix} \mathbf{A_K} & \mathbf{B_{K1}} & \mathbf{B_{K\Delta}} \\ \mathbf{C_{K1}} & \mathbf{D_{K11}} & \mathbf{D_{K1\Delta}} \\ \mathbf{C_{K\Delta}} & \mathbf{D_{K\Delta 1}} & \mathbf{D_{K\Delta\Delta}} \end{bmatrix} \tag{3.8}$$

is sufficient for the existence of such controller

$$\begin{gathered} \begin{bmatrix} \mathbf{LMI_1} & * \\ \mathbf{LMI_2} & \mathbf{LMI_3} \end{bmatrix} \prec 0, \\ \begin{bmatrix} \mathbf{R} & I & \Delta_i^T\mathbf{Q} & \Delta_i^T \\ I & \mathbf{H} & \mathbf{\Delta_{Ki}}^T & \mathbf{H}\Delta_i^T \\ \mathbf{Q}\Delta_i & \mathbf{\Delta_{Ki}} & -\mathbf{Q} & -I \\ \Delta_i & \Delta_i\mathbf{H} & -I & -\mathbf{E} \end{bmatrix} \succ 0, \quad i = 1, 2, \cdots, L. \end{gathered} \tag{3.9}$$

$$\begin{pmatrix} \mathbf{X} & I_n \\ I_n & \mathbf{Y} \end{pmatrix} \succeq 0, \tag{3.10}$$

$$\operatorname{rank}(\mathbf{X} - \mathbf{Y}^{-1}) \leq k, \tag{3.11}$$

where

$$\mathbf{LMI_1} := \begin{bmatrix} \mathbf{X}A + \mathbf{B_{K1}}C_2 + (*) & * & * & * \\ \mathbf{A_K}^T + A + B_2\mathbf{D_{K11}}C_2 & (A\mathbf{Y} + B_2\mathbf{C_{K1}}) + (*) & * & * \\ B_\Delta^T\mathbf{X} + D_{2\Delta}^T\mathbf{B_{K1}}^T & B_\Delta^T D_{2\Delta}^T\mathbf{D_{K11}}^T B_2^T & Q & * \\ \mathbf{B_{K\Delta}^T} & \mathbf{E}B_\Delta^T + \mathbf{D_{K1\Delta}}^T B_2^T & -I & \mathbf{E} \end{bmatrix}$$

$$\mathbf{LMI_2} := \begin{bmatrix} B_1^T\mathbf{X} + D_{21}^T\mathbf{B_{K1}}^T & B_1^T + D_{21}^T\mathbf{D_{K11}}^T B_2^T & 0 & 0 \\ \mathbf{R}C_\Delta + \mathbf{D_{K\Delta 1}}C_2 & \mathbf{C_{K\Delta}} & \mathbf{R}D_{\Delta\Delta} + \mathbf{D_{K\Delta 1}}D_{2\Delta} & \mathbf{D_{K\Delta\Delta}} \\ C_\Delta + D_{\Delta 2}\mathbf{D_{K11}}C_2 & C_\Delta\mathbf{Y} + D_{\Delta 2}\mathbf{C_{K1}} & D_{\Delta\Delta} + D_{\Delta 2}\mathbf{D_{K11}}D_{2\Delta} & D_{\Delta\Delta}\mathbf{E} + D_{\Delta 2}\mathbf{D_{K1\Delta}} \\ C_1 + D_{12}\mathbf{D_{K11}}C_2 & C_1\mathbf{Y} + D_{12}\mathbf{C_{K1}} & D_{1\Delta} + D_{12}\mathbf{D_{K11}}D_{2\Delta} & D_{1\Delta}\mathbf{E} + D_{12}\mathbf{D_{K1\Delta}} \end{bmatrix}$$

$$\mathbf{LMI_3} := \begin{bmatrix} -\gamma I & * & * & * \\ \mathbf{R}D_{\Delta 1} + \mathbf{D_{K\Delta 1}}D_{21} & -\mathbf{R} & * & * \\ D_{\Delta 1} + D_{\Delta 2}\mathbf{D_{K11}}D_{21} & -I & -\mathbf{H} & * \\ D_{11} + D_{12}\mathbf{D_{K11}}D_{21} & 0 & 0 & -\gamma I \end{bmatrix}. \tag{3.12}$$

Note that (3.10)-(3.11) implies that

$$\operatorname{rank}(I_n - \mathbf{X}\mathbf{Y}) \leq k.$$

Without loss of generality, assume $\text{rank}(I_n - \mathbf{XY}) = k$. Then factorize

$$I_n - \mathbf{XY} = \mathbf{MN}^T$$

with full-rank $\mathbf{M} \in R^{n\times k}$ and $\mathbf{N} \in R^{n\times k}$. Their left-inverse matrices are

$$\mathbf{M}^+ = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T, \mathbf{N}^+ = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T.$$

Also factorize

$$I - \mathbf{RH} = \mathbf{R}_{12}\mathbf{H}_{12}^T \quad \text{and} \quad I - \mathbf{QE} = \mathbf{Q}_{12}\mathbf{E}_{12}^T$$

with invertible matrices $\mathbf{R}_{12}, \mathbf{H}_{12}, \mathbf{Q}_{12}$ and $\mathbf{E}_{12}$. Accordingly, the controller (3.3) can be recovered as follows [86]:

$$D_{K11} = \mathbf{D_{K11}} \tag{3.13}$$

$$B_{K1} = \mathbf{M}^+(\mathbf{B_{K1}} - \mathbf{X}B_2D_{K11}) \tag{3.14}$$

$$C_{K1} = (\mathbf{C_{K1}} - D_{K11}C_2\mathbf{Y})(\mathbf{N}^+)^T, \tag{3.15}$$

$$A_K = \mathbf{M}^+[\mathbf{A_K} - (\mathbf{X}A\mathbf{Y} + \mathbf{M}B_{K1}C_2\mathbf{Y} + \mathbf{X}B_2C_{K1}\mathbf{N}^T + \mathbf{X}B_2D_{K11}C_2\mathbf{Y})](\mathbf{N}^+)^T \tag{3.16}$$

$$D_{k1\Delta} = (\mathbf{D_{K1\Delta}} - D_{K11}D_{2\Delta}\mathbf{E})(\mathbf{E}_{12}^{-1})^T \tag{3.17}$$

$$D_{K\Delta 1} = \mathbf{R}_{12}^{-1}(\mathbf{D_{K\Delta 1}} - \mathbf{R}D_{\Delta 2}D_{K11}) \tag{3.18}$$

$$\begin{aligned} B_{K\Delta} \;=\; & \mathbf{M}^+[\mathbf{B_{K\Delta}} - (\mathbf{X}B_\Delta\mathbf{E} + \mathbf{M}B_{k1}D_{2\Delta}\mathbf{E} + \mathbf{X}B_2D_{K11}D_{2\Delta}\mathbf{E} \\ & + \mathbf{X}B_2D_{K1\Delta}\mathbf{E}_{12}^T)](\mathbf{E}_{12}^{-1})^T \end{aligned} \tag{3.19}$$

$$\begin{aligned} C_{K\Delta} \;=\; & \mathbf{R}_{12}^{-1}[\mathbf{C_{K\Delta}} - (\mathbf{R}C_\Delta\mathbf{Y} + \mathbf{R}D_{\Delta 2}D_{K11}C_2\mathbf{Y} + \mathbf{R}_{12}D_{K\Delta 1}C_2\mathbf{Y} \\ & + \mathbf{R}D_{\Delta 2}C_{K1}\mathbf{N}^T)](\mathbf{N}^{-1})^T \end{aligned} \tag{3.20}$$

$$\begin{aligned} D_{K\Delta\Delta} \;=\; & \mathbf{R}_{12}^{-1}[\mathbf{D_{K\Delta\Delta}} - (\mathbf{R}D_{\Delta\Delta}\mathbf{E} + \mathbf{R}D_{\Delta 2}D_{K11}D_{2\Delta}\mathbf{E} + \mathbf{R}_{12}D_{K\Delta 1}D_{2\Delta}\mathbf{E} \\ & + \mathbf{R}D_{\Delta 2}D_{K1\Delta}\mathbf{E}_{12}^T)](\mathbf{E}_{12}^{-1})^T. \end{aligned} \tag{3.21}$$

It should be noted that (3.10)-(3.11) are equivalent to (3.10) and

$$\text{rank}\left(\begin{pmatrix} \mathbf{X} & I_n \\ I_n & \mathbf{Y} \end{pmatrix}\right) \leq n + k \tag{3.22}$$

which is a lower fixed rank constraint on a matrix-valued affine function of $(\mathbf{X}, \mathbf{Y})$. Although our below developed algorithms still work for this constraint (3.22), we will see that in fact the rank constraint (3.11) on a nonlinear function of $(\mathbf{X}, \mathbf{Y})$ can be more efficiently handled. The difficulty degree of formulations in [93, 3, 4] is proportional to the dimension of the control variable (A_K, B_K, C_K, D_K) in (3.3), i.e. it is proportional to the control order k. In contrast, by exploring the rank constraint (3.11), the difficulty degree in our formulation is proportional to $\min\{k, n-k\}$, i.e. the computational difficulty with k-order and $(n-k)$-order controllers is the same.

We formulate the k-order LPV-LFT $\mathcal{H}_\infty$ control as

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \ \gamma \quad \text{s.t.} \quad (3.9), (3.10), (3.11), \tag{3.23}$$

where all the nonconvexity of the problem is concentrated in the rank constraint (3.11), which is automatically satisfied for the (full) n-order control. For $k < n$, as expected (3.11) is a highly nonconvex and discontinuous constraint. Consequently, the feasibility set (3.9)-(3.11) is disconnected in general, for which locating a feasible point is already not an easy task. As an aside note, the above formulation (3.23) is based on parameter-independent Lyapunov function and static multipliers $\mathbf{R}$, $\mathbf{H}$, $\mathbf{Q}$ and $\mathbf{E}$, which may potentially be more conservative than that with either parameter-dependent Lyapunov functions (see e.g. [94]) or dynamic multipliers (see e.g. [95]). However, efficient formulations for $k-$order LPV-LFT $\mathcal{H}_\infty$ control by using parameter dependent Lyapunov functions and dynamic multipliers are still very much open for study. The interested reader is also referred to [96] and the references therein for convex relation based results for fixed-order LPV controllers for LPV systems.

The function $\text{rank}(\mathbf{X} - \mathbf{Y}^{-1})$ in (3.11) seems to be very complicated. However, we will see shortly that it can be efficiently handled from the following observation. Suppose $f_{[k]}(\mathbf{X} - \mathbf{Y}^{-1})$ is the sum of the k largest eigenvalues of $\mathbf{X} - \mathbf{Y}^{-1}$, which is positive-definite ($\mathbf{X} - \mathbf{Y}^{-1} \succeq 0$) thanks to LMI (3.10). Then the matrix rank constraint

(3.11) holds true if and only if

$$\mathsf{tr}(\mathbf{X}-\mathbf{Y}^{-1}) = f_{[k]}(\mathbf{X}-\mathbf{Y}^{-1})$$

because it implies that $\mathbf{X}-\mathbf{Y}^{-1}$ has at least $(n-k)$ zero eigenvalues. Under the LMI (3.10), the quantity $\mathsf{tr}(\mathbf{X}-\mathbf{Y}^{-1}) - f_{[k]}(\mathbf{X}-\mathbf{Y}^{-1})$ is always nonnegative and can therefore be used to measure the degree of satisfaction of the matrix rank constraint (3.11). Instead of handling nonconvex constraint (3.11) we incorporate it into the objective, resulting in the following alternative formulation to (3.23):

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} F_\mu(\mathbf{X},\mathbf{Y},\gamma) := \gamma+\mu(\mathsf{tr}(\mathbf{X}-\mathbf{Y}^{-1})-f_{[k]}(\mathbf{X}-\mathbf{Y}^{-1})) \quad \text{s.t.} \quad (3.9)-(3.10), \tag{3.24}$$

where $\mu > 0$ is a penalty parameter. Without squaring on the factor of μ, the above penalization is exact, meaning that the constraint (3.11) can be satisfied by a minimizer of (3.24) with a finite value of μ (see e.g. [97, Chapter 16]). This is generally considered as a sufficiently nice property to make such exact penalization attractive. On the other hand, any feasible $(\mathbf{X},\mathbf{Y},\gamma)$ to (3.23) is also feasible to (3.24), implying that the optimal value of (3.24) for any $\mu > 0$ is upper bounded by the optimal value of (3.23).

Suppose $(X^{(\kappa)}, Y^{(\kappa)}, \gamma^{(\kappa)})$ is a feasible point to the convex feasibility set (3.9)-(3.10). Using the following variational principle [98, p. 191]

$$f_{[k]}(\mathbf{X}-\mathbf{Y}^{-1}) = \max_{\text{orthonormal } x_1,\ldots,x_k} \sum_{i=1}^{k} x_i^H(\mathbf{X}-\mathbf{Y}^{-1})x_i$$

it follows that

$$f_{[k]}(\mathbf{X}-\mathbf{Y}^{-1}) \geq \sum_{i=1}^{k}(x_i^{(\kappa)})^H(\mathbf{X}-\mathbf{Y}^{-1})x_i^{(\kappa)}, \tag{3.25}$$

where $x_i^{(\kappa)}$, $i = 1,...,k$ are the orthonormal eigenvectors corresponding to k largest eigenvalues of $X^{(\kappa)} - (Y^{(\kappa)})^{-1}$. On the other hand, as $\mathsf{tr}(\mathbf{Y}^{-1})$ is convex in $\mathbf{Y} \succ 0$, it is true that

$$\begin{aligned} \mathsf{tr}(\mathbf{Y}^{-1}) &\geq \mathsf{tr}((Y^{(\kappa)})^{-1}) - \mathsf{tr}((Y^{(\kappa)})^{-1}(\mathbf{Y}-Y^{(\kappa)})(Y^{(\kappa)})^{-1}) \\ &= 2\mathsf{tr}((Y^{(\kappa)})^{-1}) - \mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}). \end{aligned} \tag{3.26}$$

The following convex optimization is majorant minimization for (3.24)

$$\begin{aligned}\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} & F_\mu^{(\kappa)}(\mathbf{X},\mathbf{Y},\gamma) := \gamma + \mu(\mathsf{tr}(\mathbf{X} - 2(Y^{(\kappa)})^{-1}) \\ & +\mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}) - \sum_{i=1}^{k}(x_i^{(\kappa)})^H(\mathbf{X} - \mathbf{Y}^{-1})x_i^{(\kappa)}) \quad \text{s.t.} \quad (3.9) - (3.10).\end{aligned} \tag{3.27}$$

because by (3.25) and (3.26), function $F_\mu^{(\kappa)}$ obeys the two following crucial properties

$$F_\mu^{(\kappa)}(\mathbf{X},\mathbf{Y},\gamma) \geq F_\mu(\mathbf{X},\mathbf{Y},\gamma) \; \forall \; (\mathbf{X},\mathbf{Y},\gamma) \quad \text{on} \quad (3.9) - (3.10)$$

and

$$F_\mu^{(\kappa)}(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)}) = F_\mu(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)}).$$

Therefore, for the optimal solution $(X^{(\kappa+1)},Y^{(\kappa+1)},\gamma^{(\kappa+1)})$ of the convex program (3.27), it is true that

$$\begin{aligned} F_\mu(X^{(\kappa+1)},Y^{(\kappa+1)},\gamma^{(\kappa+1)}) &\leq F_\mu^{(\kappa)}(X^{(\kappa+1)},Y^{(\kappa+1)},\gamma^{(\kappa+1)}) \\ &\leq F_\mu^{(\kappa)}(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)}) \\ &= F_\mu(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)}),\end{aligned}$$

implying that $(X^{(\kappa+1)},Y^{(\kappa+1)},\gamma^{(\kappa+1)})$ is better than $(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)})$ toward optimizing the objective in (3.24). By using [99], we can prove the following result of global convergence.

Proposition 1 *Initialized by any feasible point $(X^{(0)},Y^{(0)},\gamma^{(0)})$ of SDP (3.9)-(3.10), $\{(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)})\}$ is a sequence of improved feasible points of the nonconvex program (3.24), which converges to a point satisfying first-order necessary optimality conditions.*

Proof: The sequence $\{(X^{(\kappa)},Y^{(\kappa)},\gamma^{(\kappa)})\}$ terminates (whenever $F_\mu(X^{(\kappa+1)},Y^{(\kappa+1)},\gamma^{(\kappa+1)}) = F_\mu^{(\kappa)}(X^{(\kappa+1)},Y^{(\kappa+1)},\gamma^{(\kappa+1)})$) or convergence to $\{(\bar{X},\bar{Y},\bar{\gamma}\}$, which is the optimal solution of the convex program

$$\begin{aligned}\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki}} & \bar{F}_\mu(\mathbf{X},\mathbf{Y},\gamma) := \gamma + \mu(\mathsf{tr}(\mathbf{X} - 2\bar{Y}^{-1}) + \mathsf{tr}(\bar{Y}^{-1}\mathbf{Y}\bar{Y}^{-1}) \\ & - \sum_{i=1}^{k}(\bar{x}_i^H(\mathbf{X} - \mathbf{Y}^{-1})\bar{x}_i) \quad \text{s.t.} \quad (3.9) - (3.10).\end{aligned} \tag{3.28}$$

Algorithm 1 Nonconvex Spectral Optimization Algorithm for optimized k-order $\mathcal{H}_\infty$ controllers

1: Initialize $\kappa := 0$ and solve SDP (3.32) to find its optimal solution $(X^{(\kappa)}, Y^{(\kappa)}, \gamma^{(\kappa)})$. For k normalized eigenvectors corresponding to k largest eigenvalues of $X^{(\kappa)} - (Y^{(\kappa)})^{-1}$ stop the algorithm if

$$\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) - \sum_{i=1}^{k} (x_i^{(\kappa)})^H (X^{(\kappa)} - (Y^{(\kappa)})^{-1}) x_i^{(\kappa)}) \leq \epsilon_2$$

and accept $(X^{(0)}, Y^{(0)}, \gamma^{(0)})$ as the optimal solution of the nonconvex program (3.23). Otherwise set $\mu = 0.5$.

2: **repeat**

3: **if**

$$\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) - \sum_{i=1}^{k} (x_i^{(\kappa)})^H (X^{(\kappa)} - (Y^{(\kappa)})^{-1}) x_i^{(\kappa)}) \geq \epsilon_2 \tag{3.29}$$

for k normalized eigenvectors corresponding to k largest eigenvalues of $X^{(\kappa)} - (Y^{(\kappa)})^{-1}$ **then** reset $\mu \to 2\mu$ and solve SDP (3.27) to find the optimal solution $(X^{(\kappa+1)}, Y^{(\kappa+1)}, \gamma^{(\kappa+1)})$.

4: **else** Solve SDP (3.27) with additional convex constraint

$$\mathsf{tr}(\mathbf{X} - 2(Y^{(\kappa)})^{-1}) + \mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}) - \sum_{i=1}^{k} (x_i^{(\kappa)})^H (\mathbf{X} - \mathbf{Y}^{-1}) x_i^{(\kappa)}) \leq \epsilon_2 \tag{3.30}$$

5: **end if**

6: Set $\kappa := \kappa + 1$.

7: **until** $\gamma^{(\kappa)} - \gamma^{(\kappa-1)} \leq \epsilon_1$

8: Accept $(X^{(\kappa)}, Y^{(\kappa)}, \gamma^{(\kappa)})$ as a found solution of (3.23) if $\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) - f_{[k]}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) \leq \epsilon_2$.

where $\bar{x}_i$, $i = 1, ..., k$ are the orthonormal eigenvectors corresponding to k largest eigenvalues of $\bar{X} - \bar{Y}^{-1}$. Therefore $(\bar{X}, \bar{Y}, \bar{\gamma})$ satisfies Kuh-Tucker condition for the convex program (3.28), which is also the first-order necessary optimality condition for the nonconvex program (3.24) [99]. □

Algorithm 1 is pseudo-code for implementing the above described procedure.

Alternatively, we can also use the following formulation instead of (3.24)

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \gamma + \mu[\mathsf{tr}(\mathbf{Y} - \mathbf{X}^{-1}) - f_{[k]}(\mathbf{Y} - \mathbf{X}^{-1})] \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.31}$$

for which the Algorithm 1 can be easily adjusted for solution. Usually, the initial point $(X^{(0)}, Y^{(0)}, \gamma^{(0)})$ is taken as the optimal solution of the full-order controller program

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \gamma \quad \text{s.t.} \quad (3.9), (3.10). \tag{3.32}$$

and the preference of using (3.24) or (3.31) goes to whichever smaller among $\mathsf{tr}(X^{(0)} - Y^{(0)}) - f_{[k]}(X^{(0)} - Y^{(0)})$ and $\mathsf{tr}(Y^{(0)} - X^{(0)}) - f_{[k]}(Y^{(0)} - X^{(0)})$.

On the other hand, we can seek a k-order control to satisfy the $\mathcal{H}_\infty$-gain condition (3.5) for given γ by solving the following nonconvex program

$$\min_{\mathbf{X},\mathbf{Y},\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} F(\mathbf{X},\mathbf{Y}) := \mathsf{tr}(\mathbf{X} - \mathbf{Y}^{-1}) - f_{[k]}(\mathbf{X} - \mathbf{Y}^{-1}) \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.33}$$

where the penalty parameter μ is not needed. The pseudo-code for solving (3.33) is provided by Algorithm 2, which is terminated when the zero value (with some tolerance) of the objective in (3.33) is found so the rank condition (3.11) is fulfilled leading to the construction of k-order control.
Again, an alternative formulation to (3.33)

$$\min_{\mathbf{X},\mathbf{Y},\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} F(\mathbf{X},\mathbf{Y}) := \mathsf{tr}(\mathbf{Y} - \mathbf{X}^{-1}) - f_{[k]}(\mathbf{Y} - \mathbf{X}^{-1}) \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.34}$$

is preferred if $\mathsf{tr}(Y^{(0)} - X^{(0)}) - f_{[k]}(Y^{(0)} - X^{(0)})$ is smaller than $\mathsf{tr}(X^{(0)} - Y^{(0)}) - f_{[k]}(X^{(0)} - Y^{(0)})$, where $(X^{(0)}, Y^{(0)})$ is the optimal solution of the following full-order controller program

$$\min_{\mathbf{X},\mathbf{Y},\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \mathsf{tr}(\mathbf{X} + \mathbf{Y}) \quad \text{s.t.} \quad (3.9), (3.10). \tag{3.35}$$

Remarks on less reduced order controllers. The computational difficulty degree in the formulation in [3, 4] (for LTI systems) is proportional to the control order k. Particularly, less reduced order controllers may pose more computational challenges than highly reduced order ones. In contrast, we now show that using the rank constraints (3.11) helps us solve them at the same computational efficiency.

Algorithm 2 Nonconvex Spectral Optimization Algorithm for feasible k-order $\mathcal{H}_\infty$ controllers

1: Initialize $\kappa := 0$ and solve SDP (3.35) to find its optimal solution $(X^{(\kappa)}, Y^{(\kappa)})$. Stop the algorithm if

$$\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) - \sum_{i=1}^{k}(x_i^{(\kappa)})^H(X - Y^{-1})x_i^{(\kappa)}) \leq \epsilon_2$$

and accept $(X^{(0)}, Y^{(0)})$ as the solution of the nonconvex program (3.33).

2: **repeat**

3: Solve SDP

$$\begin{aligned}\min_{\mathbf{X},\mathbf{Y}} \ & \mathsf{tr}(\mathbf{X} - 2(Y^{(\kappa)})^{-1}) + \mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}) \\ & - \textstyle\sum_{i=1}^{k}(x_i^{(\kappa)})^H(\mathbf{X} - \mathbf{Y}^{-1})x_i^{(\kappa)}) \quad \text{s.t.} \quad (3.9), (3.10)\end{aligned} \tag{3.36}$$

to find its optimal solution $(X^{(\kappa+1)}, Y^{(\kappa+1)})$

4: Set $\kappa := \kappa + 1$.

5: **until** $\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) - \sum_{i=1}^{k}(x_i^{(\kappa)})^H(X^{(\kappa)} - (Y^{(\kappa)})-1)x_i^{(\kappa)}) \leq \epsilon_2$ or $F(X^{(\kappa-1)}, Y^{(\kappa-1)}) - F(X^{(\kappa)}, Y^{(\kappa)}) \leq \epsilon_1$

6: Accept $(X^{(\kappa)}, Y^{(\kappa)})$ as a found feasible solution of (3.23) under fixed γ if $\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) - f_{[k]}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) \leq \epsilon_2$.

Indeed, less order reduction means that $n-k$ is small. Suppose $\lambda_{[n-k]}(\mathbf{X} - \mathbf{Y}^{-1})$ is the sum of the $n-k$ smallest eigenvalues of $\mathbf{X} - \mathbf{Y}^{-1}$. Then (3.11) holds true if and only if $\lambda_{[n-k]}(\mathbf{X} - \mathbf{Y}^{-1}) = 0$. Therefore, we propose the following alternative formulation for (3.23):

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \ \gamma + \mu\lambda_{[n-k]}(\mathbf{X} - \mathbf{Y}^{-1}) \quad \text{s.t.} \quad (3.9) - (3.10). \tag{3.37}$$

Using the following variational principle

$$\lambda_{[n-k]}(X - Y^{-1}) = \min_{\text{orthornomal } x_1,\ldots,x_{n-k}} \sum_{i=1}^{n-k} x_i^H(\mathbf{X} - \mathbf{Y}^{-1})x_i$$

the following optimization is majorant optimization for (3.37)

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \ \gamma + \mu\sum_{i=1}^{n-k}(x_i^{(\kappa)})^H(\mathbf{X} - \mathbf{Y}^{-1})x_i^{(\kappa)} \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.38}$$

where $x_i^{(\kappa)}$, $i = 1, ..., n-k$ are the orthonormal eigenvectors corresponding to $(n-k)$ smallest eigenvectors of $X^{(\kappa)} - (Y^{(\kappa)})^{-1}$. Since each $(x_i^{(\kappa)})^H Y^{-1} x_i^{(\kappa)}$ is convex in $Y > 0$, the following convex optimization is majorant minimization for (3.38) and (3.37)

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \gamma + \mu \sum_{i=1}^{n-k} (x_i^{(\kappa)})^H (\mathbf{X} - 2(Y^{(\kappa)})^{-1} + (Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}) x_i^{(\kappa)}$$
$$\text{s.t.} \quad (3.9) - (3.10), \tag{3.39}$$

which provides an alternative to κ-th iteration (3.27). This iteration is more efficient than (3.27) for larger $n-k$, i.e. for lower order k of the controllers.

3.3 Static output feedback LPV-LFT $\mathcal{H}_\infty$ controller

The static output feedback LPV-LFT controller corresponds to $k = 0$, i.e. the control in (3.7) is in the form

$$u(t) = (D_{K11} + D_{K1\Delta}(I - \Delta_K(\alpha(t))D_{K\Delta\Delta})^{-1}\Delta_K(\alpha(t))D_{K\Delta 1})y(t) \tag{3.40}$$

leading to the following optimization formulation for its synthesis

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \gamma \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.41}$$
$$\mathbf{X} = \mathbf{Y}^{-1}, \tag{3.42}$$

for

$$\hat{\mathbf{K}} := \begin{bmatrix} \mathbf{D_{K11}} & \mathbf{D_{K1\Delta}} \\ \mathbf{D_{K\Delta 1}} & \mathbf{D_{K\Delta\Delta}} \end{bmatrix}$$

and setting $\mathbf{A_K} = 0$, $\mathbf{B_{K1}} = 0$, $\mathbf{B_{K_\Delta}} = 0$, $\mathbf{C_{K1}} = 0$, $\mathbf{C_{K\Delta}} =$ in (3.12). The controller (3.40) is recovered by (3.13), (3.17), (3.18) and (3.21). To the author's best knowledge such simple structured controller (3.40) has not been considered in literature so far.

The first attractive reformulation of nonlinear constraint (3.42) is given back in [100]

$$(3.10),\ \text{Trace}(\mathbf{XY}) \leq n \tag{3.43}$$

where the nonconvexity is concentrated at the last indefinite quadratic constraint in $(\mathbf{Y}, \mathbf{X})$. Alternating optimization between $\mathbf{X}$ and $\mathbf{Y}$ is applied in handling (3.43).
Later, [4] also addressed the static output feedback controller problem for LTI systems by developing the so called convex-concave inequality approach for a solution of the corresponding BMI reformulation. All these results must start from a feasible point of a nonconvex feasible set, which is not easily located.
Note that $\mathbf{X} \succeq \mathbf{Y}^{-1}$ by LMI (3.10), which yields $\mathbf{X} - \mathbf{Y}^{-1} \succeq 0$. Hence, the nonlinear equality (3.42) holds if and only if

$$\mathsf{tr}(\mathbf{X}) - \mathsf{tr}(\mathbf{Y}^{-1}) = 0.$$

In other words, the nonnegative quantity $\mathsf{tr}(\mathbf{X}) - \mathsf{tr}(\mathbf{Y}^{-1})$ can be used to measure the degree of satisfaction of the nonlinear equality (3.42). Instead of the formulation (3.24) we consider a simpler nonconvex program

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \ \gamma + \mu(\mathsf{tr}(\mathbf{X}) - \mathsf{tr}(\mathbf{Y}^{-1})) \quad \text{s.t.} \quad (3.9) - (3.10). \tag{3.44}$$

Using (3.26), at $(X^{(\kappa)}, Y^{(\kappa)})$ feasible to LMIs (3.9)-(3.10), the following convex program is a majorant minimization for the nonconvex program (3.44)

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \ \gamma + \mu(\mathsf{tr}(\mathbf{X}) - 2\mathsf{tr}(Y^{(\kappa)}) + \mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1})) \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.45}$$

The pseudo-code using (3.45) in κth iteration is given by Algorithm 3.

Alternatively, $\mathbf{Y} \succeq \mathbf{X}^{-1}$ by LMI (3.10), so whenever

$$\mathsf{tr}(Y^{(0)}) - \mathsf{tr}((X^{(0)})^{-1}) < \mathsf{tr}(X^{(0)}) - \mathsf{tr}((Y^{(0)})^{-1}) \tag{3.46}$$

for the initial point $(X^{(0)}, Y^{(0)})$ we use

$$\min_{\mathbf{X},\mathbf{Y},\gamma,\mathbf{R},\mathbf{H},\mathbf{Q},\mathbf{E},\mathbf{\Delta}_{Ki},\hat{\mathbf{K}}} \ \gamma + \mu(\mathsf{tr}(\mathbf{Y}) - \mathsf{tr}(\mathbf{X}^{-1})) \quad \text{s.t.} \quad (3.9) - (3.10) \tag{3.47}$$

instead of the formulation (3.44), for which Algorithm 3 can be easily adjusted for solution.

Algorithm 3 Nonconvex Spectral Optimization Algorithm for static LPV-LFT $\mathcal{H}_\infty$ controllers

1: Initialize $\kappa := 0$ and solve SDP (3.41) to find its optimal solution $(X^{(\kappa)}, Y^{(\kappa)}, \gamma^{(\kappa)})$. Stop the algorithm if

$$\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) \leq \epsilon_2 \tag{3.48}$$

and accept $(X^{(0)}, Y^{(0)}, \gamma^{(0)})$ as the optimal solution of the nonconvex program (3.23). Otherwise set $\mu = 0.5$.

2: **repeat**

3: **if** $\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) \geq \epsilon_2$ **then** reset $\mu \to 2\mu$ and solve SDP (3.45) to find the optimal solution $(X^{(\kappa+1)}, Y^{(\kappa+1)}, \gamma^{(\kappa+1)})$.

4: **else** Solve SDP (3.45) with additional convex constraint

$$\mathsf{tr}(\mathbf{X} - 2(Y^{(\kappa)})^{-1}) + \mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}) \leq \epsilon_2 \tag{3.49}$$

to find the optimal solution $(X^{(\kappa+1)}, Y^{(\kappa+1)}, \gamma^{(\kappa+1)})$

5: **end if**

6: Set $\kappa := \kappa + 1$.

7: **until** $\gamma^{(\kappa)} - \gamma^{(\kappa-1)} \leq \epsilon_1$

8: Accept $(X^{(\kappa)}, Y^{(\kappa)}, \gamma^{(\kappa)})$ as a found suboptimal solution of (3.41)-(3.42) if (3.48) is fulfilled.

Similarly to (3.33), we address the design of a static output feedback $\mathcal{H}_\infty$ controller to satisfy the $\mathcal{H}_\infty$-gain condition (3.5) for given γ by solving the following nonconvex program

$$\min_{\mathbf{X},\mathbf{Y}} \ \mathsf{tr}(\mathbf{X}) - \mathsf{tr}(\mathbf{Y}^{-1}) \quad \text{s.t.} \quad (3.9) - (3.10). \tag{3.50}$$

Its κth iteration is

$$\min_{\mathbf{X},\mathbf{Y}} \ \mathsf{tr}(\mathbf{X}) - 2\mathsf{tr}(Y^{(\kappa)}) + \mathsf{tr}((Y^{(\kappa)})^{-1}\mathbf{Y}(Y^{(\kappa)})^{-1}) \quad \text{s.t.} \quad (3.9) - (3.10), \tag{3.51}$$

and Algorithm 4 is the pseudo-code for the implementation.

Alternatively, whenever (3.46), we use

$$\min_{\mathbf{X},\mathbf{Y}} \ \mathsf{tr}(\mathbf{Y}) - \mathsf{tr}(\mathbf{X}^{-1}) \quad \text{s.t.} \quad (3.9) - (3.10) \tag{3.52}$$

instead of (3.50), for which Algorithm 4 can be easily adjusted for computation.

Algorithm 4 Nonconvex Spectral Optimization Algorithm for feasible static output feedback LPV-LFT $\mathcal{H}_\infty$ controllers

1: Initialize $\kappa := 0$ and solve SDP (3.41) (for fixed $\gamma = \bar{\gamma}$ to find its optimal solution $(X^{(\kappa)}, Y^{(\kappa)})$. Stop the algorithm if

$$\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) \leq \epsilon_2 \tag{3.53}$$

and accept $(X^{(0)}, Y^{(0)})$ as the optimal solution of the nonconvex program (3.23).
2: **repeat**
3: Solve SDP (3.51) to find the optimal solution $(X^{(\kappa+1)}, Y^{(\kappa+1)})$.
4: Set $\kappa := \kappa + 1$.
5: **until** $\mathsf{tr}(X^{(\kappa)} - (Y^{(\kappa)})^{-1}) \leq \epsilon_2$ or $\mathsf{tr}(X^{(\kappa-1)} - Y^{(\kappa-1)}) - \mathsf{tr}(X^{(\kappa)} - Y^{(\kappa)}) \leq \epsilon_1$
6: Accept $(X^{(\kappa)}, Y^{(\kappa)})$ as a found feasible solution of (3.41)-(3.42) under fixed γ if (3.53) is fulfilled.

3.4 Simulation results

The hardware and software facilities for our computational implementation are:

- Processor: Intel(R) Core(TM) i5-3470 CPU @3.20GHz;
- Software: Matlab version R2015b;
- Matlab toolbox: Yalmip[101] with SeDumi 1.3 [102] solver for SDP;
- Data: The data in subsection 4.1 are from [88], while the state-space data in subsections 4.2, 4.3 and 4.4 are from [103];
- Criterion: The stop and rank check criterion ϵ_1 and ϵ_2 are set as 10^{-4}.

3.4.1 RTAC control

Consider the nonlinear benchmark model [104] of rotational-translational actuator (RTAC). The regulated output is the tracking performance of the translational and angular positions and control

$$z = (0.1x_1, 0.1x_3, u)^T$$

The system can be represented by LPV-LFT (3.1)-(3.2) [88, Appendix] with the numerical values of the matrices in (3.1)-(3.2) recalled in Appendix A.
By solving SDP (3.32) for Step 1 of Algorithm 1, we found $\gamma^{(0)} = 8.1909$ with $X^0 - (Y^0)^{-1}$ of rank-four, which lead to full-order control (3.3)-(3.4) [88]. Implementing Algorithm 1 with $\mu = 1$ for the first-order controller, $\gamma = 9.3785$ was found and the following numerical data for control (3.3)-(3.4) are obtained

$$A_K = -4.2617,\ B_{K1} = \begin{bmatrix} 0.4767 & -2.0207 \end{bmatrix}, C_{K1} = -0.1327, \quad D_{K11} = \begin{bmatrix} -0.0670 & 0.0284 \end{bmatrix},$$

$$B_{K\Delta} = \begin{bmatrix} 0 & -0.0030 & -0.0132 \end{bmatrix}, D_{K1\Delta} = \begin{bmatrix} 0 & 0 & 0.0013 \end{bmatrix}, D_{K\Delta\Delta} = \begin{bmatrix} -0.0002 & -0.3509 & -1.4991 \\ 0 & -0.0003 & -0.0015 \\ 0 & 0.0002 & 0.0008 \end{bmatrix},$$

$$C_{K\Delta} = \begin{bmatrix} -30.6379 \\ 4.2867 \\ -1.6715 \end{bmatrix}, D_{K\Delta 1} = \begin{bmatrix} -11.5584 & -2.5580 \\ 0.1977 & 0.0127 \\ 0.0034 & -0.0107 \end{bmatrix},$$

$$\Delta_{K1} = \begin{bmatrix} -0.3118 & 4.8064 & -25.2717 \\ 0.3425 & 4.8514 & 64.0377 \\ -0.0932 & -27.9300 & -84.1243 \end{bmatrix}, \Delta_{K2} = \begin{bmatrix} -0.3125 & 3.5657 & 51.1251 \\ 0.3434 & 6.5093 & -39.3112 \\ -0.0934 & -28.3084 & -59.9568 \end{bmatrix},$$

$$\Delta_{K3} = \begin{bmatrix} 0.3118 & -4.8064 & 25.2717 \\ -0.3425 & -4.8514 & -64.0377 \\ 0.0932 & 27.9300 & 84.1243 \end{bmatrix}, \Delta_{K4} = \begin{bmatrix} 0.3125 & -3.5657 & -51.1251 \\ -0.3434 & -6.5093 & 39.3112 \\ 0.0934 & 28.3084 & 59.9568 \end{bmatrix},$$

Under the condition $x(0) = (0.5, 0, 0, 0)^T$, the simulation given by Figs. 3.2-3.4 clearly shows that our first-order LPV-LFT stabilizes the system well.

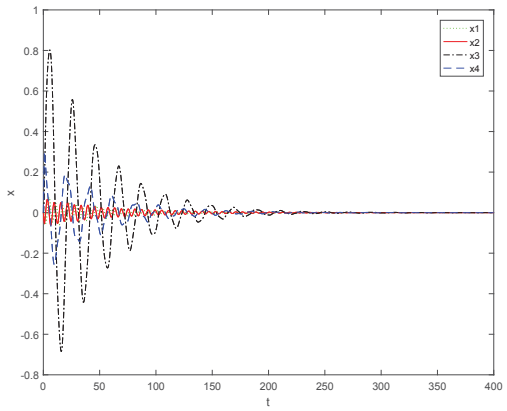

Fig. 3.2 Tracking performance of the first-order LPV-LFT controller in the absence of disturbance

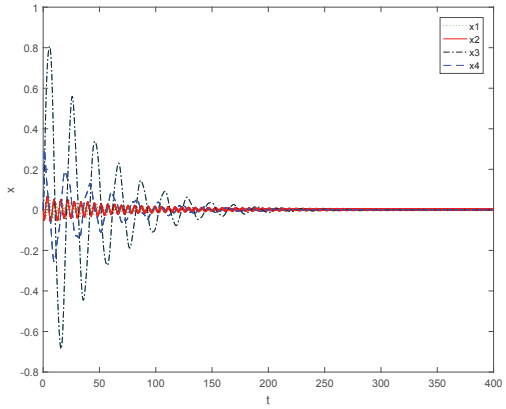

Fig. 3.3 Tracking performance of the first-order LPV-LFT controller with the disturbance $w = 0.1\sin(5\pi t)$

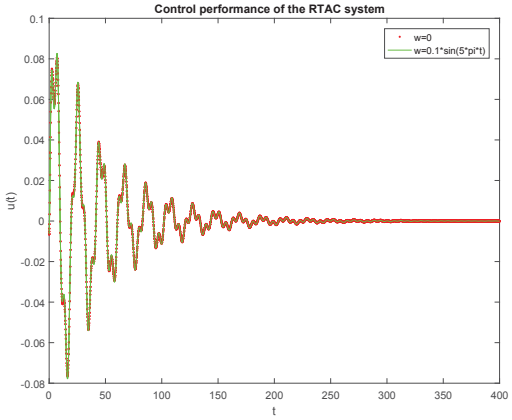

Fig. 3.4 The behaviour of the first-order LPV-LFT controller in the absence of disturbance (dot) and with disturbance $w = 0.1\sin(5\pi t)$(solid).

3.4.2 Reduced order LPV-LFT controllers

We modify the LTI examples in [3, Sec. 10] by adding the gain-scheduling channel (w_Δ, z_Δ) to have LPV-LFT system (3.1). The randomly generated matrix sets for the gain-scheduling channel are provided in Appendix B.

The computational results by implementing Algorithm 1 are provided by Table 3.1. The full order of LPV-LFT control, which is equal to the system state dimension is given in the second row with the initial γ obtained by solving (3.32) (for full-order

Table 3.1 Computational results for reduced-order LPV-LFT controllers by Algorithm 1

	Full-order	Lower bound by (3.32)	μ	γ	order k
VTOL helicopter	4	0.0871	0.05	0.2958	1
Chemical reactor	4	0.8653	0.05	0.8653	1
Transport airplane	10	1.7042	10	2.5302	1

Table 3.2 Computational results for static LPV-LFT controllers by Algorithm 3

	Full order	Lower bound by (3.32)	μ	γ
AC1	5	2.76E-08	0.1	6.68E-07
AC2	5	0.117675	1	2.113598
AC3	5	3.095253	1	4.891439
AC6	7	3.683339	5	4.948768
AC8	9	8.727383	20	9.57084
AC9	10	1.000091	5	1.002586
AC15	4	16.30176	5	18.91505
AC17	4	6.686188	5	6.688397
HE4	8	28.81522	5	37.26205

LPV-LFT control) given in the third row. The fourth column indicates the value of initial μ used. The fifth column is the found value of γ for the controller of order indicated in the sixth column.

3.4.3 Static output feedback LPV-LFT controller

We modify the LTI examples in [4, Tab. III] by adding the matrices relating to the gain-scheduling channel (w_Δ, z_Δ) as provided in Appendix C. The computational results by implementing Algorithm 3 are provided in Table 3.2, which is formatted similarly to Table 3.1. Table 3.1 and Table 3.2 reveal that very low-order (first-order or static) controllers, which lead to very efficient online control realization, work well in these examples.

3.4.4 LTI systems

In LTI systems there are no gain-scheduling channel (w_Δ, z_Δ) in system (3.1) and no gain-scheduling channel (w_K, z_K) in controller (3.3). Accordingly,

$$\hat{\mathbf{K}} := \begin{bmatrix} \mathbf{A_K} & \mathbf{B_{K1}} \\ \mathbf{C_{K1}} & \mathbf{D_{K11}} \end{bmatrix}$$

in (4.48) and LMI (3.9) becomes

$$\begin{bmatrix} \mathbf{X}A + \mathbf{B_{K1}}C_2 + (*) & * & * & * \\ \mathbf{A_K}^T + A + B_2\mathbf{D_{K11}}C_2 & (A\mathbf{Y} + B_2\mathbf{C_{K1}}) + (*) & * & * \\ B_1^T\mathbf{X} + D_{21}^T\mathbf{B_{K1}}^T & B_1^T + D_{21}^T\mathbf{D_{K11}}^T B_2^T & -\gamma I & * \\ C_1 + D_{12}\mathbf{D_{K11}}C_2 & C_1\mathbf{Y} + D_{12}\mathbf{C_{K1}} & D_{11} + D_{12}\mathbf{D_{K11}}D_{21} & -\gamma I \end{bmatrix} \prec 0 \tag{3.54}$$

with control recovered by (3.13)-(3.16) [105].
As mentioned in the Introduction, the Matlab command systune [18] is the most efficient tool for LTI systems. Our purpose in this subsection is not to show any advantage of the Lyapunov matrix rank constrained approach over the frequency approach by using the Matlab command systune [18]. We consider numerical examples from [3, Sec. 10] and [4, Tab. III] for LTI systems to only show the efficiency of our approach in handling the rank-reduced constraints.

Dynamic and static output feedback controllers in [3]

All cases in [3, Section 10] were tested. The computational results are summarized in Table 3.3, where the first column is the case name, the second column is the initial γ obtained by solving (3.32), i.e. it is the optimal $\mathcal{H}_\infty$ by the full-order controller, the third column is the objective function, which is either (3.24) or (3.31) for k-order controller and either (3.44) or (3.47) for static output feedback controller. The fourth column is the value of initial μ. The fifth column is the found value of γ. The sixth column provides the controller order. The seventh column is the iterations by our method; the last two column are the found γ and corresponding order in [3], respectively.

Compared with [3], it can be seen that, our optimal γ are better than [3] in all cases provided. For the Transport aircraft example, [3] failed to obtain 2-order controller, though it was found by our Algorithm after 6 iterations. The Piezoelectric actuator example poses the most difficulty for [3] but it is easily solved by our algorithm with 3 iterations for both order-2 controllers and statistic output feedback controller.

Table 3.3 Numerical results by Algorithms 1 and 3 compared with [3]

Cases	Lower bound	Obj	μ	γ	order	# iter	γ in [3]	order in [3]
VTOL helicopter	0.0737	(3.24)	1	0.118713	2	1	0.133	2
VTOL helicopter	0.0737	(3.44)	0.7	0.1539	0	20	0.1542	0
Chemical reactor	0.8617	(3.24)	1	0.8617	2	1	1.142	2
Chemical reactor	0.8617	(3.44)	1	0.8937	0	28	1.183	0
Transport aircraft	0.0417	(3.31)	10	0.349	1	42	2.86	1
Transport aircraft	0.0417	(3.31)	1	0.2167	2	6	failed	2
Piezoelectric actuator	3.11E-05	(3.31)	5	0.0048	2	3	0.03	2
Piezoelectric actuator	3.11E-05	(3.47)	100	0.0213	0	3	0.0578	0
Coupled springs model	0.01996	(3.24)	1	0.01997	2	4	0.0235	4

The last example in [3] for static output feedback for a plant with state dimension 82. The computational results by implementing Algorithm 2 are summarized in Table 3.4, whose format is similar to Table 3.3 but the second column is the fixed value of γ, which is better than the value provided by [3] in the 7-th column. The sixth column is the value of trace$(X - Y^{-1})$ at the last iteration.

Table 3.4 Distillation tower case with γ fixed by Algorithm 4 compared with [3]

Cases	Fixed γ	Obj	order	# iter	trace$(X - Y^{-1})$	γ in [3]	order in [3]
Distillation tower	0.8000	(3.44)	0	64	2.24E-05	1.0722	0

Static output feedback controllers in [4]

There are 45 cases in [4, Table III]. The computational results are summarized in Table 3.5 and Table 3.6 by Algorithm 3 and Algorithm 4, respectively, whose format is in similar style to Table 3.3. Any way, the best γ obtained from systune [18], which is referred to what is achievable, is also provided in the ninth column. In AC9, EB1, EB2 and EB3, the results by [4] are obviously incorrect as its provided values of γ in the seventh column is even smaller than their lower bound in the second column. It should be mentioned that, the value of μ is increased to regulate the convergence speed, but a

larger μ results in larger γ as well. According to [4], the iteration threshold to stop its solver is 300. The solver [4] is trapped by local minima in AC4 and AC12 as its found values are much bigger than that found by our Algorithm 3. The former is also heading to a wrong minima in the case AC12 as the value found after 300 iterations is still very far from that found by the latter. In AC1, AC2, AC8, AC11, AC16, AC17, HE1, HE2, REA2, DIS3, DIS4, WEC2, WEC3, IH, PSM, and NN1, the later clearly outperforms the former in both computational performance and convergence. Note that the result for HE3, DIS1, TG1, NN4, NN11 and NN16 in Table 3.6 was obtained by using Algorithm 4. Our simulation results are better than or consistent with [4]. It should be realized that the results in these Tables also depend on the setting of stopping parameters in the algorithms, i.e. the presented results are not necessarily the local minima when the algorithms stop due to slow convergence.

3.5 Conclusions

We have proposed new algorithms for solving matrix-rank constrained optimization arising in reduced-order $\mathcal{H}_\infty$ LPV-LFT controller design. Unlike the previous developments, we formulate the problem as minimization of nonconvex objective function over a convex feasibility set. The global convergence of the proposed Algorithms to a local minima follows immediately from their path-following nature, while there is no difficulty for initial solutions, which are found from a semi-definite program for full-order controller synthesis. The numerical results reported for the benchmark collections have shown their merit. Their application to solutions of reduced-order generalized $\mathcal{H}_2$ LPV-LFT controllers is obvious. Their extensions to multi-objective and structured controller design are currently under development.

Appendix A: LPV-LFT data of RTAC System

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1.0365 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0.1946 & 0 & 0 & 0 \end{bmatrix}, B_\Delta = \begin{bmatrix} 0 & 0 & 0 \\ -0.5 & 0.5 & 1.0365 \\ 0 & 0 & 0 \\ 0.5 & 0.5 & -0.1946 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 1.0365 \\ 0 \\ -0.1946 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ -0.1946 \\ 0 \\ 1.03654 \end{bmatrix}$$

$$C_\Delta = \begin{bmatrix} 1.5157 & 0 & 0 & 0 \\ 0.7088 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, D_{\Delta\Delta} = \begin{bmatrix} 1.2311 & 0 & -1.5157 \\ 0 & -0.8419 & -0.7088 \\ 0 & 0 & 0 \end{bmatrix}, D_{\Delta 1} = \begin{bmatrix} -1.5157 \\ -0.7088 \\ 0 \end{bmatrix}, D_{\Delta 2} = \begin{bmatrix} 1.5157 \\ -0.7088 \\ 0 \end{bmatrix}$$

$$C_1 = \begin{bmatrix} 0.31622 & 0 & 0 & 0 \\ 0 & 0 & 0.3162 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, D_{12} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\Delta_1 = \begin{bmatrix} 0.01224 & 0 & 0 \\ 0 & 0.01224 & 0 \\ 0 & 0 & 0.04794 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.01224 & 0 & 0 \\ 0 & 0.01224 & 0 \\ 0 & 0 & -0.04794 \end{bmatrix}$$

$$\Delta_3 = \begin{bmatrix} -0.01224 & 0 & 0 \\ 0 & -0.01224 & 0 \\ 0 & 0 & -0.04794 \end{bmatrix}, \Delta_4 = \begin{bmatrix} -0.01224 & 0 & 0 \\ 0 & -0.01224 & 0 \\ 0 & 0 & 0.04794 \end{bmatrix}.$$

$$\alpha_1(t) = \tfrac{1}{4a_2a_5}(a_2 - \delta_1(t))(a_5 - \delta_2(t)), \alpha_2(t) = \tfrac{1}{4a_2a_5}(a_2 - \delta_1(t))(a_5 + \delta_2(t))$$
$$\alpha_3(t) = \tfrac{1}{4a_2a_5}(a_2 + \delta_1(t))(a_5 + \delta_2(t)),\ \alpha_4(t) = \tfrac{1}{4a_2a_5}(a_2 + \delta_1(t))(a_5 - \delta_2(t))$$

where

$$a_1 = \epsilon\tfrac{\cos 0.5+1}{2} \quad a_2 = \epsilon\tfrac{1-\cos 0.5}{2}, a_3 = 1 - a_1, a_4 = 1 + a_1, a_5 = \epsilon 0.5 \sin 0.5$$
$$\delta_1 = \epsilon \cos x_3 - a_1, -a_2 \le \delta_1 \le a_2, \delta_2 = \epsilon x_4 \sin x_3, -a_5 \le \delta_2 \le a_5,\ \epsilon = 0.2.$$

Appendix B

Modified VTOL helicopter system

$$B_\Delta = \begin{bmatrix} 0.5243 & 0.4413 \\ 0.3440 & 0.1393 \\ 0.1109 & 0.5365 \\ 0.2478 & 0.1764 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.3031 & 0.5501 & 0.5605 & 0.0255 \\ 0.3292 & 0.2303 & 0.3367 & 0.1256 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.0740 & 0.0618 \\ 0.0217 & 0.0152 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0305 & 0.0611 \\ 0.0097 & 0.0724 \end{bmatrix}, \Delta_3 = \begin{bmatrix} 0.0477 & 0.0562 \\ 0.0371 & 0.0565 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0506 & 0.0704 \\ 0.0474 & 0.0156 \end{bmatrix}.$$

Modified chemical reactor system

$$B_\Delta = \begin{bmatrix} 0.3992 & 0.2357 & 0.1968 & 0.2560 \\ 0.3125 & 0.0069 & 0.3433 & 0.0130 \\ 0.2363 & 0.0491 & 0.0851 & 0.2498 \\ 0.3772 & 0.3506 & 0.2244 & 0.1473 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.1869 & 0.3511 & 0.2284 & 0.2440 \\ 0.3123 & 0.2240 & 0.0962 & 0.2576 \\ 0.0888 & 0.3294 & 0.2597 & 0.2844 \\ 0.1941 & 0.2879 & 0.0325 & 0.3472 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.0032 & 0.0133 & 0.0410 & 0.0322 \\ 0.0316 & 0.0095 & 0.0182 & 0.0346 \\ 0.0124 & 0.0122 & 0.0348 & 0.0287 \\ 0.0079 & 0.0028 & 0.0449 & 0.0080 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0232 & 0.0099 & 0.0097 & 0.0067 \\ 0.0403 & 0.0267 & 0.0448 & 0.0079 \\ 0.0413 & 0.0303 & 0.0039 & 0.0294 \\ 0.0128 & 0.0197 & 0.0050 & 0.0271 \end{bmatrix}$$

$$\Delta_3 = \begin{bmatrix} 0.0019 & 0.0024 & 0.0319 & 0.0148 \\ 0.0346 & 0.0320 & 0.0292 & 0.0050 \\ 0.0271 & 0.0347 & 0.0191 & 0.0011 \\ 0.0274 & 0.0366 & 0.0066 & 0.0349 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0144 & 0.0309 & 0.0408 & 0.0085 \\ 0.0141 & 0.0012 & 0.0166 & 0.0316 \\ 0.0159 & 0.0402 & 0.0213 & 0.0158 \\ 0.0223 & 0.0267 & 0.0026 & 0.0429 \end{bmatrix}.$$

Modified transport airplane system

$$B_\Delta^T = \begin{bmatrix} 0.1454 & 0.0308 & 0.2485 & 0.2151 & 0.0160 & 0.0712 & 0.1049 & 0.1605 & 0.2354 & 0.1063 \\ 0.1822 & 0.2487 & 0.1115 & 0.1627 & 0.2309 & 0.2232 & 0.0366 & 0.0514 & 0.1590 & 0.1985 \\ 0.0632 & 0.0946 & 0.0874 & 0.1169 & 0.0438 & 0.1910 & 0.0251 & 0.2280 & 0.0210 & 0.2002 \\ 0.2174 & 0.1155 & 0.1550 & 0.0908 & 0.1944 & 0.1510 & 0.1566 & 0.2160 & 0.1170 & 0.2261 \end{bmatrix},$$

$$C_\Delta = \begin{bmatrix} 0.1991 & 0.0018 & 0.2215 & 0.2496 & 0.1639 & 0.1131 & 0.0108 & 0.0643 & 0.2234 & 0.1669 \\ 0.0184 & 0.2225 & 0.1165 & 0.2177 & 0.1031 & 0.0067 & 0.1392 & 0.1817 & 0.1950 & 0.2448 \\ 0.2484 & 0.2162 & 0.0930 & 0.0000 & 0.1192 & 0.1064 & 0.1530 & 0.1785 & 0.1238 & 0.0293 \\ 0.1942 & 0.1746 & 0.0208 & 0.1702 & 0.0972 & 0.0265 & 0.1849 & 0.1261 & 0.1906 & 0.2175 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.0191 & 0.0212 & 0.0033 & 0.0055 \\ 0.0241 & 0.0303 & 0.0386 & 0.0324 \\ 0.0217 & 0.0173 & 0.0281 & 0.0281 \\ 0.0279 & 0.0353 & 0.0216 & 0.0161 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0136 & 0.0429 & 0.0119 & 0.0158 \\ 0.0156 & 0.0105 & 0.0376 & 0.0239 \\ 0.0289 & 0.0252 & 0.0393 & 0.0138 \\ 0.0342 & 0.0211 & 0.0192 & 0.0078 \end{bmatrix}$$

$$\Delta_3 = \begin{bmatrix} 0.0181 & 0.0437 & 0.0088 & 0.0238 \\ 0.0130 & 0.0085 & 0.0322 & 0.0011 \\ 0.0111 & 0.0416 & 0.0298 & 0.0010 \\ 0.0294 & 0.0260 & 0.0152 & 0.0365 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0037 & 0.0241 & 0.0188 & 0.0333 \\ 0.0234 & 0.0321 & 0.0223 & 0.0058 \\ 0.0263 & 0.0101 & 0.0321 & 0.0320 \\ 0.0269 & 0.0338 & 0.0076 & 0.0325 \end{bmatrix}$$

Appendix C

Modified AC1 system

$$B_\Delta = \begin{bmatrix} 0.2972 & 0.1612 \\ 0.4326 & 0.0854 \\ 0.1907 & 0.2157 \\ 0.4383 & 0.2318 \\ 0.5171 & 0.3084 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.2360 & 0.5001 & 0.1001 & 0.5307 & 0.0602 \\ 0.1685 & 0.2330 & 0.4901 & 0.2377 & 0.1398 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.1028 & 0.0531 \\ 0.0173 & 0.1622 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0046 & 0.1142 \\ 0.1453 & 0.0764 \end{bmatrix}, \Delta_3 = \begin{bmatrix} 0.0936 & 0.0742 \\ 0.0384 & 0.1558 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.1177 & 0.0499 \\ 0.1122 & 0.1052 \end{bmatrix}.$$

Modified AC2 system

$$B_\Delta = \begin{bmatrix} 0.1860 & 0.1471 & 0.1931 & 0.1790 & 0.3075 \\ 0.3197 & 0.2706 & 0.0325 & 0.0686 & 0.2026 \\ 0.2126 & 0.2984 & 0.0404 & 0.1789 & 0.3360 \\ 0.0559 & 0.2856 & 0.0493 & 0.0534 & 0.2518 \\ 0.0722 & 0.1151 & 0.2453 & 0.0199 & 0.2107 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.0063 & 0.2641 & 0.1275 & 0.2541 & 0.2679 \\ 0.1064 & 0.1382 & 0.2599 & 0.3052 & 0.2471 \\ 0.1364 & 0.2854 & 0.2427 & 0.1053 & 0.0538 \\ 0.0869 & 0.1257 & 0.1213 & 0.2158 & 0.2771 \\ 0.0633 & 0.2472 & 0.0694 & 0.1410 & 0.3182 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.0464 & 0.0349 & 0.0130 & 0.0420 & 0.0356 \\ 0.0500 & 0.0563 & 0.0284 & 0.0334 & 0.0376 \\ 0.0563 & 0.0300 & 0.0513 & 0.0140 & 0.0415 \\ 0.0000 & 0.0273 & 0.0327 & 0.0379 & 0.0507 \\ 0.0493 & 0.0456 & 0.0481 & 0.0048 & 0.0559 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0630 & 0.0099 & 0.0452 & 0.0041 & 0.0120 \\ 0.0476 & 0.0706 & 0.0516 & 0.0401 & 0.0155 \\ 0.0760 & 0.0396 & 0.0026 & 0.0158 & 0.0035 \\ 0.0475 & 0.0692 & 0.0503 & 0.0101 & 0.0520 \\ 0.0014 & 0.0171 & 0.0297 & 0.0168 & 0.0231 \end{bmatrix}$$

$$\Delta_3 = \begin{bmatrix} 0.0396 & 0.0091 & 0.0153 & 0.0697 & 0.0457 \\ 0.0511 & 0.0361 & 0.0415 & 0.0060 & 0.0422 \\ 0.0367 & 0.0627 & 0.0471 & 0.0078 & 0.0038 \\ 0.0394 & 0.0643 & 0.0307 & 0.0104 & 0.0685 \\ 0.0327 & 0.0199 & 0.0151 & 0.0122 & 0.0536 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0496 & 0.0578 & 0.0090 & 0.0224 & 0.0376 \\ 0.0043 & 0.0528 & 0.0021 & 0.0314 & 0.0574 \\ 0.0579 & 0.0345 & 0.0632 & 0.0436 & 0.0234 \\ 0.0628 & 0.0119 & 0.0203 & 0.0017 & 0.0300 \\ 0.0662 & 0.0268 & 0.0199 & 0.0566 & 0.0036 \end{bmatrix}.$$

Modified AC3 system

$$B_\Delta = \begin{bmatrix} 0.0421 & 0.1351 & 0.2372 & 0.3017 & 0.4218 \\ 0.0606 & 0.0929 & 0.0787 & 0.2381 & 0.0728 \\ 0.0718 & 0.1072 & 0.0905 & 0.1338 & 0.1101 \\ 0.0838 & 0.3812 & 0.0330 & 0.0710 & 0.1694 \\ 0.1355 & 0.3002 & 0.3901 & 0.2657 & 0.0316 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.2106 & 0.2908 & 0.1364 & 0.0399 & 0.2294 \\ 0.3321 & 0.1855 & 0.3337 & 0.0034 & 0.0470 \\ 0.1710 & 0.1951 & 0.3008 & 0.1827 & 0.2729 \\ 0.1179 & 0.2634 & 0.0541 & 0.2831 & 0.0546 \\ 0.0161 & 0.0257 & 0.0562 & 0.3122 & 0.0580 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.0495 & 0.0112 & 0.0254 & 0.0433 & 0.0260 \\ 0.0291 & 0.0276 & 0.0497 & 0.0243 & 0.0404 \\ 0.0712 & 0.0117 & 0.0213 & 0.0217 & 0.0538 \\ 0.0291 & 0.0549 & 0.0384 & 0.0328 & 0.0307 \\ 0.0450 & 0.0631 & 0.0603 & 0.0306 & 0.0311 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0086 & 0.0659 & 0.0523 & 0.0319 & 0.0113 \\ 0.0017 & 0.0644 & 0.0510 & 0.0146 & 0.0459 \\ 0.0200 & 0.0315 & 0.0512 & 0.0068 & 0.0616 \\ 0.0219 & 0.0166 & 0.0073 & 0.0567 & 0.0356 \\ 0.0450 & 0.0526 & 0.0469 & 0.0121 & 0.0484 \end{bmatrix}$$

$$\Delta_3 = \begin{bmatrix} 0.0107 & 0.0562 & 0.0380 & 0.0014 & 0.0640 \\ 0.0663 & 0.0520 & 0.0277 & 0.0642 & 0.0552 \\ 0.0376 & 0.0084 & 0.0289 & 0.0454 & 0.0401 \\ 0.0472 & 0.0365 & 0.0126 & 0.0648 & 0.0306 \\ 0.0025 & 0.0226 & 0.0178 & 0.0114 & 0.0179 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0471 & 0.0448 & 0.0199 & 0.0398 & 0.0395 \\ 0.0143 & 0.0402 & 0.0510 & 0.0595 & 0.0222 \\ 0.0040 & 0.0262 & 0.0494 & 0.0278 & 0.0624 \\ 0.0480 & 0.0245 & 0.0534 & 0.0038 & 0.0140 \\ 0.0420 & 0.0511 & 0.0317 & 0.0543 & 0.0409 \end{bmatrix}.$$

Modified AC6 system

$$B_\Delta^T = \begin{bmatrix} 0.0327 & 0.0447 & 0.3191 & 0.3238 & 0.1948 & 0.0202 & 0.0795 \\ 0.1196 & 0.2781 & 0.0052 & 0.0146 & 0.0572 & 0.2198 & 0.2478 \\ 0.2194 & 0.1527 & 0.1853 & 0.1004 & 0.2522 & 0.0640 & 0.2326 \\ 0.0622 & 0.1248 & 0.2119 & 0.2642 & 0.0275 & 0.3148 & 0.2627 \end{bmatrix}$$

$$C_\Delta = \begin{bmatrix} 0.3139 & 0.3081 & 0.1451 & 0.0664 & 0.3260 & 0.1220 & 0.2818 \\ 0.2094 & 0.2247 & 0.0274 & 0.0867 & 0.3412 & 0.3250 & 0.1408 \\ 0.1986 & 0.1267 & 0.0866 & 0.1507 & 0.1773 & 0.1334 & 0.0873 \\ 0.0523 & 0.1854 & 0.0445 & 0.0179 & 0.1767 & 0.0402 & 0.1459 \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0.0394 & 0.0411 & 0.0521 & 0.0284 \\ 0.0352 & 0.0413 & 0.0306 & 0.0759 \\ 0.0361 & 0.0661 & 0.0656 & 0.0708 \\ 0.0248 & 0.0643 & 0.0431 & 0.0445 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0677 & 0.0512 & 0.0246 & 0.0338 \\ 0.0638 & 0.0251 & 0.0186 & 0.1004 \\ 0.0226 & 0.0918 & 0.0248 & 0.0468 \\ 0.0328 & 0.0212 & 0.0474 & 0.0201 \end{bmatrix},$$

$$\Delta_3 = \begin{bmatrix} 0.0883 & 0.0252 & 0.0588 & 0.0290 \\ 0.0956 & 0.0399 & 0.0694 & 0.0311 \\ 0.0428 & 0.0581 & 0.0216 & 0.0414 \\ 0.0108 & 0.0256 & 0.0115 & 0.0496 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0075 & 0.0820 & 0.0209 & 0.0460 \\ 0.0232 & 0.0645 & 0.0405 & 0.0204 \\ 0.0707 & 0.0431 & 0.0850 & 0.0431 \\ 0.0026 & 0.0511 & 0.0483 & 0.0551 \end{bmatrix}.$$

Modified AC8 system

$$B_\Delta^T = \begin{bmatrix} 0.0964 & 0.3048 & 0.2852 & 0.3392 & 0.3409 & 0.1208 & 0.1274 & 0.2153 & 0.1339 \\ 0.3424 & 0.3336 & 0.2293 & 0.1083 & 0.2802 & 0.0962 & 0.1879 & 0.1583 & 0.2217 \end{bmatrix}$$

$$C_\Delta = \begin{bmatrix} 0.4599 & 0.0189 & 0.0804 & 0.3147 & 0.1684 & 0.2111 & 0.3167 & 0.0977 & 0.1679 \\ 0.2118 & 0.3876 & 0.0747 & 0.1322 & 0.2087 & 0.2006 & 0.0867 & 0.0492 & 0.3997 \end{bmatrix}$$

$$\Delta_1 = \begin{bmatrix} 0.1237 & 0.1223 \\ 0.0942 & 0.0293 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0958 & 0.1372 \\ 0.0093 & 0.1091 \end{bmatrix}, \ \Delta_3 = \begin{bmatrix} 0.1035 & 0.1122 \\ 0.1193 & 0.0495 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0001 & 0.0671 \\ 0.1746 & 0.0708 \end{bmatrix}$$

Modified AC9 system

$$B_\Delta^T = \begin{bmatrix} 0.2911 & 0.3110 & 0.1462 & 0.2964 & 0.3551 & 0.0966 & 0.0516 & 0.0896 & 0.1393 & 0.1143 \\ 0.3691 & 0.0204 & 0.2359 & 0.0648 & 0.3337 & 0.0667 & 0.1999 & 0.3977 & 0.1414 & 0.0187 \end{bmatrix}$$

$$C_\Delta = \begin{bmatrix} 0.1202 & 0.2758 & 0.2005 & 0.3071 & 0.3533 & 0.1315 & 0.1279 & 0.2939 & 0.0463 & 0.0093 \\ 0.0354 & 0.2762 & 0.1248 & 0.2039 & 0.3294 & 0.2016 & 0.2298 & 0.2752 & 0.3034 & 0.1529 \end{bmatrix}$$

$$\Delta_1 = \begin{bmatrix} 0.0704 & 0.1100 \\ 0.1312 & 0.0757 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.1194 & 0.1227 \\ 0.0871 & 0.0558 \end{bmatrix}, \ \Delta_3 = \begin{bmatrix} 0.1365 & 0.0886 \\ 0.0008 & 0.1163 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0315 & 0.1032 \\ 0.1260 & 0.1117 \end{bmatrix}$$

Modified AC15 system

$$B_\Delta^T = \begin{bmatrix} 0.0637 & 0.3367 & 0.0498 & 0.3052 \\ 0.2307 & 0.3636 & 0.3338 & 0.4218 \\ 0.4159 & 0.1560 & 0.3262 & 0.0923 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.3155 & 0.4589 & 0.4242 & 0.1216 \\ 0.1837 & 0.0175 & 0.3698 & 0.1558 \\ 0.1707 & 0.4112 & 0.0459 & 0.3157 \end{bmatrix}$$

$$\Delta_1 = \begin{bmatrix} 0.0031 & 0.0480 & 0.0618 \\ 0.0745 & 0.0905 & 0.0860 \\ 0.0500 & 0.0610 & 0.0806 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0672 & 0.1033 & 0.0196 \\ 0.0213 & 0.0033 & 0.1141 \\ 0.0280 & 0.0571 & 0.0831 \end{bmatrix},$$

$$\Delta_3 = \begin{bmatrix} 0.0726 & 0.0990 & 0.0757 \\ 0.0684 & 0.0062 & 0.0140 \\ 0.0087 & 0.0104 & 0.1187 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0804 & 0.0649 & 0.0639 \\ 0.0711 & 0.0510 & 0.0788 \\ 0.0147 & 0.0957 & 0.0447 \end{bmatrix}$$

Modified AC17 system

$$B_\Delta = \begin{bmatrix} 0.1462 & 0.2304 & 0.3153 & 0.0251 \\ 0.3333 & 0.1848 & 0.2733 & 0.1295 \\ 0.1619 & 0.3464 & 0.2865 & 0.0017 \\ 0.0925 & 0.3603 & 0.3468 & 0.3192 \end{bmatrix}, C_\Delta = \begin{bmatrix} 0.0576 & 0.1217 & 0.3314 & 0.4104 \\ 0.0607 & 0.3968 & 0.1556 & 0.3572 \\ 0.3994 & 0.0267 & 0.1841 & 0.0318 \\ 0.2767 & 0.2101 & 0.2418 & 0.1279 \end{bmatrix}$$

$$\Delta_1 = \begin{bmatrix} 0.0563 & 0.0321 & 0.0455 & 0.0393 \\ 0.0406 & 0.0076 & 0.0137 & 0.0133 \\ 0.0252 & 0.0094 & 0.0492 & 0.0632 \\ 0.0594 & 0.0076 & 0.0998 & 0.0971 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0319 & 0.0677 & 0.0523 & 0.0801 \\ 0.0038 & 0.0418 & 0.0575 & 0.0428 \\ 0.0130 & 0.0611 & 0.0803 & 0.0129 \\ 0.0070 & 0.0640 & 0.0606 & 0.0062 \end{bmatrix},$$

$$\Delta_3 = \begin{bmatrix} 0.0572 & 0.0426 & 0.0399 & 0.0193 \\ 0.0247 & 0.0573 & 0.0750 & 0.0549 \\ 0.0538 & 0.0405 & 0.0312 & 0.0673 \\ 0.0523 & 0.0391 & 0.0094 & 0.0777 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0192 & 0.0456 & 0.0270 & 0.0693 \\ 0.0679 & 0.0443 & 0.0080 & 0.0160 \\ 0.0522 & 0.0582 & 0.0738 & 0.0098 \\ 0.0360 & 0.0665 & 0.0348 & 0.0809 \end{bmatrix}$$

Modified HE4 system

$$B_\Delta^T = \begin{bmatrix} 0.1677 & 0.0014 & 0.0621 & 0.1371 & 0.1744 & 0.1203 & 0.2021 & 0.1830 \\ 0.2395 & 0.0546 & 0.1572 & 0.1196 & 0.3053 & 0.2925 & 0.2790 & 0.1386 \\ 0.2348 & 0.2842 & 0.1166 & 0.0363 & 0.1641 & 0.1066 & 0.2654 & 0.0582 \\ 0.0652 & 0.1657 & 0.0263 & 0.0752 & 0.1983 & 0.0347 & 0.1898 & 0.2780 \end{bmatrix}$$

$$C_\Delta = \begin{bmatrix} 0.1227 & 0.2765 & 0.1709 & 0.1227 & 0.0710 & 0.0542 & 0.0562 & 0.1699 \\ 0.2542 & 0.1658 & 0.2155 & 0.1823 & 0.1005 & 0.0276 & 0.1711 & 0.2431 \\ 0.1595 & 0.1418 & 0.1693 & 0.0498 & 0.0855 & 0.1956 & 0.2001 & 0.1906 \\ 0.2060 & 0.1499 & 0.2037 & 0.2471 & 0.2728 & 0.0271 & 0.2510 & 0.2213 \end{bmatrix}$$

$$\Delta_1 = \begin{bmatrix} 0.0433 & 0.0570 & 0.0391 & 0.0547 \\ 0.0277 & 0.0428 & 0.0431 & 0.0079 \\ 0.0673 & 0.0387 & 0.0760 & 0.0099 \\ 0.0842 & 0.0164 & 0.0632 & 0.0508 \end{bmatrix}, \Delta_2 = \begin{bmatrix} 0.0472 & 0.0643 & 0.0811 & 0.0197 \\ 0.0557 & 0.0492 & 0.0298 & 0.0826 \\ 0.0039 & 0.0164 & 0.0519 & 0.0068 \\ 0.0592 & 0.0771 & 0.0243 & 0.0200 \end{bmatrix},$$

$$\Delta_3 = \begin{bmatrix} 0.0838 & 0.0086 & 0.0144 & 0.0314 \\ 0.0495 & 0.0756 & 0.0556 & 0.0346 \\ 0.0502 & 0.0767 & 0.0155 & 0.0077 \\ 0.0817 & 0.0184 & 0.0100 & 0.0588 \end{bmatrix}, \Delta_4 = \begin{bmatrix} 0.0444 & 0.0540 & 0.0086 & 0.0591 \\ 0.0801 & 0.0604 & 0.0718 & 0.0232 \\ 0.0173 & 0.0307 & 0.0362 & 0.0482 \\ 0.0746 & 0.0386 & 0.0543 & 0.0301 \end{bmatrix}$$

Table 3.5 Numerical results of static output feedback controllers by Algorithm 3 compared with [4]

Cases	Lower bound	Obj	μ	γ	# iter	γ in [4]	# iter	γ by systune [18]	# iter
AC1	2.68E-06	(3.47)	0.1	3.42E-05	7	0.0177	93	6.97E-05	85
AC2	0.1115	(3.47)	0.1	0.1115	1	0.1140	99	0.1115	53
AC3	2.9675	(3.47)	0.15	3.4696	300	3.4859	210	3.4056	111
AC4	0.5573	(3.44)	0.3	1.0064	300	69.9900	2	0.9355	45
AC6	3.4275	(3.47)	0.1	4.1208	132	4.1954	167	4.114	59
AC7	0.0396	(3.47)	0.1	0.0657	150	0.0548	300	0.0651	28
AC8	1.6165	(3.47)	2	2.0508	16	3.052	247	2.005	37
AC9	1.0000	(3.47)	1	1.003	1	0.9237 (wrong)	300	1.0006	102
AC11	2.8079	(3.44)	10	2.9261	300	3.0104	68	2.818	95
AC12	0.0225	(3.44)	1	0.4706	14	2.3025	300	0.0537	300
AC15	14.8628	(3.47)	0.2	15.1730	116	15.1995	105	15.1689	54
AC16	14.8556	(3.44)	0.09	15.0012	24	14.9881	186	14.858	54
AC17	6.6124	(3.47)	1	6.6124	1	6.6373	129	6.6124	28
HE1	0.0737	(3.44)	0.7	0.1539	20	0.1807	300	0.1538	43
HE2	2.4181	(3.47)	10	4.4162	272	6.7846	177	3.8958	58
HE4	22.8382	(3.47)	3	22.8431	203	22.8713	252	22.8382	69
REA1	0.8617	(3.47)	0.2	0.8911	189	0.8815	96	0.8656	54
REA2	1.1341	(3.44)	1	1.1895	1	1.4188	300	1.149	45
REA3	74.2513	(3.47)	1	74.2513	4	74.5478	2	74.2513	24
DIS1	4.1593	(3.44)	5	4.5625	276	4.1943	93	4.1606	43
DIS3	1.0423	(3.47)	0.1	1.0933	150	1.1382	285	1.0624	117
DIS4	0.7315	(3.44)	0.1	0.7556	64	0.7498	126	0.7353	76
AGS	8.1732	(3.47)	1	8.1732	5	8.1732	24	8.1732	20
WEC2	3.5981	(3.47)	100	5.9166	128	6.6082	300	4.2483	95
WEC3	3.7685	(3.47)	100	6.2305	107	6.8402	300	4.4496	101
BDT1	0.2653	(3.47)	0.1	0.331	195	0.8562	29	0.2662	30
MFP	4.1865	(3.47)	300	31.5978	300	31.6079	171	31.5899	42
IH	1.26E-06	(3.47)	1	1.40E-05	1	1.1858	114	0.002	300
CSE1	0.02	(3.44)	1	0.02	1	0.022	3	0.02	34
PSM	0.9202	(3.44)	0.1	0.9206	15	0.9227	87	0.9202	18
EB1	3.1041	(3.44)	20	3.142	1	2.2076 (wrong)	300	3.1225	21
EB2	1.7676	(3.44)	1	2.0205	24	0.8148 (wrong)	84	2.0201	22
EB3	1.7976	(3.44)	1	2.058	26	0.8153 (wrong)	84	2.0575	22
NN1	13.1299	(3.47)	1	17.2732	4	18.4813	300	13.8474	45
NN2	1.7645	(3.44)	1	2.2217	27	2.2216	9	2.2216	20
NN8	2.3576	(3.47)	0.47	3.074	312	2.9345	180	2.8854	47
NN9	13.6461	(3.47)	40	30.0387	300	32.1222	300	28.6673	77
NN11	0.0181	(3.47)	50	0.1981	648	0.1566	9	0.0914	92
NN15	0.0977	(3.44)	1	0.0993	2	0.1194	6	0.0981	38
NN17	2.6386	(3.44)	15	11.2182	165	11.2381	117	11.2182	26

Table 3.6 Numerical results of static output feedback controllers by Algorithm 4 compared with [4]

Cases	Lower Bound	Fixed γ	Obj	# iter	γ in [4]	# iter	γ by systune [18]	# iter
HE3	0.7990	0.9200	(3.44)	257	0.9243	105	0.8052	70
DIS1	4.1593	4.1700	(3.47)	10	4.1943	93	4.1606	43
TG1	3.4652	12.8000	(3.44)	264	12.9336	45	12.8462	36
NN4	1.2862	1.3600	(3.44)	283	1.3802	156	1.3598	79
NN11	0.0181	0.1000	(3.44)	178	0.1566	9	0.0914	92
NN16	0.9556	0.9600	(3.44)	53	0.9656	48	0.9558	65

Chapter 4

Parameterized Bilinear Matrix Inequality Techniques in $\mathcal{H}_\infty$ Fuzzy PID Control Design

4.1 Introduction

Tagaki-Sugeno (T-S) fuzzy model [20] has been proved as one of the most practical tools for representing complex nonlinear systems by gain-scheduling systems, which are easily implemented online. Treating T-S fuzzy models as gain-scheduling systems allows the application of advanced gain-scheduling control techniques in tackling state feedback and output feedback stabilization of nonlinear systems [41, 42]. Until now, most of the gain-scheduling controllers are assumed structure-free and full-rank to admit computationally tractable parameterized linear matrix inequality (PLMI) or linear matrix inequality (LMI) formulations [43, 41, 42].

Meanwhile, proportional-integral-derivative (PID) structured controller is the indispensable component of industrial control so that PID control theory is still the subject of recent research [44–49], mainly concerning with linear time-invariant systems in the frequency domain. PID controller for fuzzy systems has been considered in [13]. Reference [12] proposed an LMI based iterative algorithm for a proportional-integral (PI) controller in T-S systems under the specific structure of both system and controller.

A recent work [14] transformed the fuzzy diagonal PID controller into a static output feedback problem with the dimension of controller dramatically increased. That is why all its testing examples are restricted on single input and single output systems with two states.

This chapter is concerned with the PID parallel distribution compensation (PDC) for T-S fuzzy models. The control design problem is formulated as a parameterized bilinear matrix inequality (PBMI) optimization problem that is in contrast to the PLMI formulation for the structure-free PDC design [41]. This is quite expected because the PID controller design for linear time-invariant systems is already nonconvex, which is equivalent to a BMI optimization problem in the state space. In our approach, PBMI is then relaxed to a bilinear matrix inequality (BMI) for more tractable computation. It should be noted that BMI optimization constitutes one of the most computational challenging problems, for which there is no efficient computational methodology. The state-of-the-art BMI solvers [3, 4] in addressing the structure-constrained stabilizing controllers for linear time-invariant systems must initialize from a feasible controller and then move within a convex feasibility subset containing this initialized point. Usually their convergence is very slow [4]. Furthermore, finding a feasible structure-constrained stabilizing controller is still a NP-hard problem [39]. The most efficient method to find such a feasible controller is via the so-called spectral abscissa optimization [17], which seeks a controller such that the state matrix of the closed loop system has only eigenvalues with negative real parts. This spectral abscissa optimization-based approach cannot be extended to gain-scheduling systems, whose stability does not quite depend on the spectrum of the time-varying state matrix. The main contribution of the present chapter is to develop efficient computational procedures for the BMI arisen from the PBMI optimization, which generate a sequence of unstabilizing controllers that rapidly converges to the optimal stabilizing controller.

The rest of this chapter is organized as follows. Section 4.2 is devoted to formulating the H_∞ fuzzy PID control in T-S system by a PLMI, which is then relaxed by a system of BMIs. Several nonconvex optimization techniques for addressing this BMI system are developed in Section 4.3. Simulation for benchmark systems is provided in Section 4.4 to support the solution development of the previous sections. Section 4.5 concludes the chapter.

Notation. Notation used in this chapter is standard. Particularly, $X \succeq 0$, $X \succ 0$, $X \preceq 0$ and $X \prec 0$ mean that a symmetric matrix X is positive semi-definite, positive definite, negative semi-definite and negative definite, respectively. Trace(X) represents the trace of X, while $||X||^2 = \text{Trace}(XX^T)$ is its square norm. In symmetric block matrices or long matrix expressions, we use $*$ as an ellipsis for terms that are induced by symmetry, e.g.,

$$K \begin{bmatrix} S + (*) & * \\ M & Q \end{bmatrix} * \equiv K \begin{bmatrix} S + S^T & M^T \\ M & Q \end{bmatrix} K^T.$$

All matrix variables are boldfaced. Denote by I_n the identity matrix of dimension $n \times n$ and by $0_{n\times m}$ the zero matrix of dimension $n \times m$. The subscript $n \times m$ is omitted when it is either not important or is clear in context.

4.2 H_∞ fuzzy PID PDS for T-S systems

Suppose that x is the state vector with dimension n_x, u is the control input with dimension n_u, y is the measurement output with dimension n_y, w and z are the disturbance and controlled output of the system with the same dimension n_∞, and L denotes the number of IF-THEN rules. In T-S fuzzy modeling, each $i-$th plant rule is the form

$$\begin{aligned} &\text{IF} \quad z_1(t) \text{ is } N_{i1} \text{ and } \ldots \ z_p(t) \text{ is } N_{ip} \\ &\text{THEN} \quad \begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix} = \begin{bmatrix} A_i & B_{1i} & B_{2i} \\ C_{1i} & D_{11i} & D_{12i} \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix}. \end{aligned} \tag{4.1}$$

Here z_i are premise variables, which are assumed independent of the control u, and N_{ij} are fuzzy sets. Denoting by $N_{ij}(z_i(t))$ the grade of membership of $z_i(t)$ in N_{ij}, the weight $w_i(t) = \prod_{j=1}^{p} N_{ij}(z_i(t))$ of each $i-$th IF-THEN rule is then normalized by

$$\begin{aligned} \alpha_i(t) &= \frac{w_i(t)}{\sum_{j=1}^{L} w_j(t)} \geq 0, \quad i = 1, 2, \ldots, L \\ \Rightarrow \alpha(t) &= (\alpha_1(t), \ldots, \alpha_L(t)) \in \Gamma, \end{aligned} \tag{4.2}$$

with

$$\Gamma := \{\alpha \in \mathbb{R}^L : \sum_{i=1}^{L} \alpha_i = 1, \alpha_i \geq 0\}. \tag{4.3}$$

In the state space, the T-S model is thus represented by the following gain-scheduling system

$$\begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix} = \begin{bmatrix} A(\alpha(t)) & B_1(\alpha(t)) & B_2(\alpha(t)) \\ C_1(\alpha(t)) & D_{11}(\alpha(t)) & D_{12}(\alpha(t)) \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix} \tag{4.4}$$

where

$$\begin{bmatrix} A(\alpha(t)) & B_1(\alpha(t)) & B_2(\alpha(t)) \\ C_1(\alpha(t)) & D_{11}(\alpha(t)) & D_{12}(\alpha(t)) \\ C_2 & D_{21} & 0 \end{bmatrix} = \sum_{i=1}^{L} \alpha_i(t) \begin{bmatrix} A_i & B_{1i} & B_{2i} \\ C_{1i} & D_{11i} & D_{12i} \\ C_2 & D_{21} & 0 \end{bmatrix}. \tag{4.5}$$

In this chapter, we seek the output feedback controller in the class of PID PDC with each i-th plant rule inferred by

$$\begin{aligned} &\text{IF} \quad z_1(t) \text{ is } N_{i1} \text{ and } \ldots \; z_p(t) \text{ is } N_{ip} \\ &\text{THEN} \quad \begin{bmatrix} \dot{x_K} \\ u \end{bmatrix} = \left[\begin{array}{cc|c} 0_{n_u \times n_u} & 0_{n_u \times n_u} & \mathbf{R}_{I,i} \\ 0_{n_u \times n_u} & -\tau I_{n_u} & \mathbf{R}_{D,i} \\ \hline I_{n_u} & I_{n_u} & \mathbf{R}_{P,i} \end{array}\right] \begin{bmatrix} x_K \\ y \end{bmatrix} \end{aligned} \tag{4.6}$$

for a given $\tau > 0$, where

$$\mathbf{R}_{x,j} \in \mathcal{R}^{n_u \times n_y}, x \in \{I, D, P\}.$$

Note that the transfer function of this i-th plant rule is

$$\mathbf{K}_i(s) = \mathbf{R}_{P,i} + \frac{\mathbf{R}_{I,i}}{s} + \frac{\mathbf{R}_{D,i}}{s+\tau} \tag{4.7}$$

$$= \mathbf{K}_{P,i} + \frac{\mathbf{K}_{I,i}}{s} + \frac{\mathbf{K}_{D,i} s}{1+\varepsilon s}, \tag{4.8}$$

with $\mathbf{K}_{P,i} := \mathbf{R}_{P,i} + \varepsilon\mathbf{R}_{D,i}$, $\mathbf{K}_{I,i} := \mathbf{R}_{I,i}$, $\mathbf{K}_{D,i} := -\varepsilon^2\mathbf{R}_{D,i}$, and $\varepsilon = 1/\tau$. It is clear from (4.8) that $\mathbf{K}_{P,i}$, $\mathbf{K}_{I,i}$ and $\mathbf{K}_{D,i}$ respectively are the proportional, integral and derivative gain matrices, while ϵ is a small tuning scalar which determines how close the last term in (4.6) comes to a pure derivative action [106]. In other words, (4.6) is the state-space representation of multi-input multi-output PID structured controllers.

The PID PDC with rule set by (4.6) and the membership function defined by (4.2) is thus represented by the following gain-scheduling PID controller

$$\begin{bmatrix} \dot{x_K} \\ u \end{bmatrix} = (\sum_{i=1}^{L} \alpha_i(t)\mathcal{K}_i(\mathbf{R}_i)) \begin{bmatrix} x_K \\ y \end{bmatrix} \tag{4.9}$$

where

$$\mathcal{K}_i(\mathbf{R}_i) = \left[\begin{array}{cc|c} 0_{n_u\times n_u} & 0_{n_u\times n_u} & \mathbf{R}_{I,i} \\ 0_{n_u\times n_u} & -\tau I_{n_u} & \mathbf{R}_{D,i} \\ \hline I_{n_u} & I_{n_u} & \mathbf{R}_{P,i} \end{array}\right], \quad i = 1, \ldots, L, \tag{4.10}$$

for

$$\mathbf{R}_i = \begin{bmatrix} \mathbf{R}_{I,i} \\ \mathbf{R}_{D,i} \\ \mathbf{R}_{P,i} \end{bmatrix} \in \mathbb{R}^{(3n_u)\times n_y}, i = 1, \ldots, L.$$

The $\mathcal{H}_\infty$ control problem consists of finding the stabilizing controller (4.9) for (4.4) to solve

$$\gamma \to \min : \gamma > 0, \tag{4.11a}$$

$$\int_0^T ||z(t)||^2 dt \leq \gamma^2 \int_0^T ||w(t)||^2 dt \tag{4.11b}$$

$$\forall w, \quad \forall T > 0, \quad x(0) = 0.$$

Using the shorthand

$$A_{0i} = \begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{(n_x+2n_u)\times(n_x+2n_u)},$$
$$B_{01i} = \begin{pmatrix} B_{1i} \\ 0 \end{pmatrix} \in \mathbb{R}^{(n_x+2n_u)\times n_\infty},$$
$$C_{01i} = \begin{pmatrix} C_{1i} & 0 \end{pmatrix} \in \mathbb{R}^{n_\infty\times(n_x+2n_u)},$$
$$\mathcal{B}_i = \begin{pmatrix} 0 & B_{2i} \\ I_{2n_u} & 0 \end{pmatrix} \in \mathbb{R}^{(n_x+2n_u)\times 3n_u},$$
$$\mathcal{C} = \begin{pmatrix} 0_{2n_u\times n_x} & I_{2n_u} \\ C_2 & 0_{n_y\times 2n_u} \end{pmatrix} \in \mathbb{R}^{(2n_u+n_y)\times(n_x+2n_u)},$$
$$\mathcal{D}_{12i} = \begin{pmatrix} 0 & D_{12i} \end{pmatrix} \in \mathbb{R}^{n_\infty\times 3n_u},$$
$$\mathcal{D}_{21} = \begin{pmatrix} 0 \\ D_{21} \end{pmatrix} \in \mathbb{R}^{(2n_u+n_y)\times n_\infty}, x_{cl} = \begin{pmatrix} x \\ x_K \end{pmatrix},$$

and then defining

$$\begin{bmatrix} A_0(\alpha) & \mathcal{B}(\alpha) \\ C_{01}(\alpha) & \mathcal{D}_{12}(\alpha) \end{bmatrix} := \sum_{i=1}^{L} \alpha_i \begin{bmatrix} A_{0i} & \mathcal{B}_i \\ C_{01i} & \mathcal{D}_{12i} \end{bmatrix},$$
$$\mathcal{K}(\alpha(t)) := \sum_{i=1}^{L} \alpha_i(t)\mathcal{K}_i(\mathbf{R}_i),$$
$$x_{cl} = (x^T, x_K^T)^T$$

the closed-loop system (4.4), (4.9) is rewritten by

$$\begin{bmatrix} \dot{x}_{cl} \\ z \end{bmatrix} = \left[\begin{array}{c} A_0(\alpha(t)) + \mathcal{B}(\alpha(t))\mathcal{K}(\alpha(t))\mathcal{C} \\ C_{01}(\alpha(t)) + \mathcal{D}_{12}(\alpha(t))\mathcal{K}(\alpha(t))\mathcal{C} \end{array}\right.$$
$$\left.\begin{array}{c} B_{01} + \mathcal{B}(\alpha(t))\mathcal{K}(\alpha(t))\mathcal{D}_{21} \\ D_{11}(\alpha(t)) + \mathcal{D}_{12}(\alpha(t))\mathcal{K}(\alpha(t))\mathcal{D}_{21} \end{array}\right] \begin{bmatrix} x_{cl} \\ w \end{bmatrix}. \tag{4.12}$$

Using the quadratic Lyapunov function $V(t) := x_{cl}^T(t)\mathbf{X}x_{cl}(t)$, $0 \prec \mathbf{X} \in \mathbb{R}^{(n_x+2n_u)\times(n_x+2n_u)}$ to make (4.11c) fulfilled by forcing

$$\dot{V}(t) + \gamma^{-1}||z(t)||^2 - \gamma||w(t)||^2 \leq 0,$$

one can easily see that (4.11c) is fulfilled by the following parameterized matrix inequality

$$\begin{bmatrix} (A_0(\alpha) + \mathcal{B}(\alpha)\mathcal{K}(\alpha)\mathcal{C})\mathbf{X} + (*) & * & * \\ (B_{01}(\alpha) + \mathcal{B}(\alpha)\mathcal{K}(\alpha)\mathcal{D}_{21})^T & -\gamma I & * \\ (C_{01}(\alpha) + \mathcal{D}_{12}(\alpha)\mathcal{K}(\alpha)\mathcal{C})\mathbf{X} & D_{11}(\alpha) + \mathcal{D}_{12}(\alpha)\mathcal{K}(\alpha)\mathcal{D}_{21} & -\gamma I \end{bmatrix} \prec 0, \tag{4.13a}$$

$$\mathbf{X} \succ 0, \quad \forall \alpha \in \Gamma, \tag{4.13b}$$

Set

$$\mathbf{W}_i := \mathcal{K}_i(\mathbf{R}_i)\mathcal{C}\mathbf{X}, i = 1, \ldots, L. \tag{4.14}$$

Then

$$\mathcal{K}(\alpha)\mathcal{C}\mathbf{X} = \sum_{i=1}^{L} \alpha_i \mathbf{W}_i.$$

For

$$\mathcal{M}_{ij}(\mathbf{X}, \mathbf{R}_j, \mathbf{W}_j, \gamma) := \begin{bmatrix} (A_{0i}\mathbf{X} + \mathcal{B}_i\mathbf{W}_j) + * & * & * \\ (B_{01i} + \mathcal{B}_i\mathcal{K}_j(\mathbf{R}_j)\mathcal{D}_{21})^T & -\gamma I & * \\ C_{01i}\mathbf{X} + \mathcal{D}_{12i}\mathbf{W}_j & D_{11i} + \mathcal{D}_{12i}\mathcal{K}_j(\mathbf{R}_j)\mathcal{D}_{21} & -\gamma I \end{bmatrix}, \tag{4.15}$$

which is linear in its variables, the parameterized matrix inequality (4.13a) is written by

$$\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i\alpha_j\mathcal{M}_{ij}(\mathbf{X},\mathbf{R}_j,\mathbf{W}_j,\gamma)\prec 0 \quad \forall\ \alpha\in\Gamma. \tag{4.16}$$

It follows from [41, Theorem 2.2] that (4.16) is guaranteed by the following matrix inequalities

$$\mathcal{M}_{ii}(\mathbf{X},\mathbf{R}_i,\mathbf{W}_i,\gamma)\prec 0, \quad i=1,2,\cdots,L \tag{4.17}$$

$$\frac{1}{L-1}\mathcal{M}_{ii}(\mathbf{X},\mathbf{R}_i,\mathbf{W}_i,\gamma)+\frac{1}{2}(\mathcal{M}_{ij}(\mathbf{X},\mathbf{R}_j,\mathbf{W}_j,\gamma) +\mathcal{M}_{ji}(\mathbf{X},\mathbf{R}_i,\mathbf{W}_i,\gamma))\prec 0, \quad 1\le i\neq j\le L. \tag{4.18}$$

Thus the upper bound of (4.11) is provided by the following optimization problem:

$$\min_{\gamma,\mathbf{X},\mathbf{R},\mathbf{W}}\ \gamma \quad \text{s.t.} \quad (4.13b),(4.14),(4.17),(4.18). \tag{4.19}$$

which is a BMI optimization in the decision variables $\mathbf{X}$, $\mathbf{R}=(\mathbf{R}_1,\ldots,\mathbf{R}_L)$ and $\mathbf{W}=(\mathbf{W}_1,\ldots,\mathbf{W}_L)$ due to the bilinear constraints (4.14).

We address this optimization problem through the following bisection procedure for a given computational tolerance $0<\eta<<1$.

Bisection procedure. Start from γ_u such that the BMI system

$$(4.13b),(4.14),(4.17),(4.18) \tag{4.20}$$

is feasible for $\gamma=\gamma_u$. Check the feasibility of BMI (4.20) for $\gamma=(1-\eta)\gamma_u$. If BMI (4.20) is feasible, reset $\gamma_u=\gamma$. Otherwise, reset $\gamma_l=\gamma$. Stop until $(\gamma_u-\gamma_l)/\gamma_u\le\eta$ and accept γ_u as the optimal H_∞ gain.

The next section is devoted to address the BMI feasibility problem (4.20). Its outcome is also a simple method to find an initial γ_u to start the above bisection procedure.

4.3 Nonconvex spectral optimization techniques for solving BMIs

The sparse structure of matrix $\mathcal{C}$ in (4.14) suggests that (4.14) is a sparse nonlinear constraint in the sense that there are not so many nonlinear terms in its right hand side. Indeed, by partitioning

$$\begin{gathered} 0 \prec \mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{X}_{13} \\ * & \mathbf{X}_{22} & \mathbf{X}_{23} \\ * & * & \mathbf{X}_{33} \end{bmatrix}, \\ \mathbf{X}_{11} \in \mathbb{R}^{n_x \times n_x}, \mathbf{X}_{1j} \in \mathbb{R}^{n_x \times n_u}, j = 1, 2; \\ \mathbf{X}_{22} \in \mathbb{R}^{n_u \times n_u}, \mathbf{X}_{23} \in \mathbb{R}^{n_u \times n_u}, \mathbf{X}_{33} \in \mathbb{R}^{n_u \times n_u} \end{gathered} \tag{4.21}$$

with $\mathbf{X}_{ii}$ symmetric, it can be checked that

$$\begin{aligned} \mathcal{K}_j(\mathbf{R}_j)\mathcal{C}\mathbf{X} \;=\; & \begin{bmatrix} \mathbf{R}_{I,j}C_2\mathbf{X}_{11} \\ -\tau\mathbf{X}_{13}^T + \mathbf{R}_{D,j}C_2\mathbf{X}_{11} \\ \mathbf{X}_{12}^T + \mathbf{X}_{13}^T + \mathbf{R}_{P,j}C_2X_{11} \end{bmatrix} \\ & \begin{bmatrix} \mathbf{R}_{I,j}C_2\mathbf{X}_{12} \\ -\tau\mathbf{X}_{23}^T + \mathbf{R}_{D,j}C_2\mathbf{X}_{12} \\ \mathbf{X}_{22} + \mathbf{X}_{23}^T + \mathbf{R}_{P,j}C_2\mathbf{X}_{12} \end{bmatrix} \\ & \begin{bmatrix} \mathbf{R}_{I,j}C_2\mathbf{X}_{13} \\ -\tau\mathbf{X}_{33} + \mathbf{R}_{D,j}C_2\mathbf{X}_{13} \\ \mathbf{X}_{23} + \mathbf{X}_{33} + \mathbf{R}_{P,j}C_2\mathbf{X}_{13} \end{bmatrix} \;= \\ & \begin{bmatrix} 0 & 0 & 0 \\ -\tau\mathbf{X}_{13}^T & -\tau\mathbf{X}_{23}^T & -\tau\mathbf{X}_{33} \\ \mathbf{X}_{12}^T + \mathbf{X}_{13}^T & \mathbf{X}_{22} + \mathbf{X}_{23}^T & \mathbf{X}_{23} + \mathbf{X}_{33} \end{bmatrix} \\ & + \mathbf{R}_j C_2 \mathbf{X}_1 \end{aligned} \tag{4.22}$$

for

$$\mathbf{X}_1 = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \mathbf{X}_{13} \end{bmatrix} \in \mathbb{R}^{n_x \times (n_x + 2n_u)}. \tag{4.23}$$

Therefore, the bilinear constraints (4.14) are expressed by the linear constraints

$$\mathbf{W}_j = \begin{bmatrix} 0 & 0 & 0 \\ -\tau \mathbf{X}_{13}^T & -\tau \mathbf{X}_{23}^T & -\tau \mathbf{X}_{33} \\ \mathbf{X}_{12}^T + \mathbf{X}_{13}^T & \mathbf{X}_{22} + \mathbf{X}_{23}^T & \mathbf{X}_{23} + \mathbf{X}_{33} \end{bmatrix} + \mathbf{Y}_j, \quad j = 1, \ldots, L \tag{4.24}$$

plus the bilinear constraints

$$\mathbf{Y}_j = \mathbf{R}_j C_2 \mathbf{X}_1, j = 1, \ldots . L. \tag{4.25}$$

In other words, the BMI feasibility problem (4.20) in $\mathbf{X}$, $\mathbf{R}$ and $\mathbf{W}$ is now equivalently transformed to the following BMI feasibility problem in $\mathbf{X}$, $\mathbf{R}$, $\mathbf{W}$ and $\mathbf{Y} := (\mathbf{Y}_1, \ldots, \mathbf{Y}_L)$:

$$(4.13b), (4.17), (4.18), (4.24), (4.25), \tag{4.26}$$

where (4.13b), (4.17) and (4.18) are linear matrix inequality (LMI) constraints, while (4.24) is linear constraints. The difficulty is now concentrated at L bilinear constraints in (4.25), in which only $\mathbf{X}_1$ is considered as a complicating variable that makes L constraints in (4.25) nonlinear. Based on this observation, our strategy is to decouple this complicating variable $\mathbf{X}_1$ from (4.25) for a better treatment. Let us recall an auxiliary result.

Lemma 1 *[107] For given matrix W_{12}, W_{22} of sizes $n \times m$ and $m \times m$ with $W_{22} \succeq 0$, one has*

$$\begin{pmatrix} 0 & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \succeq 0 \tag{4.27}$$

if and only if $W_{12} = 0$.

Using the above Lemma, we are now in a position to state the following result, which is a cornerstone in handling bilinear constraints like (4.25), which share a common complicating variable.

Theorem 1 *L bilinear constraints in (4.25) are equivalently expressed by the following L LMI constraints*

$$\begin{pmatrix} \mathbf{W}_{11,j} & \mathbf{Y}_j & \mathbf{R}_j \\ \mathbf{Y}_j^T & \mathbf{W}_{22} & \mathbf{X}_1^T C_2^T \\ \mathbf{R}_j^T & C_2\mathbf{X}_1 & I_{n_y} \end{pmatrix} \succeq 0, \quad j = 1, \ldots, L, \tag{4.28}$$

plus the single bilinear constraint

$$\mathbf{W}_{22} = \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1. \tag{4.29}$$

Proof. It can be easily seen that those $\mathbf{Y}_j$, $\mathbf{R}_j$ and $\mathbf{X}_1$ that are constrained by (4.25) together with $\mathbf{W}_{11,j} = \mathbf{R}_j(\mathbf{R}_j)^T$ and $\mathbf{W}_{22} = C_2\mathbf{X}_1(C_2\mathbf{X}_1)^T$ are feasible for (4.28) and (4.29), showing the implication (4.25)⇒(4.28) & (4.29).

On the other hand, by Shur's complement, it follows from (4.28) that

$$\begin{aligned} 0 \;\preceq\; & \begin{pmatrix} \mathbf{W}_{11,j} & \mathbf{Y}_j \\ \mathbf{Y}_j^T & \mathbf{W}_{22} \end{pmatrix} - \begin{pmatrix} \mathbf{R}_j \\ \mathbf{X}_1^T C_2^T \end{pmatrix} \begin{pmatrix} \mathbf{R}_j^T & C_2\mathbf{X}_1 \end{pmatrix} \\ = \; & \begin{pmatrix} \mathbf{W}_{11,k} & \mathbf{Y}_k - \mathbf{R}_k C_2 \mathbf{X}_1 \\ \mathbf{Y}_k^T - \mathbf{X}_1^T C_2^T \mathbf{R}_k^T & \mathbf{W}_{22} - \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1 \end{pmatrix} \\ = \; & \begin{pmatrix} \mathbf{W}_{11,k} & \mathbf{Y}_k - \mathbf{R}_k C_2 \mathbf{X}_1 \\ \mathbf{Y}_k^T - \mathbf{X}_1^T C_2^T \mathbf{R}_k^T & 0 \end{pmatrix}, \end{aligned} \tag{4.30}$$

where we also used (4.29) in obtaining the last equality (4.30). Then applying Lemma 1 yields (4.25), showing the implication (4.28) & (4.29)⇒(4.25).

□

Now, the problem's nonconvexity is concentrated on the single constraint (4.29) that involves only $\mathbf{X}_1$.

Theorem 2 *Under LMI constraints (4.28), the bilinear constraint (4.29) is equivalent to any from the two following constraints:*
(i) *The matrix rank constraint*

$$rank(\mathcal{Q}) = n_y, \tag{4.31}$$

for

$$\mathcal{Q} := \begin{pmatrix} \mathbf{W}_{22} & \mathbf{X}_1^T C_2^T \\ C_2\mathbf{X}_1 & I_{n_y} \end{pmatrix}; \tag{4.32}$$

(ii) *The quadratic constraint*

$$Trace(\mathbf{W}_{22}) = ||C_2\mathbf{X}_1||^2 \tag{4.33}$$

Proof. Note that (4.28) implies

$$\mathcal{Q} \succeq 0 \tag{4.34}$$

which also yields

$$\mathbf{W}_{22} \succeq \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1 \tag{4.35}$$

by Shur's complement. Also,

$$\begin{aligned} \text{rank}(\mathcal{Q}) &= \text{rank}(I_{n_y}) + \text{rank}(\mathbf{W}_{22} - \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1) \\ &= n_y + \text{rank}(\mathbf{W}_{22} - \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1), \end{aligned}$$

so (4.31) holds true if and only if $\text{rank}(\mathbf{W}_{22} - \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1) = 0$, which is (4.29). Next, it follows from (4.35) that (4.29) holds true if and only if

$$\begin{aligned} & \text{Trace}(\mathbf{W}_{22} - \mathbf{X}_1^T C_2^T C_2 \mathbf{X}_1) = 0 \\ \Leftrightarrow\ & \text{Trace}(\mathbf{W}_{22}) - ||C_2\mathbf{X}_1||^2 = 0 \\ \Leftrightarrow\ & (4.33). \end{aligned}$$

This completes the proof of Theorem 2.

□

The rank constraint (4.31) is discrete and absolutely intractable in general. However, under condition (4.34), this rank constraint is equivalent to the following continuous matrix-spectral constraint

$$\text{Trace}(\mathcal{Q}) - \lambda_{[n_y]}(\mathcal{Q}) = 0, \tag{4.36}$$

where $\lambda_{[n_y]}(\mathcal{Q})$ is the summation of the n_y largest eigenvalues of $\mathcal{Q}$. Indeed, $\text{rank}(\mathcal{Q}) \geq n_y$ but (4.36) means $\mathcal{Q}$ has at most n_y nonzero eigenvalues so its rank is n_y.
On the other hand, as

$$\text{Trace}(\mathcal{Q}) - \lambda_{[n_y]}(\mathcal{Q}) \geq 0,$$

it follows from (4.36) that

$$\text{Trace}(\mathcal{Q}) - \lambda_{[n_y]}(\mathcal{Q}) \tag{4.37}$$

can be used to measure the degree of satisfaction of the rank constraint (4.31). Instead of handling the nonconvex constraint (4.36) we incorporate it into the objective, resulting in the following alternative formulation to (4.26)

$$\min_{\mathbf{X},\mathbf{W},\mathbf{R},\mathbf{Y}} \; F(\mathcal{Q}) := \text{Trace}(\mathcal{Q}) - \lambda_{[n_y]}(\mathcal{Q}) \tag{4.38a}$$

$$\text{s.t.} \quad (4.13b), (4.17), (4.18), (4.24), (4.28). \tag{4.38b}$$

Suppose $X_1^{(\kappa)}$ and $W_{22}^{(\kappa)}$ are feasible for (4.38). Set

$$\mathcal{Q}^{(\kappa)} := \begin{pmatrix} W_{22}^{(\kappa)} & (X_1^{(\kappa)})^T C_2^T \\ C_2 X_1^{(\kappa)} & I_{n_y} \end{pmatrix}$$

Function $\lambda_{[n_y]}(\mathcal{Q})$ is nonsmooth but is lower bounded by the linear function

$$\sum_{i=1}^{n_y} (w_i^{(\kappa)})^T \mathcal{Q} w_i^{(\kappa)}, \tag{4.39}$$

where $w_1^{(\kappa)}, ..., w_{n_y}^{(\kappa)}$ are the normalized eigenvectors corresponding to n_y largest eigenvalues of $\mathcal{Q}^{(\kappa)}$. Thus, the following convex optimization problem provides an upper

bound for the nonconvex optimization problem (4.38),

$$\begin{aligned} \min_{\mathbf{X},\mathbf{W},\mathbf{R},\mathbf{Y}} \ & F^{(\kappa)}(\mathcal{Q}) := \mathrm{Trace}(\mathcal{Q}) - \sum_{i=1}^{n_y} (w_i^{(\kappa)})^T \mathcal{Q} w_i^{(\kappa)} \\ \text{s.t.} \quad & (4.38b). \end{aligned} \tag{4.40}$$

Suppose that $(X_1^{(\kappa+1)}, W_{22}^{(\kappa+1)})$ is the optimal solution of (4.40) and

$$\mathcal{Q}^{(\kappa+1)} := \begin{pmatrix} W_{22}^{(\kappa+1)} & (X_1^{(\kappa+1)})^T C_2^T \\ C_2 X_1^{(\kappa+1)} & I_{n_y} \end{pmatrix}.$$

Then

$$\begin{aligned} F(\mathcal{Q}^{(\kappa+1)}) &\leq F^{(\kappa)}(\mathcal{Q}^{(\kappa+1)}) \\ &< F^{(\kappa)}(\mathcal{Q}^{(\kappa)}) \\ &= F(\mathcal{Q}^{(\kappa)}), \end{aligned}$$

as far as $\mathcal{Q}^{(\kappa+1)} \neq \mathcal{Q}^{(\kappa)}$, implying that $\mathcal{Q}^{(\kappa+1)}$ is better than $\mathcal{Q}^{(\kappa)}$ towards optimizing (4.40). Similarly to [108], we establish the following result.

Proposition 2 *Initialized by any feasible point $\mathcal{Q}^{(0)}$ for the convex constraints (4.38b), $\{\mathcal{Q}^{(\kappa)}\}$ is a sequence of improved feasible points of the nonconvex optimization problem (4.38), which converges to a point satisfying the first-order necessary optimality conditions.*

In Algorithm 10 we propose a convex programming based computational procedure for the nonconvex optimization problem (4.38).

So far, in solving (4.38) we are based on (4.37) as the satisfaction degree of the rank constraint (4.31) and thus of the bilinear constraint (4.29). For larger value of n_y, Algorithm 10 may converge slowly. We now use

$$1 - \frac{||C_2 \mathbf{X}_1||^2}{\mathrm{Trace}(\mathbf{W}_{22})} \tag{4.43}$$

as an alternative degree for satisfaction of the bilinear constraint (4.29) because according to (4.35), (4.43) is positive and by (4.33), it is zero if and only if the bilinear constraint (4.29) is satisfied. Accordingly, instead of (4.38) we use the following

Algorithm 5 Nonconvex Spectral Optimization Algorithm for Solving BMI feasibility

1: **Initialization.** Set $\kappa := 0$ and solve the LMI (4.38b) to find a feasible point $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$. Given computational tolerance $\epsilon > 0$, stop the algorithm and accept $(X^{(0)}, W^{(0)}, R^{(0)}, Y^{(0)})$ as the solution of BMI (4.20) if

$$F(\mathcal{Q}^{(\kappa)}) \leq \epsilon. \tag{4.41}$$

2: **repeat**

3: Solve the convex optimization problem (4.40), to find the optimal solution $(X^{(\kappa+1)}, W^{(\kappa+1)}, R^{(\kappa+1)}, Y^{(\kappa+1)})$

4: Set $\kappa := \kappa + 1$.

5: **until**

$$\frac{F(\mathcal{Q}^{(\kappa-1)}) - F(\mathcal{Q}^{(\kappa)})}{F(\mathcal{Q}^{(\kappa-1)})} \leq \epsilon. \tag{4.42}$$

6: Accept $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$ as the solution of (4.38). Accept $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$ as the solution of BMI (4.20) if $F(\mathcal{Q}^{(\kappa)}) \leq \epsilon$. Otherwise declare that BMI (4.20) is infeasible.

optimization problem:

$$\min_{\mathbf{X},\mathbf{W},\mathbf{R},\mathbf{Y}} \; -\frac{||C_2\mathbf{X}_1||^2}{\text{Trace}(\mathbf{W}_{22})} \quad \text{s.t.} \quad (4.38b). \tag{4.44}$$

Note that function $g(\mathbf{X}_1, \mathbf{W}_{22}) := ||C_2\mathbf{X}_1||^2/\text{Trace}(\mathbf{W}_{22})$ is convex in $\mathbf{X}_1$ and $\mathbf{W}_{22} \succeq 0$ [89], so

$$\begin{aligned} g(\mathbf{X}_1, \mathbf{W}_{22}) \;&\geq\; g(X_1^{(\kappa)}, W_{22}^{(\kappa)}) + \langle \nabla g(X_1^{(\kappa)}, W_{22}^{(\kappa)}), \\ &\quad\; (\mathbf{X}_1, \mathbf{W}_{22}) - (X_1^{(\kappa)}, W_{22}^{(\kappa)}) \rangle \\ &=\; 2\frac{\text{Trace}((X_1^{(\kappa)})^T C_2^T C_2 \mathbf{X}_1)}{\text{Trace}(W_{22}^{(\kappa)})} \\ &\quad\; -\frac{||C_2 X_1^{(\kappa)}||^2 \text{Trace}(\mathbf{W}_{22})}{(\text{Trace}(W_{22}^{(\kappa)}))^2}. \end{aligned}$$

Thus, instead of (4.40), we solve the following convex optimization problem, which is an upper bound for the nonconvex optimization problem (4.44), to generate

Algorithm 6 Fractional Optimization Algorithm for Solving BMI feasibility

1: **Initialization.** Set $\kappa := 0$ and solve the LMI (4.38b) to find a feasible point $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$. Given computational tolerance $\epsilon > 0$, stop the algorithm and accept $(X^{(0)}, W^{(0)}, R^{(0)}, Y^{(0)})$ as the solution of BMI (4.20) if

$$1 - g(X_1^{(\kappa)}, W_{22}^{(\kappa)}) \leq \epsilon. \tag{4.46}$$

2: **repeat**

3: Solve the convex optimization problem (4.45), to find the optimal solution $(X^{(\kappa+1)}, W^{(\kappa+1)}, R^{(\kappa+1)}, Y^{(\kappa+1)})$

4: Set $\kappa := \kappa + 1$.

5: **until**

$$\frac{g(X_1^{(\kappa)}, W_{22}^{(\kappa)}) - g(X_1^{(\kappa-1)}, W_{22}^{(\kappa-1)})}{g(X_1^{(\kappa-1)}, W_{22}^{(\kappa-1)})} \leq \epsilon. \tag{4.47}$$

6: Accept $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$ as the solution of (4.38). Accept $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$ as the solution of BMI (4.20) if $1 - g(X_1^{(\kappa)}, W_{22}^{(\kappa)}) \geq \epsilon$. Otherwise declare that BMI (4.20) is infeasible.

$(X^{(\kappa+1)}, W^{(\kappa+1)}, R^{(\kappa+1)}, Y^{(\kappa+1)})$ at the κ-th iteration:

$$\begin{aligned} \min_{\mathbf{X},\mathbf{W},\mathbf{R},\mathbf{Y}} \quad & -2\frac{\text{Trace}((X_1^{(\kappa)})^T C_2^T C_2 \mathbf{X}_1)}{\text{Trace}(W_{22}^{(\kappa)})} \\ & + \frac{||C_2 X_1^{(\kappa)}||^2 \text{Trace}(\mathbf{W}_{22})}{(\text{Trace}(W_{22}^{(\kappa)}))^2} \\ \text{s.t.} \quad & (4.38b). \end{aligned} \tag{4.45}$$

A pseudo-code for the computational procedure, which is based on computation for (4.45) at each iteration, is described by Algorithm 11.

4.4 Simulation results

An important step is to check if there is a controller (4.6) to stabilize system (4.4). Define the block $(1,1)$ in (4.15) as

$$\widetilde{\mathcal{M}}_{ij}(\mathbf{X}, \mathbf{R}_j, \mathbf{W}_j) := (A_{0i}\mathbf{X} + \mathcal{B}_i \mathbf{W}_j) + (*).$$

Then the existing of a stabilizing controller (4.6) is guaranteed by the feasibility of the system consisting of (4.13b), (4.14) and

$$\widetilde{\mathcal{M}}_{ii}(\mathbf{X}, \mathbf{R}_i, \mathbf{W}_i) \prec 0, \quad i = 1, \ldots, L, \tag{4.48}$$

$$\frac{1}{L-1}\widetilde{\mathcal{M}}_{ii}(\mathbf{X}, \mathbf{R}_i, \mathbf{W}_i) + \frac{1}{2}(\widetilde{\mathcal{M}}_{ij}(\mathbf{X}, \mathbf{R}_j, \mathbf{W}_j) + \widetilde{\mathcal{M}}_{ji}(\mathbf{X}, \mathbf{R}_i, \mathbf{W}_i)) \prec 0, \quad 1 \leq i \neq j \leq L. \tag{4.49}$$

Thus, we can use Algorithm 10 or Algorithm 11 to check its feasibility, which invokes either the convex optimization problem

$$\begin{aligned} \min_{\mathbf{X},\mathbf{W},\mathbf{R},\mathbf{Y}} \ & \mathrm{Trace}(\mathcal{Q}) - \sum_{i=1}^{n_y} (w_i^{(\kappa)})^T \mathcal{Q} w_i^{(\kappa)} \\ \text{s.t.} \quad & (4.13b), (4.14), (4.48), (4.49), \end{aligned} \tag{4.50}$$

or the convex optimization problem

$$\begin{aligned} \min_{\mathbf{X},\mathbf{W},\mathbf{R},\mathbf{Y}} \ & -2\frac{\mathrm{Trace}((X_1^{(\kappa)})^T C_2^T C_2 \mathbf{X}_1)}{\mathrm{Trace}(W_{22}^{(\kappa)})} \\ & + \frac{||C_2 X_1^{(\kappa)}||^2 \mathrm{Trace}(\mathbf{W}_{22})}{(\mathrm{Trace}(W_{22}^{(\kappa)}))^2} \\ \text{s.t.} \quad & (4.13b), (4.14), (4.48), (4.49) \end{aligned} \tag{4.51}$$

instead of (4.40) or (4.45) at the κ-th iteration to generate the next iterative point $(X^{(\kappa+1)}, W^{(\kappa+1)}, R^{(\kappa+1)}, Y^{(\kappa+1)})$.

Whenever, a feasible point $(X^{(\kappa)}, W^{(\kappa)}, R^{(\kappa)}, Y^{(\kappa)})$ of (4.13b), (4.14), (4.48) and (4.49) is found, we solve the following convex optimization problem to determine the initial γ_u for the bisection procedure:

$$\begin{aligned} \min_{\gamma} \gamma \quad \text{s.t.} \quad & \mathcal{M}_{ii}(X^{(\kappa)}, R_i^{(\kappa)}, W_i^{(\kappa)}, \gamma) \prec 0, \\ & \frac{1}{L-1}\mathcal{M}_{ii}(X^{(\kappa)}, R_i^{(\kappa)}, W_i^{(\kappa)}, \gamma) \\ & + \frac{1}{2}(\mathcal{M}_{ij}(X^{(\kappa)}, R_j^{(\kappa)}, W_j^{(\kappa)}, \gamma) \\ & + \mathcal{M}_{ji}(X^{(\kappa)}, R_i^{(\kappa)}, W_i^{(\kappa)}, \gamma)) \prec 0, \\ & 1 \leq i \neq j \leq L. \end{aligned}$$

4.4.1 Inverted pendulum control

The motion of an inverted pendulum system with a point mass of mass $m = 2$ kg, a rigid rod of the length $\ell = 0.5m$ and a cart of mass $M = 8$ kg can be described by (4.5) [109] with $L = 2$ and

$$A_1 = \begin{bmatrix} 0 & 1 \\ 17.2941 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 12.6305 & 0 \end{bmatrix},$$
$$B_{11} = B_{12} = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}, \quad B_{21} = \begin{bmatrix} 0 \\ -0.1765 \end{bmatrix},$$
$$B_{22} = \begin{bmatrix} 0 \\ -0.0779 \end{bmatrix}, \quad C_{11} = C_{12} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 3 & 0 \end{bmatrix},$$
$$D_{11,i} \equiv 0.1, \quad D_{12,i} \equiv 0, \quad D_{21} = 0.$$

The system state is $x = (x_1, x_2)^T$, where x_1 is the angle measured from the inverted equilibrium position (angular position) and x_2 is the angular velocity. The membership functions in (4.2) are

$$\begin{aligned} \alpha_1(t) = \; & (1 - (1 + e^{(-7(x_1(t) - \pi/4))})^{-1}) \cdot \\ & (1 + e^{(-7(x_1(t) + \pi/4))})^{-1} \\ \alpha_2(t) = \; & 1 - \alpha_1(t), \quad x_1(t) \in [-\pi/3, \pi/3]. \end{aligned} \tag{4.52}$$

Based on the measured output $y = x_1(t)$ the task of the PID control is to minimize the effect of the disturbance in stabilizing the system. Therefore, the controlled output is set as $z = x_1 + x_2$.

In this example, $\tau = 6$ is set for (4.6). The minimal $\gamma = 0.12$ is obtained by using the bisection procedure. At $\gamma = 0.12$, Algorithm 10 needs 4 iterations to arrive the following numerical values for implementing PID PDC (4.6): $R_{P1} = 72.3777$, $R_{P2} = 99.2379$, $R_{I1} = 0.1449$, $R_{I2} = 0.1028$, $R_{D1} = 5.0864$ and $R_{D2} = 8.8573$. Figs. 4.1-4.2 respectively show the behavior of the system state and control with disturbance $w = 3\sin(5\pi t)$ and with no disturbance. The initial state is $x(0) = (\pi/4, -\pi/4)^T$. The obtained PID PDC stabilizes the inverted pendulum system well in the both scenarios. The system state motion and control load are very smooth compared with [14, Fig. 2]

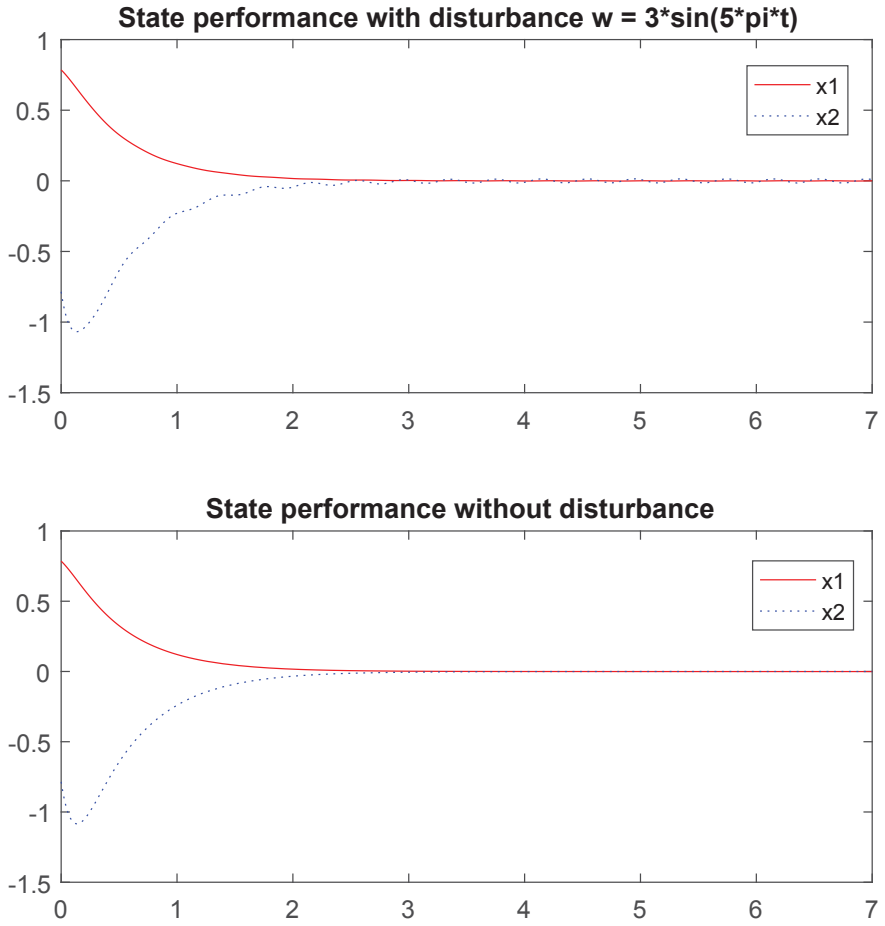

Fig. 4.1 The state behaviour with and without disturbance

Algorithm 11 achieves worse $\gamma = 0.13$ and needs 5 iterations for convergence for $\gamma = 0.13$. Fig.4.3 show the convergence behaviour of Algorithm 10 (for $\gamma = 0.12$) and Algorithm 11 (for $\gamma = 0.13$).

4.4.2 Duffing forced-oscillation

By [110], the Duffing forced-oscillation equation

$$\ddot{\mathbf{x}} + 0.2\dot{\mathbf{x}} + \mathbf{x}^3 - 10\cos t - u(t) = 0$$

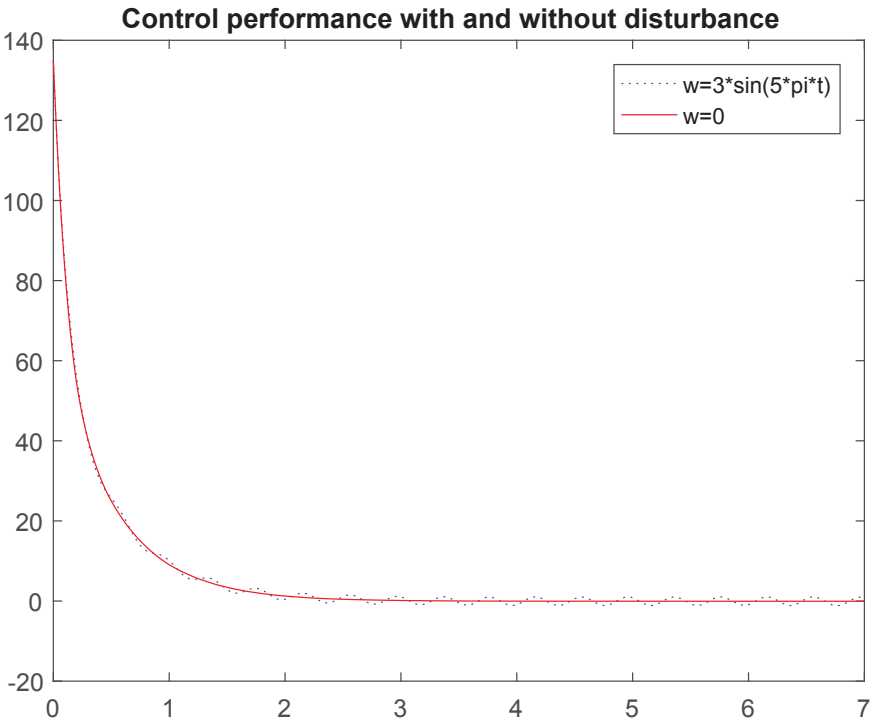

Fig. 4.2 PID PDC behaviour with and without the disturbance

with control input $u(t)$ and measured output $\mathbf{x}$ can be described by (4.5) with $L = 2$ and

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & -0.2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ -d^2 & -0.2 \end{bmatrix},$$
$$B_{11} = B_{12} = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}, \quad B_{21} = B_{22} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$
$$C_{11} = C_{12} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1 & 0 \end{bmatrix},$$
$$D_{11,i} \equiv 0.1, \quad D_{12,i} \equiv 0, \quad D_{21} = 0.$$

The membership functions in (4.2) are

$$\alpha_1(t) = 1 - \frac{x_1^2(t)}{d^2}, \alpha_2(t) = \frac{x_1^2(t)}{d^2}, x_1(t) \in [-d, d].$$

The system state is $x = (\mathbf{x}, \dot{\mathbf{x}})^T$ but only $\mathbf{x}$ is measurable so $y = \mathbf{x}$. The task is to minimize the effect of the disturbance $w(t)$ in stabilizing the system, so the controlled output is set as $z = \mathbf{x} + \dot{\mathbf{x}}$. The reader is also referred to [42, IV.B] for a different form

of fuzzy systems for this oscillation. Without the control input $u(t)$ the system state behaviour is chaotic as Fig. 4.4 shows. Since x_1 is always in the region $[-4, 4]$ we can set $d = 4$.

We set $\tau = 2$ for (4.6) in this example. The minimal $\gamma = 1.1$ is obtained by the bisection procedure. For this value of γ, Algorithm 10 need 10 iterations to arrive the following numerical values for implementing PID PDC (4.6): $R_{P1} = -96.8448$, $R_{P2} = 6.4360$, $R_{I1} = -1.4964$, $R_{I2} = -1.4984$, $R_{D1} = -0.7271$ and $R_{D2} = -0.0094$.

Fig. 4.5 represents the state plane with PID PDC. The initial state condition $x(0) = (0.1, 0)^T$. Fig. 4.6 depicts the behavior of the state and PID PDC. Again the PID PDC stabilizes the Duffing forced-oscillation system well.

Meanwhile Algorithm 11 achieves worse $\gamma = 1.4$ and needs 20 iterations for converge for this value of γ. Fig.4.7 shows the convergence behaviour of Algorithm 10 (for $\gamma = 1.1$) and Algorithm 11 (for $\gamma = 1.4$). Their convergence is dependent on initial points. Algorithm 10 converges not rapidly until seventh iteration, while Algorithm 10 converges rapidly after the first iteration.

4.4.3 TORA

By [111] and [41], the eccentric rotational proof mass actuator (TORA) system can be represented by T-S model (4.5) with

$$L = 4, \quad \alpha = 0.99, \quad \phi = 0.1, \quad c = 4,$$

$$B_{1i} \equiv 0, \quad D_{11i} \equiv 0, \quad D_{12i} \equiv 0, \quad D_{21i} \equiv 0,$$

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & \epsilon \sin(\alpha\pi)/(\alpha\pi) & 0 \\ 0 & 0 & 0 & 1 \\ -\phi/(1-\phi^2) & 0 & 0 & 0 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 2\phi/\pi & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & \phi & 0 \\ 0 & 0 & 0 & 1 \\ \phi/(1-\phi^2) & 0 & -\phi^2/(1-\phi^2) & 0 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & \phi & 0 \\ 0 & 0 & 0 & 1 \\ \phi/(1-\phi^2) & 0 & -\phi^2(1-c^2)/(1-\phi^2) & 0 \end{bmatrix},$$

$$B_{21} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/(1-\phi^2) \end{bmatrix}, B_{22} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

$$B_{23} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/(1-\phi^2) \end{bmatrix}, B_{24} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/(1-\phi^2) \end{bmatrix},$$

$$C_{1i} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

The membership functions in (4.2) are

$$\alpha_1(t) = \tfrac{x_1^2(t)}{a^2}, \quad \alpha_2(t) = \tfrac{1}{2} - \alpha_1(t),$$
$$\alpha_3(t) = \tfrac{b\sin(x_3(t)) - x_3(t)\sin(b)}{x_3(t)(b-\sin(b))}, \quad \alpha_4(t) = \tfrac{1}{2} - \alpha_3(t),$$

with $a = 0.8$, $b = 0.6$, and $x_1(t) \in [-a, a]$ and $x_3(t) \in [-b, b]$. The system state is $x = (x_1, x_2, x_3, x_4)$, where $x_3 = \theta$ and $x_4 = \dot{\theta}$ are the angular position and angular

velocity of the rotational proof mass, and $x_1 = \bar{x}_1 + \epsilon \sin x_3$, $x_2 = \bar{x}_2 + \epsilon x_4 \cos x_3$ with $\bar{x}_1 = q$ and $\bar{x}_2 = \dot{q}$ the translational position and velocity of the cart. In this application, only the translation position and angular position are measurable so $y = (x_1, x_3)^T$. The main task is to minimize the effect of the disturbance w in regulating the translation and angular positions to the equilibrium so the controlled output is set as $z = (x_1, x_3)^T$.

We set $\tau = 1$ for (4.6) in this example. The minimal $\gamma = 9.9$ is obtained by the bisection procedure. For this value of γ, Algorithm 11 needs 4 iterations to arrive the following numerical values for implementing PID PDC (4.6):

$$
\begin{aligned}
R_{P1} &= [-7.1101, -16.1981, 11.42817], \\
R_{P2} &= [-5.5390, -11.9724, 8.5207], \\
R_{P3} &= [-5.7119, -12.9499, 9.0553], \\
R_{P4} &= [-5.7189, -12.9240, 9.0397], \\
R_{I1} &= [-0.3471, -1.0139, 0.6820], \\
R_{I2} &= [-0.3450, -1.01858, 0.6830], \\
R_{I3} &= [-0.4091, -1.1184, 0.7669], \\
R_{I4} &= [-0.4337, -1.1552, 0.7969], \\
R_{D1} &= [0.8038, 2.0883, -1.2315], \\
R_{D2} &= [0.6084, 1.3740, -0.7847], \\
R_{D3} &= [0.5537, 1.4268, -0.7742], \\
R_{D4} &= [0.5048, 1.4486, -0.7513].
\end{aligned}
$$

Figs. 4.8-4.9 respectively show the behavior of system state and control with disturbance $w = 10sin(\pi t)$ and with no disturbance. The initial state condition is $x(0) = (0, 0, 0.5, 0)^T$. The TORA system is smoothly stabilized well by PID PDC.

Algorithm 10 achieves worse $\gamma = 10.3$ and needs 11 iterations for converge for this value of γ. Fig.4.3 shows the convergence behaviour of Algorithm 10 (for $\gamma = 10.3$ and Algorithm 11 (for $\gamma = 9.9$).

4.5 Conclusions

This chapter has addressed the problem of designing H_∞ PID PDC for T-S systems based on a parameterized bilinear matrix inequality (PLMI), which is a system of infinitely many bilinear matrix inequalities. Efficient computational procedures for this PLMI have been developed. Their merit has been analysed through the benchmark examples. In the end, the effectiveness of PID PDC in smoothly stabilizing nonlinear systems has been confirmed.

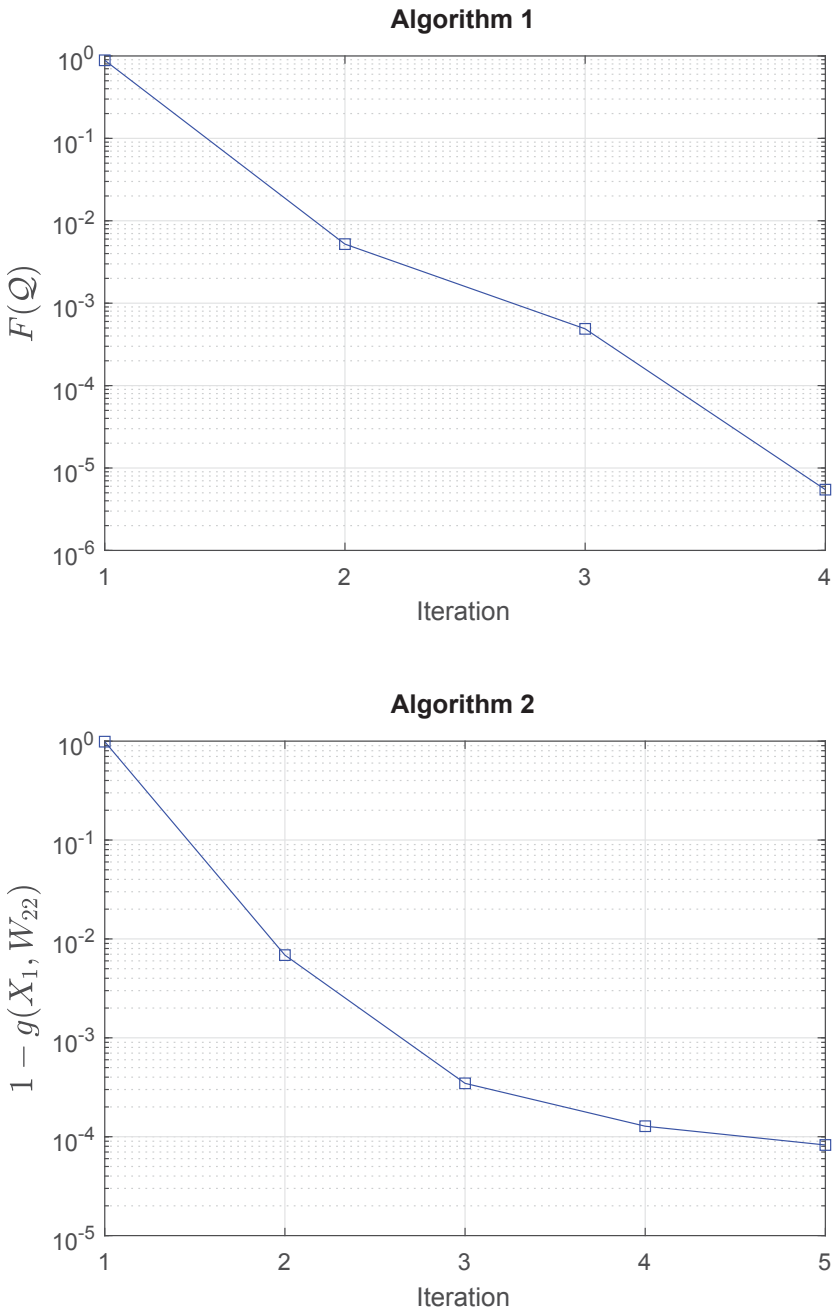

Fig. 4.3 Convergence performance by Algorithm 10 and Algorithm 11 for the inverted pendulum system

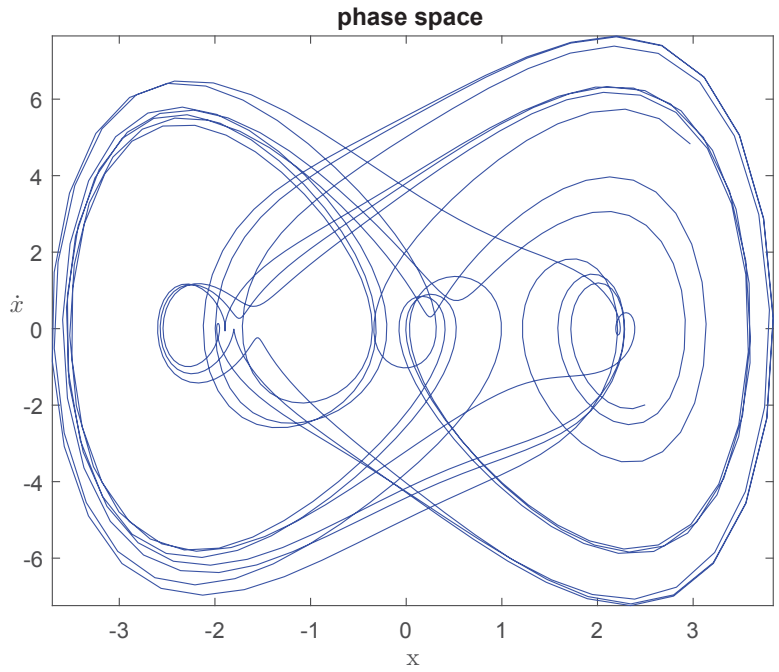

Fig. 4.4 The system state behavior without control

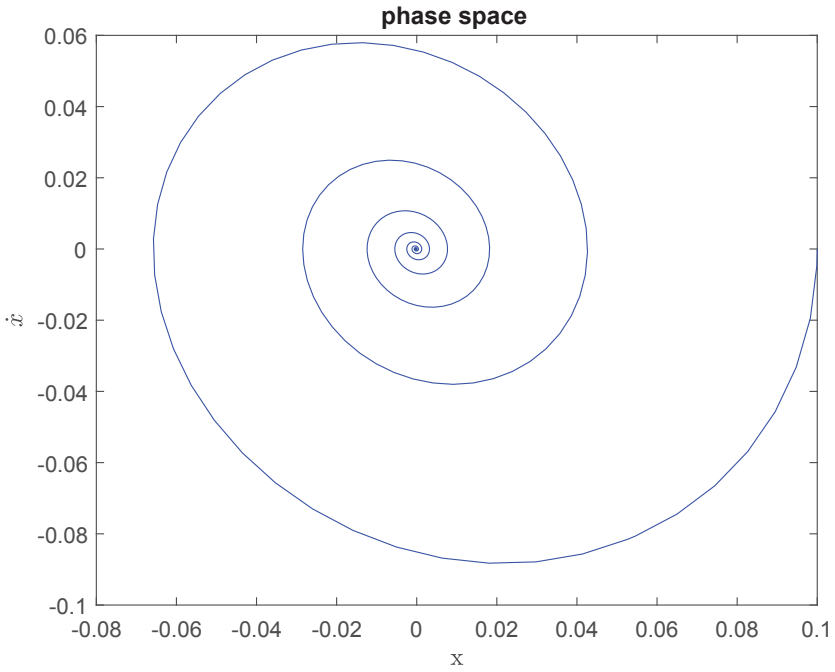

Fig. 4.5 The system state under PID PDC control

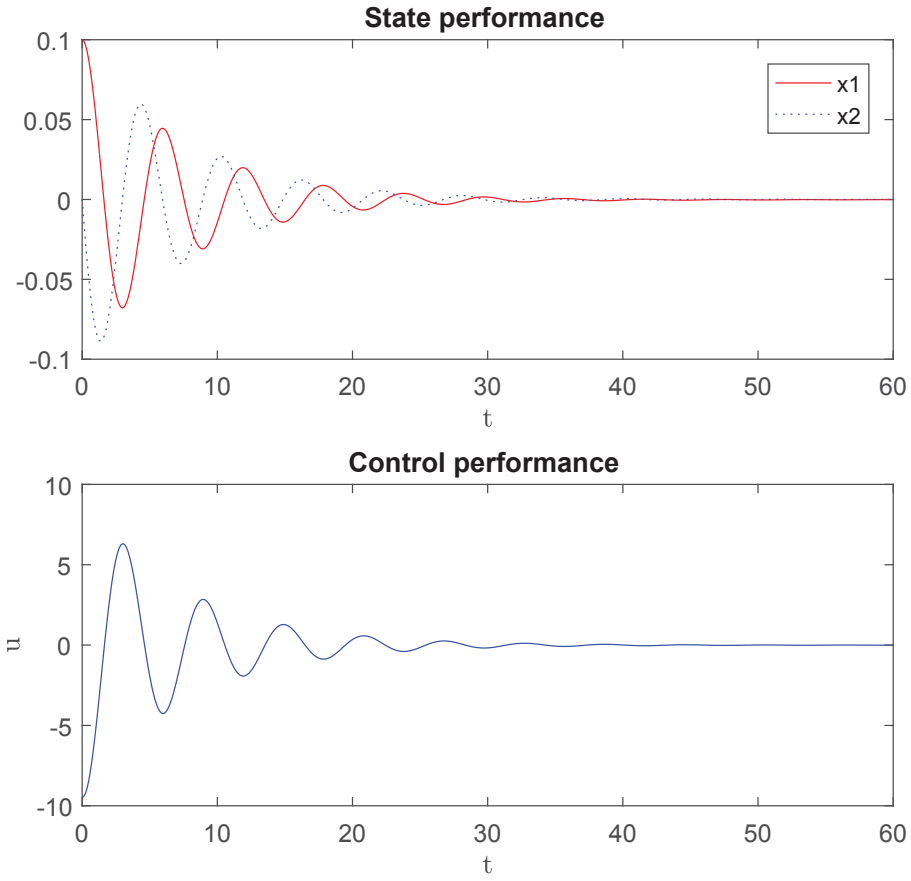

Fig. 4.6 The state and PID PDC behavior of the Duffing forced-oscillation system

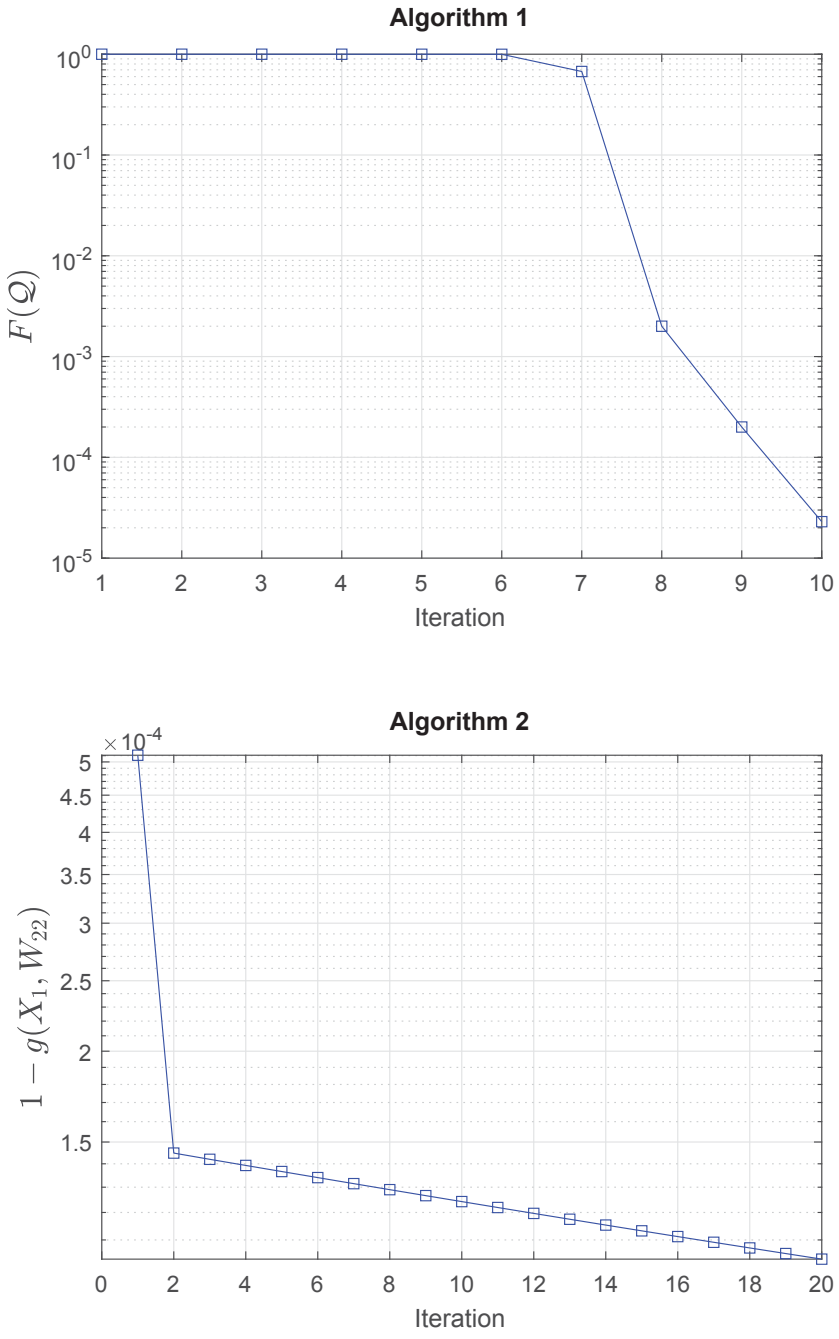

Fig. 4.7 Convergence performance by Algorithm 10 and Algorithm 11 for the Duffing forced-oscillation system

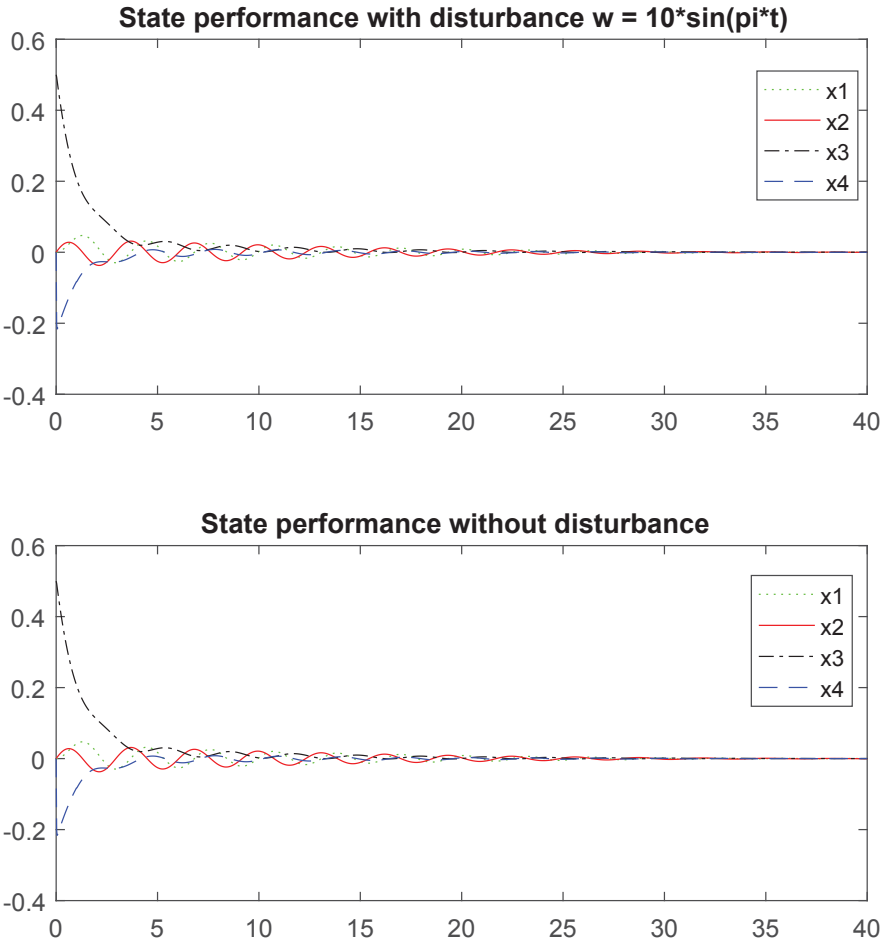

Fig. 4.8 The state behaviour with and without disturbance

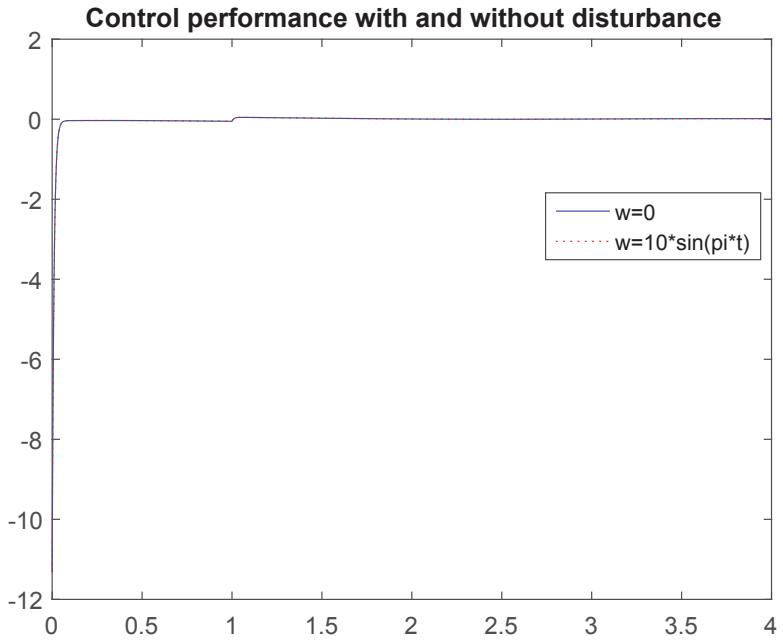

Fig. 4.9 The PID PDC behaviour with and without the disturbance

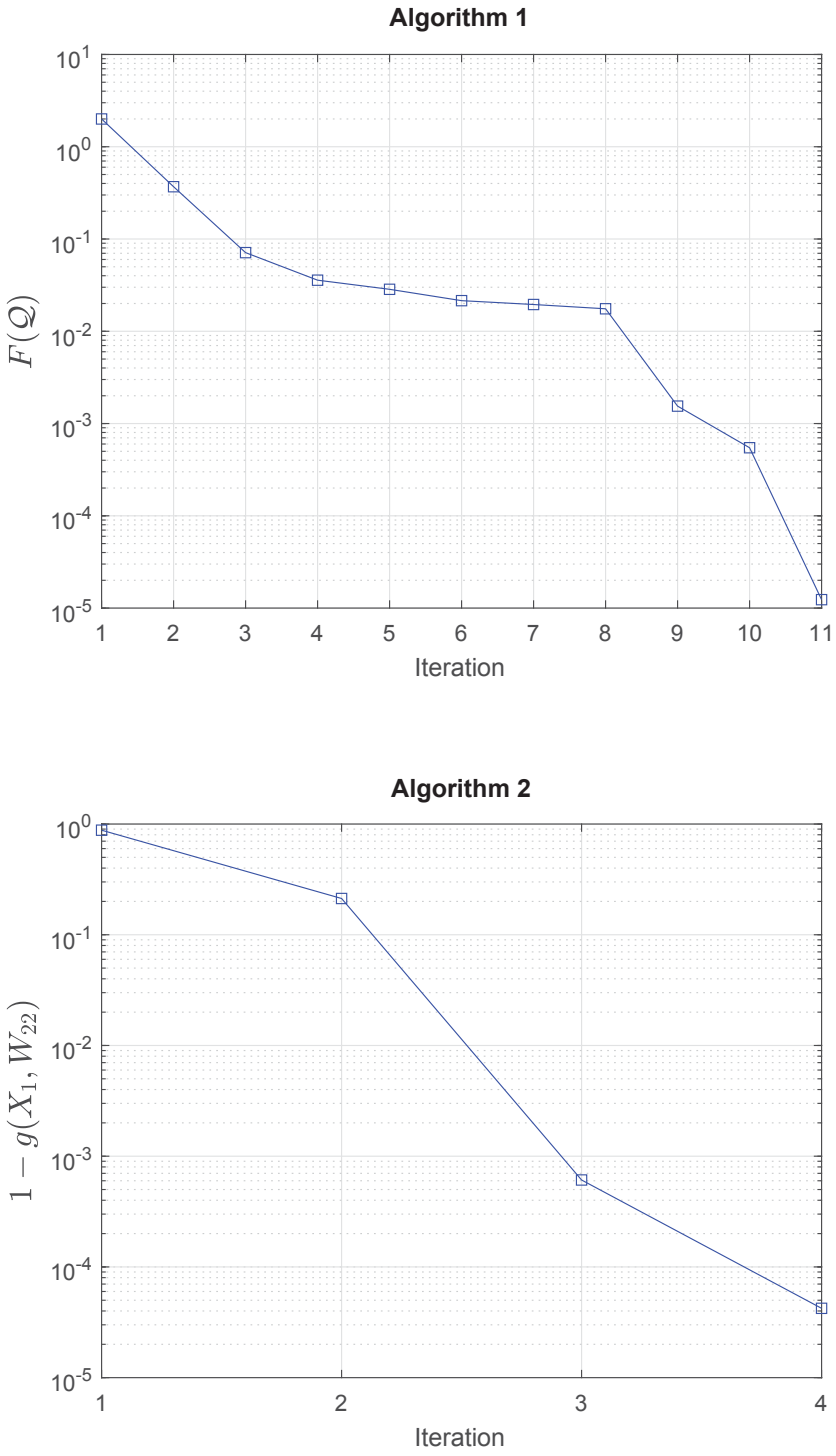

Fig. 4.10 Convergence performance by Algorithm 10 and Algorithm 11 for the TORA system

Chapter 5

Global Optimization for Optimal Power Flow over Transmission Networks

5.1 Introduction

Smart grids are operated by the advanced distribution management system (DMS), which is responsible for supervisory control and data acquisition in reactive dispatch, voltage regulation, contingency analysis, capability maximization and other smart operations. The optimal power flow (OPF) problem, which determines a steady state operating point that minimizes the cost of electric power generation or the transmission loss is the backbone of DMS (see e.g. [22–25] and references therein). Since its introduction by[22], OPF has received considerable interest (see e.g. [23–25] and references therein) but its solution remains largely open. As the basic quantities in power networks can be expressed in terms of the local bus voltages from Kirchhoff's voltage law, OPF can be represented by highly nonlinear optimization problems in voltage complex variables, whose NP-hard computational complexity has been particularly shown in [26]. The underlying difficulty of OPF lies on the multiple nonlinear constraints on the voltages variables due to the bus interconnections, hardware operating capacity and balance between power demand and supply. These nonlinear constraints are difficult

so the state-of-the-art nonlinear optimization solvers may converge to just stationary points (see [27] and references therein).

There has been recently renewed interest in the reformulation of the nonlinear constraints as convex constraints plus the nonconvex rank-one constraints on the matrix of outer product of voltage vector variables for the solution of OPF [50, 26]. For instance, by modifying numerical settings of some IEEE benchmark networks, [26] found that the matrix solution of the semi-definite relaxation (SDR) by dropping the rank-one matrix constraint, is of rank-one and hence the global solution of the OPF problem is found. Such SDR, dropping the rank-one matrix constraint, also provides the rank-one matrix solution in the so-called load over-satisfaction conditions, under which the quadratic equality constraints are essentially relaxed to loose one-sided inequality constraints. However, the point found from SDR is not necessarily feasible for the nonlinear equality constraints.

Indefinite quadratic programs find the most difficulty from the multiple quadratic equality constraints, which lead to nonzero duality gap and thus rank more than one of the matrix solution of SDR [51]. Another important contribution in SDR is provided in [7], where SDR has been shown to provide the rank-one matrix solution in the power distribution networks. When the networks are transmissions, [7] aims at a theoretical low-rank matrix solution of SDR, which could not lead to a feasible point for the original nonconvex OPF problem. The common drawback of the SDR approaches in [26, 7, 52] is that once a rank-one matrix solution is not found there is no way to find even a feasible point of the original OPF problem. It should also be noted that in many examples, SDRs might have the rank-one matrix solution among their multiple matrix solutions, however the SDP solvers can output a matrix solution of rank-more-than-one. The modification in [26] adds a small resistance for the lines of zero resistance, leading to the rank-one matrix solution of the SDR, is implicitly related to the features of its used software.

Meanwhile, [5] adopts the high-order SDR method of [53] to find the global solution of the OPF problem over power transmission networks. Theoretically, such method is able to generate a sequence of higher-order convex approximations, which converges to the original nonconvex problem in terms of the solution, regardless whether the former

is unique (as required in [53]) or not [54]. However, the dimension of these convex approximation problems grows dramatically in terms of decision variables, size and number of semi-definite constraints, making this approach suitable for only networks with a very small number of buses. Not surprisingly, [5] tests the performance of this method for networks with 2, 3 and 5 buses, where there are only 2, 3 and 5 bus voltages variables with the objective function linear in generation power. Another drawback of the high-order SDR method is that it works for real variables only so the dimension of the complex voltage variables become double for its utilization.

Power transmission networks in modern smart grids are often devised with a few thousand buses [55–57]. Under a such large number n of buses it is impossible to use the single matrix $W \in \mathbb{C}^{n\times n}$, which involves $n(n+1)/2 \approx O(10^7)$ complex variables. On the other hand, the number of the flow lines for bus connection is relatively moderate so only a small portion of the crossed nonlinear terms $V_k V_m^*$ appears in the nonlinear constraints. The common approach is to use the outer products of overlapped groups of the voltage variables to cover them [58, 59, 52]. All rank-one constraints on these outer products are then dropped for SDR. Obviously, the optimal solution of this SDR usually is not of rank-one and thus does not have any physical meaning. There is no technique to retrieve a feasible rank-one point from the rank-more-than-one solution of SDR. Multiple matrix rank constrained optimization has received a great attention due to its potential application in robust control synthesis [31, 3] but to our best knowledge there is no effective computation so far.

In this chapter, we firstly follow the approach of [108] to tackle this nonconvex OPF problem over power transmission networks with moderate size. Namely, the OPF problem, which is minimization of a nonlinear objective function over nonconvex constraints, is shown to be equivalent to a problem of minimization of d.c. (difference of convex functions) function [89] over convex constraints. Since this d.c. objective is nonsmooth (non-differentiable) we develop an iterative process of nonsmooth optimization to generate a sequence of improved points, which converges often to a solution of the OPF problem. Each iteration solves a semi-definite problem, whose dimension is moderate and unchanged during the whole process. Then to handle the large-scale OPF problems, an effective decomposition is proposed. It involves reduced numbers of the rank-one constraints on matrices of moderate size for expressing the network

nonlinear constraints. Accordingly, A new iterative procedure for rank-one constrained optimization, which is practical for computational solutions of large-scale indefinite quadratic programming. Our simulation demonstrates the computational efficiency of our approach.

The rest of this chapter is structured as follows. Section 5.2 is devoted to the problem formulation and its challenges. Its computational solution for OPF problem and large-scale OPF problem are respectively developed in Section 5.3 and Section 5.4. Section 5.5 provides simulation to show the efficiency of our methods. The conclusions are drawn in Section 5.6.

The notation used in this chapter is standard. More specifically, j denotes the imaginary unit, $M \succeq 0$ means that the Hermitian symmetric matrix M is positive semi-definite, $\mathsf{rank}(M)$ is the rank of the matrix M; $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of a complex quantity. $a \leq b$ for two complex numbers a and b is componentwise understood, i.e. $\Re(a) \leq \Re(b)$ and $\Im(a) \leq \Im(b)$. $\langle ., . \rangle$ is the dot product of matrices, while $\{A_i\}_{i=1,\ldots,n}$ denotes the matrix with diagonal blocks A_i and zero off-diagonal blocks.

5.2 Optimal power flow problem and challenges

Consider an AC electricity transmission network with a set of buses $\mathcal{N} := \{1, 2, \cdots, n\}$. The buses are connected through a set of flow lines $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$, i.e. bus m is connected to bus k if and only if $(m, k) \in \mathcal{L}$. Accordingly, $\mathcal{N}(k) := \{m \in \mathcal{N} \ : \ (m, k) \in \mathcal{L}\}$. The power demanded at bus $k \in \mathcal{N}$ is

$$S_{L_k} = P_{L_k} + jQ_{L_k},$$

where P_{L_k} and Q_{L_k} are the real and reactive power. A subset $\mathcal{G} \subseteq \mathcal{N}$ of buses is supposed to be connected to generators. Any bus $k \in \mathcal{N} \setminus \mathcal{G}$ is thus not connected to generators. Other notations are:

- $Y = [y_{km}]_{(k,m)\in\mathcal{N}\times\mathcal{N}} \in \mathbb{C}^{n\times n}$ is the admittance matrix [6]. Each y_{km} is the mutual admittance between bus k and bus m, so $y_{km} = y_{mk}\ \forall\ (k, m) \in \mathcal{L}$.

- V is the complex voltage vector, $V = [V_1, V_2, \cdots, V_n]^T \in \mathbb{C}^n$, where V_k is the complex voltage at bus $k \in \mathcal{N}$.
- I is the complex current vector, $I = YV = [I_1, I_2, \cdots, I_n]^T \in \mathbb{C}^n$, where I_k is the complex current injected to bus $k \in \mathcal{N}$.
- I_{km} is the complex current in the power line $(k, m) \in \mathcal{L}$, $\sum_{m \in \mathcal{N}(k)} I_{km} = I_k = \sum_{m \in \mathcal{N}(k)} y_{km} V_m$.
- $S_{km} = P_{km} + jQ_{km}$ is the complex power transferred from bus k to bus m, where P_{km} and Q_{km} represent the real and reactive transferred power.
- $S_{G_k} = P_{G_k} + jQ_{G_k}$ is the complex power injected by bus $k \in \mathcal{G}$, where P_{G_k} and Q_{G_k} represent the real and reactive generated power.

For each bus k, it is obvious that

$$\begin{aligned} S_{G_k} - S_{L_k} &= (P_{G_k} - P_{L_k}) + j(Q_{G_k} - Q_{L_k}) \\ &= V_k I_k^* \\ &= V_k \sum_{m \in \mathcal{N}(k)} I_{km}^* \\ &= V_k \sum_{m \in \mathcal{N}(k)} V_m^* y_{km}^*. \end{aligned} \tag{5.1}$$

Therefore, we can express the real generated power P_{G_k} and reactive generated power Q_{G_k} at bus k as the following nonconvex quadratic functions of the bus voltage vector variable $V := (V_1, V_2, ..., V_n)^T \in \mathbb{C}^n$,

$$\begin{aligned} P_{G_k} &= P_{L_k} + \Re\Big(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*\Big), \\ Q_{G_k} &= Q_{L_k} + \Im\Big(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*\Big). \end{aligned} \tag{5.2}$$

The objective of OPF is to minimize the following cost function of real active generated power P_G

$$f(P_G) = \sum_{k \in \mathcal{G}} (c_{k2} P_{G_k}^2 + c_{k1} P_{G_k} + c_{k0}), \tag{5.3}$$

where $c_{k2} > 0, c_{k1}, c_{k0}$ are given, which by (5.2) is a function of the bus voltages V:

$$\begin{aligned} f(V) &= \sum_{k\in\mathcal{G}}[c_{k2}(P_{L_k} + \Re(\sum_{m\in\mathcal{N}(k)} V_k V_m^* y_{km}^*))^2 \\ &\quad + c_{k1}(P_{L_k} + \Re(\sum_{m\in\mathcal{N}(k)} V_k V_m^* y_{km}^*)) + c_{k0}]. \end{aligned} \tag{5.4}$$

Accordingly, the following OPF problem is formulated

$$\min_{V\in\mathbb{C}^n} f(V) \quad \text{s.t.} \tag{5.5a}$$

$$-P_{L_k} - jQ_{L_k} = \sum_{m\in\mathcal{N}(k)} V_k V_m^* y_{km}^*, k \in \mathcal{N}\setminus\mathcal{G}, \tag{5.5b}$$

$$P_{G_k}^{min} \le P_{L_k} + \Re(\sum_{m\in\mathcal{N}(k)} V_k V_m^* y_{km}^*) \le P_{G_k}^{max}, k \in \mathcal{G}, \tag{5.5c}$$

$$Q_{G_k}^{min} \le Q_{L_k} + \Im(\sum_{m\in\mathcal{N}(k)} V_k V_m^* y_{km}^*) \le Q_{G_k}^{max}, k \in \mathcal{G}, \tag{5.5d}$$

$$V_k^{min} \le |V_k| \le V_k^{max}, k \in \mathcal{N}, \tag{5.5e}$$

$$|S_{km}| = |V_k V_m^* y_{km}^*| \le S_{km}^{max}, \forall (k,m) \in \mathcal{L} \tag{5.5f}$$

$$|V_k - V_m| \le V_{km}^{max}, (k,m) \in \mathcal{L}, \tag{5.5g}$$

$$|\arg(V_k) - \arg(V_m)| \le \theta_{km}^{max}, (k,m) \in \mathcal{L}, \tag{5.5h}$$

where

- (5.5b) is the equation of the balance between the demand and supply power at bus $k \in \mathcal{N}\setminus\mathcal{G}$;
- (5.5c)-(5.5d) are the power generation bounds, where $P_{G_k}^{min}$, $Q_{G_k}^{min}$ and $P_{G_k}^{max}$, $Q_{G_k}^{max}$ are the lower bound and upper bound of the real power reactive power generations, respectively;
- (5.5e) are the voltage amplitude bounds;
- (5.5f)-(5.5h) are capacity limitations, where the line currents between the connected buses are constrained by (5.5f), while (5.5g)-(5.5h) guarantee the voltage balance in terms of their magnitude and phases [6].

It is obvious that (5.5) is minimization of nonconvex objective function over quadratic equality constraints (5.5b) and (nonconvex) indefinite quadratic constraints (5.5c)-(5.5h).

Define the Hermitian symmetric matrix of outer product $W = VV^H \in \mathbb{C}^{n\times n}$, which must satisfy $W \succeq 0$ and $\mathsf{rank}(W) = 1$. By replacing $W_{km} = V_k V_m^*$, $(k,m) \in \mathcal{N} \times \mathcal{N}$, in (5.4) and (5.5b)-(5.5h), we reformulate (5.5) to the following optimization in matrix $W \in \mathbb{C}^{n\times n}$,

$$\min_{W\in\mathbb{C}^{n\times n}} \quad F(W) \quad \text{s.t.} \tag{5.6a}$$

$$-P_{L_k} - jQ_{L_k} = \sum_{m\in\mathcal{N}(k)} W_{km} y_{km}^* \, k \in \mathcal{N} \setminus \mathcal{G}, \tag{5.6b}$$

$$P_{G_k}^{min} \le P_{L_k} + \Re(\sum_{m\in\mathcal{N}(k)} W_{km} y_{km}^*) \le P_{G_k}^{max}, k \in \mathcal{G}, \tag{5.6c}$$

$$Q_{G_k}^{min} \le Q_{L_k} + \Im(\sum_{m\in\mathcal{N}(k)} W_{km} y_{km}^*) \le Q_{G_k}^{max}, k \in \mathcal{G}, \tag{5.6d}$$

$$(V_k^{min})^2 \le W_{kk} \le (V_k^{max})^2, k \in \mathcal{N}, \tag{5.6e}$$

$$|W_{km} y_{km}^*| \le S_{km}^{max}, (k,m) \in \mathcal{L}, \tag{5.6f}$$

$$W_{kk} + W_{mm} - W_{km} - W_{mk} \le (V_{km}^{max})^2, (k,m) \in \mathcal{L}, \tag{5.6g}$$

$$\Im(W_{km}) \le \Re(W_{km}) \tan\theta_{km}^{max}, (k,m) \in \mathcal{L}, \tag{5.6h}$$

$$W \succeq 0, \tag{5.6i}$$

$$\mathsf{rank}(W) = 1, \tag{5.6j}$$

where

$$\begin{aligned} F(W) \;=\; & \sum_{k\in\mathcal{G}} [c_{k2}(P_{L_k} + \Re(\sum_{m\in\mathcal{N}(k)} W_{km} y_{km}^*))^2 \\ & + c_{k1}(P_{L_k} + \Re(\sum_{m\in\mathcal{N}(k)} W_{km} y_{km}^*)) + c_{k0}], \end{aligned} \tag{5.7}$$

which is convex quadratic in W.

There is only one nonconvex constraint (5.6j) in (5.6) because other constraints are either linear (as (5.6b)-(5.6e), (5.6g)-(5.26b)) or convex quadratic (as (5.6f)). Also, by (5.26b)

$$\begin{bmatrix} W_{kk} & W_{km} \\ W_{mk} & W_{mm} \end{bmatrix} \succeq 0,$$

which implies

$$\begin{array}{rcl} W_{kk} + W_{mm} - W_{km} - W_{mk} & = & \\ \begin{bmatrix} 1 \\ -1 \end{bmatrix}^H \begin{bmatrix} W_{kk} & W_{km} \\ W_{mk} & W_{mm} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} & \geq & 0 \end{array}$$

so (5.26b), (5.6j) and (5.6g) are an equivalent expression to (5.5g).

SDR approach (see e.g. [26]) is to drop the only nonconvex constraint (5.6j) to have the SDP (5.6a)-(5.26b). If the solution of this relaxed SDP is of rank-one, i.e. it satisfies the nonconvex rank-one constraint (5.6j) then it obviously leads to the global solution of the nonconvex optimization problem (5.6). Actually, [26] modified some numerical setting for some IEEE benchmark networks so the corresponding relaxed SDP indeed have the rank-one matrix solution. However, it is known that (5.6a)-(5.26b) does not have the rank-one matrix solution in general, even for networks with a small number of buses. By relaxing the equality constraint (5.5b) by the one-sided inequality constraint

$$-P_{L_k} - jQ_{L_k} \leq \sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*, k \in \mathcal{N} \setminus \mathcal{G}, \tag{5.8}$$

which is expressed in terms of W as

$$-P_{L_k} - jQ_{L_k} \leq \sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*, k \in \mathcal{N} \setminus \mathcal{G}, \tag{5.9}$$

instead of (5.6b), [26] also found out that the SDP

$$\min_{W \in C^{n \times n}} \; F(W) \quad \text{s.t.} \quad (5.9), (5.6c) - (5.26b) \tag{5.10}$$

has the rank-one matrix solution, which however is not necessarily feasible for (5.6). Furthermore, [5] employed the high-order semi-definite relaxation [53] for (5.5), which leads to explosive growth of decision variables in the relaxed SDP. The first-order semi-definite relaxation involves the first-order moment matrix W of dimension $2n \times 2n$ with $\eta(n) := n(2n+1)$ entries[1], which are the decision variables, while linear inequality constraints (5.6c)-(5.26b) are replaced by semi-definite constraints in W. Each linear equality in (5.6b) is replaced by two semi-definite constraints in W. The second-

[1] W is obtained by replacing $w_{kn} = \bar{V}_k \bar{V}_\ell$ for $k = 0, 1, ..., 2n$ and $\ell = 0, 1, ..., 2n$ and $\bar{V} = (\Re\{V\}^T, \Im\{V\}^T)^T$ with setting $\Im\{V_n\} = 1$

order semi-definite relaxation involves the second-order moment matrix of dimension $\eta(n) \times \eta(n)$ with $\eta(n)(\eta(n)+1)/2$ entries and so on. The number of the decision variables and the size of semi-definite constraints on moment matrices are dramatically grown up, making this SDR approach quickly impractical.
In the next section we will provide an efficient computational method for the optimal solution of the nonconvex problem (5.6).

5.3 Nonsmooth optimization algorithm for OPF

Following [108], we address (5.6) as follows. It is obvious that a positive semi-definite matrix is of rank-one if and only if it has only one nonzero positive eigenvalue. Under the positive semi-definiteness condition (5.26b), the matrix rank-one constraint (5.6j) is thus equivalent to the spectral constraint

$$\text{Trace}(W) - \lambda_{\max}(W) = 0, \tag{5.11}$$

where $\lambda_{\max}(W)$ stands for the maximal eigenvalue of W. Then

$$W = \lambda_{\max}(W) w_{\max} w_{\max}^H,$$

where $w_{\max}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(W)$ of W. Accordingly, $V = \sqrt{\lambda_{\max}(W)} w_{\max}$ is feasible for the nonconvex constraints (5.5b)-(5.5h) in (5.5).
Therefore, (5.6) is equivalent to

$$\min_{W \in \mathbb{C}^{n \times n}} \; F(W) \quad \text{s.t.} \quad (5.6b) - (5.26b), (5.11). \tag{5.12}$$

Under the positive semi-definiteness condition (5.26b), the quantity $\text{Trace}(W) - \lambda_{\max}(W)$ is always nonnegative and can therefore be used to measure the degree of satisfaction of the matrix rank-one constraint (5.6j). Instead of handling nonconvex constraint (5.11) we incorporate it into the objective, resulting in the following alternative

formulation to (5.12)

$$\min_{W\in\mathbb{C}^{n\times n}} \; F_\mu(W) := F(W) + \mu(\text{Trace}(W) - \lambda_{\max}(W)) \quad \text{s.t.} \quad (5.6b) - (5.26b), \tag{5.13}$$

where $\mu > 0$ is a penalty parameter. Without squaring $\text{Trace}(W) - \lambda_{\max}(W)$, the above penalization is exact, meaning that the constraint $\text{Trace}(W) = \lambda_{\max}(W)$ can be satisfied by a minimizer of (5.13) with a finite value of μ (see e.g. [97, Chapter 16]). This is generally considered as a sufficiently nice property to make such exact penalization attractive. On the other hand, any feasible W for (5.12) is also feasible for (5.13), for which $F_\mu(W) = F(W)$, implying that the optimal value F_μ of (5.13) for any $\mu > 0$ is upper bounded by the optimal value F_{opt} of (5.12).

The objective function F_μ in (5.13) is the difference of two convex functions $g(W) := F(W) + \mu\text{Trace}(W)$ and $h(W) = \mu\lambda_{\max}(W)$, so (5.13) belongs to the class of d.c. (difference of convex functions) optimization [89]. Following [99, 112–114] we now adapt the d.c. iterations (DCI) for the solution of (5.13).
Function $\lambda_{\max}(W)$ is nonsmooth, which is lower bounded by

$$\lambda_{\max}(W) = \max_{||w||=1} w^H W w \geq (w_{\max}^{(\kappa)})^H W w_{\max}^{(\kappa)}, \tag{5.14}$$

where $w_{\max}^{(\kappa)}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(W^{(\kappa)})$, i.e. $\lambda_{\max}(W^{(\kappa)}) = (w_{\max}^{(\kappa)})^H W^{(\kappa)} w_{\max}^{(\kappa)}$.
Therefore, for any $W^{(\kappa)}$ feasible in the convex constraints (5.6b)-(5.26b), the following convex optimization problem provides an upper bound for the nonconvex optimization problem (5.13)

$$\min_{W\in C^{n\times n}} \; F^{(\kappa)}(W) := F(W) + \mu[\text{Trace}(W) - (w_{\max}^{(\kappa)})^H W w_{\max}^{(\kappa)}] \quad \text{s.t.} \quad (5.6b) - (5.26b) \tag{5.15}$$

because $F^{(\kappa)}(W) \geq F_\mu(W) \quad \forall\, W \succeq 0$ according to (5.29).
Suppose that $W^{(\kappa+1)}$ is the optimal solution of SDP (5.15). Since $W^{(\kappa)}$ is also feasible to (5.15) with $F_\mu(W^{(\kappa)}) = F^{(\kappa)}(W^{(\kappa)})$, it is true that

$$F_\mu(W^{(\kappa)}) = F^{(\kappa)}(W^{(\kappa)}) \geq F^{(\kappa)}(W^{(\kappa+1)}) \geq F_\mu(W^{(\kappa+1)}), \tag{5.16}$$

so $W^{(\kappa+1)}$ is a better feasible point for (5.13) than $W^{(\kappa)}$. Similarly to [112, 108, 113], the following result holds.

Proposition 3 *Initialized by any feasible point $W^{(0)}$ for SDP (5.6a)-(5.26b), $\{W^{(\kappa)}\}$ is a sequence of improved feasible points of the nonconvex optimization problem (5.13), which converges to a point satisfying first-order necessary optimality conditions.*

Proof: It follows from (5.6e) that the sequence $\{W^{(\kappa)}\}$ is bounded. By Cauchy theorem there is a subsequence $\{W^{(\kappa_\nu)}\}$, which converges to a point $\bar{W}$.
For every κ there is ν such that $\kappa_\nu \leq \kappa$ and $\kappa + 1 \leq \kappa_{\nu+1}$. By (5.16)

$$\begin{aligned} 0 &= \lim_{\nu\to\infty}[F_\mu(W^{(\kappa_\nu)}) - F_\mu(\bar{W})] \\ &\leq \lim_{\kappa\to\infty}[F_\mu(W^{(\kappa)}) - F_\mu(\bar{W})] \\ &\leq \lim_{\nu\to\infty}[F_\mu(W^{(\kappa_{\nu+1})}) - F_\mu(\bar{W})] \\ &= 0, \end{aligned}$$

yielding $\lim_{\kappa\to\infty} F_\mu(W^{(\kappa)}) = F_\mu(\bar{W})$, i.e the sequence $\{F_\mu(W^{(\kappa)})\}$ monotonically converges to the unique value $F_\mu(\bar{W})$. Furthermore, $\bar{W}$ is a solution of the optimization problem

$$\min_{W\in C^{n\times n}} \; F(W) + \mu[\mathrm{Trace}(W) - \bar{w}_{\max}^H W \bar{w}_{\max}] \quad \text{s.t.} \quad (5.6b) - (5.26b), \tag{5.17}$$

where $\bar{w}_{\max}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(\bar{W})$ of $\bar{W}$. Therefore,

$$F(W) - F(\bar{W}) + \mu[\mathrm{Trace}(W - \bar{W}) - \bar{w}_{\max}^H (W - \bar{W})\bar{w}_{\max}] \geq 0,$$

or equivalently,

$$g(W) - g(\bar{W}) - \langle \mu \bar{w}_{\max} \bar{w}_{\max}^H, W - \bar{W}\rangle \geq 0$$

for all feasible points W in (5.6b)-(5.26b). This means $\bar{W}$ is the optimal solution of the following convex program

$$\min_{W\in\mathbb{C}^{n\times n}} \; g(W) - \langle \mu \bar{w}_{\max} \bar{w}_{\max}^H, W\rangle \quad \text{s.t.} \quad (5.6b) - (5.26b).$$

Therefore,

$$\langle \nabla g(\bar{W}) - \mu \bar{w}_{\max} \bar{w}_{\max}^H, W - \bar{W} \rangle \geq 0$$

for all feasible points W for (5.6b)-(5.26b), which is also the first order necessary optimality condition for (5.13) because $\mu \bar{w}_{\max} \bar{w}_{\max}^H$ is a subgradient of h at $\bar{W}$ by (5.29).[2]

□

The nonsmooth optimization algorithm for the nonconvex optimization problem (5.6) is as follows.

Nonsmooth Optimization Algorithm (NOA).

Initialization. Solve SDP (5.6a)-(5.26b) to find its solution $W^{(0)}$. If $\mathsf{rank}(W^{(0)}) = 1$ stop: $W^{(0)}$ is the global solution of the nonconvex optimization problem (5.6a)-(5.6j). Otherwise determine μ such that $F(W^{(0)})$ and $\mu(\mathrm{Trace}(W^{(0)}) - \lambda_{\max}(W^{(0)}))$ are of similar magnitude and go to the next step.

Step 1. For $\kappa = 0, 1, \cdots$, solve (5.15) to find its solution $W^{(\kappa+1)}$. Stop if

$$\mathrm{Trace}(W^{(\kappa+1)}) - \lambda_{\max}(W^{(\kappa+1)}) \leq \epsilon \tag{5.18}$$

for given tolerance ϵ.

The reader is also referred to [108] for the practical convergence of the above algorithm to the solution of nonconvex optimization problem (5.6).

Remark 1. Considering NOA as an infeasible point algorithm for solving (5.6), it is observed in our simulation that a feasible point of found by NOA is almost optimal solution of (5.6). That's why the termination criterion (5.18) is set.

[2] the interested reader is also referred e.g. to [115, VI.5.1] for the full characterization for the set $\partial\lambda_{\max}(\bar{W})$ of all subgradients of $\lambda_{\max}(.)$ at $\bar{W}$

The above initialization also checks if SDR (5.6a)-(5.26b) admits the rank-one matrix solution. If it is not the case, one may initialize $W^{(0)}$ as the solution of SDP

$$\min_{W\in C^{n\times n}} \; F(W)+\mu\text{Trace}(W) \quad \text{s.t.} \quad (5.6b)-(5.26b), \tag{5.19}$$

instead of that by SDP (5.6a)-(5.26b).
On the other hand, we can stabilize the NOA by replacing the equality constraint (5.5b) by its computationally tolerant inequality

$$|S_{L_k}+\sum_{m\in\mathcal{N}(k)} V_kV_m^*y_{km}^*|^2\le\tilde{\epsilon}, \quad \forall\; k\in\mathcal{N}\setminus\mathcal{G} \tag{5.20}$$

for some given $\tilde{\epsilon}=10^{-4}$ or 10^{-6} for instance. Accordingly, the linear equality constraint (5.6b) is replaced by the convex quadratic constraint

$$|S_{L_k}+\sum_{m\in\mathcal{N}(k)} W_{km}y_{km}^*|^2\le\tilde{\epsilon}, \quad \forall\; k\in\mathcal{N}\setminus\mathcal{G}. \tag{5.21}$$

By doing this, we directly regulate the linear equality constraints to have stability feature of the applied SDP solvers as well. Accordingly, initialized from $W^{(0)}$ as the solution of SDP

$$\min_{W\in C^{n\times n}} \; F(W) \quad \text{s.t.} \quad (5.6c)-(5.26b), (5.21) \tag{5.22}$$

or SDP

$$\min_{W\in C^{n\times n}} \; F(W)+\mu\text{Trace}(W) \quad \text{s.t.} \quad (5.6c)-(5.26b), (5.21) \tag{5.23}$$

step 1 of NOA is to solve the following SDP instead of (5.15) to generate $W^{(\kappa+1)}$ for $\kappa=0,1,\cdots,$

$$\min_{W\in C^{n\times n}} \; F(W)+\mu[\text{Trace}(W)-(w_{\max}^{(\kappa)})^HWw_{\max}^{(\kappa)}] \quad \text{s.t.} \quad (5.6c)-(5.26b), (5.21). \tag{5.24}$$

5.4 Decomposed nonsmooth optimization for large-scale OPF

The first issue is to decompose large-size matrix W in (5.6) into matrices of smaller size to make the problem tractable even with limited computational power. This is also prompted by the fact that there is only a small portion of the crossed terms $V_k V_m^*$ appearing in the nonlinear constraints (5.5b)-(5.5h) so the large-size matrix variable $W \in \mathbb{C}^{n\times n}$ contains many redundant terms $V_k V_m^*$. The main result of [58, 59, 52] is to decompose the set $\mathcal{N} := \{1, 2, ..., n\}$ of buses into $\mathcal{I}$ overlapped subsets $\mathcal{N}_i = \{i_1, ..., i_{N_i}\}$ of buses, called bags, such that $i_\ell \in \mathcal{N}(i_{\ell+1}), \ell = 1, ..., N_i - 1 \quad \text{and} \quad i_{N_i} \in \mathcal{N}(i_1)$, for each $i = 1, 2, ..., \mathcal{I}$, i.e. the buses in the same bag are serially connected. The set of bags can be reset to make bags of relatively same size. Define the Hermitian symmetric matrix variables

$$W^i = [W_{i_k i_m}]_{k,m=1,..,N_i} \in \mathbb{C}^{N_i\times N_i}, i = 1, 2, ..., \mathcal{I}. \tag{5.25}$$

By replacing $W_{km} = V_k V_m^*$ in (5.6) we have the following equivalent reformulation for (5.6)

$$\min_{W=\mathsf{diag}\{W^i\}} F(W) \quad \text{s.t.} \quad (5.6b) - (5.6h), \tag{5.26a}$$

$$W^i \succeq 0, i = 1, ..., \mathcal{I}, \tag{5.26b}$$

$$\mathsf{rank}(W^i) = 1, i = 1, ..., \mathcal{I}. \tag{5.26c}$$

Reference [58, 59] just dropped all rank-one constraints in (5.26c) for SDR without any justification. Reference [52] also dropped all rank-constraints in (5.26c) but then used a penalized SDR for locating low-rank semi-definite matrices W^i in (5.25). Based on these low-rank matrices, [52, Sec. IV] also proposed to find rank-one matrices, which however are not necessarily feasible to (5.26).

The variable number in (5.26) is $\sum_{i=1}^{\mathcal{I}} N_i(N_i + 1)/2$. To keep this number reasonably moderate, it is desired that both $\mathcal{I}$ and N_i are sufficiently moderate. However, one can see that the above described decomposition [58, 59, 52] leads to a large number $\mathcal{I}$ of

bags as well as few large size N_i that result in many rank-one constraints in (5.26c), which are much less probably satisfied by solving SDR.

Our first step toward to computation of (5.5) is to develop a new decomposition with many fewer bags involved. Recalling that $\mathcal{N}(k)$ is the set of the buses that are connected to bus k, the cardinality $|\mathcal{N}(k)|$ is small in large-scale networks. We resort $\mathcal{N} = \{1, 2..., n\}$ as $\mathcal{N} = \{N_1, ...,$
$N_n\}$ such that the cardinality $|\mathcal{N}(N_k)|$ is in decreased order:

$$|\mathcal{N}(N_1)| \geq |\mathcal{N}(N_2)| \geq ... \geq |\mathcal{N}(N_n)|.$$

Accordingly, the first bag of buses is defined as $\mathcal{N}_1 = \mathcal{N}(N_1)$. The second bag is defined as

$$\mathcal{N}_2 = \{i \in \mathcal{N}(N_2) : \ \{i, N_2\} \not\subset \mathcal{N}_1\}.$$

Note that the crossed term $V_i V_{N_2}^*$ is already treated in the previous bag $\mathcal{N}_1$ whenever $\{i, N_2\} \subset \mathcal{N}_1$ so we exclude such bus i in defining bag $\mathcal{N}_2$.
Similarly, for $\ell \geq 3$ the ℓ-th bag is defined as

$$\mathcal{N}_\ell = \{i \in \mathcal{N}(N_\ell) : \{i, N_\ell\} \not\subset \mathcal{N}_{\ell'} \ \forall 1 \leq \ell' \leq \ell - 1\}$$

to exclude those buses i, whose crossed term $V_i V_{N_\ell}^*$ already is treated in a previous bag. As each $|\mathcal{N}_i|$ is obviously small, such decomposition is very efficient, leading to a substantial reduction of involved bags in comparison with that used in [58, 59, 52].

Our next step is to tackle the numerous difficult rank-one constraints in (5.26c), not dropping them for SDR as in all the previous works.
Firstly we express $\mathcal{I}$ rank-one constraints in (5.26c) by the following single spectral constraint

$$\sum_{i=1}^{\mathcal{I}} (\text{Trace}(W^i) - \lambda_{\max}(W^i)) = 0, \tag{5.27}$$

where $\lambda_{\max}(W^i)$ stands for the maximal eigenvalue of $W^{(i)}$. Indeed, (5.26b) implies $\text{Trace}(W^i) - \lambda_{\max}(W^i) \geq 0 \ \forall \ i$, so (5.27) means that $\text{Trace}(W^i) = \lambda_{\max}(W^i)$, i.e. W^i has only one nonzero eigenvalue so it is of rank-one. The nonnegative quantity $\sum_{i=1}^{\mathcal{I}} (\text{Trace}(W^i) - \lambda_{\max}(W^i))$ can therefore be used to measure the degree of

satisfaction of the rank-one constraints (5.27). Without squaring, the penalization $\sum_{i=1}^{\mathcal{I}}(\text{Trace}(W^i) - \lambda_{\max}(W^i))$ is exact, meaning that (5.27) can be satisfied by a minimizer of the problem

$$\min_{W=\mathsf{diag}\{W^i\}} \; F_\mu(W) := F(W) + \mu \sum_{i=1}^{\mathcal{I}}(\text{Trace}(W^i) - \lambda_{\max}(W^i)) \quad \text{s.t.} \quad (5.6b) - (5.6h), (5.26b), \tag{5.28}$$

with a finite value of $\mu > 0$ (see e.g. [97, Chapter 16]). This is generally considered as a sufficiently nice property to make such exact penalization attractive.
For any $W^{i,(\kappa)}$ feasible for the convex constraints (5.6b)-(5.6h), (5.26b), function $\lambda_{\max}(W^i)$ is nonsmooth and is lower bounded by

$$\lambda_{\max}(W^i) = \max_{||w||=1} w^H W^i w \geq (w^{i,(\kappa)}_{\max})^H W^i w^{i,(\kappa)}_{\max}, \tag{5.29}$$

where $w^{i,(\kappa)}_{\max}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(W^{i,(\kappa)})$, i.e.

$$\lambda_{\max}(W^{i,(\kappa)}) = (w^{i,(\kappa)}_{\max})^H W^{i,(\kappa)} w^{i,(\kappa)}_{\max}. \tag{5.30}$$

Accordingly, $\mu\lambda_{\max}(W^i) - \mu\lambda_{\max}(W^{i,(\kappa)}) \geq \mu((w^{i,(\kappa)}_{\max})^H W^i$
$.w^{i,(\kappa)}_{\max} - (w^{i,(\kappa)}_{\max})^H W^{i,(\kappa)} w^{i,(\kappa)}_{\max}) = \mu\langle w^{i,(\kappa)}_{\max}(w^{i,(\kappa)}_{\max})^H, W^i - W^{i,(\kappa)}\rangle$, so $\mu w^{i,(\kappa)}_{\max}(w^{i,(\kappa)}_{\max})^H$ is a subgradient of the function $\mu\lambda_{\max}(W^i)$ at $W^{i,(\kappa)}$. Then $\mu\mathsf{diag}\{w^{i,(\kappa)}_{\max}(w^{i,(\kappa)}_{\max})^H\}$ is a subgradient of the function $\mu\sum_{i=1}^{\mathcal{I}}\lambda_{\max}(W^i)$ at $\mathsf{diag}\{W^{i,(\kappa)}\}$.

The following SDP provides an upper bound for the nonconvex optimization problem(5.13)

$$\min_{W=\mathsf{diag}\{W^i\}} \; F^{(\kappa)}(W) := F(W) + \mu \sum_{i=1}^{\mathcal{I}}(\text{Trace}(W^i) - (w^{i,(\kappa)}_{\max})^H W^i w^{i,(\kappa)}_{\max}) \quad \text{s.t.} \quad (5.6b) - (5.6h), (5.26b) \tag{5.31}$$

because $F^{(\kappa)}(\mathsf{diag}\{W^i\}) \geq F_\mu(\mathsf{diag}\{W^i\}) \quad \forall\ W^i \succeq 0$ according to (5.29). Suppose that $W^{(\kappa+1)} = \mathsf{diag}\{W^{i,(\kappa+1)}\}$ is the optimal solution of SDP (5.15). Since $W^{(\kappa)} =$

$\mathsf{diag}\{W^{i,(\kappa)}\}$ is also feasible to (5.15) with $F_\mu(W^{(\kappa)}) = F^{(\kappa)}(W^{(\kappa)})$, it is true that

$$F_\mu(W^{(\kappa+1)}) \leq F^{(\kappa)}(W^{(\kappa+1)}) \leq F^{(\kappa)}(W^{(\kappa)}) = F_\mu(W^{(\kappa)}),$$

so $W^{(\kappa+1)}$ is a better feasible point of (5.13) than $W^{(\kappa)}$. Initialized by any feasible point $W^{(0)} = \mathsf{diag}\{W^{i,(0)}\}$ of SDP constraint (5.26a)-(5.26b), the sequence $\{W^{(\kappa)}\} = \{\mathsf{diag}\{W^{i,(\kappa)}\}\}$ with $W^{(\kappa+1)} = \mathsf{diag}\{W^{i,(\kappa+1)}\}$ iteratively generated as the optimal solution of SDP (5.15) is a sequence of improved feasible points of the nonconvex optimization problem (5.13). Since $W^{(\kappa)}$ are uniformly bounded, the sequence $\{W^{(\kappa)}\}$ has a limit point $\bar{W} = \mathsf{diag}\{\bar{W}^i\}$, which is the optimal solution of the optimization problem

$$\begin{aligned}\min_{W=\mathsf{diag}\{W^i\}} \; & F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\mathrm{Trace}(W^i) \\ -(\bar{w}^i_{\max})^H W^i \bar{w}^i_{\max}) \quad & \text{s.t.} \quad (5.6b) - (5.6h), (5.26b), \end{aligned} \tag{5.32}$$

where $\bar{w}^i_{\max}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(\bar{W}^i)$ of $\bar{W}^i$. Particularly,

$$\begin{aligned} & F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\mathrm{Trace}(W^i) - (\bar{w}^i_{\max})^H W^i \bar{w}^i_{\max}) \quad \geq \\ & F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\mathrm{Trace}(\bar{W}^i) - (\bar{w}^i_{\max})^H \bar{W}^i \bar{w}^i_{\max}), \end{aligned}$$

or equivalently, under the definition $g(W) = F(W) + \mu \sum_{i=1}^{\mathcal{I}} \mathrm{Trace}(W^i)$,

$$g(W) - g(\bar{W}) - \langle \mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}, W - \bar{W} \rangle \geq 0$$

for all feasible points $W = \mathsf{diag}\{W^i\}$ in (5.6b)-(5.6h), (5.26b). As a result, $\bar{W}$ is the optimal solution of the convex optimization problem

$$\begin{aligned}\min_{W=\mathsf{diag}\{W^i\}} \; & g(W) - \langle \mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}, W - \bar{W} \rangle \\ & \text{s.t.} \quad (5.6b) - (5.6h), (5.26b), \end{aligned}$$

so it must satisfy the optimality condition

$$\langle \nabla g(\bar{W}) - \mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}, W - \bar{W}\rangle \geq 0$$

for all feasible points $\mathsf{diag}\{W^i\}$ in (5.6b)-(5.6h), (5.26b). The latter is also the first order necessary optimality condition for (5.13) because $\mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}$ is a subgradient of the function $\mu \sum_{i=1}^{\mathcal{I}} \lambda_{\max}(W^i)$ at $\bar{W}$. As our simulations will show, $\bar{W}$ is indeed the global optimal solution of (5.13) and (5.26).

However, unlike [92] with only a single rank-one constrained matrix, although quantity

$$\sum_{i=1}^{\mathcal{I}} (\mathrm{Trace}(W^{i,(\kappa)}) - \lambda_{\max}(W^{i,(\kappa)})) \tag{5.33}$$

in (5.13) is iteratively decreased, not all individual quantities

$$\mathrm{Trace}(W^{i,(\kappa)}) - \lambda_{\max}(W^{i,(\kappa)}) \tag{5.34}$$

are iteratively decreased so the rank of each matrix $W^{i,(\kappa)}$ is no longer iteratively reduced to one as expected. Worse, $W^{i,(\kappa)}$ is rank-one but the rank of $W^{i,(\kappa+1)}$ in the next iteration may turn to be more than one with

$$\begin{aligned} &\mathrm{Trace}(W^{i,(\kappa+1)}) - \lambda_{\max}(W^{i,(\kappa+1)}) \quad > \\ &\mathrm{Trace}(W^{i,(\kappa)}) - \lambda_{\max}(W^{i,(\kappa)}). \end{aligned}$$

Consequently, it is very difficult to achieve rank-one for all $W^{i,(\kappa)}$ as desired. It is also impossible to add a "weight" to each term under the sum in the objective in (5.15) to handle the individual convergence of $\mathrm{Trace}(W^i) - \lambda_{\max}(W^i)$.
We now develop a systematic way to resolve this issue as follows. For $\kappa = 0, 1, ...,$ and $W^{(\kappa)} = \mathsf{diag}\{W^{(i,(\kappa))}\}$ define

$$\mathcal{L}^{(\kappa)} = \{i \in \{1, ..., \mathcal{I}\} \ : \ \mathsf{rank}(W^{i,(\kappa)}) = 1\} \tag{5.35}$$

and generate $W^{(\kappa+1)} = \mathsf{diag}\{W^{i,(\kappa+1)}\}$ as the optimal solution of the following SDP instead of SDP (5.15)

$$\min_{W=\mathsf{diag}\{W^i\}} \; F(W) + \mu\sum_{i=1}^{\mathcal{I}}[\mathrm{Trace}(W^i) - (w_{\max}^{i,(\kappa)})^H W^i w_{\max}^{i,(\kappa)}] \quad \text{s.t.} \quad (5.6b)-(5.6h), (5.26b), \tag{5.36a}$$

$$\mathrm{Trace}(W^i) - (w_{\max}^{i,(\kappa)})^H W^i w_{\max}^{i,(\kappa)} \le \epsilon_{tol}, i \in \mathcal{L}^{(\kappa)}. \tag{5.36b}$$

Note that $\mathrm{Trace}(W^i) \ge w^H W w$ for all $||w|| = 1$ and it is obvious that $\mathsf{rank}(W^i) = 1$ if and only if $\mathrm{Trace}(W^i) - w_{max}^H W^i w_{\max} = 0$ for some normalized $w_{\max}$. Therefore, the constraint (5.36b) for some tolerance ϵ_{tol} is introduced to warrant the rank-one of all $W^{i,(\kappa+1)}$, $i \in \mathcal{L}^{(\kappa)}$. As a result $\mathcal{L}^{(\kappa)} \subset \mathcal{L}^{(\kappa+1)}$ and $\mathcal{L}^{(\kappa)} \to \{1, ..., \mathcal{I}\}$ is expected to have all $W^{i,(\kappa)}$ of rank-one. Unlike (5.15), the iterations (5.36) leads to achieving rank-one of all W^i while the objective function F_μ is still decreased.

In summary, we propose the following *Large-Scale Non-smooth Optimization Algorithm (Large-scale NOA)* for the multiple rank-one constrained optimization problem (5.13).

Initialization. Solve SDP

$$\min_{W=\mathsf{diag}\{W^i\}} \; F(W) \quad \text{s.t.} \quad (5.6b)-(5.6h), (5.26b) \tag{5.37}$$

to generate $W^{(0)} := \mathsf{diag}\{W^{i,(0)}\}$. If $\mathsf{rank}(W^{i,(0)}) \equiv 1$ stop: $W^{(0)}$ is the global solution of the nonconvex optimization problem (5.26). Otherwise set $\kappa = 0$ and define $\mathcal{L}^{(\kappa)}$ by (5.35).

κ-th iteration. For $\kappa = 0, 1, ..$, solve (5.36) to generate $W^{(\kappa+1)} := \mathsf{diag}\{W^{i,(\kappa+1)}\}$. Reset $\kappa = \kappa + 1$ and define $\mathcal{L}^{(\kappa)}$ by (5.35). Stop whenever $\mathcal{L}^{(\kappa)} = \{1, ..., \mathcal{I}\}$. Otherwise go to the next iteration.

5.5 Simulation results

5.5.1 Simulation results for OPF problem

The hardware and software facilities for our computational implementation are:

- Processor: Intel(R) Core(TM) i5-3470 CPU @3.20GHz;
- Software: Matlab version R2013b;
- Matlab toolbox: Matpower version 5.1[6] to compute the admittance matrix Y from the power system data; Yalmip[101] with SeDumi 1.3 [102] solver for SDP (5.15).

The tolerance $\epsilon = 10^{-5}$ is set for the stop criterion (5.18) of NOA, which is applied to solutions of all cases, except the IEEE300 case, for which the above description adaptation of NOA is used. We also provide the value of μ used in each simulation so the reader can easily reproduce our numerical result.[3]

WB2mod Network [1]

WB2mod is a power network with constraints modified from WB2, in which there are $n = 2$ buses with 1 generator, 1 load bus and 1 transmission line, which lead to only one nonlinear equality constraint in (5.5b). The size of the Hermitian symmetric matrix variable W is 2×2. The SDP (5.6a)-(5.26b) of SDR does not admit the rank-one solution with V_2^{max} in (5.6d) exceeding 0.976 [5].
The computational results are summarized in Table 5.1, where the first column is the different values of V_2^{max} (unit p.u.), the second column is the corresponding optimal value (unit \$/h) of SDP (5.6a)-(5.6h), which provides a lower bound (LB) for (5.6a)-(5.26b). The solution of SDP (5.6a)-(5.6h) is not of rank-one as the third column lists all its nonzero eigenvalues. The fourth column indicates the value of the penalty parameter μ in (5.13), while the fifth column provides the iteration number of NOA for solution of (5.13). The sixth column is the optimal value (unit \$/h) of

[3]The code for each simulation is provided upon request to the corresponding author H.D. Tuan

Table 5.1 WB2mod [5]

V_2^{max}	SDR LB	SDR's nonzero eig.	μ	# Iter.	NOA	$\lambda_{\max}$
0.983	903.12	0.0008,1.8680	10^4	2	905.73	1.8552
0.989	900.84	0.0012,1.8794	10^4	2	905.73	1.8552
0.996	898.17	0.0016,1.8929	10^4	2	905.73	1.8552
1.002	895.86	0.0018,1.9047	3×10^4	2	905.73	1.8552
1.009	893.16	0.0018,1.9188	2×10^2	1	905.73	1.8552
1.015	890.82	0.0016,1.9311	2×10^2	1	905.73	1.8552
1.022	888.08	0.0012,1.9458	2×10^2	1	905.73	1.8552
1.028	885.71	0.0007,1.9586	2×10^2	1	905.73	1.8552

(5.6a)-(5.26b), which is consistent with that in [5, Table 1]. The last column represents the only nonzero eigenvalue of the matrix solution of (5.6a)-(5.26b) found by NOA. The second-order and third-order SDR used in [5] for obtaining these optimal values involve 55 and 1540 decision variables, while (5.15) used for our iterations involves only 3 decision variables.

For $V_2^{max} = 1.009, 1.015, 1.022$ and 1.028, NOA is initialized by the solution $W^{(0)}$ of (5.19).

WB5 and WB5mod Network [1]

WB5 is a power network with 5 buses, 2 generators and 6 transmission lines, which lead to 3 nonlinear equality constraints in (5.5b). The size of the Hermitian symmetric matrix variable W is 5×5. The second row in Table 5.2, whose format is similar to that of Table 5.1, reveals that the optimal value found by NOA is almost the same as its lower bound by SDR, so the global optimality of the former is obvious. This result is also consistent with that in [1].

WB5mod [5, Subsection IV.C] is modified from WB5 by changing the values of Q_5^{min}, for which the computational results in Table 5.2 are also consistent with [5, Table III]. NOA involving the matrix variable W of dimension 5×5 is initialized by the solution $W^{(0)}$ of (5.19). The second-order SDR used in [5] involves 1540 decision variables.

Case39mod1 and Case118mod Network [1]

Case39mod1 Network is a modification of New-England test network with 39 buses, 10 generators and 46 transmission lines, which lead to 29 nonlinear equality constraints in

Table 5.2 WB5 [1] and WB5mod [5]

Q_5^{min}	SDR LB	SDR's nonzero eig.	μ	# Iter.	NOA	$\lambda_{\max}$
-30.00	946.53	0.0005,4.9823	10^3	2	946.58	4.9842
-20.51	954.82	0.0065,4.9647	2×10^3	22	1146.48	4.9720
-10.22	963.83	0.0129,4.9457	2×10^3	27	1209.11	4.9732
0.07	972.85	0.0192,4.9267	2×10^3	31	1267.79	4.9748
10.36	981.89	0.0255,4.9078	2×10^3	34	1323.86	4.9767
20.65	990.95	0.0318,4.8889	2×10^3	36	1377.97	4.9787
30.94	1005.13	0.0460,4.8546	2×10^3	39	1430.54	4.9809
41.23	1033.07	0.0711,4.8152	2×10^3	41	1481.81	4.9832
51.52	1070.39	0.1033,4.7748	2×10^3	43	1531.97	4.9856

(5.5b). On the other hand, Case118mod network is a modification of IEEE 118 bus network, with 118 buses, 54 generators and 46 transmission lines, which lead to 64 nonlinear equality constraints in (5.5b). The computational results are summarized in Table 5.3 while the convergence speed of NOA for Case39mod1 is plotted by Figure 5.1.

Table 5.3 Case39mod1 and Case118mod [1]

Network	Case39mod1	Case118mod
SDR LB	10804.08	129682.50
SDR's nonzero eig.	0.0295, 0.0358, 0.0431, 0.0705, 41.1777	0.0121,126.4189
μ	4.5×10^6	2×10^6
# Iter.	11	2
NOA	11216.50	129686.03
$\lambda_{\max}$	39.5586	126.3758

With such large n, the higher-order SDR used by [5] is not applicable as the second-order SDR already involves such number of decision variables that cannot be handled by the present PCs. The second row of Table 5.3 reveals that the matrix solution of SDR for Case39mod1 is of rank-five. The optimal value 11216.50 found by NOA is better than the value 11219.00 provided by [1]. For Case118mod, the third row in Table 5.3 shows that the matrix solution of SDR is rank-two. The optimal value 129686.03 found by NOA is almost same with the value 129625.03 provided by [1].

Case9, 14, 30, 57 Network [6]

Data of Case9, 14, 30 and 57 networks are modified from the Matpower test data files [6] by setting all voltage bounds $V_k^{min} = 0.95$ and $V_k^{max} = 1.05$, $\forall k \in \mathcal{N}$. In

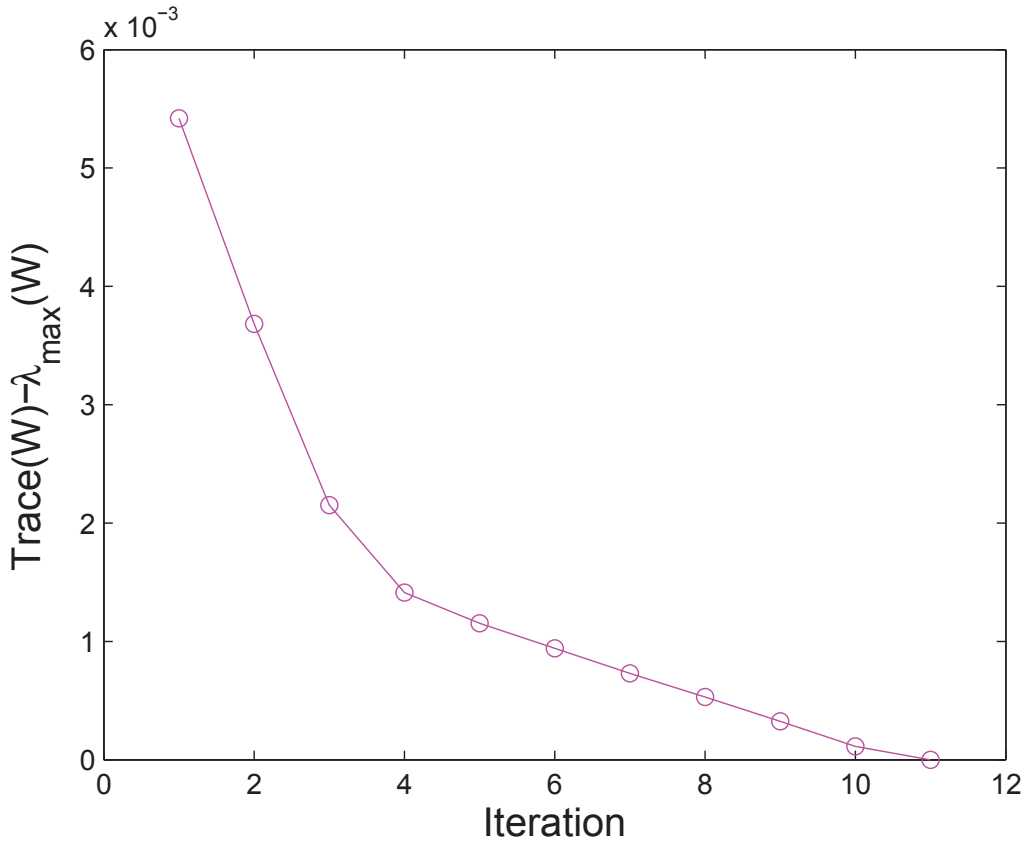

Fig. 5.1 Case39mod1 [1]

Case9 network, there are 9 buses with 3 generators and 9 transmission lines, which lead to 6 nonlinear equality constraints in (5.5b). In Case14 network, there are 14 buses with 5 generators and 20 transmission lines, which lead to 9 nonlinear equality constraints in (5.5b). In Case30 network, there are 30 buses with 6 generators and 41 transmission lines, which lead to 24 nonlinear equality constraints in (5.5b). In Case57 network, there are 57 buses with 7 generators and 80 transmission lines, which lead to 50 nonlinear equality constraints in (5.5b).

The computational results summarized in Table 5.4 reveal that the matrix solution of

Table 5.4 Case9, 14, 30, 57 [6]

Network	SDR LB	SDR's nonzero eig.	μ	# Iter.	NOA	$\lambda_{\max}$
Case9	5305.56	0.0115,9.7464	10^3	2	5305.68	9.7328
Case14	8088.71	14.8299	-	-	-	-
Case30	574.87	31.5534	-	-	-	-
Case57	41213.99	0.0077,56.7939	10^4	47	41313.72	56.7778

SDR in Case14 and Case30 is of rank-one, while it is of rank-two in Case9 and Case57. For Case57, constraint (5.5b) was replaced by constraint (5.21) to regulate NOA for $\tilde{\epsilon} = 6 \times 10^{-6}$. For Case57, the matrix solution of SDR looks like nearly rank-one as its second largest eigenvalue is only 0.0077. However, it is not the case because it leads to an infeasible solution of (5.5), where a number of equality constraints in (5.5b) is violated at unacceptable margin.

Table 5.5 provides the difference between the left hand side and the right side of (5.5b) for some buses. [26, 7] also consider these cases but their data seem to be slightly

Table 5.5 SDR feasibility check for Case57

Bus #	Difference between LHS and RHS of (5.5b)
13	0.0017 + 0.0139i
14	0.0525 + 0.3223i
46	0.0261 + 0.1747i
47	0.0055 + 0.0093i

different from [6] with the charging capacitance b_c not incorporated. Also, the objective functions (5.3) in [26, 7] are linear in P_{G_k}, i.e. $c_{k2} = 0$ in (5.3) so they also lead to different ranks of the matrix solution of SDR (5.6a)-(5.26b). Particularly, [7] found that the matrix solution of SDR (5.6a)-(5.26b) are of rank-two for Case14, Case30 and Case57. Then [7] proposes to find a rank-one matrix point of (5.6) by solving the so-called penalized SDR

$$\min_{W \in C^{n \times n}} \; F(W) + \epsilon_1 \sum_{k \in \mathcal{G}} (Q_{L_k} + \Im(\sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*)) \quad \text{s.t.} \quad (5.6b) - (5.26b) \qquad (5.38)$$

with ϵ_1 increased till (5.38) has an rank-one matrix solution. We also verified this simulation and found some computational inconsistencies between [7] and ours. The numerical results are summarized in Table 5.6. According to the third row of Table 5.6 for Case30mod, the matrix solution of SDR is of rank-one but not of rank-two as in [7]. With the same $\epsilon_1 = 1.5$ in (5.38) as [7, Table II(c)] the solution of (5.38) for Case57mod is found to have rank-two with two nonzero eigenvalues 0.0604 and 58.1802 shown in the fourth row of Table 5.6 but not rank one with the nonzero eigenvalue 56.887 as in [7, Table II(c)]. Our NOA with using (5.21) with $\tilde{\epsilon} = 6 \times 10^{-6}$ outputs the rank-one matrix solution with the eigenvalue 56.67404 shown in the same row of the Table.

Table 5.6 Case14, 30, 57mod [7]

Network	Case14mod	Case30mod	Case57mod
SDR (LB)	316.46	388.98	366.06
SDR's nonzero eig.	0.0090,14.8784	31.6008	0.0608, 58.1756
P-SDR [7]	316.49	-	366.8731 (infeas.)
P-SDR's nonzero eig.	14.8590	-	0.0604, 58.1802
μ	10^3	-	10^4
# Iter	2	-	41
NOA	316.49	-	479.59
$\lambda_{\max}$	14.8610	-	56.6740

Results compared with Matpower6.0[6]

In the section, we make a comparison between the results obtained by Matpower6.0 and NOA for the cases tested in the previous section. The voltage bounds are all set as $V_k^{min} = 0.95$ and $V_k^{max} = 1.05$, $\forall k \in \mathcal{N}$. The computational results are summarized in Table 5.7, where the first column is the network name, the second and third column respectively provides the lower bound and optimum of SDP (5.6a)-(5.26b), while the last column is the optimum obtained by Matpower6.0 [6].

Table 5.7 Results compared with Matpower6.0 [6]

Network	SDR LB	NOA	Matpower
WB2mod	877.78	877.78	Failed
WB5	946.53	946.58	1082.33
Case9	5305.56	5305.68	5305.56
Case14	8088.71	8088.71	8088.71
Case30	574.87	574.87	577.69
Case39mod1	10804.08	11216.50	11221.00
Case57	41213.99	41313.72	44669.53
Case118mod	1.2968×10^5	1.2968×10^5	1.2968×10^5
IEEE300	7.1971×10^5	7.1974×10^5	7.4325×10^5

From Table 5.7, we can see that our optimum are better than or at least consistent with Matpower6.0's. For WB2mod, Matpower6.0 can not locate a feasible point; for WB5, Case30, Case39mod1, Case57 and IEEE300, our optimum are much better than Matpower6.0's; for other cases, our results are consistent with Matpower6.0's.

5.5.2 Simulation results for large-scale OPF problem

We set the tolerances $\epsilon = \epsilon_{tol} = 10^{-5}$ and the penalty parameter $\mu = 10^6$ which makes the penalty term $\mu \sum_{i=1}^{\mathcal{I}} (\text{Trace}(W^{(i,(0))}) - \lambda_{\max}(W^{i,(0)}))$ at similar magnitude with the objective $F(W^{(0)})$. The data source for all examples is Matpower version 5.1 [6]. All examples were considered in [58, 59, 52] by SDR only. We recall that $\mathcal{T}$ is the number of matrix variables W^i in the OPF problem (5.26) and $\mathcal{L}^{(\kappa)}$ is defined by (5.35) is the set of rank-one matrices W^i found after κ-th iteration. The capability of our large-scale NOA in locating the global optimal solution of the OPF problems is demonstrated by showing that the global optimality tolerance (GOT) of its found solution defined as

$$\frac{\text{the found value- lower bound}}{\text{lower bound}}$$

is almost zero.

The numerical examples are presented as follows.

Polish-2383wp system

There are $n = 2383$ buses, 327 generators and 2896 transmission lines, leading to 2056 nonlinear constraints in (5.5b).
Initialization. A lower bound 1.8490×10^6 of (5.26) is found by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1210$ and there are 32 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 10×10 (2×2, resp.).
Stage 1. $|\mathcal{L}^{(10)}| = 1234$ is achieved. There are 8 matrices $W^{i,(10)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 10×10 and (3×3, resp.).
Stage 2. $|\mathcal{L}^{(19)}| = 1237$ is achieved. There are 5 matrices $W^{i,(20)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 9×9 (3×3, resp.).
Stage 3. $|\mathcal{L}^{(25)}| = \mathcal{I} = 1242$ is achieved. The found value of the objective is 1.8408×10^6 with GOT $4.3267e-04$.

Table 5.8 Comparison of bags number $\mathcal{I}$, largest bag size M_i and number of variables

System	$\mathcal{I}$	$\mathcal{I}$ by [52]	Max. N_i	Max. N_i by [52]	Var. # in (5.6)	Var. # in (5.26)	Var. # in (5.26) by [52]	Found value	Found by [52]
Polish-2383wp	1242	2383	10	23	2,840,536	23,199	89,893	1.8408×10^6	1.8742×10^6
Polish-2736sp	1538	2736	10	23	3,744,216	27,298	104,388	1.3042×10^6	1.3082×10^6
Polish-2737sop	1538	2737	10	23	3,746,953	27,034	103,720	7.7572×10^5	7.7766×10^5
Polish-2746wop	1546	2746	10	23	3,771,631	29,024	108,950	1.2040×10^6	1.2085×10^6
Polish-2746wp	1547	2746	10	24	3,771,631	28,257	107,148	1.6266×10^6	1.6324×10^6
Polish-3012wp	1689	3012	10	24	4,537,578	30,996	116,799	2.5727×10^6	2.6089×10^6
Polish-3120sp	1757	3120	10	24	4,868,760	32,637	121,869	2.1391×10^6	2.1608×10^6

Table 5.9 Performance comparison

System	Found value	Found by [52]	Found by [6]
Polish-2383wp	1.8408×10^6	1.8742×10^6	1.8685×10^6
Polish-2736sp	1.3042×10^6	1.3082×10^6	1.3078×10^6
Polish-2737sop	7.7572×10^5	7.7766×10^5	7.7763×10^5
Polish-2746wop	1.2040×10^6	1.2085×10^6	1.2083×10^5
Polish-2746wp	1.6266×10^6	1.6324×10^6	1.6317×10^6
Polish-3012wp	2.5727×10^6	2.6089×10^6	2.5917×10^6
Polish-3120sp	2.1391×10^6	2.1608×10^6	2.1427×10^6

Polish-2736sp system

There are $n = 2736$ buses, 420 generators and 3504 transmission lines, which lead to 2316 nonlinear constraints in (5.5b).

Initialization. A lower bound 1.3041×10^6 of (5.26) is obtained by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1534$ and there are 4 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 6×6 (4×4, resp.).

Stage 1. $|\mathcal{L}^{(9)}| = \mathcal{I} = 1538$ is achieved. The found value of the objective is 1.3042×10^6 with GOT $7.6681e-05$.

Polish-2737sop system

There are $n = 2737$ buses, 399 generators and 3506 transmission lines, which lead to 2338 nonlinear constraints in (5.5b).

Initialization. A lower bound 7.7571×10^5 of (5.26) is obtained by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1532$ and there are 6 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 6×6 (3×3, resp.).

Stage 1. $|\mathcal{L}^{(2)}| = \mathcal{I} = 1538$ is achieved. The found value of the objective is 7.7572×10^5 with GOT $1.2891e-05$.

Polish-2746wop system

There are $n = 2746$ buses, 514 generators and 3514 transmission lines, which lead to 2232 nonlinear constraints in (5.5b).
Initialization. A lower bound 1.2039×10^6 of (5.26) is obtained by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1538$ and there are 8 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 6×6 (3×3, resp.).
Stage 1. $|\mathcal{L}^{(2)}| = \mathcal{I} = 1546$ is achieved. The found value of the objective is 1.2040×10^6 with GOT $8.3063e-05$.

Polish-2746wp system

There are $n = 2746$ buses, 520 generators and 3514 transmission lines, which lead to 2226 nonlinear constraints in (5.5b).
Initialization. A lower bound 1.626590×10^6 of (5.26) is obtained by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1545$ and there are 2 matrices $W^{i,(0)}$ of rank-more-than-one. Their size is 4×4.
Stage 1. $|\mathcal{L}^{(1)}| = \mathcal{I} = 1547$ is achieved. The found value of the objective is 1.626591×10^6 with GOT $6.1478e-07$.

Polish-3012wp system

There are $n = 3012$ buses, 502 generators and 3572 transmission lines, which lead to 2510 nonlinear constraints in (5.5b).
Initialization. A lower bound 2.5717×10^6 of (5.26) is obtained by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1682$ and there are 7 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 7×7 (2×2, resp.).
Stage 1. $|\mathcal{L}^{(4)}| = \mathcal{I} = 1689$ is achieved. The found value of the objective is 2.5727×10^6 with GOT $3.8885e-04$.

Polish-3120sp system

There are $n = 3120$ buses, 505 generators and 3693 transmission lines, which lead to 2615 nonlinear constraints in (5.5b).
Initialization. A lower bound 2.1314×10^6 of (5.26) is obtained by solving SDP (5.37). $|\mathcal{L}^{(0)}| = 1749$ and there are 8 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 8×8 (2×2, resp.).
Stage 1. $|\mathcal{L}^{(9)}| = \mathcal{I} = 1757$ is achieved. The found value of the objective is 2.1391×10^6 with GOT 0.0036.

Numerical summary

One can observe that GOT of the solutions computed by the large-scale NOA is very small, proving its capability to provide the global solution of (5.26). Table 5.8 and Table 5.9 summarize the main points in our simulation. The second and third columns of Table 5.8 are the number $\mathcal{I}$ of bags in (5.25) by our decomposition and by that in [52], while the fourth and fifth columns give the maximum size N_i in (5.25). One can see that both $\mathcal{I}$ and the maximum N_i by our decomposition are substantially smaller than their counterparts by [52]. This leads to far smaller numbers of variables in (5.26), which are provided in the seventh and eighth columns. The number $n(n+1)/2$ of complex variables in (5.6) is also provided in the sixth column to contrast to the number of complex variables in (5.26) in the seventh column. Furthermore, the second column of Table 5.9 provides the best values of (5.5) found by our large-scale NOA, which are far smaller than ones in the third and fourth columns found by [52] and Matpower6.0 [6] (using an interior point method), respectively. In short, our computation approach to the OPF problem (5.26) outperforms other existing approaches in terms of computational efficiency and performance.

5.6 Conclusions

OPF over power transmission networks is a difficult nonconvex optimization problem with multiple nonlinear equality and inequality constraints. We have shown that

the proposed nonsmooth optimization algorithm (NOA) is able to overcome the shortcomings of the existing methods to compute its optimal solution efficiently and even for networks with reasonably large numbers of buses.

Chapter 6

Model Predictive Control for Smart Grids with Multiple Electric-Vehicle Charging Stations

6.1 Introduction

Electrical vehicles (EVs) have emerged as a promising solution to resolve both the economic and environmental concerns in the transportation industry [60]. Using a smart power grid in concurrently serving residences and charging EVs constitutes one of the most important applications of the smart grid technology. However, the massive integration of plug-in EVs (PEVs) into the grid causes many potential impacts such as voltage deviation, increased load variations and power loss of the grid [61], which requires different strategies for load shifting and energy trading and storage in the grid [62–65]. The main difficulty in scheduling of PEV charging to manage the cost and impact of PEV integration is that individual PEVs randomly arrive for charging with their individual demands on charging load and deadlines, which cannot be known before hand. In other words, the future charging demand of PEVs cannot be known a priori. Many existing works consider a simple smart grid with a single charging station (CS) to exclusively serve PEVs. For instance, [66] sets no charging deadlines for PEVs, whose arrival process follows a probability distribution, while [67] assumes

that the future load demand is perfectly known a priori. The future load demand is also assumed to be known in [68] as all PEVs are assumed to arrive at the same time with no charging deadline. It is assumed in [69] that only statistics of demand are known but the PEVs can be fully charged in a single time slot [69, (30)]. It should be realized that serving PEVs is typically considered during a 12-hour time period (for instance from 8:00 pm to 8:00 am), where the integration of a massive number of PEVs has a sizable effect on the power grid, and as such, the length of a time slot is rationally set by 30 minutes or one hour. In other words, the charging scheduling should be considered over a finite horizon of 12-24 time slots, but not over an infinite horizon as considered in [70]. Due to their physical limitations, PEVs are rarely able to be fully charged just during a single time slot.

In this chapter, we consider joint PEV charging scheduling and power control to save service costs for PEVs and the power generation costs in meeting both residential and PEV power demands. Such a problem was considered in [116] but only a small number of PEVs and with each CS serving only one PEV, whose power demand is very small compared with the inelastic demand, so that the integration of PEVs into the grid has almost no effect on the grid. Note that the optimal power flow problems posed in [116] cannot be solved exactly by semi-definite programming relaxation (SDR) [117]. Therefore, it is not known if the objective in PEVs charging scheduling is convex and as such, it is not known if its proposed valley-filling solution is optimal. Larger PEVs' penetration in a few CSs was considered in [118, 119] under the assumption of known arrival and departure times of PEVs. In the present chapter, we are interested in more practical scenarios of a massive number of PEVs arriving randomly at different CSs. No assumption on the probability distribution of their arrival is made, so the conventional model predictive control (MPC) [120, 121] is not applicable. Our contribution is to develop a novel MPC-based approach to address this problem.

The rest of this chapter is structured as follows. Section 6.2 is devoted to the system modeling for this problem and analyzing its computational challenges. An online computational solution using the proposed MPC-based approach is developed in Section 6.3. An off-line computational solution is considered in Section 6.4, which is then compared with the online computational solution in Section 6.5 to show the optimality of the later. Section 6.6 concludes this chapter.

Notation. The notation used in this chapter is standard. Particularly, j is the imaginary unit, X^H is Hermitian transpose of a vector/matrix X, $M \succeq 0$ for a Hermitian symmetric matrix M means that it is positive semi-definite, $\mathsf{rank}(M)$ and $\mathrm{Trace}(M)$ are the rank and trace of a matrix M, respectively. $\Re(\cdot)$ and $\Im(\cdot)$ are the real and imaginary parts of a complex quantity, and $a \leq b$ for two complex numbers a and b is componentwise understood, i.e. $\Re(a) \leq \Re(b)$ and $\Im(a) \leq \Im(b)$. The cardinality of a set $\mathcal{C}$ is denoted by $|\mathcal{C}|$.

6.2 Problem statement and computational challenges

Consider an electricity grid with a set of buses $\mathcal{N} := \{1, 2, ..., N\}$ connected through a set of flow lines $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$, i.e. bus k is connected to bus m if and only if $(k, m) \in \mathcal{L}$. Accordingly, $\mathcal{N}(k)$ is the set of other buses connected to bus k. There is a subset $\mathcal{G} \subseteq \mathcal{N}$, whose elements are connected to distributed generators (DGs). Any bus $k \notin \mathcal{G}$ is thus not connected to DGs. Any bus $k \in \mathcal{G}$ also has a function to serve PEVs and in what follow is also referred to CS k. By defining $M = |\mathcal{G}|$, there are M CSs in the grid. Denote by $\mathcal{H}_k$ the set of those PEVs that arrive at CS k. Accordingly, k_n is the n-th PEV that arrives at CS k. Figure 6.1 provides a general structure of PEV charging in a smart grid system.

The serving time period of the grid is divided into T time slots of length δ_t, which usually varies from 30 minutes to an hour. Under the definition $\mathcal{T} := \{1, 2, \ldots, T\}$, PEV k_n arrives at $t_{a,k_n} \in \mathcal{T}$ and needs to depart at $t_{k_n,d} \in \mathcal{T}$. The constraint

$$t_{k_n,d} - t_{k_n,a} \leq T_{k_n}, \tag{6.1}$$

expresses the PEV k_n's time demand. Suppose that C_{k_n} and $s^0_{k_n}$ are the battery capacity and initial state of charge (SOC) of PEV k_n. It must be fully charged by the departure time $t_{k_n,d}$, i.e.

$$\sum_{t'=t_{k_n,a}}^{t_{k_n,d}} u_h P_{k_n}(t') = C_{k_n}(1 - s^0_{k_n}), \tag{6.2}$$

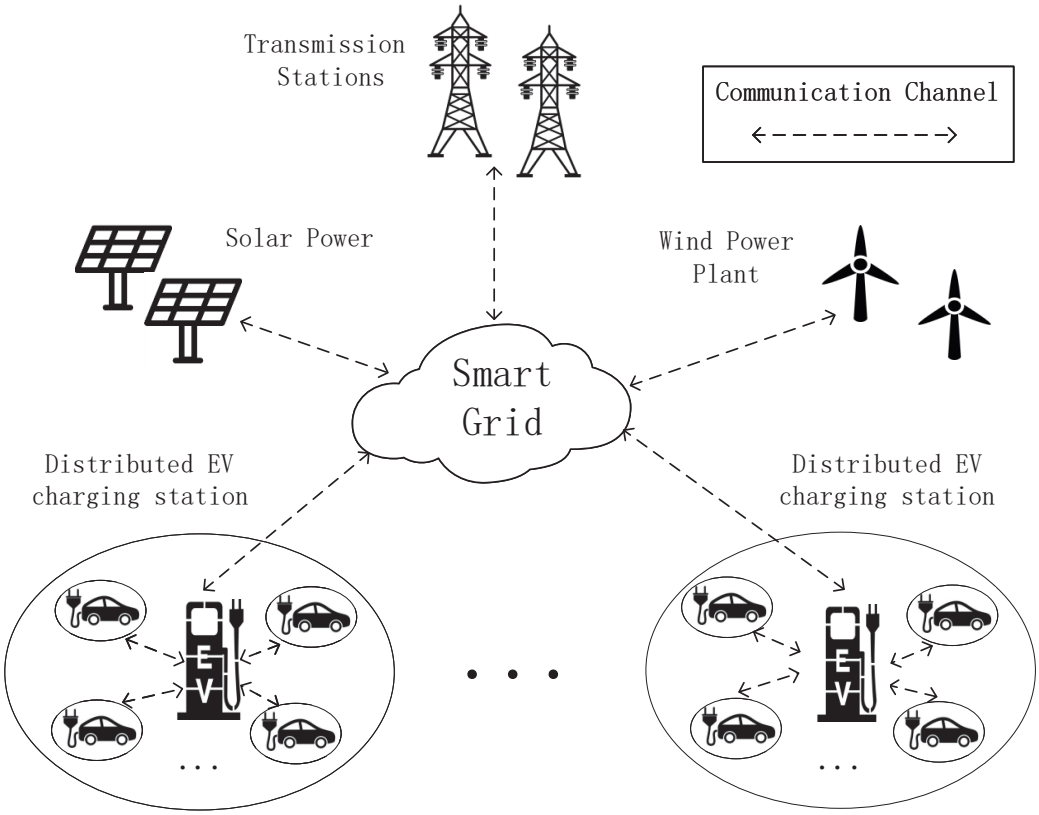

Fig. 6.1 System architecture of PEV charging in smart grid

where u_h is the charging efficiency of the battery and $P_{k_n}(t')$ is a decision variable representing the power charging rate of PEV $k_n \in \mathcal{H}_k$ at time t'.

Due to the limited capacity of the hardware, the following constraint must be imposed:

$$0 \leq P_{k_n}(t') \leq \overline{P}_{k_n}, t_{k_n,a} \leq t' \leq t_{k_n,d}, \tag{6.3}$$

for a given $\overline{P}_{k,n}$. For ease of presentation, we set

$$P_{k_n}(t') = 0, t' \notin [t_{k_n,a}, t_{k_n,b}]. \tag{6.4}$$

Let $y_{km} \in \mathbb{C}$ be the admittance of line (k, m). The current $I_k(t')$ at node $k \in \mathcal{N}$ is,

$$\begin{aligned} I_k(t') &= \sum_{m \in \mathcal{N}(k)} I_{km}(t') \\ &= \sum_{m \in \mathcal{N}(k)} y_{km}(V_k(t') - V_m(t')), \end{aligned}$$

where $V_k(t')$ is the complex voltage at bus k during the time slot t'. For $k \in \mathcal{G}$, the total supply and demand energy is balanced as,

$$V_k(t')(I_k(t'))^* = (P_{g_k}(t') - P_{l_k}(t') - \sum_{n \in \mathcal{H}_k} P_{k_n}(t')) + j(Q_{g_k}(t') - Q_{l_k}(t')),$$

where $P_{g_k}(t')$ and $Q_{g_k}(t')$ are the real and reactive powers generated by DG k, and $P_{l_k}(t')$ and $Q_{l_k}(t')$ are respectively known real and reactive price-inelastic demands at bus k to express the residential power demand. By using the last two equations, we obtain

$$V_k(t')(\sum_{m \in \mathcal{N}(k)} y_{km}(V_k(t') - V_m(t')))^* = (P_{g_k}(t') - P_{l_k}(t') - \sum_{n \in \mathcal{H}_k} P_{k_{n(t')}}) + j(Q_{g_{k(t')}} - Q_{l_k}(t')), k \in \mathcal{G}. \tag{6.5}$$

Similarly,

$$V_k(t')(\sum_{m \in \mathcal{N}(k)} y_{km}(V_k(t') - V_m(t')))^* = -P_{l_k}(t') - jQ_{l_k}(t'), k \notin \mathcal{G}. \tag{6.6}$$

The next constraints relate to the acceptable range of generated power by the DGs:

$$\underline{P}_{g_k} \leq P_{g_k}(t') \leq \overline{P}_{g_k} \ \& \ \underline{Q}_{g_k} \leq Q_{g_k}(t') \leq \overline{Q}_{g_k}, \ k \in \mathcal{G}, \tag{6.7}$$

where $\underline{P}_{g_k}$, $\underline{Q}_{g_k}$ and $\overline{P}_{g_k}$, $\overline{Q}_{g_k}$ are the the lower limit and upper limit of the real generated and reactive generated powers, respectively.

The constraints of voltage are

$$\underline{V}_k \leq |V_k(t')| \leq \overline{V}_k, \quad k \in \mathcal{N}, \tag{6.8}$$

$$|\arg(V_k(t')) - \arg(V_m(t'))| \leq \theta_{km}^{\max}, (k,m) \in \mathcal{L}, t' \in \mathcal{T}, \tag{6.9}$$

where $\underline{V}_k$ and $\overline{V}_k$ are the lower limit and upper limit of the voltage amplitude, while $\theta_{k,m}^{\max}$ are given to express the voltage phase balance.

The constraints of line capacity are

$$|V_k(t')(V_k(t')^* - V_m(t')^*)y_{km}^*| \leq S_{km}, \quad (k,m) \in \mathcal{L}, t' \in \mathcal{T}, \tag{6.10}$$

where S_{km} is the upper limit of capacity for line (k,m).

The problem of interest is to minimize both the energy cost to DGs and charging cost for PEVs. Thus, by defining

$$V(t') = (V_1(t'), \ldots, V_N(t')), \quad \mathcal{V} = \{V(t')\}_{t' \in \mathcal{T}},$$

$$P_g(t') = (P_{g_1}(t'), \ldots, P_{g_M}(t')),$$

$$Q_g(t') = (Q_{g_1}(t'), \ldots, Q_{g_M}(t')),$$

$$R(t') = \{P_g(t'), Q_g(t')\}, \quad \mathcal{R} = \{R(t')\}_{t' \in \mathcal{T}},$$

$$\mathcal{P}^{PEV} = \{P^{PEV}(t')\}_{t' \in \mathcal{T}},$$

$$P^{PEV}(t') = \{P_{k_n}(t')\}_{k_n \in \mathcal{H}_k, k=1,\ldots,M},$$

the objective function is given by

$$\begin{aligned} F(\mathcal{R}, \mathcal{P}^{PEV}) &= \sum_{t' \in \mathcal{T}} \sum_{k \in \mathcal{G}} f(P_{g_k}(t')) \\ &\quad + \sum_{t' \in \mathcal{T}} \sum_{k \in \mathcal{N}} \sum_{n \in \mathcal{H}_k} \beta_t P_{k_n}(t'), \end{aligned}$$

where $f(P_{g_k}(t'))$ is the cost function of real power generation by DGs, which is linear or quadratic in $P_{g_k}(t')$, and β_t is the known PEV charging price during the time interval $(t', t'+1]$.

The joint PEV charging scheduling and power control is then mathematically formulated as

$$\min_{\mathcal{V}, \mathcal{R}, \mathcal{P}^{PEV}} \; F(\mathcal{R}, \mathcal{P}^{PEV}) \quad \text{s.t.} \quad (6.2) - (6.10). \tag{6.11}$$

The above problem (6.11) is very computationally challenging because the quadratic equality constraints (6.5) and (6.6) and nonlinear inequality constraints (6.8) and (6.9)

constitute nonconvex constraints. Moreover, the arrival time $t_{k_n,a}$ of each individual PEV k_n, its charging demand and its departure time $t_{k_n,d}$ are unknown.

6.3 Model predictive control (MPC)-based computational solution

Considering $(R(t'), P^{PEV}(t'))$ and $V(t')$ as the plant state and control, respectively, equations (6.5), (6.6), and (6.7) provide state behavioral equations [122] with the end constraint (6.2) while equations (6.8) and (6.9) provide control constraints. On the surface, (6.11) appears to be a control problem over the finite horizon $[1, T]$. However, all equations in (6.11) are unpredictable beforehand, preventing the application of conventional model predictive control (MPC) [120, 121]. We now follow the idea of [123] to address (6.11).

The conventional MPC relies on the two key steps at time t: predicting future events and minimizing a reference-based cost function by considering the plant over a short receding horizon $[t, t+T]$. Both these steps cannot be implemented for problem (6.11) because the PEVs' arrivals cannot be anticipated while there is no reference for PEV charging. Our proposed method, which does not need prediction for PEVs' arrivals or reference for PEV charging is described as follow. At each time t denote by $C(t)$ the set of PEVs that need to be charged. For each $k_n \in C(t)$, let $\mathcal{P}_{k_n}(t)$ be its remaining demand for charging by the departure time $t_{k_n,d}$. Define

$$\Psi(t) = \max_{k_n \in C(t)} t_{k_n,d}. \tag{6.12}$$

At time t we solve the following optimal power flow (OPF) problem over the prediction horizon $[t, \Psi(t)]$ but then take only $V(t), P_{k_n}(t), R(t)$ for online updating solution of (6.11):

$$\min_{V(t'),R(t'),P_{k_n}(t'),t' \in [t,\Psi(t)],k_n \in C(t)} F_{[t,\Psi(t)]} \tag{6.13a}$$

$$\text{s.t.} \quad \text{network balance and bound constraints in} \quad [t, \Psi(t)], \tag{6.13b}$$

$$V_k(t')(\sum_{m\in\mathcal{N}(k)} y_{km}(V_k(t')-V_m(t')))^* = (P_{g_k}(t') -$$
$$P_{l_k}(t') - \sum_{k_n\in C(t)} P_{k_n}(t')) + j(Q_{g_k}(t') - Q_{l_k}(t')), \tag{6.13c}$$
$$\sum_{t'=t}^{t_{k_n,d}} u_h P_{k_n}(t') = \mathcal{P}_{k_n}(t), \tag{6.13d}$$

with

$$F_{[t,\Psi(t)]} := \sum_{t'=t}^{\Psi(t)} \sum_{k\in\mathcal{G}} f(P_{g_k}(t')) + \sum_{t'=t}^{\Psi(t)} \sum_{k_n\in C(t)} \beta_t P_{k_n}(t').$$

One can notice that (6.13) includes only what is known at the present time t. Of course, (6.13) is a still difficult nonconvex optimization and in the end we need only its solution at t, so we propose the following approach in tackling its solution at t.

Define the Hermitian symmetric matrix $W(t') = V(t')V^H(t') \in \mathbb{C}^{N\times N}$, which must satisfy $W(t') \succeq 0$ and $\mathsf{rank}(W(t')) = 1$. By replacing $W_{km}(t') = V_k(t')V_m^*(t')$, $(k,m) \in \mathcal{N}\times\mathcal{N}$, in (6.13c)-(6.13b), we reformulate (6.13) to the following optimization problem in matrices $W(t') \in \mathbb{C}^{N\times N}$, $t' \in [t, \Psi(t)]$:

$$\min_{W(t'),R(t'),P_{k_n}(t'),t'\in[t,\Psi(t)],k_n\in C(t)} F_{[t,\Psi(t)]} \tag{6.14a}$$
$$\text{s.t.} \quad \text{network balance and bound constraints in} \quad [t,\Psi(t)], \tag{6.14b}$$
$$\sum_{m\in\mathcal{N}(k)} (W_{kk}(t') - W_{km}(t'))y_{km}^* = (P_{g_k}(t') - P_{l_k}(t')$$
$$- \sum_{k_n\in C(t)} P_{k_n(t')}) + j(Q_{g_k}(t') - Q_{l_k}(t')), \quad k\in\mathcal{G}, \tag{6.14c}$$
$$\sum_{m\in\mathcal{N}(k)} (W_{kk}(t') - W_{km}(t'))y_{km}^* =$$
$$-P_{l_k}(t') - jQ_{l_k}(t'), k\notin\mathcal{G}, \tag{6.14d}$$
$$\underline{V}_k^2 \le W_{kk}(t') \le \overline{V}_k^2, \quad k\in\mathcal{N}, \tag{6.14e}$$
$$\Im(W_{km}(t')) \le \Re(W_{km}(t'))\tan(\theta_{km}^{max}), (k,m)\in\mathcal{L}, \tag{6.14f}$$
$$|(W_{kk}(t') - W_{km}(t'))y_{km}^*| \le S_{km}, \tag{6.14g}$$
$$W(t') \succeq 0, \tag{6.14h}$$
$$\text{rank}(W(t')) = 1. \tag{6.14i}$$

Instead of (6.14), which is difficult due to multiple nonconvex matrix rank-one constraints in (6.14h), we solve its semi-definite relaxation (SDR)

$$\min_{W(t'),R(t'),P_{k_n}(t')} F_{[t,\Psi(t)]} \quad \text{s.t. } (6.14b)-(6.14h). \tag{6.15}$$

Suppose that $\hat{W}(t')$ and $(\hat{R}(t'),\hat{P}_{k_n}(t'))$, $t'\in[t,\Psi(t)]$ are the optimal solution of (6.15). If $\text{rank}(\hat{W}(t'))\equiv 1$, $t'\in[t,\Psi(t)]$, then $\hat{V}(t')$ such that $\hat{W}(t')=\hat{V}(t')\hat{V}^H(t')$ together with $\hat{R}(t')$ and $\hat{P}_{k_n}(t')$ constitute the optimal solution of the nonconvex optimization problem (6.13). Otherwise, we consider the following problem:

$$\min_{W(t),R(t)} \; F(P_g(t))) := \sum_{k\in\mathcal{G}} f(P_{g_k}(t)) \tag{6.16a}$$

$$\text{s.t.} \quad (6.3)-(6.4),(6.7),(6.14d)-(6.14h) \quad \text{for } t'=t, \tag{6.16b}$$

$$\sum_{m\in\mathcal{N}(k)} (W_{kk}(t)-W_{km}(t))y^*_{km} = (P_{g_k}(t)-P_{l_k}(t) - \sum_{k_n\in C(t)} \hat{P}_{k_n}(t)) + j(Q_{g_k}(t)-Q_{l_k}(t)), \quad k\in\mathcal{G}, \tag{6.16c}$$

$$\text{rank}(W(t))=1. \tag{6.16d}$$

Note that in contrast to (6.14) involving $\Psi(t)-t$ matrix variables $W(t')$, $t'\in[t,\Psi(t)]$ and also variables $P_{k_n}(t')$, $k_n\in C(t)$ and $t'\in[t,\Psi(t)]$, there is only single matrix variable $W(t)$ in (6.16). The power generation variable $R(t)$ in (6.16) is latent as it is inferred from $W(t)$ in equation (6.16c).

Following our previous works [124, 92, 125, 117, 126], a nonsmooth optimization algorithm (NOA) is proposed to deal with the discontinuous matrix rank-one constraint (6.16d) in the optimization problem (6.16). Under condition (6.14h) in (6.16b),

$$\text{Trace}(W(t))-\lambda_{\max}(W(t))\geq 0,$$

where $\lambda_{\max}(W(t))$ stands for the maximal eigenvalue of $W(t)$. The discontinuous matrix rank-one constraint (6.16d) is then equivalently expressed by the following continuous spectral constraint:

$$\text{Trace}(W(t))-\lambda_{\max}(W(t))=0, \tag{6.17}$$

because it means that $W(t)$ has only one nonzero eigenvalue. Thus the quantity $\text{Trace}(W(t)) - \lambda_{\max}(W(t))$ expresses the degree of the matrix rank-one constraint satisfaction (6.17), which is incorporated into the objective (6.16a), leading to the following penalized optimization problem:

$$\begin{aligned}\min_{W(t),R(t)} F_\mu(W(t),P_g(t)) &:= F(P_g(t)) \\ &+\mu(\text{Trace}(W(t)) - \lambda_{\max}(W(t))) \\ &\text{s.t.} \quad (6.16b) - (6.16c),\end{aligned} \tag{6.18}$$

where $\mu > 0$ is a penalty parameter. The above penalized optimization is exact because the constraint (6.16b) can be satisfied by a minimizer of (6.18) with a finite value of μ. On the other hand, any $W(t)$ feasible for (6.18) is also feasible for (6.16), implying that the optimal value of (6.18) for any $\mu > 0$ is upper bounded by the optimal value of (6.16).

For any $W^{(\kappa)}(t)$ feasible for the convex constraints (6.16c)-(6.16b), let $w_{\max}^{(\kappa)}(t)$ be the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(W^{(\kappa)}(t))$. Then

$$\begin{aligned}\lambda_{\max}(W(t)) &= \max_{||w||^2=1} w^H W(t) w \\ &\geq (w_{\max}^{(\kappa)}(t))^H W(t) w_{\max}^{(\kappa)}(t),\end{aligned} \tag{6.19}$$

i.e. the function $\lambda_{\max}(W(t))$ is lower bounded by the linear function $(w_{\max}^{(\kappa)}(t))^H W(t) w_{\max}^{(\kappa)}(t)$. Accordingly, the following semi-definite program (SDP) provides an upper bound for the nonconvex optimization problem (6.18):

$$\begin{aligned}\min_{W(t),R(t)} F_\mu^{(\kappa)}(W(t),R(t)) &:= F(P_g(t)) \\ &+\mu(\text{Trace}(W(t)) - (w_{\max}^{(\kappa)}(t))^H W(t) w_{\max}^{(\kappa)}(t)) \\ &\text{s.t.} \quad (6.16b) - (6.16c),\end{aligned} \tag{6.20}$$

because

$$F_\mu^{(\kappa)}(W(t),R(t)) \geq F_\mu(W(t),R(t))$$

according to (6.19).

Algorithm 7 NOA 10 for solving (6.16)

1: **Set** $\kappa = 0$ and $(W^{(0)}(t), R^{(0)}(t)) = (\hat{W}(t), \hat{R}(t))$.
2: **Until** $\text{Trace}(W^{(\kappa)}(t)) - (w_{\max}^{(\kappa)}(t))^H W^{(\kappa)}(t) w_{\max}^{(\kappa)}(t) \leq \epsilon$, **solve** (6.20), to find the optimal solution $(W^{(\kappa+1)}(t), R^{(\kappa+1)}(t))$ and **reset** $\kappa + 1 \to \kappa$.
3: **Accept** $(W^{(\kappa)}(t), R^{(\kappa)}(t))$ as the optimal solution of the nonconvex optimization problem (6.16).

Suppose that $(W^{(\kappa+1)}(t), R^{(\kappa+1)}(t))$ is the optimal solution of SDP (6.20). Since $(W^{(\kappa)}(t), R^{(\kappa)}(t))$ is also feasible for (6.20), it is true that

$$\begin{aligned} & F_\mu(W^{(\kappa)}(t), R^{(\kappa)}(t)) \\ = \; & F_\mu^{(\kappa)}(W^{(\kappa)}(t), R^{(\kappa)}(t)) \\ \geq \; & F_\mu^{(\kappa)}(W^{(\kappa+1)}(t), R^{(\kappa)+1}(t)) \\ \geq \; & F_\mu(W^{(\kappa+1)}(t), R^{(\kappa+1)}(t)), \end{aligned}$$

so $W^{(\kappa+1)}(t)$ is a better feasible point of (6.18) than $W^{(\kappa)}(t)$.

In Nonsmooth Optimization Algorithm (NOA) 10 we propose an iterative procedure, which is initialized by the solution $\hat{W}(t)$ of SDR (6.15) and generates a feasible point $W^{(\kappa+1)}(t)$ at the κ-th iteration for $\kappa = 0, 1, \ldots,$ as the optimal solution of SDP (6.18). As proved in [117], this algorithm converges at least to a local minimizer of (6.18). Note that the procedure terminates at

$$\begin{aligned} 0 \; & \leq \; \text{Trace}(W^{(\kappa)}(t)) - \lambda_{\max}(W^{(\kappa)}(t)) \\ & \leq \; \text{Trace}(W^{(\kappa)}(t)) - (w_{\max}^{(\kappa)}(t))^H W^{(\kappa)}(t) w_{\max}^{(\kappa)}(t) \\ & \leq \; \epsilon, \end{aligned}$$

so the spectral constraint (6.17) for the matrix rank-one is satisfied with the computational tolerance ϵ. In summary, our proposed MPC-based computation for (6.11) is based on solving SDP (6.15) for online coordinating PEV charge $\hat{P}_{k_n}(t)$ and solving (6.18) by NOA 10 for online updating the generated voltage $\hat{V}(t)$ for the generated power $\hat{R}(t)$ by

$$\hat{V}(t) = \sqrt{\lambda_{\max}(W^{(\kappa)})} w_{\max}^{(\kappa)}(t), \tag{6.21}$$

whenever the solution $\hat{W}(t)$ of SDR (6.15) is not of rank-one. If $\text{rank}(\hat{W}(t)) = 1$, it is obvious that $\hat{V}(t) = \sqrt{\lambda_{\max}(\hat{W}(t))}\hat{w}_{\max}(t)$ with the normalized eigenvector $\hat{w}_{\max}(t)$ corresponding to $\lambda_{\max}(\hat{W}(t))$ is the optimal solution of (6.13), which is what we need.

6.4 Lower bound by off-line optimization

To investigate the optimality of the MPC-based online computation proposed in the previous section, in this section we address an off-line computation for (6.11), which provides a lower bound for the optimal value of its online computation. Under the additional definition $\mathcal{W} = \{W(t')\}_{t' \in \mathcal{T}}$, we reformulate (6.11) as

$$\min_{\mathcal{W},\mathcal{R},\mathcal{P}^{PEV}} \; F(\mathcal{R}, \mathcal{R}^{PEV}) \quad \text{s.t.} \tag{6.22a}$$

$$(6.2) - (6.4), (6.7), (6.14d) - (6.14h) \text{ for } t' \in \mathcal{T}, \tag{6.22b}$$

$$\sum_{m \in \mathcal{N}(k)} (W_{kk}(t') - W_{km}(t'))y^*_{km} = (P_{g_k}(t') - P_{l_k}(t') - \sum_{n \in \mathcal{H}_k} P_{k_n}(t')) + j(Q_{g_k}(t') - Q_{l_k}(t')), \; k \in \mathcal{G}, \tag{6.22c}$$

$$\mathsf{rank}(W(t')) = 1, t' \in \mathcal{T}. \tag{6.22d}$$

First, we solve its SDR by dropping the matrix rank-one constraints in (6.22d):

$$\min_{\mathcal{W},\mathcal{R},\mathcal{P}^{PEV}} \; F(\mathcal{R}, \mathcal{R}^{PEV}) \quad \text{s.t.} \quad (6.22b) - (6.22c). \tag{6.23}$$

Suppose that $\hat{\mathcal{W}}$ and $\hat{\mathcal{P}}^{PEV}$ are the optimal solution of SDP (6.23). If $\text{rank}(\hat{W}(t)) \equiv 1$, $t \in \mathcal{T}$ then a global solution of the original problem (6.11) is found as $\hat{\mathcal{P}}^{PEV}$, $\hat{\mathcal{R}}$ and $\hat{\mathcal{V}}$ and with $\hat{V}(t)\hat{V}^H(t) = \hat{W}(t)$, $t \in \mathcal{T}$. However, such a matrix rank-one condition is rarely achieved. In what follows we propose two methods to address the matrix rank-one constraints in (6.22d).

Again, under condition (6.14h) for $t' \in \mathcal{T}$ in (6.22b), the rank-one constraints in (6.22d) are equivalently expressed by the single spectral constraint $\sum_{t \in \mathcal{T}}(\text{Trace}(W(t)) - \lambda_{\max}(W(t))) = 0$, which is incorporated into the objective function in (6.22a) for the

Algorithm 8 NOA 11 for solving (6.22)

1: **Set** $\kappa = 0$ and $\mathcal{W}^{(0)} = \hat{\mathcal{W}}$.
2: **Until** $\sum_{t\in\mathcal{T}}(\text{Trace}(W^{(\kappa)}(t)) - (w_{\max}^{(\kappa)}(t))^H W^{(\kappa)}(t) w_{\max}^{(\kappa)}(t)) \leq \epsilon$ **solve** (6.25) to generate $\mathcal{W}^{(\kappa+1)}$, $\mathcal{R}$ and $\mathcal{P}^{PEV}$ and **reset** $\kappa + 1 \to \kappa$.
3: **Accept** $\mathcal{W}^{(\kappa)}$, $\mathcal{R}$ and $\mathcal{P}^{PEV}$ as the optimal solution of the nonconvex optimization problem (6.11).

following penalized function optimization:

$$\min_{\mathcal{W},\mathcal{R},\mathcal{P}^{PEV}} F(\mathcal{R},\mathcal{R}^{PEV}) + \mu \sum_{t\in\mathcal{T}} (\text{Trace}(W(t)) - \lambda_{\max}(W(t))) \quad \text{s.t.} \quad (6.22b) - (6.22c), \tag{6.24}$$

with a penalty parameter $\mu > 0$. Initialized by $\mathcal{W}^{(0)} = \hat{\mathcal{W}}$, the following SDP is solved at κ-th iteration to generate $\mathcal{W}^{(\kappa+1)}$ and $\mathcal{P}^{PEV}$:

$$\min_{\mathcal{W},\mathcal{R},\mathcal{P}^{PEV}} F(\mathcal{R},\mathcal{R}^{PEV}) + \mu \sum_{t\in\mathcal{T}} (\text{Trace}(W(t)) - (w_{\max}^{(\kappa)}(t))^H W(t) w_{\max}^{(\kappa)}(t)) \quad \text{s.t.} \quad (6.22b) - (6.22c). \tag{6.25}$$

This computational procedure is summarized in NOA 11.

Alternatively, we propose the following scalable algorithm for computing (6.22). By replacing $P_{k_n}(t)$ by $\hat{P}_{k_n}(t)$, which was found by solving from (6.23), in (6.22) at every $t \in \mathcal{T}$, we obtain the following optimization problem in $W(t)$ and $R(t)$ only:

$$\min_{W(t),R(t)} \sum_{k\in\mathcal{G}} f(P_{g_k}) \text{ s.t. } (6.22b) - (6.22c) \text{ for } t' = t, \tag{6.26a}$$

$$\mathsf{rank}(W(t)) = 1, \tag{6.26b}$$

which is computed by the distributed NOA Algorithm (DNOA) 9.

Algorithm 9 DNOA 9 for solving (6.26)

1: **Set** $\kappa = 0$ and $W^{(0)}(t) = \hat{W}(t)$, where $\hat{W}(t)$ is found by solving (6.23).
2: **Until** $\text{Trace}(W^{(\kappa)}(t)) - (w_{\max}^{(\kappa)}(t))^H W(t) w_{\max}^{(\kappa)}(t) \leq \epsilon$ **solve** $\{\min_{W(t),R(t)} \sum_{k\in\mathcal{G}} f(P_{g_k})\delta_t + \mu(\text{Trace}(W(t)) - (w_{\max}^{(\kappa)}(t))^H W(t) w_{\max}^{(\kappa)}(t))$ s.t. (6.22b)-(6.22c)$\}$ to generate $W^{(\kappa+1)}(t)$ and $R(t)$, and **reset** $\kappa + 1 \to \kappa$.
3: **Accept** $W^{(\kappa)}(t)$ and $R(t)$ as a found solution of (6.26).

6.5 Simulation results

6.5.1 Simulation setup

The energy sources of the charging stations come from the transmission stations or from self-generated sources such as photovoltaics. Without loss of generality, the charging stations are set at generator buses, which however are not necessarily far from the residential neighborhood. The proposed method still works whenever the charging stations are set at other buses in the grid.

Roughly speaking, distribution system operators (DSOs) are divided into three segments in terms of customer care [127]: DSOs connected to the transmission system, regional DSO and local DSO. Like [116], DSOs in our set-up thus belongs to the first segment, where CSs and PEVs serve as distribution systems connected to transmission networks. Each CS broadcasts the charging commands and communicates the charging demands with connected PEVs. Thus, the proposed method should be conducted by DSOs.

The SDPs (6.15), (6.20), (6.23) and (6.25) are computed using Sedumi[128] interfaced by CVX [129] on a Core i5-3470 processor. Four power networks from Matpower [6] are chosen. The tolerance $\epsilon = 10^{-4}$ is set for the stop criterions.

Generally, PEVs are charged after their owners' working hours. We focus on the charging period from 6:00 pm to 6:00 am of the next day, which is then uniformly divided into 24 time slots of 30 minute length [130]. Accordingly, the charging time horizon is $\mathcal{T} = \{1, 2, \ldots, 24\}$. It is also reasonable to assume that the PEVs arrive during the time period from 6:00 pm to midnight. The PEVs must be fully charged after being plugged into the grid. During this time period, PEV charging demands

usually have one peak [131, Fig. 3], [132, Fig. 2] and [133, Fig. 3], so the arrival times of PEVs are assumed to be independent and can be generated by a truncated normal distribution $(20, 1.5^2)$, which is depicted by Fig. 6.2.

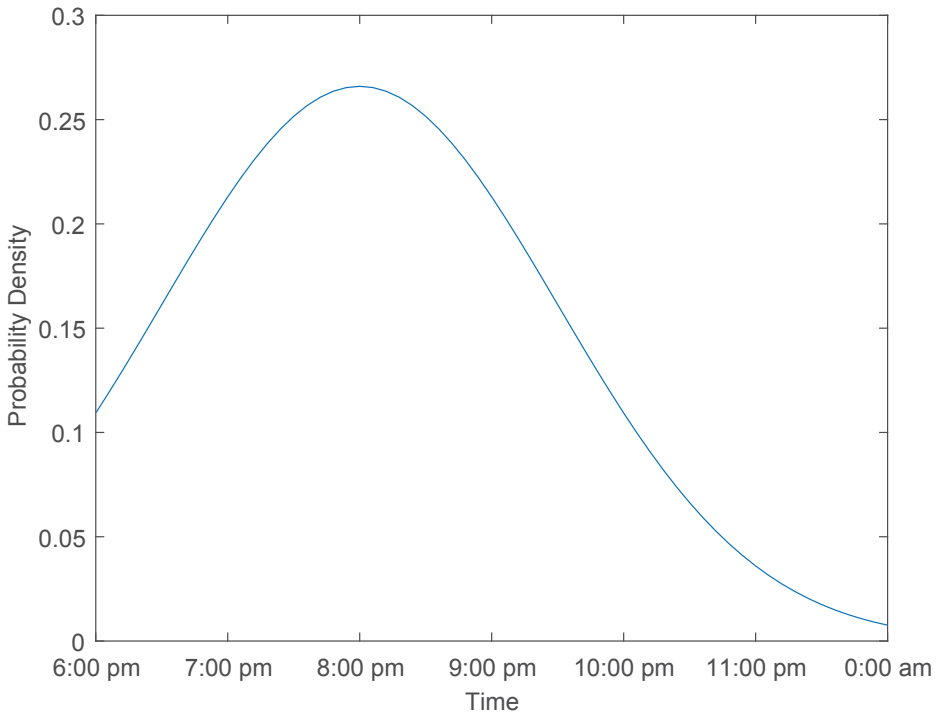

Fig. 6.2 The probability density of PEVs' arrivals

We assume that the PEVs are Tesla Model S's, which have a battery capacity of 100 KWh [134]. The SOC of all PEVs is set as 20%. The structure and physical limits of the considered grids are given in the Matpower library [6] together with the specific cost functions $f(P_{g_k}(t))$.

The price-inelastic load $P_{l_k}(t)$ is calculated as

$$P_{l_k}(t) = \frac{l(t) \times \bar{P}_{l_k} \times T}{\sum_{t=1}^{24} l(t)}, \quad t \in \mathcal{T}, \tag{6.27}$$

where $\bar{P}_{l_k}$ is the load demand specified by [6] and $l(t)$ is the residential load demand taken from [135]. Four profiles are taken from different days in 2017 [135]. Profile 1 is the residential load and energy price from 6:00 pm on February 5th to 6:00 am on February 6th, Profile 2 is from 6:00 pm on March 5th to 6:00 am on March 6th, Profile 3 is from 6:00 pm on April 5th to 6:00 am on April 6th, and Profile 4 is from 6:00 pm

on May 5th to 6:00 am on May 6th. Fig. 6.3 and Fig. 6.4 provide the residential load demand and energy price for these profiles.

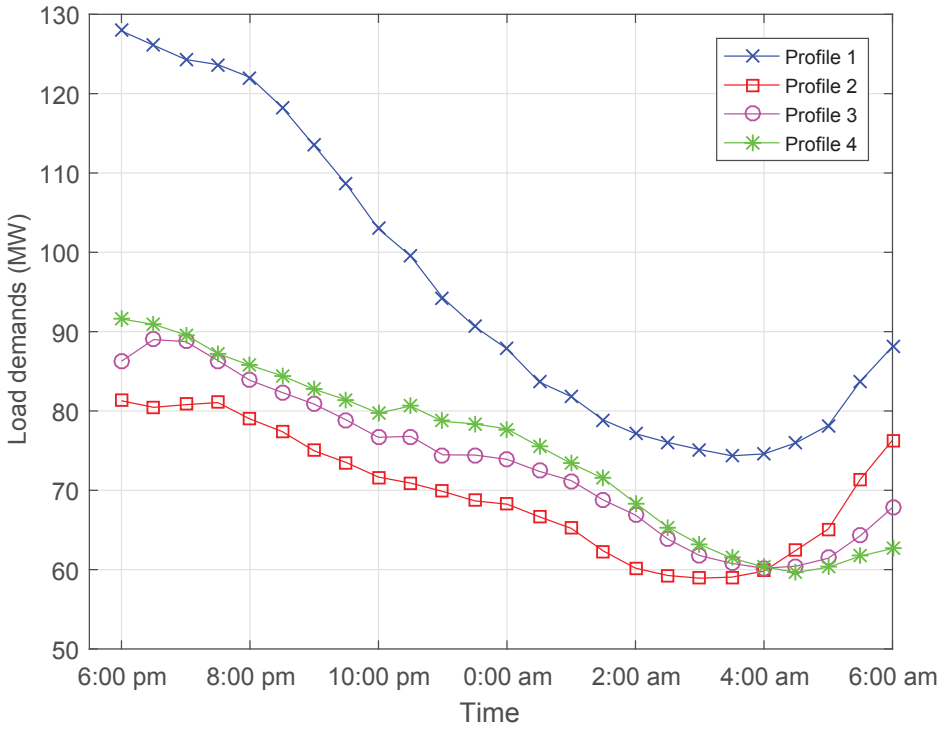

Fig. 6.3 Residential load demands of four profiles

6.5.2 MPC-based online computational results

Four network simulation

We test MPC-based online computation for Case9, Case14, Case30 and Case118mod from [6] and profile 2 of the residential data. The information on these networks is given in Table. 7.1, where the first column is the name of network, the second column indicates the numbers of buses, generators and branches. The dimension of $W(t)$ is given in the third column, while the total number of PEVs is shown in the last column. The computational results are summarized in Table 6.2. Again, the first column is the network name. The second column presents the initial rank of the optimal solution $\hat{W}(t)$ of SDR (6.15). It is observed that the rank of $\hat{W}(t)$ is the same for all $t \in \mathcal{T}$. The value of the penalty parameter μ in (6.18) is given in the third column. As the initial rank of Case14 and Case30 are all rank-one, SDR (6.15) already outputs the

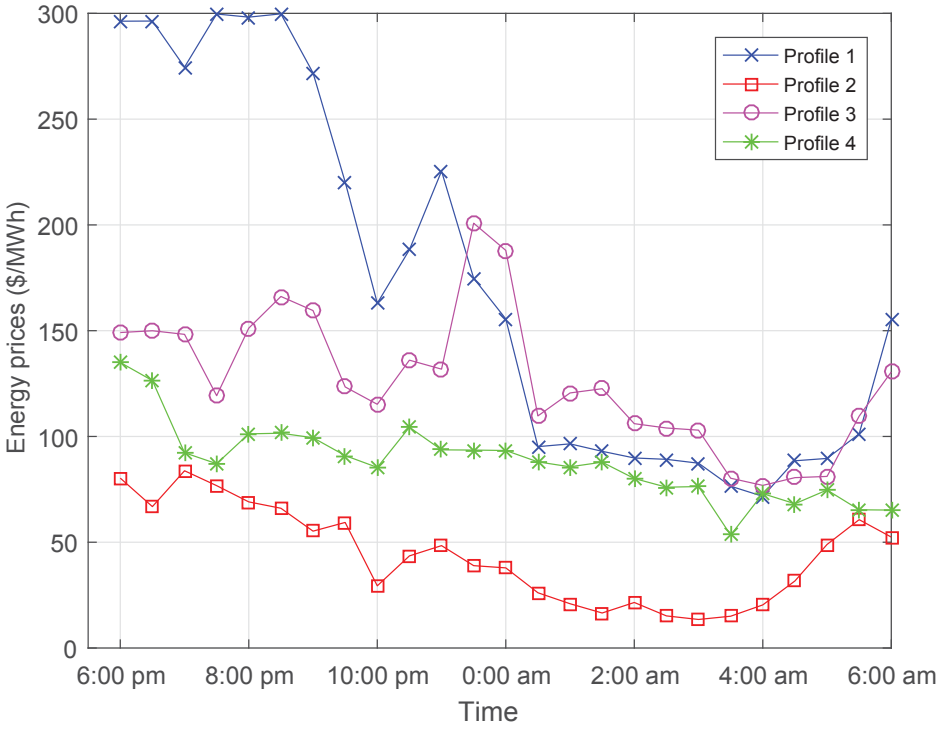

Fig. 6.4 Energy prices for four profiles

Table 6.1 Information on four networks

	Buses/Generators/Branches	Dim. of $W(t)$	PEVs
Case9	9/3/9	$\mathbb{C}^{9\times 9}$	291
Case14	14/5/20	$\mathbb{C}^{14\times 14}$	485
Case30	30/6/24	$\mathbb{C}^{30\times 30}$	582
Case118mod	118/54/186	$\mathbb{C}^{118\times 118}$	5238

optimal solution for (6.13). Comparing the lower bound (LB) in the fourth column by solving SDR (6.23) at each time and the value found by the proposed MPC-based computation with using NOA 10 in computing (6.11) in each time reveals the capability of the MPC-based computation for (6.11). These values are either the same (for Case14 and Case30) or almost the same (for Case9 and Case118mod), so indeed the proposed MPC-based computation could exactly locate a globally optimal solution. The average running time for solving (6.13) to implement the proposed MPC-based computation is provided in the sixth column, which is very short compared with the 30 minute time slot and thus is practical for this particular online application.

Table 6.2 MPC results

	Rank	μ	LB	Comp.value	Time(s)
Case9	9	10	27991.4	27992.3	7.4
Case14	1	-	40824.1	40824.1	8.5
Case30	1	-	4935.6	4935.6	8.7
Case118mod	2	100	644245.9	644278.5	432.1

The voltage profile for the four networks during the charging period are shown in Fig. 6.5. For all cases, the voltage bound constraints (6.8) are satisfied. The voltage behavior is stable and smooth.

Four residential profile simulation

We consider Case30 together with four different residential profiles. The computational results are provided in Table 6.3, whose format is similar to Table 6.2. It can be seen

Table 6.3 MPC results for Case30 with four different residential profiles

	Rank	μ	LB	Comp. value	Time(s)
Profile 1	2	10	31961.2	31963.5	10.9
Profile 2	1	-	4963.3	4963.3	8.7
Profile 3	2	10	10771.3	10774.7	8.7
Profile 4	1	-	8139.3	8139.3	8.1

that, even for the same network, the rank of the optimal solution $\hat{W}(t)$ of SDR (6.15) may be different depending on the residential profiles. For profile 2 and profile 4, the initial rank is one and SDR (6.15) has located a globally optimal solution. However, for profile 1 and profile 3, NOA 10 is needed for obtaining the rank-one solution. The convergence speed is fast, and the optimum values are all equal or close to the lower bound, which clearly shows the global efficiency of the proposed MPC-based computation.

The aggregated active powers generated at each time are shown in Fig. 6.6, from which the trends of generated power are seen to be similar to the residential load demand in Fig. 6.3.

The stable and smooth voltage profile for these 4 residential profiles during the charging period are shown in Fig. 6.7.

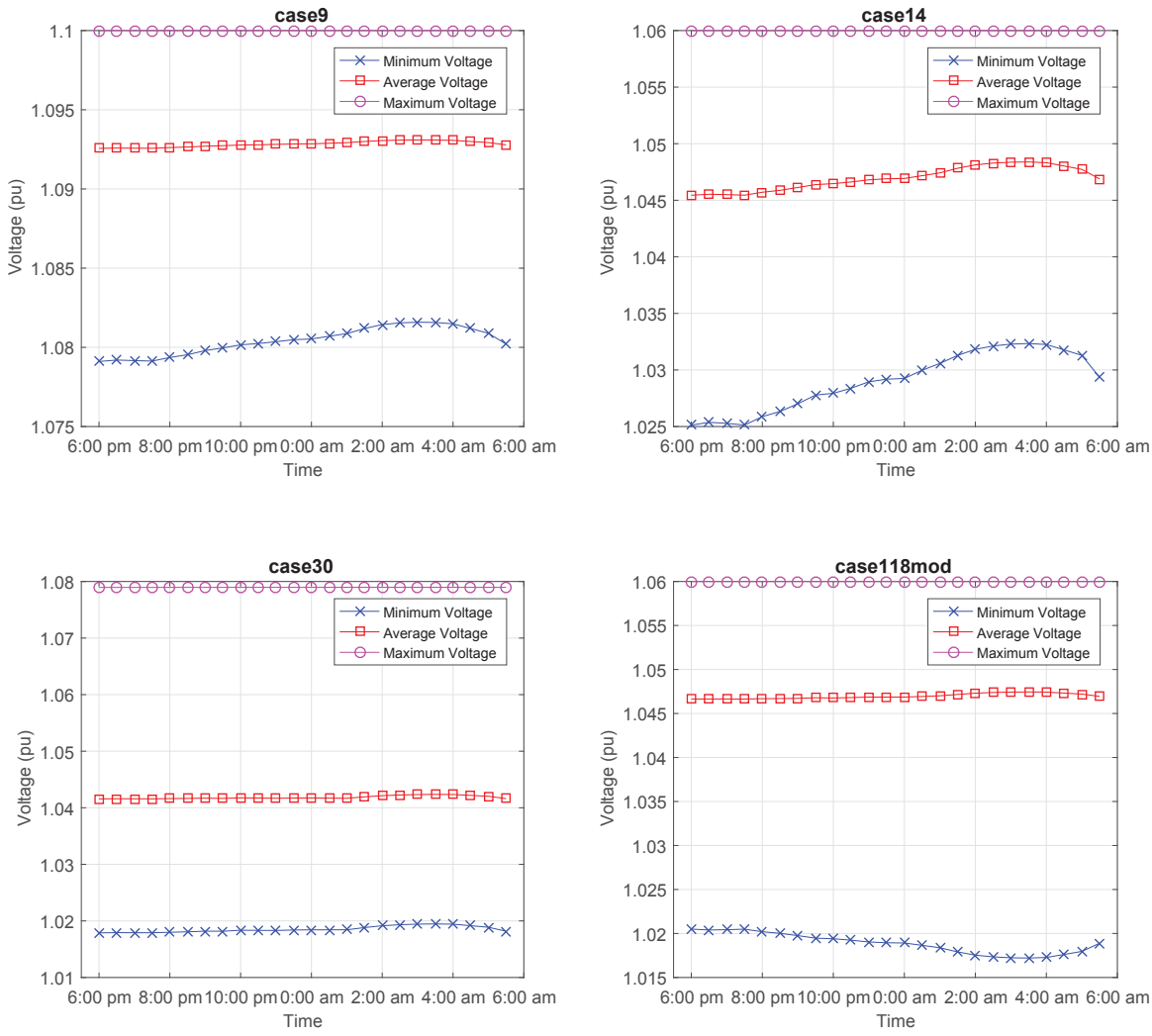

Fig. 6.5 Voltage profile for four networks during the charging period

Fig.6.8 plots the SoC of four PEVs randomly taken from case30 under profile 2, which arrive at different times.

6.5.3 Off-line computation and comparison with MPC-based online computation

Firstly, Case9, Case14, Case30 and Case118mod are tested with the residential data of profile 2 to analyze the efficiency of off-line computation by using Algorithm 11 and Algorithm 9. The computational results are summarized in Table 6.4. The initial rank in the second column is the rank of the optimal solution $\hat{W}(t)$ of SDR (6.23),

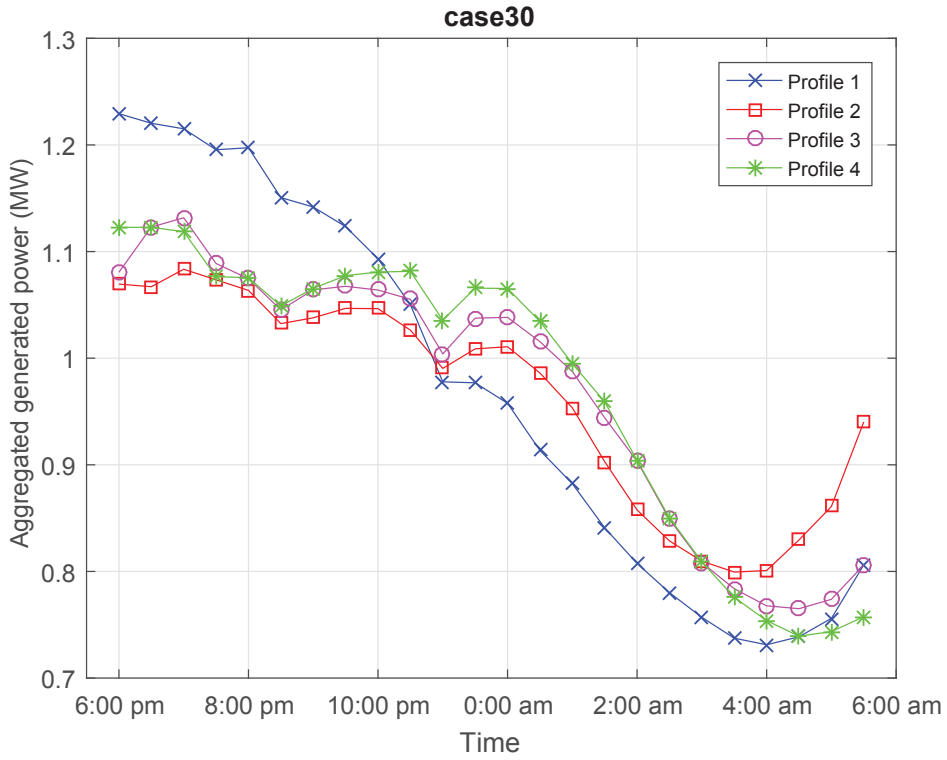

Fig. 6.6 Aggregated active power of online charging for Case30 under four residential profiles

Table 6.4 Offline results of optimal PEV charging for four networks

	Rank	μ	Lower bound	Computed value	Opt. degree	Iterations	NOA time(s)	DNOA time(s)
Case9	9	1	27978.1	27978.1	100%	2	11.2	23.2
Case14	1	-	40800.7	40800.7	100%	1	8.9	8.9
Case30	1	-	4935.6	4935.6	100%	1	24.5	36.3
Case118mod	2	50	644225.3	644233.9	99.999%	3	1094.8	363.5

which is the same for all $t \in \mathcal{T}$. The value of penalty parameter μ in (6.25) is in the third column. The fourth column provides the lower bound by computing SDR (6.23). The values found by solving (6.24) and (6.26) by Algorithm 11 and Algorithm 9 are in the fifth column as they are the same and are either exactly same as their lower bounds in the fourth column (Case9, Case14 and Case30) or almost the same (Case118mod). According to the seventh column both Algorithm 11 and Algorithm 9 converge in two and three iterations for Case9 and Case188mod, while for Case9 and Case30, SDR (6.23) already outputs the optimal rank-one solution. The running times of Algorithm 11 and Algorithm 9 are provided in the eighth and ninth column, respectively. Algorithm 11 requires less running time for small-scale networks such as

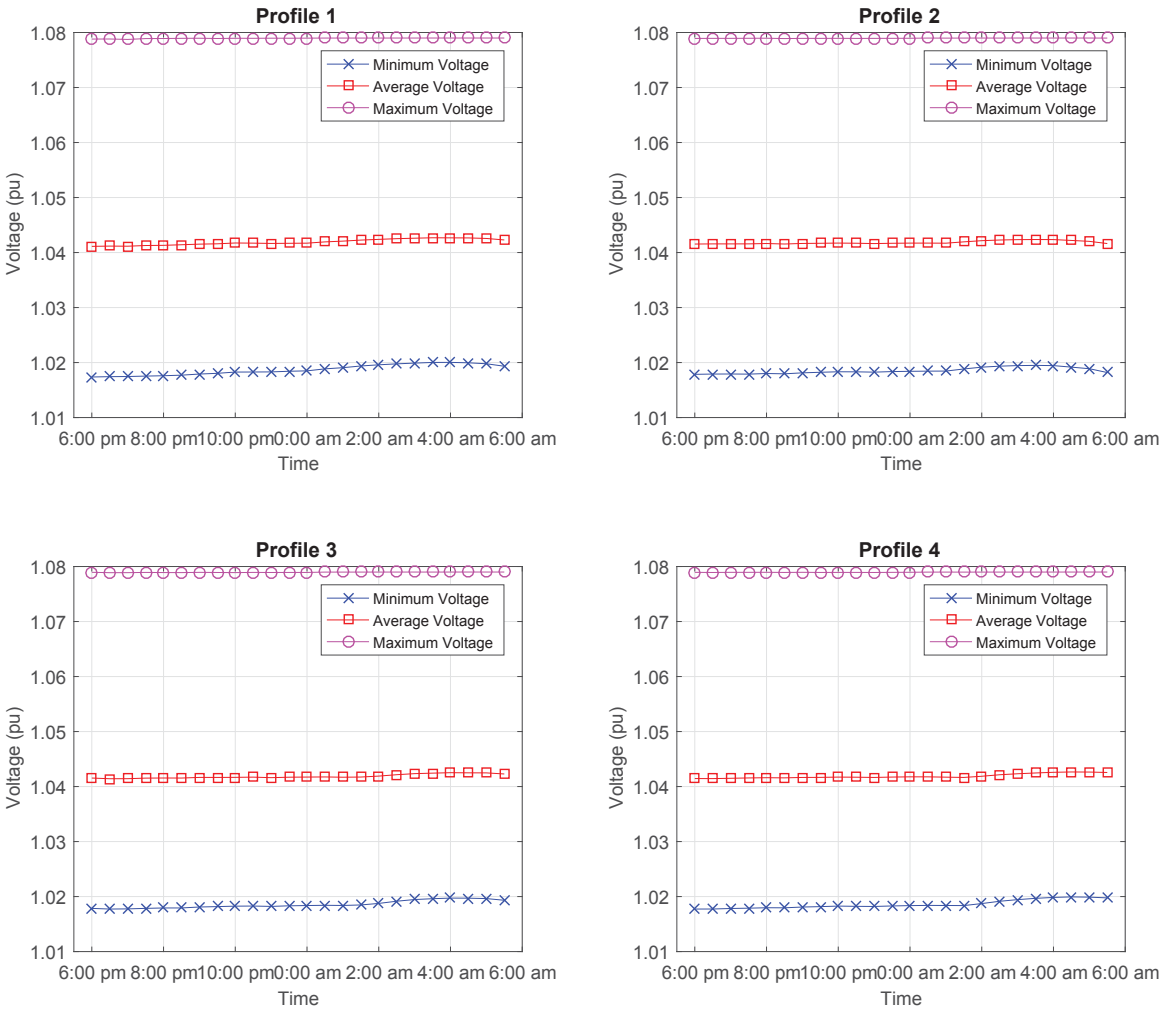

Fig. 6.7 Voltage profile of online charging for Case30 under four residential profiles

Case9, Case14 and Case118mod. However, its running time increases dramatically for large-scale networks such as Case118mod, for which the scalable Algorithm 9 is clearly advantageous.

A performance comparison between MPC-based computation and off-line computation for Case9 and Case30 with the four mentioned residential profiles is provided in Table 6.5, which clearly shows the global optimality of the proposed MPC-based computation as it attains objective values very close to the lower bounds provided by the off-line computation.

Fig. 6.9 plots online and offline power generations in Case30 with four residential profiles, while Fig. 6.10 plots the corresponding PEV charging scheduling. The charging

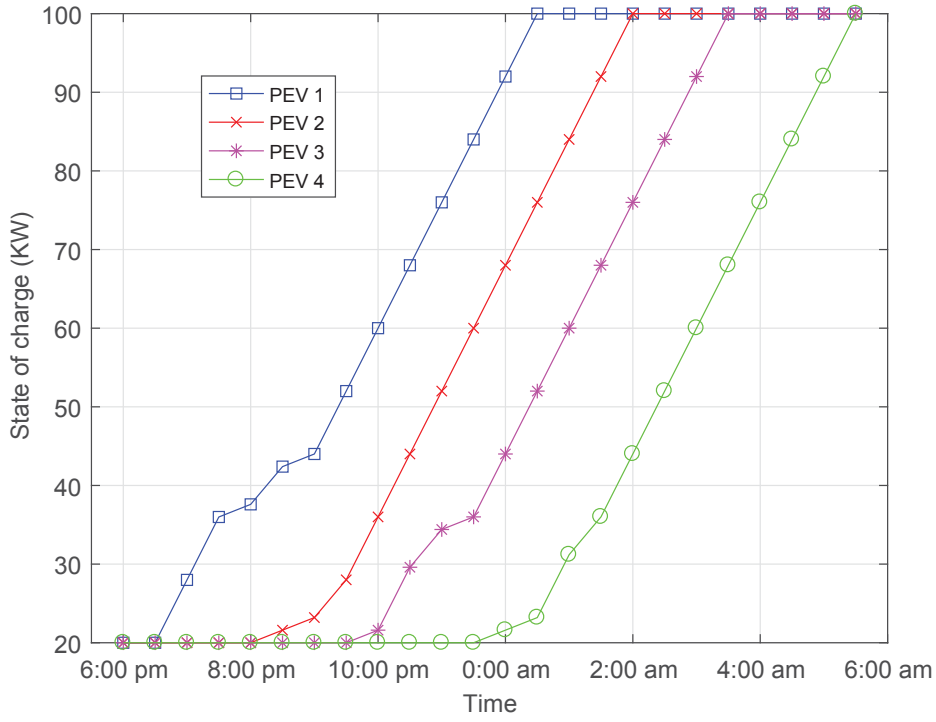

Fig. 6.8 The SOC of PEVs during the charging period

Table 6.5 Performance comparison under MPC-based and off-line computations

		Rank	MPC	Offline	Offline/MPC	MPC times(s)	Offline times(s)
Case9	Profile 1	9	31963.5	31963.1	99.99%	161.1	15.2
	Profile 2	9	27992.3	27978.1	99.94%	177.4	15.1
	Profile 3	9	31102.9	30885.1	99.29%	173.7	14.8
	Profile 4	9	29896.2	29870.9	99.91%	178.4	15.0
Case30	Profile 1	2	31963.5	31963.1	99.99%	262.4	37.4
	Profile 2	1	4963.3	4935.6	99.43%	209.8	12.6
	Profile 3	2	10774.7	10330.8	95.70%	208.2	24.5
	Profile 4	1	8139.3	8087.2	99.35%	194.9	12.9

load drops dramatically after 0 : 00 am, by which all PEVs have been integrated into the grid but some of them have already been fully charged. Obviously, the charging load is sensitive to the energy price. For example in profile 3, the increase of the energy price at 11 : 30 pm and 0 : 00 am leads to a significant drop of the charging load. It should be noted that, energy price is not the only impact factor for the aggregating charging rate in Fig. 6.10. The power balance and residential demands also have significant effects on the aggregating charging rate. From 6:00 pm to 9:00 pm PEVs are continuingly connected to the grid, so the aggregating charging rate during that

time is increased. The charging load under MPC-based and off-line simulation are the same after 0 : 00 am because there are no new PEVs arriving after that time.

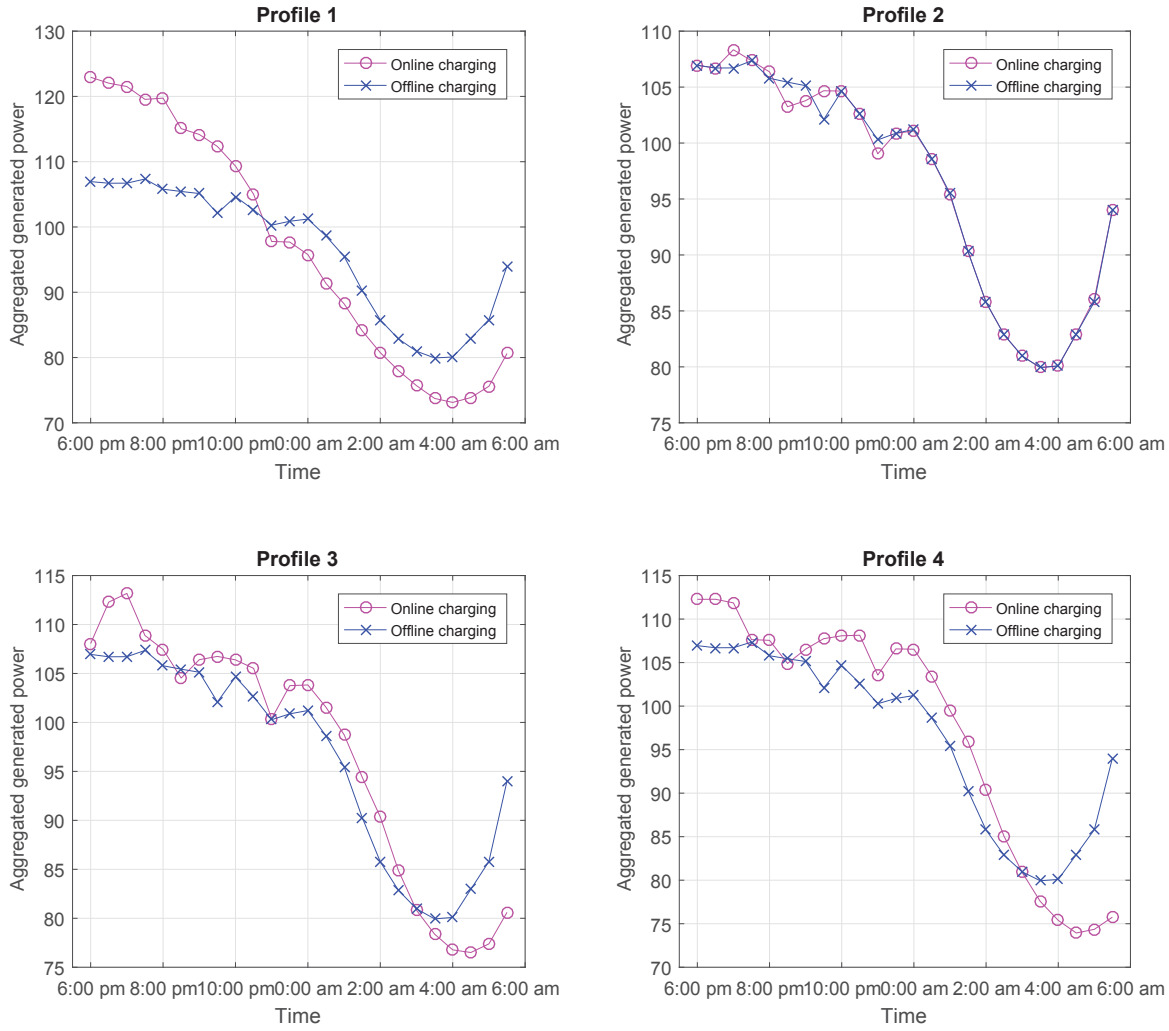

Fig. 6.9 Power generation under MPC-based (online) and offline computation for Case30 with four residential profiles

Fig.6.11 presents the SoC of four PEVs randomly taken from Case30 with profile 2. It can be seen that, though the objective values shown in Table. 6.5 look similar, the SoC are different by online and offline algorithms.

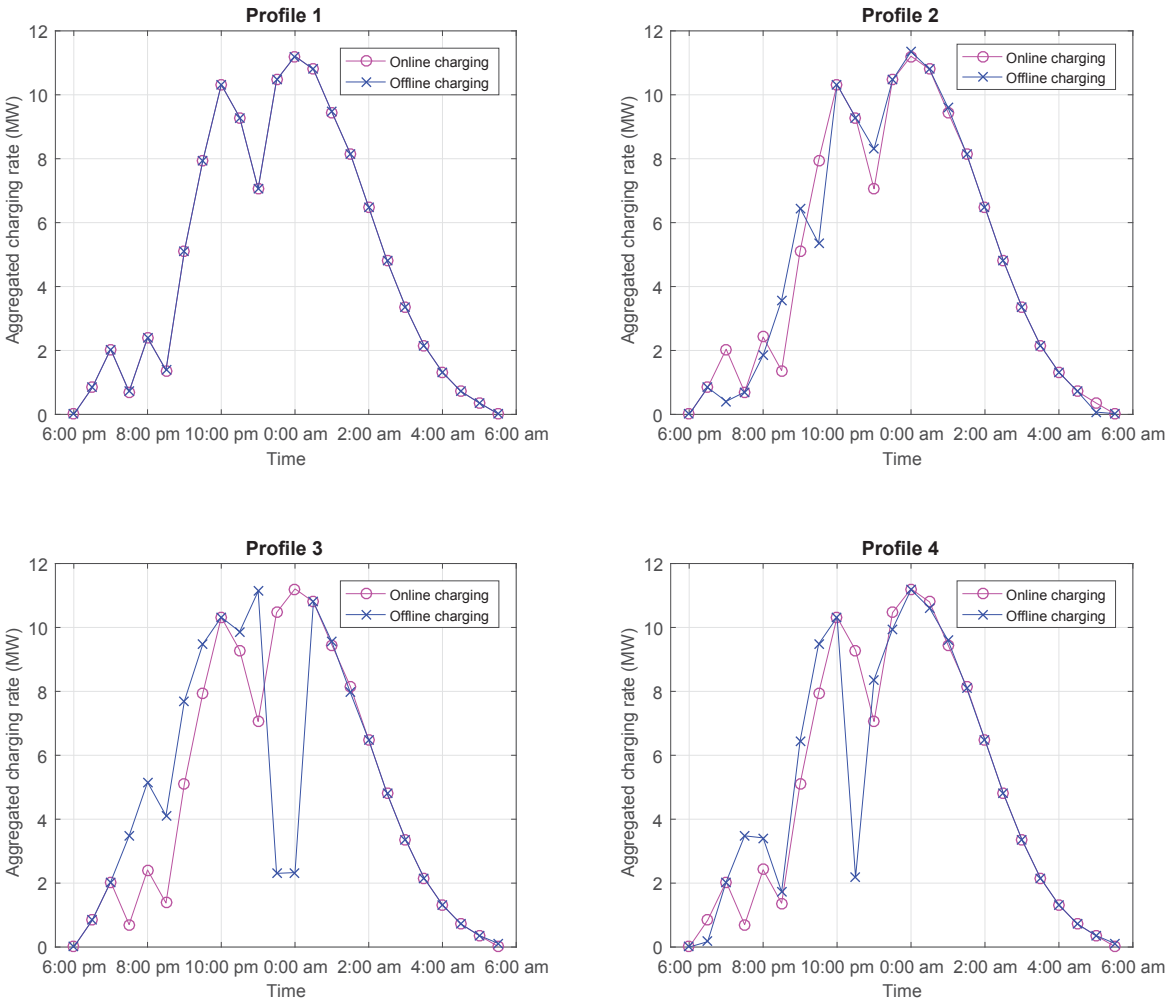

Fig. 6.10 PEVs charging load under MPC-based and offline computation for Case30 with four residential profiles

6.6 Conclusions

Joint PEV charging scheduling and power control for power grids to serve both PEVs at a competitive cost and residential power demands at a competitive operating cost is very difficult due to the random nature of PEVs' arrivals and demands. We have proposed a novel and easily-implemented MPC-based computational algorithm that can achieve a globally optimal solution.

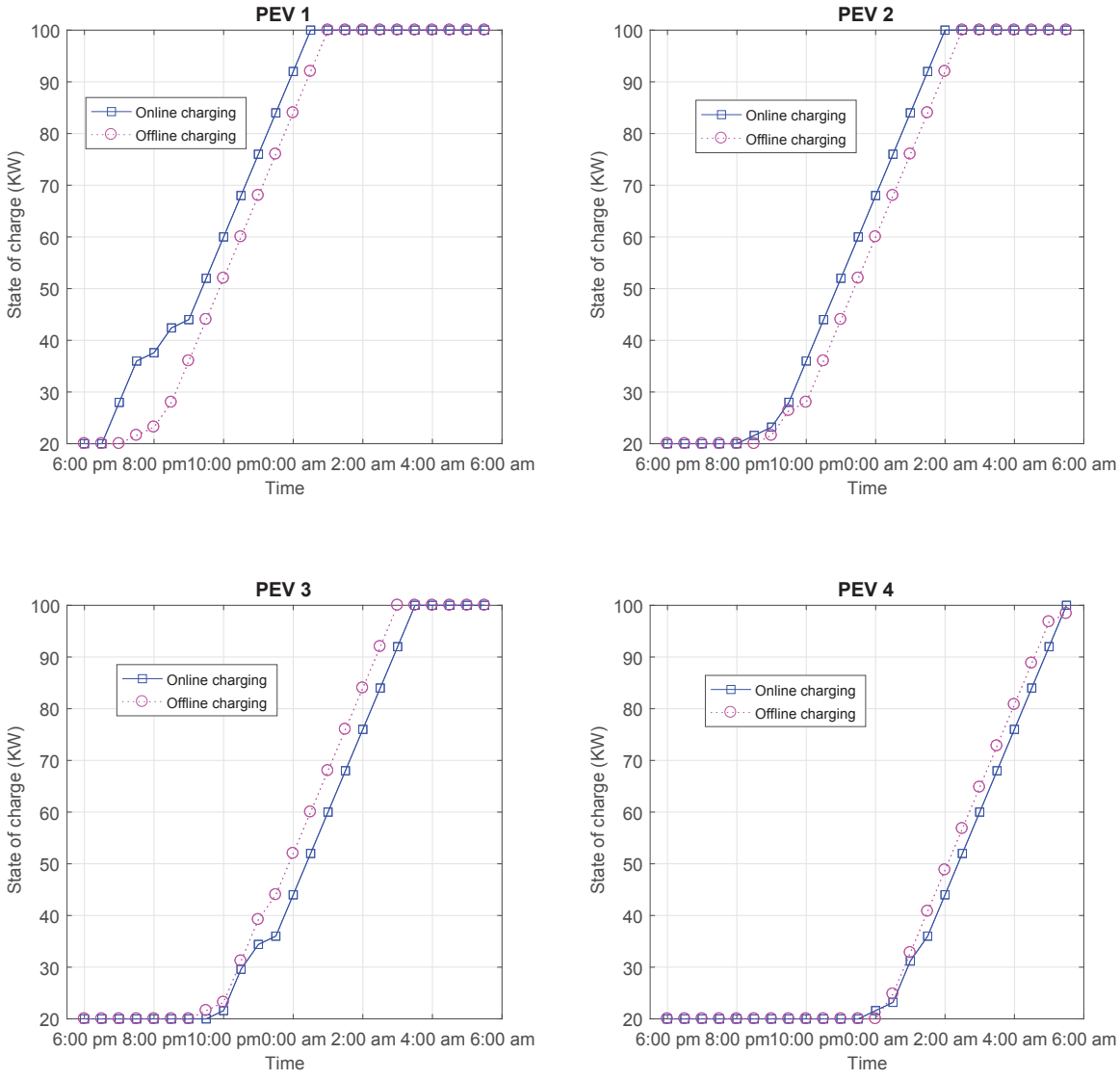

Fig. 6.11 SOC of PEVs randomly taken from Case30 with profile 2

Chapter 7

PMU Placement Optimization for Smart Grid Obvervability and State Estimation

7.1 Introduction

Phasor measurement unit (PMU) is an advanced digital meter, which is used in smart power grids for real-time monitoring of grid operations [71]. By installing it at a buse, the state-of-the-art PMU can measure not only the phasor of the bus voltage but also the current phasors of incident power branches with high accuracy [72]. These measurements are explored by the modern energy management systems (EMSs) for critical applications such as optimal power flow, contingency analysis, and cyber security, etc. [73–75].

As phasor measurement units (PMUs) are costly, there is a vast amount of literature on PMU placement optimization to target the minimal number of PMUs. Under different degrees of observability, the mission is accomplished by binary linear programming (BLP) [76, 77]. Here, the complete observability means that there is no bus left unobserved by the placed PMUs, while depth-of-n unobservability means that there are at most n connecting buses left unobserved by the placed PMUs [78], making as many states as possible observed by restricted number of PMUs. An exhaustive binary

search was proposed in [79] to deal with this objective under the complete observability condition and additional operating conditions such as the single branch outage and the presence of zero power injections. A binary particle swarm optimization algorithm was proposed in [80] to deal with it while maintaining the complete observability conditions under the contingencies of PMU loss or branch outage. Binary quadratic programming and BLP were respectively used in [81] and [82] to study the effect of conventional measurements and zero bus injections to the complete observability.

Apparently, observability alone does not necessarily lead to a meaningful state estimate or an informative PMU configuration. In fact, PMU configurations, which use the same number of PMUs to make the grid completely observable, can result in quite different estimation accuracies [83]. Intuitively, a better estimator can be obtained by appropriately employing more PMUs. PMU placement optimization to minimize the mean squared error (of grid state estimation) or to maximize the mutual information between the measurement output and grid state under a fixed allowable number of PMUs was considered in [2] and [84], respectively. Obviously, these placement tasks are mathematically modelled by optimization of nonlinear objective functions of binary variables subject to a simple linear constraint for fixing the number of PMUs. A convex relaxation with the binary constraint $\{0, 1\}$ for binary variables relaxed to the box constraint $[0, 1]$ was proposed in [2], which not only fails to provide even a local optimal solution in general but also is not scalable in the grid dimension as it involves an additional large-size semi-definite matrix variable. A greedy algorithm proposed in [84] does not provide a local optimal solution either. More importantly, both [2] and [84] ignored observability constraints for computational tractability. It was argued in [84] that its proposed mutual information criterion includes the grid complete observability, which is obviously not right simply because as shown later in the chapter, the latter differentiates the state estimate from its unconditional mean, which is the trivial estimate, while the former does not.

To fill the gap due to disconnected considerations for the grid state observability and state estimation in the existing approaches, this chapter considers PMU placement to optimize the estimation performance under different degrees of observability and with a fixed number of PMUs. These problems are posed as binary nonlinear optimization

problems, which are computationally much challenging. To the authors' best knowledge, such optimization problems are still quite open for research.

The rest of the chapter is structured as follows. Section II is devoted to the problem statement, which also particularly shows the importance of imposing observability constraints in optimization formulations. Section III develops two scalable algorithms for PMU placement optimization to minimize the mean squared error (of grid state estimation) or maximize the mutual information between the measurement outputs and phasor states under a fixed number of PMUs and different degrees of observability. Section IV presents tailored path-following discrete optimization algorithms for the problems without observability constraint. Simulations are provided in Section V, which demonstrates the efficiency of our algorithms. Section VI concludes the chapter. The fundamental inequalities used in Section III are given in the Appendix.

Notation. The notation used in this chapter is standard. Particularly, $A \succ 0$ ($A \succeq 0$, resp.) for a Hermitian symmetric matrix A means that it is positive definite (semi-definite, resp.). Trace(.) and $|.|$ are the trace and determinant operator. 1_N is an N-dimensional vector of ones. I_N is the identity matrix of size N. $a \leq b$ for two real vectors $a = (a_1, \ldots, a_n)^T$ and $b = (b_1, \ldots, b_n)^T$ is componentwise understood, i.e. $a_i \leq b_i$, $i = 1, \ldots, n$. The cardinality of a set $\mathcal{C}$ is denoted by $|\mathcal{C}|$. $\mathbb{E}(.)$ denotes expectation, so the mean $\bar{u}$ of a random variable (RV) u is $\bar{u} = \mathbb{E}(u)$. For two random variables u and v, their cross-covariance matrix R_{uv} is $\mathbb{E}((u-\bar{u})(v-\bar{v})^T)$. Accordingly, the autocovariance $\mathcal{R}_u$ of u is $\mathbb{E}((u-\bar{u})(u-\bar{u})^T)$. $u \sim \mathcal{N}(\bar{u}, \mathcal{R}_u)$ means u is a Gaussian random variable with means $\bar{u}$ and autocovariance $\mathcal{R}_u$, which represent the first moment of u. The entropy of u is $\mathcal{H}(u) = \frac{1}{2}\log_2|\mathcal{R}_u| = \frac{1}{2\ln 2}\ln|\mathcal{R}_u|$. Finally, denote by $u|v$ a RV u conditioned on the RV v.

7.2 Problem statement

Consider a power grid with a set of buses indexed by $\mathcal{N} := \{1, 2, \cdots, N\}$, where buses are connected through a set of transmission lines $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$, i.e. bus k is connected to bus m if and only if $(k, m) \in \mathcal{L}$. Accordingly, $\mathcal{N}(k)$ is the set of other buses connected

to bus k. In a DC power model, the power injection at bus k is approximated by

$$P_k = B_{kk}\theta_k + \sum_{m\in\mathcal{N}(k)} B_{km}\theta_m, \tag{7.1}$$

where P_k is the power injection at bus k and θ_m is the voltage phasor angle at bus m, while B_{km} is the imaginary part of the (k,m)-entry of the grid's admitance matrix Y. Let $P := (P_1,\ldots,P_N)^T \in \mathbb{R}^N$ be the power injection vector and $\theta := (\theta_1,\ldots,\theta_N)^T \in \mathbb{R}^N$ be the voltage phasor vector. Then (7.1) can be re-written as $P = B\theta$, where $B \in \mathbb{R}^{N\times N}$ is the so called susceptance matrix with the entries $B(k,k) = B_{kk}$ and $B(k,m) = B_{km}$, if $m \in \mathcal{N}(k)$, while $B(k,m) = 0$, otherwise. The susceptance matrix B is invertible under the assumption that the grid is fully connected [136]. Since P can be assumed to be $\mathcal{N}(u_p,\Sigma_P)$ [137], it is obvious that $\theta \sim \mathcal{N}(B^{-1}u_p, B^{-1}\Sigma_p(B^{-1})^T)$.

On the other hand, the measurement equation of a PMU installed at bus k in the linear DC power flow model [138] is [72, 139, 84],

$$\begin{aligned} \zeta_k &= \theta_k + \vartheta_k, \\ \zeta_{km} &= \theta_k - \theta_m + \vartheta_{km}, \quad k\in\mathcal{N}, m\in\mathcal{N}(k), \end{aligned} \tag{7.2}$$

with noises $\vartheta_k \sim \mathcal{N}(0,r_k)$ and $\vartheta_{km} \sim \mathcal{N}(0,\rho_k)$. The number of incident lines of bus k is the cardinality $|\mathcal{N}(k)|$. Accordingly, the measurement vector $z_k := (\zeta_k,\zeta_{k1},\ldots,\zeta_{k|\mathcal{N}(k)|})^T$ is of dimension $M_k = |\mathcal{N}(k)|+1$. For simplicity, (7.2) is rewritten in regression form as:

$$z_k = H_k\theta + w_k, \tag{7.3}$$

where $H_k \in \mathbb{R}^{M_k\times N}$ is the associated regression matrix, $w_k := (\vartheta_k,\vartheta_{k1},\ldots,\vartheta_{k|\mathcal{N}(k)|})^T \sim \mathcal{N}(0,R_{w_k})$ with diagonal covariance R_{w_k}.

To describe the presence or absence of PMU at bus k, we introduce a selection vector $\mathbf{x} = (x_1,\cdots,x_N)^T \in \{0,1\}^N$, where $x_k = 1$ if a PMU is installed at bus k, and $x_k = 0$ otherwise. Let us assume that we have S PMUs in total for installation, so

$$\sum_{k\in\mathcal{N}} x_k = S. \tag{7.4}$$

Define

$$\mathcal{D}_S := \{\mathbf{x} \in \{0,1\}^N \ : \ \sum_{k\in\mathcal{N}} x_k = S\} \tag{7.5}$$

and $\mathbf{X} = \text{diag}[x_k\mathcal{I}_k]_{k=1,\ldots,N}$, $\mathcal{R}_w = \text{diag}[R_{w_k}]_{k\in\mathcal{N}}$, where $\mathcal{I}_k$ is the identity matrix of size $M_k \times M_k$.

For every $\mathbf{x} \in \mathcal{D}_S$, let $k_j \in \mathcal{N}$, $j = 1, \ldots, S$ for which $x_{k_j} = 1$. Define accordingly, $\mathcal{R}_w(\mathbf{x}) = \text{diag}[\mathcal{R}_{w_{k_j}}]_{j=1,\ldots,S}$, and

$$z(\mathbf{x}) = \begin{bmatrix} z_{k_1} \\ \cdots \\ z_{k_S} \end{bmatrix}, \ w(\mathbf{x}) = \begin{bmatrix} w_{k_1} \\ \cdots \\ w_{k_S} \end{bmatrix}, \ \bar{H}(\mathbf{x}) = \begin{bmatrix} H_{k_1} \\ \cdots \\ H_{k_S} \end{bmatrix}.$$

The multi-input-multi-output PMU measurement equation is

$$z(\mathbf{x}) = \bar{H}(\mathbf{x})\theta + w(\mathbf{x}).$$

It is obvious that $\mathcal{R}_{z(\mathbf{x})\theta} = \bar{H}(\mathbf{x})\mathcal{R}_\theta$ while $\mathcal{R}_{z(\mathbf{x})} = \bar{H}(\mathbf{x})R_\theta\bar{H}(\mathbf{x})^T + R_{w(\mathbf{x})}$. Let $\theta|z(\mathbf{x})$ be the RV θ conditioned on the RV $z(\mathbf{x})$. By [140]

$$\theta|z(\mathbf{x}) \sim \mathcal{N}(\hat{\theta}, \mathcal{R}_e(\mathbf{x})), \tag{7.6}$$

where

$$\begin{aligned} \hat{\theta} &= \bar{\theta} + \mathcal{R}^T_{z(\mathbf{x})\theta}\mathcal{R}^{-1}_{z(\mathbf{x})}(z(\mathbf{x}) - \overline{z(\mathbf{x})}) \\ &= \bar{\theta} + \mathcal{R}_\theta\bar{H}(\mathbf{x})^T(\bar{H}(\mathbf{x})\mathcal{R}_\theta\bar{H}(\mathbf{x})^T + \mathcal{R}_{w(\mathbf{x})})^{-1} \\ &\quad \times(z(\mathbf{x}) - \bar{H}(\mathbf{x})\bar{\theta}), \end{aligned}$$

which is the minimum mean squared error (MMSE) estimate of θ based on PMU output $z(\mathbf{x})$, and

$$\begin{aligned} \mathcal{R}_e(\mathbf{x}) &= \mathcal{R}_\theta - \mathcal{R}^T_{z(\mathbf{x})\theta}\mathcal{R}^{-1}_{z(\mathbf{x})}\mathcal{R}_{z(\mathbf{x})\theta} \\ &= \mathcal{R}_\theta - \mathcal{R}_\theta\bar{H}(\mathbf{x})^T\left(\bar{H}(\mathbf{x})\mathcal{R}_\theta\bar{H}(\mathbf{x})^T + \mathcal{R}_{w(\mathbf{x})}\right)^{-1} \\ &\quad \times\bar{H}(\mathbf{x})\mathcal{R}_\theta \\ &= \left(\mathcal{R}^{-1}_\theta + \bar{H}(\mathbf{x})^T\mathcal{R}^{-1}_{w(\mathbf{x})}\bar{H}(\mathbf{x})^T\right)^{-1} \end{aligned}$$

$$= \left(\mathcal{R}_\theta^{-1} + \sum_{j=1}^{S} H_{k_j}^T \mathcal{R}_{w_{k_j}}^{-1} H_{k_j}\right)^{-1} \tag{7.7}$$

$$= \left(B^T \Sigma_P^{-1} B + \sum_{k\in\mathcal{N}} x_k H_k^T \mathcal{R}_{w_k}^{-1} H_k\right)^{-1}. \tag{7.8}$$

The mean squared error (MSE) $\mathbb{E}(||\theta - \hat{\theta}||^2)$ is

$$f_e(\mathbf{x}) := \text{Trace}(\mathcal{R}_e(\mathbf{x})),$$

which obviously is an analytical function of the PMU selection vector $\mathbf{x}$.

Further, the mutual information (MI) $I(\theta; z(\mathbf{x}))$ between RVs θ and $z(\mathbf{x})$ is [141, formula (6)]

$$\begin{aligned} I(\theta; z(\mathbf{x})) &= \mathcal{H}(\theta) - \mathcal{H}(\theta|z(\mathbf{x})) \\ &= \frac{1}{2\ln 2}(\ln|\mathcal{R}_\theta| - \ln|\mathcal{R}_e(\mathbf{x})|). \end{aligned}$$

Maximizing the MI $I(\theta; z(\mathbf{x}))$ is thus equivalent to maximizing $f_{MI}(\mathbf{x})$ for

$$f_{MI}(\mathbf{x}) := -\ln|\mathcal{R}_e(\mathbf{x})| = \ln|B^T\Sigma_P^{-1}B + \sum_{k\in\mathcal{N}} x_k H_k^T R_{w_k}^{-1} H_k|.$$

It should be realized that either the MSE $f_e(\mathbf{x})$ or MI $f_{MI}(\mathbf{x})$ does not indicate the depth of the placed PMUs in reaching the measurement for the whole phasor state. One needs either the constraint

$$\mathcal{A}\mathbf{x} \geq 1_N, \tag{7.9}$$

of the complete observability to assure that the phasor state θ is completely observable [139, 142, 143], where $\mathcal{A}$ is the bus-to-bus incidence matrix defined by $\mathcal{A}_{km} = 1$ if $k = m$ or bus k is adjacent to bus m, and $\mathcal{A}_{km} = 0$ otherwise, or the constraint

$$\mathcal{B}\mathcal{A}\mathbf{x} \geq 1_{N_B}, \tag{7.10}$$

of the depth-of-one unobservability to assure that there are no two connecting buses that are unobservable [78]. Here and after $\mathcal{B}$ is the branch-to-bus incident matrix and N_B is the total number of branches. The general case of dept-of-n unobservability with an arbitrary n is treated similarly though its practicability is unknown.

Let us analyse the constraints (7.9) and (7.10) from the information-theoretic view point. The constraint (7.9) guarantees that all state components θ_m are observable, i.e. each θ_m appears at least once in the measurement equations (7.2), which implies $\theta_m|z(\mathbf{x}) \neq \theta_m$, making the measurement equations (7.2) meaningful for estimating θ_m. When some θ_m is not observable, i.e. it does not appear in the measurement equations (7.2), it follows that $\theta_m|z(\mathbf{x}) = \theta_m$ so the measurement equations in (7.2) are useless for estimating θ_m. In this case, the estimate for θ_m is its unconditional mean $\bar{\theta}_m$ with $\mathbb{E}((\theta_m - \bar{\theta}_m)^2) = \mathcal{R}_\theta(m,m)$ and $I(\theta_m; z(\mathbf{x})) = \mathcal{H}(\theta) - \mathcal{H}(\theta|z(\mathbf{x})) = 0$. In other words, the optimization problem for maximizing $I(\theta; z(\mathbf{x}))$ does not reveal a nontrivial estimate for θ_m that is a contradiction to [84, statement 1), page 448, 2nd column] which states that the mutual information metric includes the complete observability condition (7.9) as a special case. Of course, the number of PMUs, S, needs to be sufficient enough to make the constraint (7.9) fulfilled. When S is not allowed to be sufficient, one may go for more relaxed constraint (7.10), which forces all neighboring buses of any unobservable bus to be observable and thus essentially makes as many states as possible be observable by the PMUs.

Thus, we can state the problem of PMU placement optimization to minimize the MMSE or to maximize the MI between the measurement output and phasor state under a fixed number of PMUs and observability/depth-of-one unobservability as the following binary nonlinear optimization problem

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{D}_S, (7.9)/(7.10), \tag{7.11}$$

where $f(\mathbf{x}) \in \{f_e(\mathbf{x}), -f_{MI}(\mathbf{x})\}$, which is a convex function.

7.3 Scalable Penalty algorithms for optimal PMU selection

It is obvious that the main issue is regarding how to handle the discrete constraint $\mathbf{x} \in \mathcal{D}_S$ in (7.11). The following result establishes the equivalence of this discrete constraint and a continuous constraint.

Lemma 2 *For a polytope* $\mathsf{Poly}(\mathcal{D}_S) = \{\mathbf{x} \in [0,1]^N : \sum_{k\in\mathcal{N}} x_k = S\}$, *the discrete constraint* $\mathbf{x} \in \mathcal{D}_S$ *in (7.11) is equivalent to the continuous constraint*

$$\mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), g(\mathbf{x}) \geq S, \tag{7.12}$$

for $g(\mathbf{x}) := \sum_{k\in\mathcal{N}} x_k^L$ *with* $L > 1$.

Proof: Note that $x_k^L \leq x_k \ \forall \ x_k \in [0,1]$, so $g(\mathbf{x}) \leq \sum_{k\in\mathcal{N}} x_k = S \ \forall \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S)$. Therefore constraint (7.12) forces $g(\mathbf{x}) = S$, which is possible if and only if $x_k^L = x_k$, $k \in \mathcal{N}$, i.e $x_k \in \{0,1\}$, $k \in \mathcal{N}$, implying $\mathbf{x} \in \mathcal{D}_S$. □

Since $g(\mathbf{x})$ is convex in $\mathbf{x}$, the constraint $g(\mathbf{x}) \geq S$ in (7.12) is a reverse convex constraint [89]. As such $\mathcal{D}_S = \mathsf{Poly}(\mathcal{D}_S) \setminus \{\mathbf{x} \ : \ g(\mathbf{x}) < S\}$, i.e. $\mathcal{D}_S$ is difference of two convex sets $\mathsf{Poly}(\mathcal{D}_S)$ and $\{\mathbf{x} \ : \ g(\mathbf{x}) < S\}$. Also as L decreases, $g(\mathbf{x})$ tends to approach a linear function $\sum_{k\in\mathcal{N}} x_k$ and thus, the constraint $g(\mathbf{x}) \geq S$ approaches the linear constraint $\sum_{k\in\mathcal{N}} x_k \geq S$. However, it does not mean that choosing L closer to 1 is effective because the function $g(\mathbf{x}) - S$ also approaches zero very quickly, making the constraint $g(\mathbf{x}) \geq S$ highly artificial. In our previous works [144, 145], $L = 2$ was chosen. However, as we will see shortly, $L = 1.5$ is a much better choice, accelerating the convergence of the iterative computational processes. The following result is a direct consequence of Lemma 2.

Proposition 4 *The function*

$$\tilde{g}(\mathbf{x}) = 1/g(\mathbf{x}) - 1/S$$

can be used to measure the degree of satisfaction of the discrete constraint $\mathbf{x} \in \mathcal{D}_S$ *in the sense that* $\tilde{g}(\mathbf{x}) \geq 0 \ \forall \ \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S)$ *and* $\tilde{g}(\mathbf{x}) = 0$ *if and only if* $\mathbf{x} \in \mathcal{D}_S$. □

Following our previous developments in [144] and [145], instead of handling constraint (7.12), we incorporate the degree of its satisfaction into the objective in (7.11), leading to the following penalized optimization problem:

$$\min_{\mathbf{x}} F_\mu(\mathbf{x}) := f(\mathbf{x}) + \mu(1/g(\mathbf{x}) - 1/S)$$

$$\text{s.t.} \quad \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), (7.9)/(7.10), \tag{7.13}$$

where $\mu > 0$ is a penalty parameter. This penalized optimization problem is exact with a sufficiently large μ. Note that (7.13) is a minimization of a nonconvex function over a convex set. We now develop a path-following computational procedure for its solution. For this purpose, we firstly develop an upper bounding approximation for (7.13), at some feasible point $\mathbf{x}^{(\kappa)}$ (at κ-th iteration). As the function $g(\mathbf{x})$ is convex, it is true that [89],

$$\begin{aligned} g(\mathbf{x}) &\geq g^{(\kappa)}(\mathbf{x}) \\ &:= g(\mathbf{x}^{(\kappa)}) + \langle \nabla g(\mathbf{x}^{(\kappa)}), \mathbf{x} - \mathbf{x}^{(\kappa)} \rangle \\ &= -(L-1) \sum_{k \in \mathcal{N}} (\mathbf{x}_k^{(\kappa)})^L + L \sum_{k \in \mathcal{N}} (\mathbf{x}_k^{(\kappa)})^{L-1} \mathbf{x}_k. \end{aligned}$$

Therefore, an upper bounding approximation at $\mathbf{x}^{(\kappa)}$ for $1/g(\mathbf{x})$ can be easily obtained as $1/g(\mathbf{x}) \leq 1/g^{(\kappa)}(\mathbf{x})$ over the trust region

$$g^{(\kappa)}(\mathbf{x}) > 0. \tag{7.14}$$

At the κ-th iteration we are supposed to solve the following convex optimization problem to generate the next iterative point $\mathbf{x}^{(\kappa+1)}$:

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x}) + \mu(1/g^{(\kappa)}(\mathbf{x}) - 1/S) \\ &\text{s.t. } \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), (7.9)/(7.10), (7.14). \end{aligned} \tag{7.15}$$

Although function $f(\mathbf{x})$ is convex, it is not easy to optimize it. For instance, when $f = f_e$, usually (7.15) is solved via the following semi-definite optimization problem with the introduction of slack symmetric $N \times N$ matrix variable $\mathbf{T}$:

$$\begin{aligned} &\min\nolimits_{\mathbf{x},\mathbf{R}} \ \text{Trace}(\mathbf{R}) + \mu(1/g^{(\kappa)}(\mathbf{x}) - 1/S) \\ &\text{s.t.} \ \ \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), (7.9)/(7.10), (7.14), \begin{bmatrix} \mathcal{R}_e^{-1}(\mathbf{x}) & I_N \\ I_N & \mathbf{R} \end{bmatrix} \succeq 0, \end{aligned}$$

which is not scalable to $\mathbf{x}$. For $f = -f_{MI}$, (7.15) is

$$\begin{aligned} \max_{\mathbf{x}\in[0,1]^N} \ln|\mathcal{R}_e^{-1}(\mathbf{x})| - \mu(1/g^{(\kappa)}(\mathbf{x}) - 1/S) \\ \text{s.t.} \quad \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), (7.14), \end{aligned}$$

with no known convex solver of polynomial complexity.

In the following, we propose a different approach to provide scalable iterations for (7.11). Obviously, there is $\epsilon > 0$ such that

$$\mathcal{A}_\epsilon := B^T\Sigma_P^{-1}B - \epsilon\sum_{k\in\mathcal{N}} H_k^T R_{w_k}^{-1} H_k \succ 0.$$

For $f = f_e$, applying inequality (7.23) in the Appendix for

$$A_0 \to \mathcal{A}_\epsilon, x_k \to x_k + \epsilon, \bar{x}_k \to x_k^{(\kappa)} + \epsilon, \tag{7.16}$$

yields $f_e(\mathbf{x}) \geq f_e^{(\kappa)}(\mathbf{x}) := a_0^{(\kappa)} + \sum_{k\in\mathcal{N}} \dfrac{a_k^{(\kappa)}}{x_k+\epsilon}$ for $0 < a_0^{(\kappa)} := \text{Trace}((\mathcal{R}_e(\mathbf{x}^{(\kappa)}))^2\mathcal{A}_\epsilon)$ and

$$\begin{aligned} 0 < a_k^{(\kappa)} := (x_k^{(\kappa)} + \epsilon)^2 \text{Trace}((\mathcal{R}_e(\mathbf{x}^{(\kappa)}))^2 H_k^T R_{w_k}^{-1} H_k), \\ k \in \mathcal{N}. \end{aligned}$$

Accordingly, initialized by a feasible point $\mathbf{x}^{(0)}$ for (7.13), at the κ-th iteration for $\kappa = 0, 1, \ldots$, we solve the following convex optimization problem to generate the next iterative point $\mathbf{x}^{(\kappa+1)}$, instead of (7.15):

$$\begin{aligned} \min_{\mathbf{x}} F_\mu^{(\kappa)}(\mathbf{x}) := f_e^{(\kappa)}(\mathbf{x}) + \mu(1/g^{(\kappa)}(\mathbf{x})1 - 1/S) \\ \text{s.t.} \quad \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), (7.9)/(7.10), (7.14). \end{aligned} \tag{7.17}$$

Note that $F_\mu(\mathbf{x}) \leq F_\mu^{(\kappa)}(\mathbf{x})\ \forall\ \mathbf{x}$, and $F_\mu(\mathbf{x}^{(\kappa)}) = F_\mu^{(\kappa)}(\mathbf{x}^{(\kappa)})$, and $F_\mu^{(\kappa)}(\mathbf{x}^{(\kappa+1)}) < F_\mu^{(\kappa)}(\mathbf{x}^{(\kappa)})$ (because $\mathbf{x}^{(\kappa+1)}$ and $\mathbf{x}^{(\kappa)}$ are the optimal solution and a feasible point for (7.17)). Therefore,

$$F_\mu(\mathbf{x}^{(\kappa+1)}) \leq F_\mu^{(\kappa)}(\mathbf{x}^{(\kappa+1)}) < F_\mu^{(\kappa)}(\mathbf{x}^{(\kappa)}) = F_\mu(\mathbf{x}^{(\kappa)}),$$

i.e. $\mathbf{x}^{(\kappa+1)}$ is a better feasible point than $\mathbf{x}^{(\kappa)}$ for (7.13). For a sufficient large $\mu > 0$, $\tilde{g}(\mathbf{x}^{(\kappa)}) \to 0$ as well, yielding an optimal solution of the binary nonlinear optimization problem (7.11) for the case $f = f_e$. Algorithm 10 provides a pseudo-code for the proposed computational procedure.

Algorithm 10 Scalable Penalized MMSE Algorithm

1: **Initialization.** Set $\kappa = 0$. Take any feasible point $\mathbf{x}^{(0)} \in (0,1)^N$ for (7.13). Choose μ such that $f_e(\mathbf{x}^{(0)})$ and $(1/S - 1/g(\mathbf{x}^{(0)}))$ achieve similar magnitude.
2: **Repeat**
3: Solve the convex optimization problem (7.17) to generate the next feasible point $\mathbf{x}^{(\kappa+1)}$.
4: Set $\kappa := \kappa + 1$.
5: **Until** convergence.

Analogously, based on inequality (7.24) in the Appendix, for A_0, x_k, and $\bar{x}_k$ from (7.16), at the κ-th iteration we solve the following convex optimization problem to generate the next iterative point $\mathbf{x}^{(\kappa+1)}$, instead of (7.15), when $f = -f_{MI}$:

$$\begin{aligned} \max_{\mathbf{x}} \left[\alpha_0^{(\kappa)} - \sum_{k\in\mathcal{N}} \frac{\alpha_k^{(\kappa)}}{x_k + \epsilon} - \mu\left(\frac{1}{g^{(\kappa)}(\mathbf{x})} - \frac{1}{S}\right) \right] \\ \text{s.t.} \quad \mathbf{x} \in \mathsf{Poly}(\mathcal{D}_S), (7.9)/(7.10), (7.14), \end{aligned} \tag{7.18}$$

for

$$\begin{aligned} \alpha_0^{(\kappa)} &:= -\ln|\mathcal{R}_e(\mathbf{x}^{(\kappa)})| \\ &\quad +\mathrm{Trace}(\mathcal{R}_e(\mathbf{x}^{(\kappa)})(\textstyle\sum_{k\in\mathcal{N}}(\epsilon + x_k^{(\kappa)})H_k^T R_{w_k}^{-1} H_k)), \\ \alpha_k^{(\kappa)} &:= (x_k^{(\kappa)} + \epsilon)^2 \mathrm{Trace}(\mathcal{R}_e(\mathbf{x}^{(\kappa)}) H_k^T R_{w_k}^{-1} H_k), \\ &\quad k \in \mathcal{N}. \end{aligned}$$

Algorithm 11 is a pseudo-code for solution of the binary nonlinear optimization problem (7.11) for the case $f = -f_{MI}$.

Algorithm 11 Scalable Penalized MI Algorithm

1: **Initialization.** Set $\kappa = 0$. Take any feasible point $\mathbf{x}^{(0)} \in (0,1)^N$ for (7.13). Choose μ such that $f_{MI}(\mathbf{x}^{(0)})$ and $(1/S - 1/g(\mathbf{x}^{(0)}))$ achieve similar magnitude.
2: **Repeat**
3: Solve the convex optimization problem (7.18) to generate the next feasible point $\mathbf{x}^{(\kappa+1)}$.
4: Set $\kappa := \kappa + 1$.
5: **Until** convergence.

7.4 Tailored path-following discrete optimization algorithms

In this section, we address problem (7.11) without the observability constraint (7.9)/(7.10):

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{D}_S. \tag{7.19}$$

which was considered in [146, 2] for $f = f_e$ with the help of semi-definite relaxation (SDR). The reader is referred to [147] for capacity of SDR to address discrete optimization problems such as (7.19). We now develop a simple but very efficient path-following discrete optimization algorithm that explores a simple structure of the discrete constraint $\mathbf{x} \in \mathcal{D}_S$ to address (7.19).

Lemma 3 *$\mathcal{D}_S$ is the set of vertices of* $\mathsf{Poly}(\mathcal{D}_S)$.

Proof: For $\mathbf{x} \in \mathcal{D}_S$ define

$$J(\mathbf{x}) = \{k_1 < k_2 < .... < k_S | x_{k_j} = 1,\ j = 1, 2, ..., S\}. \tag{7.20}$$

Suppose $\bar{\mathbf{x}} \in \mathcal{D}_S$. It suffices to show that if $\bar{\mathbf{x}} = \mu\mathbf{a} + (1-\mu)\mathbf{b}$ for $\mathbf{a}, \mathbf{b} \in \mathsf{Poly}(\mathcal{D}_S)$ and $0 < \mu < 1$ then $\mathbf{a} = \mathbf{b} = \bar{\mathbf{x}}$. Indeed, for $i \in J(\bar{\mathbf{x}})$ we have $\bar{x}_i = 1 = \mu a_i + (1-\mu)b_i$ and since $a_i \in [0,1]$ and $b_i \in [0,1]$ it follows that $a_i = b_i = 1$. For $i \notin J(\bar{\mathbf{x}})$ we have $\bar{x}_i = 0 = \mu a_i + (1-\mu)b_i$ and since $a_i \in [0,1]$, and $b_i \in [0,1]$ it follows that $a_i = b_i = 0$. Hence $\mathbf{a} = \mathbf{b} = \bar{\mathbf{x}}$ as asserted. □

Recall that point $\mathbf{x}$ is a vertex neighbouring the vertex $\bar{\mathbf{x}}$ if and only if there exists a pair i and j with $i \in \{S+1, \ldots, N\}$ and $j \in \{1, \ldots, S\}$ such that $x_i = 1, x_j = 0$ and $x_k = \bar{x}_k = 1$ for all $k \in \mathcal{N} \setminus \{j\}$ and $x_k = \bar{x}_k = 0$ for all $k \in \{N+1, \ldots, M\} \setminus \{i\}$.

A $\bar{\mathbf{x}} \in \mathcal{D}_S$ is a minimizer of f over $\mathsf{Poly}(\mathcal{D}_S)$ if and only if $f(\bar{\mathbf{x}}) \leq f(\mathbf{v})$ for every $\mathbf{v} \in \mathcal{D}_S$ neighbouring $\bar{\mathbf{x}}$.

Algorithm 12 Path-following discrete optimization algorithm

Initialization. Start from a $\mathbf{x}^{(0)} \in \mathcal{D}_S$. Set $\kappa = 0$.
κ-th iteration. If there is a $\bar{\mathbf{x}} \in \mathcal{D}_S$ neighbouring $\mathbf{x}^{(\kappa)}$ such that $f(\bar{\mathbf{x}}) < f(\mathbf{x}^{(\kappa)})$ then reset $\kappa+1 \to \kappa$ and $\mathbf{x}^{(\kappa)} \to \bar{\mathbf{x}}$. Otherwise, if $f(\mathbf{x}) \geq f(\mathbf{x}^{(\kappa)})$ for all $\mathbf{x} \in \mathcal{D}_S$ neighbouring $\mathbf{x}^{(\kappa)}$ then stop: $\mathbf{x}^{(\kappa)}$ is a local optimal solution of (7.11).

The proposed Algorithm 12 looks like the Dantzig simplex method for linear programming, which is of the 20th century's top ten algorithms [148] although its polynomial complexity cannot be proved (in contrast to the polynomial complexity of the interior points methods for linear programming).[1] Based on this powerful algorithm, we propose Algorithm 13 for the following problem of choosing the minimum number of PMUs to satisfy MMSE or MI constraint:

$$\min_{\mathbf{x}} \sum_{k \in \mathcal{N}} x_k \ : \ \mathbf{x} \in \{0,1\}^N, f(\mathbf{x}) \leq \epsilon. \tag{7.21}$$

Algorithm 13 Iterative Procedure

Initialization. Start from $1 < S_0 < N$ and use Algorithm 12 to find the optimal solution $\mathbf{x}^{(0)}$ of (7.19) for $S = S_0$.
κ-th iteration. Reset $S \to S-1$ if $f(\mathbf{x}_{opt}) < \epsilon$ and $S \to S+1$ if $f(\mathbf{x}_{opt}) > \epsilon$.
Set $\kappa := \kappa + 1$.
Until $f(\mathbf{x}^{(\kappa)}) \leq \epsilon$ but $f(\mathbf{x}^{(\kappa-1)}) > \epsilon$.

7.5 Simulation results

In the simulation, the real power injections P are normally distributed and independent across different buses [137]. Similarly to the simulation setup in [84], the mean vector

[1] Conceptually, Dantzig simplex method is very simple: starting from any vertex of a simplex it moves to a better neighbouring vertex until there is no better neighbouring vertex found

of real power injection $u_p = (u_p(1), \ldots, u_p(N))^T$ is obtained by properly scaling the power profiles in [6], while the diagonal entries of power injection covariance matrix are assumed to be 10% of the mean values, i.e. Σ_P is a diagonal matrix with diagonal entries $\Sigma_P(k,k) = 0.1u_p(k)$. The deviation of measurement noise for bus voltage and current branch are set as $r_k = 0.01$ and $\rho_k = 0.02$, respectively. All algorithms are solved by Matlab on a Core i7-7600U processor. Sedumi[128] interfaced by CVX is used to solve the convex optimization problems (7.17) and (7.18). The commonly used benchmark power networks IEEE 30-bus, IEEE 39-bus, IEEE 57-bus and IEEE 118-bus with their structure and susceptance matrix obtained from Matpower [6] are tested.

It is observed in [77] that the minimum number of PMUs for the network complete observability (CO) or depth-of-one unobservability (DoOU) can be found by solving the following binary linear program

$$\min_{\mathbf{x}} \sum_{k\in\mathcal{N}} x_k \ : \ \mathbf{x} \in \{0,1\}^N, (7.9)/(7.10). \tag{7.22}$$

Table 7.1 provides the minimum number of PMUs needed for the network's CO and DoOU (obtained by solving (7.22 by CPLEX [149]) given in the third and fourth columns.

Table 7.1 The minimum number of PMUs needed for two observability conditions

IEEE	# Branch	# PMUs for CO	# PMUs for DoOU
30-bus	41	10	4
39-bus	46	13	7
57-bus	80	17	11
118-bus	186	32	18

Fig.7.1 depicts the MMSE obtained by different methods versus the number of placed PMUs. The curve "Algorithm 10" is the theoretical MMSE by solving problem (7.17) under the constraint (7.9) of the complete observability, while the curve "Monte-Carlo" is obtained through Monte-Carlo simulation. The MMSEs by Algorithm 10 and Monte-Carlo simulation are seen consistent with the increase in the number of placed PMUs leading to a better MMSE. The curve "Observable" is the MMSE at feasible points for (7.17) that is found by CPLEX [149]. Algorithm 10 is seen to achieve

much better MMSE. The last curve "Algorithm 3" is the MMSE by solving (7.19) by Algorithm 12. Obviously, the Algorithm 12 achieves better MMSE due to the absence of constraints (7.9) and (7.10). The curves in Fig. 7.2 provide normalized MI results in a similar format to Fig.7.1. The capability and efficiency of Algorithm 11 and Algorithm 12 to obtain informative PMU placements are quite clear.

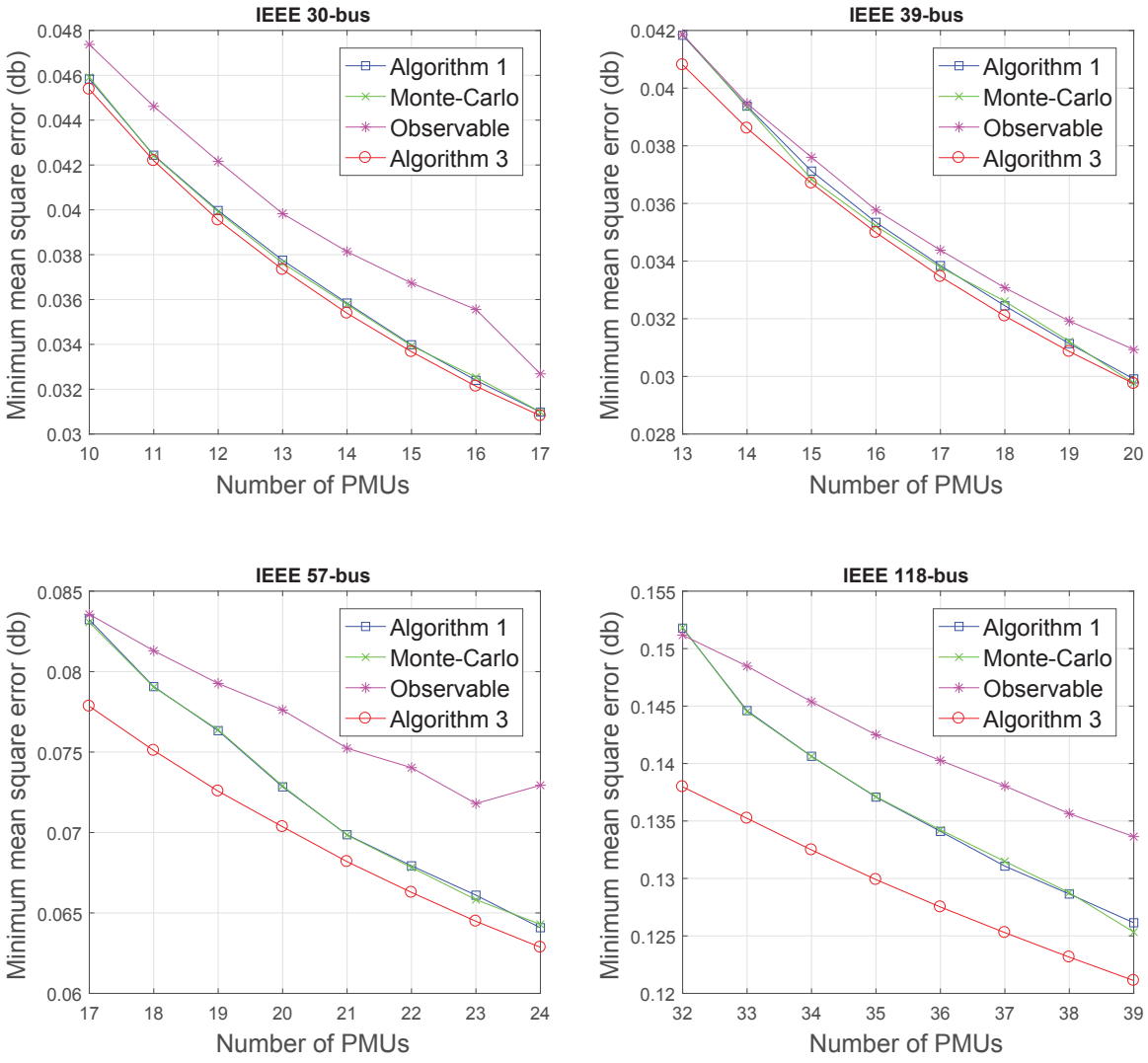

Fig. 7.1 MMSE by different methods

Table 7.2 provides numerical details of Algorithm 10, Algorithm 11 and Algorithm 12. The value of the penalized parameter μ in implementing Algorithm 10 and Algorithm 11 is given by the second and fourth column, while the average CPU time is given by the third and fifth column. The last two columns provide average CPU time

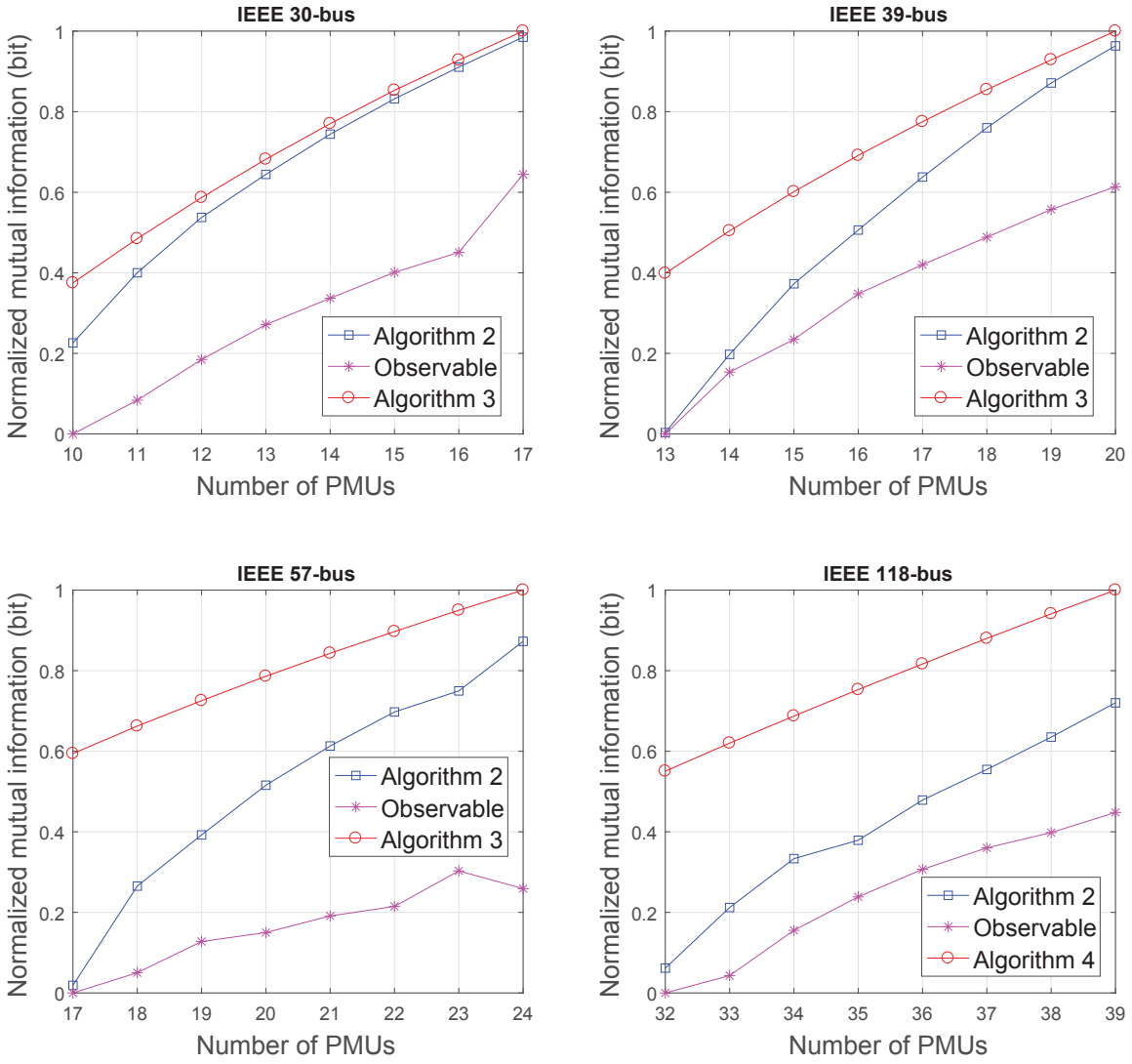

Fig. 7.2 MI by different methods

by Algorithm 12 in solving MMSE and MI. Algorithm 12 needs much less time for small-scale networks but its computational cost increases dramatically with the growth of network size. On the other hand, the CPU time of Algorithm 10 and Algorithm 11 increases moderately when the size of networks grows, demonstrating their scalability and superiority in addressing large-scale networks.

For problem (7.19), Kekatos et al [2] relaxed the integer constraint $\mathbf{x} \in \{0,1\}^N$ to the box constraint $\mathbf{x} \in [0,1]^N$ to formulate a convex problem and then round the S largest values of the solution of this convex program to 1. Obviously, their solution is hardly optimal in any sense. Fig. 7.3 compares the MMSE values of problem (7.19)

Table 7.2 Numerical details of Algorithm 10, Algorithm 11 and Algorithm 12

IEEE	Alg. 10		Alg. 11		CPU (s) of Alg. 12	
	μ	Avg. T. (s)	μ	CPU (s)	MMSE	MI
30-bus	0.1	65.78	1	62.94	4.01	3.17
39-bus	0.1	79.73	1	77.25	11.58	7.98
57-bus	1	80.47	10	81.14	49.09	46.03
118-bus	1	216.31	10	193.24	1222.11	2142.08

founded by Algorithm 12 and by Kekatos et al [2]. The former clearly outperforms the latter, especially for large scale networks.

Due to space limitation, only IEEE 30-bus and IEEE-39 networks are selected for MMSE results solved by Algorithm 10 under the constraint (7.10) of depth-of-one unobservability. Fig. 7.4 provides MMSE performance obtained via Algorithm 10 (under the constraint (7.10)) and Algorithm 12 (without any observability constraints), while Fig. 7.5 provides the number of bus left unobservable (for IEEE 30-bus). As expected, Algorithm 12 achieves a better MMSE but leaves more buses unobservable because it sacrifices buses to achieve the averaged performance.

For IEEE 57-bus network and IEEE 118-bus network, Fig.7.6 presents the number of iterations needed for the convergence of Algorithm 3 for MMSE and MI, respectively.

Given different tolerances ϵ, the required minimum number of PMUs can be obtained by Algorithm 13. For the case of $f = F_e$, the results are presented in Fig.7.7.

7.6 Conclusions

In this chapter, we have considered PMU placement optimization to minimize the mean squared error or maximize the mutual information between the measurement outputs and phasor states under a fixed number of PMUs and different observability conditions. These binary optimization problems are very computationally challenging due to high nonlinearity of the objective functions. Nevertheless, we have developed the scalable algorithms for their computational solution, which result at least in local optimal solutions. We also developed extremely efficient algorithms of very low computational

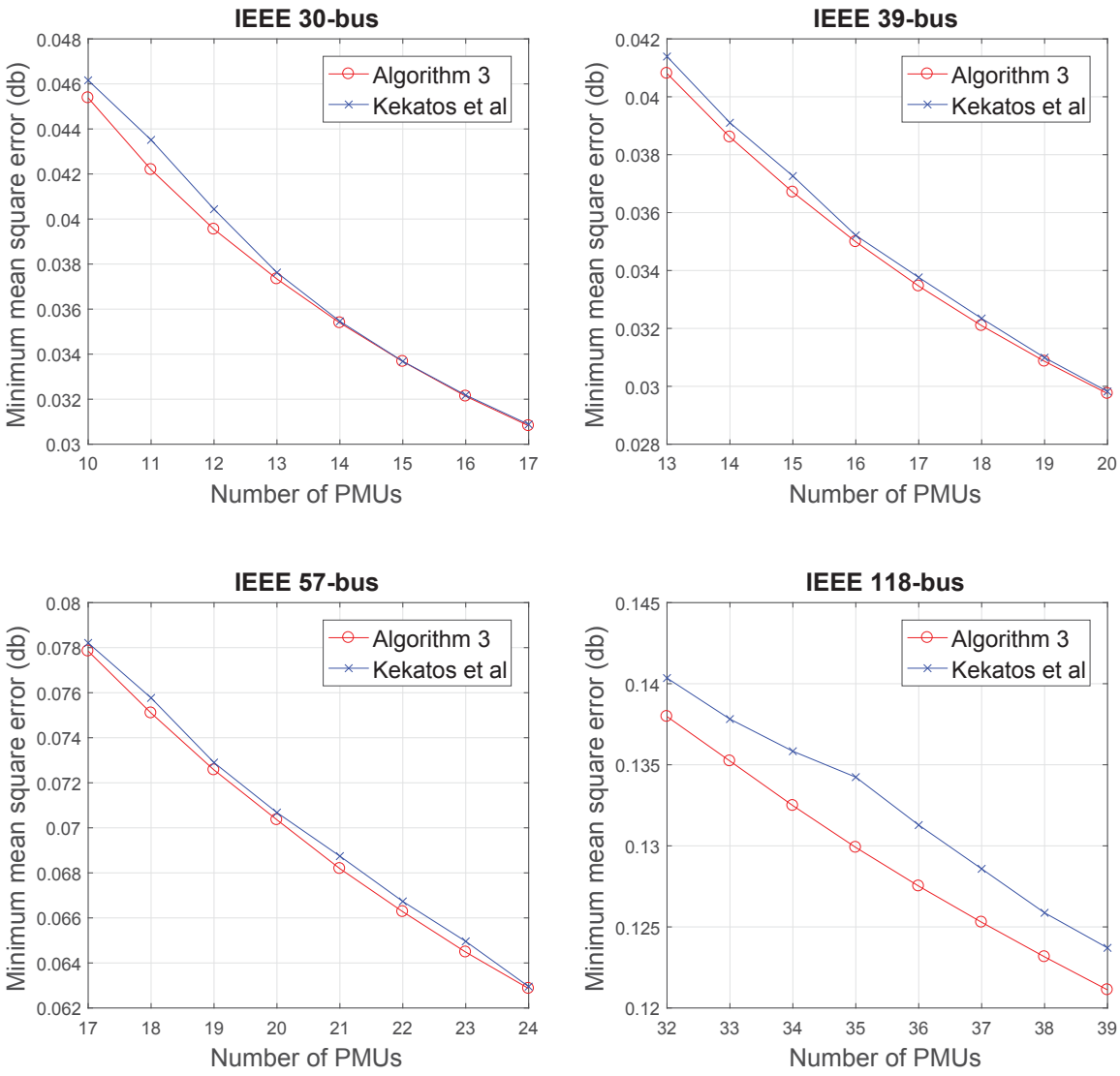

Fig. 7.3 MMSE found by Algorithm 3 and by [2]

complexity for cases of absent observability. The viability of our proposed algorithms has been confirmed through simulations with benchmark IEEE grids. The algorithmic developments for PMU placement optimization involving other practical constraints such as branch outages are under way.

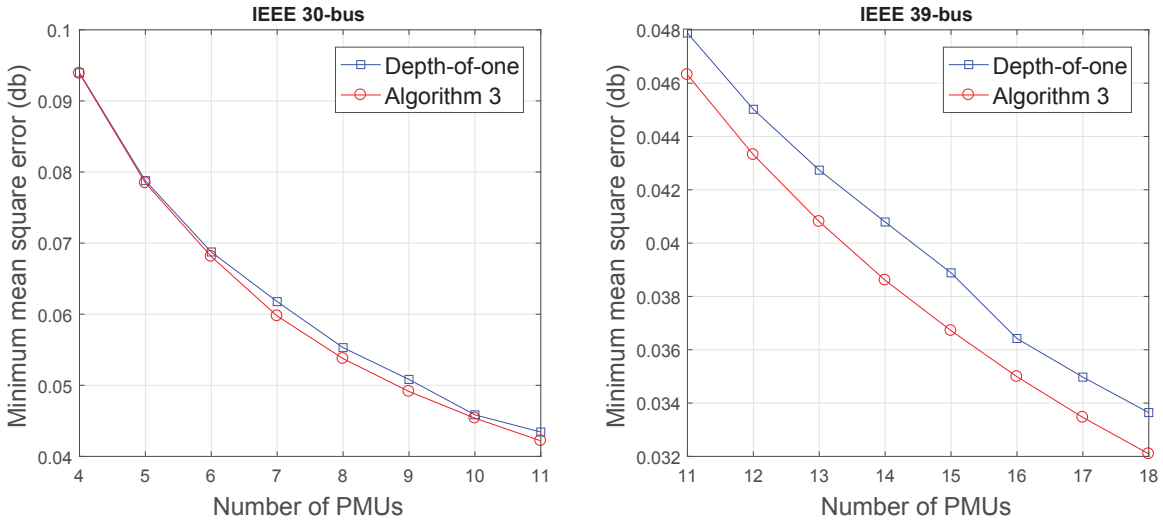

Fig. 7.4 MMSE found by Algorithm 10 under depth-of-one unobservability condition and Algorithm 12 without any observability constraints

Appendix: Fundamental Inequalities

Let $\mathbb{R}^N_+ := \{x \in \mathbb{R}^N : x_k \geq 0, k \in \mathcal{N}\}$ and $\mathsf{int}(\mathbb{R}^N_+) := \{x \in \mathbb{R}^N : x_k > 0, k \in \mathcal{N}\}$. For $A_0 \succ 0$ and $A_k \succeq 0$, $k \in \mathcal{N}$ let $\Phi(\mathbf{x}) := (A_0 + \sum_{k\in\mathcal{N}} \frac{1}{x_k} A_k)^{-1}$, and $\Psi(\mathbf{x}) := (A_0 + \sum_{k=1}^{N} x_k A_k)^{-1}$. Recall the following result [150, Th.1]:

Theorem 3 *Function $\varphi(\mathbf{x}) = Trace(\Phi(\mathbf{x}))$ is concave in the domain $\mathsf{int}(\mathbb{R}^N_+)$, so for all $\mathbf{x} \in \mathsf{int}(\mathbb{R}^N_+)$ and $\bar{\mathbf{x}} \in \mathsf{int}(\mathbb{R}^N_+)$ one has*

$$\begin{aligned} \varphi(\mathbf{x}) &\leq \varphi(\bar{\mathbf{x}}) + \langle \nabla\varphi(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle \\ &= Trace\left(\Phi^2(\bar{\mathbf{x}})A_0\right) + \sum_{k\in\mathcal{N}} \frac{x_k}{\bar{x}_k^2} Trace\left(\Phi^2(\bar{\mathbf{x}})A_k\right). \end{aligned}$$

Therefore,

$$\begin{aligned} Trace(\Psi(\mathbf{x})) &\leq Trace\left((\Psi(\bar{\mathbf{x}}))^2 A_0\right) \\ &\quad + \sum_{k\in\mathcal{N}} \frac{\bar{x}_k^2}{x_k} Trace\left((\Psi(\bar{\mathbf{x}}))^2 A_k\right). \end{aligned} \tag{7.23}$$

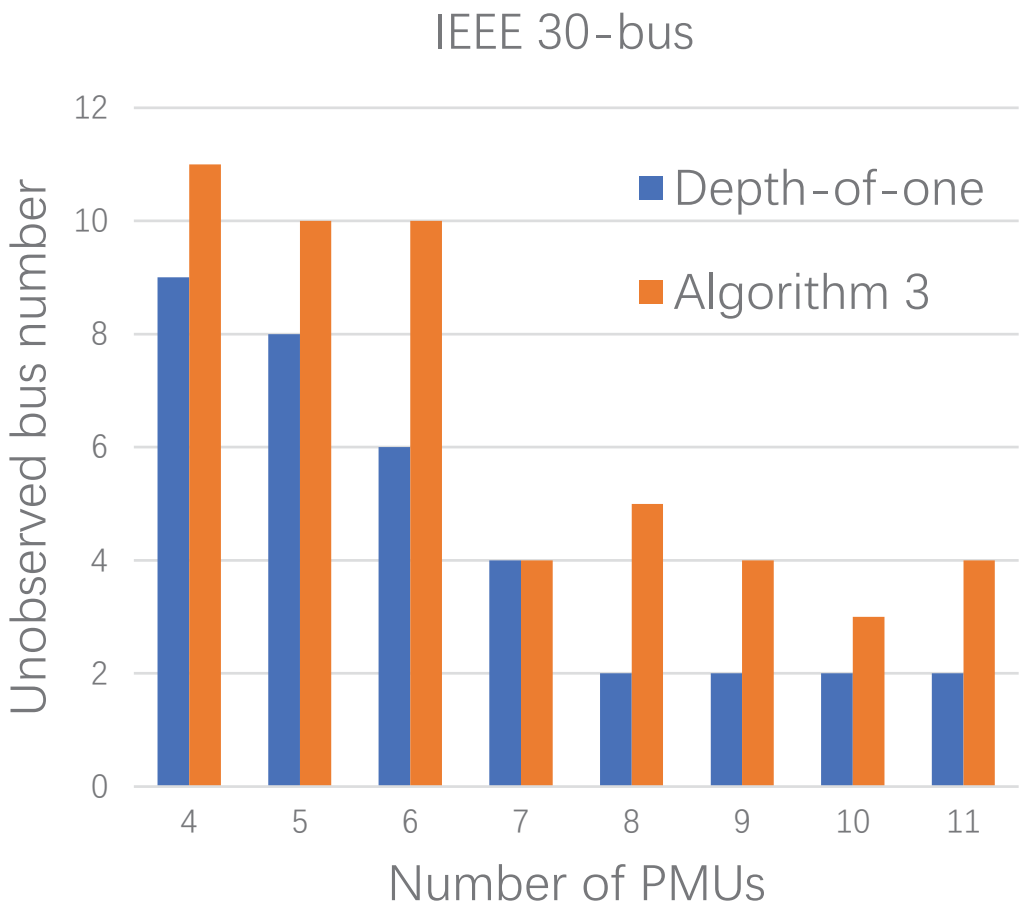

Fig. 7.5 Number of buses left unobserved by Algorithm 10 under depth-of-one unobservability condition and Algorithm 12 without any observability constraints for IEEE 30-bus network

Next,

Theorem 4 *For* $A \succ 0$ *function* $\ln|A + H\mathbf{X}^{-1}H^H|$ *is convex in* $\mathbf{X} \succ 0$.

Proof: Since $(A + H\mathbf{X}^{-1}H^H)^{-1} = A^{-1} - A^{-1}(H^H A^{-1}H + \mathbf{X})^{-1}A^{-1}$, by [151, Appendix B], function

$$f(\mathbf{X}) := A^{-1} - A^{-1}(H^H A^{-1}H + \mathbf{X})^{-1}A^{-1}$$

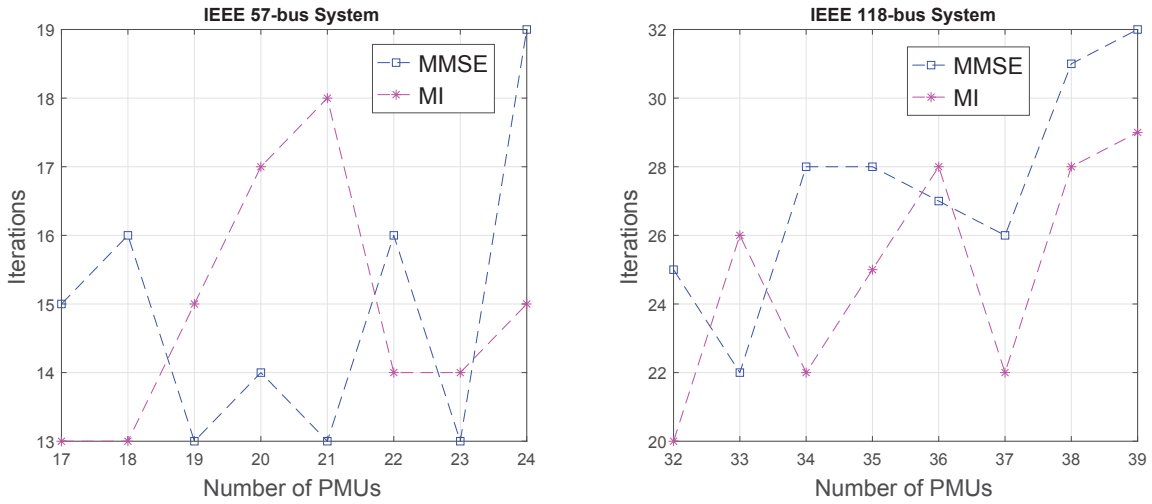

Fig. 7.6 Number of iterations required for the convergence of Algorithm 3

is concave, i.e. $f(\alpha\mathbf{X}+\beta\mathbf{Y}) \succ \alpha f(\mathbf{X})+\beta f(\mathbf{Y}) \quad \forall\, \mathbf{X} \succ 0, \mathbf{Y} \succ 0$, and $\alpha \geq 0, \beta \geq 0, \alpha+\beta = 1$. Therefore $\ln|f(\alpha\mathbf{X}+\beta\mathbf{Y})| \geq \ln|\alpha f(\mathbf{X})+\beta f(\mathbf{Y})| \geq \alpha \ln|f(\mathbf{X})| + \beta \ln|f(\mathbf{Y})|$, showing that $\ln|A+H\mathbf{X}^{-1}H^H|^{-1} = -\ln|A+H\mathbf{X}^{-1}H^H|$ is concave in $\mathbf{X}$. □

The following Theorem is a direct consequence of Theorem 4.

Theorem 5 *Function* $\phi(\mathbf{x}) = -\ln|\Phi(\mathbf{x})|$ *is convex in the domain* $\mathsf{int}(\mathbb{R}^N_+)$*, so for all* $\mathbf{x} \in \mathsf{int}(\mathbb{R}^N_+)$ *and* $\bar{\mathbf{x}} \in \mathsf{int}(\mathbb{R}^N_+)$ *one has*

$$\begin{aligned}\phi(\mathbf{x}) &\geq \phi(\bar{\mathbf{x}}) + \langle \nabla\phi(\bar{\mathbf{x}}), \mathbf{x}-\bar{\mathbf{x}}\rangle \\ &= -\ln|\Phi(\bar{\mathbf{x}})| + \mathit{Trace}\left((\Phi(\bar{\mathbf{x}}))^{-1}(\sum_{k\in\mathcal{N}}\frac{1}{\bar{x}_k}A_k)\right) \\ &\quad -\sum_{k\in\mathcal{N}}\frac{x_k}{\bar{x}_k^2}\mathit{Trace}\left((\Phi(\bar{\mathbf{x}}))^{-1}A_k\right).\end{aligned}$$

Therefore,

$$\begin{aligned}-\ln|\Psi(\mathbf{x})| \geq -\ln|\Psi(\bar{\mathbf{x}})| + \mathit{Trace}&\left(\Psi(\bar{\mathbf{x}})(\sum_{k\in\mathcal{N}}\bar{x}_k A_k)\right) \\ &-\sum_{k\in\mathcal{N}}\frac{\bar{x}_k^2}{x_k}\mathit{Trace}\left(\Psi(\bar{\mathbf{x}})A_k\right). \qquad (7.24)\end{aligned}$$

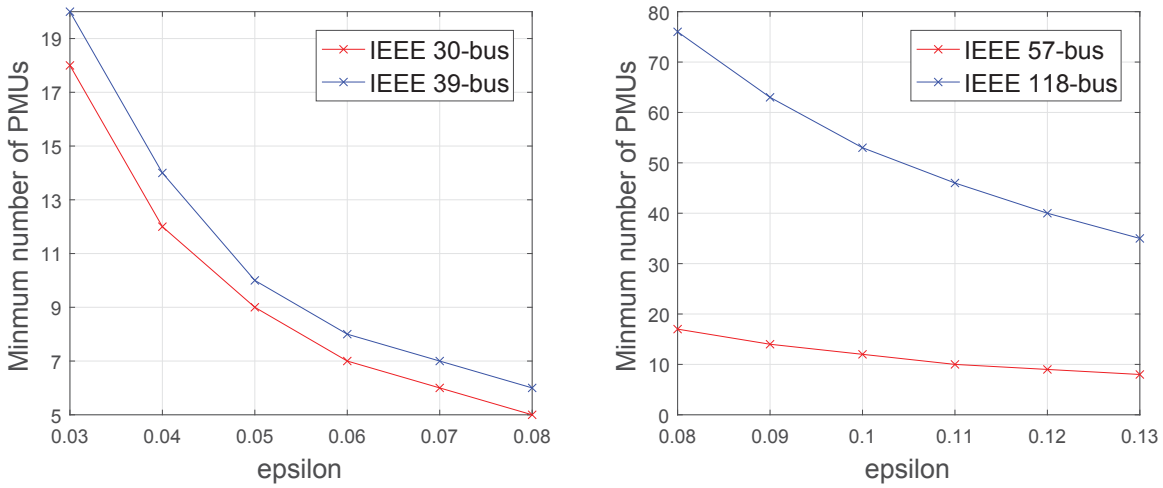

Fig. 7.7 Minimum number of PMUs required versus different values of tolerance level ϵ for MMSE

Chapter 8

Conclusions and Future Work

8.1 Conclusions

In Chapter 3, we proposed new algorithms for solving matrix-rank constrained optimization arising in reduced-order $\mathcal{H}_\infty$ LPV-LFT controller design. Unlike the previous developments, we formulated the problem as minimization of nonconvex objective function over a convex feasibility set. The global convergence of the proposed Algorithms to a local minima follows immediately from their path-following nature, while there is no difficulty for initial solutions, which are found from a semi-definite program for full-order controller synthesis. The numerical results reported for the benchmark collections have shown their merit. Their application to solutions of reduced-order generalized $\mathcal{H}_2$ LPV-LFT controllers is obvious. Their extensions to multi-objective and structured controller design are currently under development.

In Chapter 4, we have addressed the problem of designing H_∞ PID PDC for T-S systems based on a parameterized bilinear matrix inequality (PLMI). Efficient computational procedures for this PLMI have been developed. Their merit has been analysed through the benchmark examples. In the end, the effectiveness of PID PDC in smoothly stabilizing nonlinear systems has been confirmed.

In Chapter 5, we considered the nonconvex OPF problem over power transmission networks. We have shown that the proposed nonsmooth optimization algorithm (NOA) is able to overcome the shortcomings of the existing methods to compute its optimal

solution efficiently and practically even for networks with reasonably large numbers of buses.

In Chapter 6, we considered the joint PEV charging scheduling and power control for smart grids. This problem is to serve both PEVs at a competitive cost and residential power demands at a competitive operating cost, which is very difficult due to the random nature of PEVs' arrivals and demands. We have proposed a novel and easily-implemented MPC-based computational algorithm that can achieve a globally optimal solution.

In Chapter 7, we considered PMU placement optimization to minimize the mean squared error or maximize the mutual information between the measurement outputs and phasor states under a fixed number of PMUs and different observability conditions. These binary optimization problems are very computationally challenging due to high nonlinearity of the objective functions. Nevertheless, we have developed the scalable algorithms for their computational solution, which result at least in local optimal solutions. We also developed extremely efficient algorithms of very low computational complexity for cases of absent observability. The viability of our proposed algorithms has been confirmed through simulations with benchmark IEEE grids. The algorithmic developments for PMU placement optimization involving other practical constraints such as branch outages are under way.

8.2 Future work

Interesting directions for future work are presented in the following.

- In Chapter 3 and Chapter 4, we considered the reduced-order $\mathcal{H}_\infty$ LPV-LFT controller design and H_∞ PID PDC design, respectively. It is interesting to consider reduced-order $\mathcal{H}_2$ LPV-LFT controller design and H_2 PID PDC design in the future

- In Chapter 5, we considered the balanced single phase OPF problem over power transmission networks. The interesting future work can be carried out on unbalanced three-phase power network.

- In Chapter 6, we considered joint PEV charging scheduling and power control for smart grids. The control variable of PEV charging rate is setting continuously random. However, it is easy and efficient for online implementation with bang-bang charging strategy, under which PEVs either charge the maximal power rate or do not charge at all at each time slot. In future, the interesting PEV charging scheduling under the bang-bang charging strategy can be considered.

- In Chapter 7, we considered PMU placement optimization to minimize the mean squared error or maximize the mutual information between the measurement outputs and phasor states under a fixed number of PMUs and different observability conditions. But the contingency of PMU and line outage has not been addressed. The effect of PMU and line outage on the PMU placement problem will be investigated in the future.

References

[1] W. Bukhsh, A. Grothey, K. McKinnon, and P. Trodden, "Test case archive of optimal power flow (OPF) problems with local optima." http://http://www.maths.ed.ac.uk/optenergy/LocalOpt/introduction.html.

[2] V. Kekatos, G. B. Giannakis, and B. Wollenberg, "Optimal placement of phasor measurement units via convex relaxation," *IEEE Trans. Power Systems*, vol. 27, no. 3, pp. 1521–1530, 2012.

[3] P. Apkarian, D. Noll, and O. Prot, "A trust region spectral bundle method for nonconvex eigenvalue optimization," *SIAM J. Optimization*, vol. 19, no. 1, pp. 281–306, 2008.

[4] Q. T. Dinh, S. Gumussoy, W. Michiels, and M. Diehl, "Combining convex-concave decompositions and linearization approaches for solving BMIs, with application to static outputfeedback," *IEEE Trans. Automat. Control*, vol. 57, pp. 1377–1390, 2012.

[5] C. Jocz, J. Maeght, P. Panciatici, and J. C. Gilbert, "Application of the Moment-SOS approach to global optimization of the OPF problem," *IEEE Trans. Power Systems*, vol. 30, pp. 463–470, Jan. 2015.

[6] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Systems*, vol. 26, pp. 12–19, Feb 2011.

[7] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Trans. Power Systems*, vol. 30, pp. 199–211, Jan. 2015.

[8] K. Zhou and J. C. Doyle, *Essentials of robust control*, vol. 104. Prentice hall Upper Saddle River, NJ, 1998.

[9] P. Dorato, "A historical review of robust control," *IEEE Control Systems Magazine*, vol. 7, no. 2, pp. 44–47, 1987.

[10] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. automatic control*, vol. 26, no. 2, pp. 301–320, 1981.

[11] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *Int. J. Robust and Nonlinear control*, vol. 4, no. 4, pp. 421–448, 1994.

[12] F. Zheng, Q.-G. Wang, T. H. Lee, and X. Huang, "Robust PI controller design for nonlinear systems via fuzzy modeling approach," *IEEE Trans. Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 31, no. 6, pp. 666–675, 2001.

[13] P. Gil, C. Lucena, A. Cardoso, and L. B. Palma, "Gain tuning of fuzzy PID controllers for MIMO systems: A performance-driven approach," *IEEE Trans. Fuzzy System*, vol. 23, no. 4, pp. 757–768, 2015.

[14] K. Cao, X. Gao, H. K. Lam, and A. Vasilakos, "H_∞ fuzzy PID control synthesis for Tagaki-Sugeno fuzzy systems," *IET Control Theory & Applications*, vol. 10, no. 6, pp. 607–616, 2016.

[15] J. G. VanAntwerp and R. D. Braatz, "A tutorial on linear and bilinear matrix inequalities," *Journal of process control*, vol. 10, no. 4, pp. 363–385, 2000.

[16] F. Leibfritz, "Compleib: Constraint matrix optimization problem library – a collection of test examples for nonlinear semidefinite programs, control system design and related problems," tech. rep., Univ. Trier, Dept. Math, Trier, Germany, 2004.

[17] P. Apkarian and D. Noll, "Nonsmooth H_∞ synthesis," *IEEE Trans. Automatic Control*, vol. 51, no. 1, pp. 71–86, 2006.

[18] MathWorks, *Robust Control Toolbox Release 2015b*. The MathWorks, Inc., Natick, Massachusetts, United State, 2015.

[19] J. Mohammadpour and C. W. Scherer (Eds.), *Control of Linear Parameter Varying Systems with Applications*. Springer, 2012.

[20] T. Tagaki and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC-15, pp. 116–132, Jan 1985.

[21] A. C. Mary B. Cain, Richard P. O'Neill, "History of optimal power flow and formulations." https://www.ferc.gov/industries/electric/indus-act/market-planning/opf-papers/acopf-1-history-formulation-testing.pdf.

[22] J. Carpentier, "Contribution to the economic dispatch problem," *Bulletin Society Francaise Electriciens*, vol. 3, no. 8, pp. 431–447, 1962.

[23] M. Huneault and F. D. Galiana, "A survey of the optimal power flow literature," *IEEE Trans. Power Systems*, vol. 6, no. 2, pp. 762–770, 1991.

[24] J. A. Momoh, M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. part I: Nonlinear and quadratic programming approaches," *Solar Energy*, vol. 14, no. 1, pp. 96–104, 1999.

[25] K. S. Pandya and S. K. Joshi, "A survey of optimal power flow methods," *J. of Theoretical and Applied Information Technology*, vol. 4, no. 5, pp. 450–458, 2008.

[26] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Systems*, vol. 27, pp. 92–107, Jan. 2012.

[27] W. Bukhsh, A. Grothey, K. McKinnon, and P. Trodden, "Local solutions of the optimal power flow problem," *IEEE Trans. Power Systems*, vol. 28, pp. 4780–4788, Nov 2013.

[28] P. Apkarian and H. D. Tuan, "Concave programming in control theory," *J. Global Optimization*, vol. 15, pp. 243–270, Apr. 1999.

[29] M. Mesbahi, "On the rank minimization problem and its control applications," *Syst. Control Lett.*, vol. 33, pp. 31–36, 1998.

[30] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, "Hankel matrix rank minimization with applications to system identification and realization," *SIAM. J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 946–977, 2013.

[31] P. Apkarian and H. D. Tuan, "A sequential SDP/Gauss-Newton algorithm for rank-constrained LMI problems," in *Proc. of the 38-th Conference on Decision and Control (CDC), Phoenix, Arizona*, pp. 2238–2245, Dec. 1999.

[32] R. Orsi, U. Helmke, and J. B. Moore, "A Newton-like method for solving rank constrained linear matrix inequalities," *Automatica*, vol. 42, pp. 1875–1882, 2006.

[33] B. Vandereycken, P. A. Absil, and S. Vandewalle, "A Riemannian geometry with cmplete geodesics for the set of positive semi-definite matrices of fixed rank," *IMA J. Numerical Analysis*, vol. 33, pp. 481–514, 2013.

[34] P. Apkarian and H. D. Tuan, "Robust control via concave minimization: local and global algorithms," *IEEE Trans. Automatic Control*, vol. 45, pp. 299–305, 2000.

[35] H. D. Tuan, P. Apkarian, and Y. Nakashima, "A new Lagrangian dual global optimization algorithm for solving bilinear matrix inequalities," *Int. J. Nonlinear Robust Controls*, vol. 10, pp. 561–578, 2000.

[36] H. D. Tuan and P. Apkarian, "Low nonconvexity-rank bilinear matrix inequalities: algorithm and applications in robust controller and structure designs," *IEEE Trans. Automat. Control*, vol. 45, pp. 2111–2117, 2000.

[37] D. Henrion, J. Lefberg, M. Kocvara, and M. Stingl, "Solving polynomial static output feedback problems with PENBMI," in *Proc. joint IEEE Conf. Decision and Control and Europ. Control Conf.*, pp. 7581–7586, 2005.

[38] S. Gumussoy, D. Henrion, M. Millstone, and M. L. Overton, "Multi-objective robust control with HIFOO 2.0," in *Proc. IFAC Symp. on Robust Control Design*, 2009.

[39] V. Blondel and J. N. Tsitsiklis, "NP-hardness of some linear control design problems," *SIAM J. Control Optimization*, vol. 35, pp. 2118–2127, 1997.

[40] Q. T. Dinh, W. Michiels, S. Gros, and M. Diehl, "An inner convex approximation algorithm for BMI optimization and applications in control," in *Proc. 51th IEEE Conf. Decision and Control*, pp. 3576–3581, 2012.

[41] H. D. Tuan, P. Apkarian, T. Narikiyo, and Y. Yamamoto, "Parameterized linear matrix inequality techniques in fuzzy control system design," *IEEE Trans. Fuzzy Systems*, vol. 9, no. 2, pp. 324–332, 2001.

[42] H. D. Tuan, P. Apkarian, T. Narikiyo, and M. Kanota, "New fuzzy control model and dynamic output feedback paralell distributed compensation," *IEEE Trans. Fuzzy Systems*, vol. 12, no. 2, pp. 13–21, 2004.

[43] P. Apkarian and H. D. Tuan, "Robust control via concave optimization: Local and global algorithms," *IEEE Trans. Autom. Control*, vol. 45, pp. 299–305, Feb. 2000.

[44] K. H. Ang, G. Chong, and Y. Li, "PID control system analysis, design, and technology," *IEEE Trans. Control Syst. Tech.*, vol. 13, no. 4, pp. 559–576, 2005.

[45] M. Araki and H. Taguchi, "Two-degree-of-freedom PID controller," *Int. J. Control, Automation and Systems*, vol. 4, pp. 401–411, 2003.

[46] P. Gahinet and P. Apkarian, "Structured H_∞ synthesis in MATLAB," in *Proc. IFAC 2011, Milan*, pp. 1–5, Jun. 2011.

[47] O. Garpinger and T. Hagglund, "Software-based optimal PID design with robustness and noise sensitivity constraints," *J. Process Control*, vol. 33, pp. 90–101, 2015.

[48] S. Boyd, M. Hast, and K. J. Astrom, "MIMO PID tuning via iterated LMI restriction," *Int. J. Robust and Nonlinear Control*, vol. 26, pp. 1718–1731, 2016.

[49] S. Hosoe, H. D. Tuan, and T. N. Nguyen, "2D bilinear programming for robust PID/DD controller design," *Int. J. Robust Nonlinear Control*, vol. 27, pp. 461–482, 2017.

[50] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Elect. Power Energy Syst.*, vol. 30, no. 6-7, pp. 383–392, 2008.

[51] H. Tuy and H. D. Tuan, "Generalized S-lemma and strong duality in nonconvex quadratic programming," *J. Global Optimization*, vol. 56, pp. 1045–1072, 2013.

[52] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contigency-constrained optimal power flow problem," *IEEE Trans. Power Systems*, vol. 30, 2015.

[53] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[54] H. D. Tuan, N. T. Hoang, H. Q. Ngo, H. Tuy, and B. Vo, "A dual frequency-selective bounded real lemma and its applications to iir filter design," in *Proc. 45th IEEE Conf. Control and Decision (CDC)*, pp. 1478–1485, 2006.

[55] Bamberger and et al, *Vision and Strategy for Europe's Electricity Networks of the Future*. Office for Official Publications of the European Communities, 2006.

[56] P. Hallberg and et al, *Smart Grids and Networks of the Future - EURELECTRIC Views*. Ref: 2009-030-0440, May 2009.

[57] M. Hashmi, S. Hänninen, and K. Mäki, "Survey of smart grid concepts, architectures, and technological demonstrations worldwide," in *Proc. 2011 IEEE PES Conf. on Innovative Smart Grid Technologies (ISGT Latin America)*, pp. 1–7, Oct 2011.

[58] D. K. Molzahn, J. T. Holzer, B. C. Lesieutre, and C. L. DeMarco, "Implementation of a large-scale optimal power flow solver based on semidefinite programming," *IEEE Trans. Power Systems*, vol. 28, pp. 3987–3998, Nov 2013.

[59] M. S. Andersen, A. Hansson, and L. Vandenberghe, "Reduced-complexity semidefinite relaxations of optimal power flow problems," *IEEE Trans. Power Systems*, vol. 29, pp. 1855–1863, Jul 2014.

[60] W. Tusha, C. Yuen, S. Huang, D. B. Smith, and H. V. Poor, "Cost minimization of charging stations with photovoltaics: An approach with EV classification," *IEEE Trans. Intell. Transport. Systs.*, vol. 17, no. 1, pp. 156–169, 2016.

[61] W. Tang, S. Bi, and Y. J. Zhang, "Online charging scheduling algorithms of electric vehicles in smart grid: An overview," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 76–83, 2016.

[62] Y. Wang, W. Saad, Z. Han, H. V. Poor, and T. Basar, "A game-theoretic approach to energy trading in the smart grid," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1439–1450, 2014.

[63] L. Yang, J. Zhang, and H. V. Poor, "Risk-aware day-ahead scheduling and real-time dispatch for electric vehicle charging," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 693–702, 2014.

[64] Y. Wang, W. Saad, N. B. Mandayam, and H. V. Poor, "Load shifting in the smart grid: To participate or not?," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2604–2614, 2016.

[65] S. Lakshminarayana, Y. Xu, H. V. Poor, and T. Q. S. Quek, "Cooperation of storage operation in a power network with renewable generation," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2108–2122, 2016.

[66] G. Li and X.-P. Zhang, "Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 492–499, 2012.

[67] H. Mohsenian-Rad and M. Ghamkhari, "Optimal charging of electric vehicles with uncertain departure times: A closed-form solution," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 940–942, 2015.

[68] H. Xing, M. Fu, Z. Lin, and Y. Mou, "Decentralized optimal scheduling for charging and discharging of plug-in electric vehicles in smart grids," *IEEE Trans. Power Systems*, vol. 31, no. 5, pp. 4118–4127, 2016.

[69] W. Tang and Y. J. A. Zhang, "A model predictive control approach for low-complexity electric vehicle charging scheduling: optimality and scalability," *IEEE Trans. Power Systems*, vol. 32, no. 2, pp. 1050–1063, 2017.

[70] Y. Kim, J. Kwak, and S. Chong, "Dynamic pricing, scheduling, and energy management for profit maximization in PHEV charging stations," *IEEE Trans. Vehic. Tech.*, vol. 66, no. 2, pp. 1011–1026, 2017.

[71] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Trans. smart grid*, vol. 1, no. 1, pp. 20–27, 2010.

[72] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York: Springer, 2008.

[73] J. A. Momoh, R. Adapa, and M. E. El-Hawary, "A review of selected optimal power flow literature to 1993. i. nonlinear and quadratic programming approaches," *IEEE Trans. on Power Systems*, vol. 14, no. 1, pp. 96–104, 1999.

[74] J. Zhao, G. Zhang, K. Das, G. N. Korres, N. M. Manousakis, A. K. Sinha, and Z. He, "Power system real-time monitoring by using PMU-based robust state estimation method," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 300–309, 2016.

[75] C.-W. Ten, A. Ginter, and R. Bulbul, "Cyber-based contingency analysis," *IEEE Trans. Power Systems*, vol. 31, no. 4, pp. 3040–3050, 2016.

[76] B. Gou, "Optimal placement of PMUs by integer linear programming," *IEEE Trans. Power Systems*, vol. 23, no. 3, pp. 1525–1526, 2008.

[77] B. Gou, "Generalized integer linear programming formulation for optimal PMU placement," *IEEE Trans. Power Systems*, vol. 23, no. 3, pp. 1099–1104, 2008.

[78] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," *IEEE Trans. Power Delivery*, vol. 20, no. 4, pp. 2381–2388, 2005.

[79] S. Chakrabarti and E. Kyriakides, "Optimal placement of phasor measurement units for power system observability," *IEEE Trans. Power Systems*, vol. 23, no. 3, pp. 1433–1440, 2008.

[80] M. Hajian, A. M. Ranjbar, T. Amraee, and B. Mozafari, "Optimal placement of PMUs to maintain network observability using a modified bpso algorithm," *Int. J. Elect. Power Energy Syst.*, vol. 33, no. 1, pp. 28–34, 2011.

[81] S. Chakrabarti, E. Kyriakides, and D. G. Eliades, "Placement of synchronized measurements for power system observability," *IEEE Trans. Power Delivery*, vol. 24, no. 1, pp. 12–19, 2009.

[82] K. G. Khajeh, E. Bashar, A. M. Rad, and G. B. Gharehpetian, "Integrated model considering effects of zero injection buses and conventional measurements on optimal PMU placement," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 1006–1013, 2017.

[83] M. J. Rice and G. T. Heydt, "Power systems state estimation accuracy enhancement through the use of PMU measurements," in *IEEE PES Transmission and Distribution Conf. and Exposition*, 2006.

[84] Q. Li, T. Cui, Y. Weng, R. Negi, F. Franchetti, and M. D. Ilic, "An information-theoretic approach to PMU placement in electric power systems," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 446–456, 2013.

[85] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, vol. 15. SIAM, 1994.

[86] P. Apkarian, P. Pellanda, and H. D. Tuan, "Mixed H_2/H_∞ multi-channel linear parameter-varying control in discrete time," *Syst. Control Lett.*, vol. 41, p. 333–346, 2000.

[87] P. Apkarian, P. Pellanda, and H. D. Tuan, "Missile autopilot design via a multi-channel LFT/LPV control method," *Int. J. Robust Nonlinear Control*, vol. 12, pp. 1–20, 2002.

[88] H. D. Tuan, P. Apkarian, T. Narikiyo, and M. Kanota, "New fuzzy control model and dynamic output feedback parallel distributed compensation," *IEEE Trans. Fuzzy Systems*, vol. 13-21, pp. 2975–2984, Feb 2004.

[89] H. Tuy, *Convex Analysis and Global Optimization.* Kluwer Academic, 1997.

[90] A. H. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Signal Processing*, vol. 60, pp. 2941–2951, Jun. 2012.

[91] A. H. Phan, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Beamforming optimization in multi-user amplify-and-forward wireless relay networks," *IEEE Trans. Wireless Communications*, vol. 11, pp. 1510–1520, Apr. 2012.

[92] Y. Shi, H. D. Tuan, S. Su, and H. H. M. Tam, "Nonsmooth optimization for optimal power flow over transmission networks," in *Proc. 2015 IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, Dec. 2015.

[93] P. Apkarian and D. Noll, "Nonsmooth optimization for multiband frequency domain control design," *Automatica*, vol. 43, no. 4, pp. 724–731, 2007.

[94] H. Pfifer and P. Seiler, "Less conservative robustness analysis of linear parameter varying systems using integral quadratic constraints," *Int. J. Robust Nonlinear Control*, vol. 26, 2016.

[95] J. Veenman and C. W. Scherer, "IQC-synthesis with general dynamic multipliers," *Int. J. Robust Nonlinear Control*, vol. 24, pp. 3027–3056, Jul. 2014.

[96] Z. Emedi and A. Karimi, "Fixed-order LPV controller design for LPV systems by convex optimization," in *Proc. of 5th IFAC Symp. System Stricture and Control*, pp. 1–6, 2013.

[97] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, and C. Sagastizábal, *Numerical Optimization – Theoretical and Practical Aspects (second edition)*. Springer, 2006.

[98] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[99] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programms," *Operations Research*, vol. 26, pp. 681–683, Jul. 1978.

[100] L. E. Ghaoui, F. Oustry, and M. AitRami, "A cone complementary linearization algorithm for static output feedback and related problems," *IEEE Trans. Automat. Control*, vol. 42, pp. 1171–1176, 1997.

[101] J. Lofberg, "Yalmip : a toolbox for modeling and optimization in matlab," in *IEEE Inter. Symp. on Computer Aided Control Systems Design,*, pp. 284–289, Sept 2004.

[102] J. F. Sturm, "Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones," *Optimiz. Meth. Softw*, vol. 11, pp. 625–653, 1999.

[103] F. Leibfritz, "Compleib: Constrained matrix optimization problem library," 2006.

[104] R. T. Bupp, D. S. Bernstein, and V. T. Coppola, "A benchmark problem for nonlinear control design: Problem, experimental testbed, and passive nonlinear compensation," in *Proc. 1995 Amer. Control Conf., Seattle, WA*, p. 4363–4367, 1995.

[105] C. Scherer, P. Gahinet, and M. Chilali, "Multi-objective output-feedback control via LMI optimization," *IEEE Trans. Automat. Control*, vol. 42, pp. 896–911, 1997.

[106] P. Apkarian, V. Bompart, and D. Noll, "Non-smooth structured control design with application to PID loop-shaping of a process," *Int. J. Nonlinear Robust Control*, vol. 17, pp. 1320–1342, 2007.

[107] U. Rashid, H. D. Tuan, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Communications*, vol. 62, pp. 488–499, 2014.

[108] H. A. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Signal Processing*, vol. 60, pp. 2941–2951, Jun. 2012.

[109] S. G. Cao, N. W. Rees, and G. Feng, "Stability analysis and design for a class of continuous-time fuzzy control systems," *Int. J. Control*, vol. 64, no. 6, pp. 1069–1087, 1996.

[110] K. Tanaka and H. O. Wang, *Fuzzy control systems design and analysis: a linear matrix inequality approach*. John Wiley & Sons, 2004.

[111] K. Tanaka, T. Ikeda, and H. O. Wang, "Fuzzy regulators and fuzzy observers: relaxed stability conditions and LMI-based designs," *IEEE Trans. Fuzzy System*, vol. 6, no. 2, pp. 250–265, 1998.

[112] H. D. Tuan, P. Apkarian, S. Hosoe, and H. Tuy, "D.c. optimization approach to robust controls: the feasibility problems," *Inter. J. of Control*, vol. 73, pp. 89–104, Feb. 2000.

[113] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local d.c. programming," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 510–515, Feb. 2012.

[114] L. T. H. An and P. D. Tao, "The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems," *Annals of Operations Research*, vol. 133, no. 104, pp. 23–45, 2005.

[115] J. B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms (second edition)*. Springer, 1996.

[116] N. Chen, C. W. Tan, and T. Q. Quek, "Electric vehicle charging in smart grid: Optimality and valley-filling algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1073–1083, 2014.

[117] Y. Shi, H. D. Tuan, H. Tuy, and S. Su, "Global optimization for optimal power flow over transmission networks," *J. Global Optimization. (to appear)*, 2017.

[118] J. F. Franco, M. J. Rider, and R. Romero, "A mixed-integer linear programming model for the electric vehicle charging coordination problem in unbalanced electrical distribution systems," *IEEE Trans. Smart Grid*, vol. 6, pp. 2200–2210, Sept 2015.

[119] L. Zhang, V. Kekatos, and G. B. Giannakis, "Scalable electric vehicle charging protocols," *IEEE Trans. Power Systems*, vol. 32, no. 2, pp. 1451–1462, 2017.

[120] E. F. Camacho and C. Bordons, *Model Predictive Control*. Springer: Springer-Verlag, 2004.

[121] A. Mesbah, "Stochastic model predictive control: An overview and perspectives for future research," *IEEE Control Systems Mag.*, vol. 36, no. 6, pp. 30–44, 2016.

[122] J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach, 2nd Edition*. Springer-Verlag New York, 1998.

[123] H. D. Tuan, A. Savkin, T. Nguyen, and H. T. Nguyen, "Decentralised model predictive control with stability constraints and its application in process control," *J. of Process Control*, vol. 26, pp. 73–89, 2015.

[124] A. H. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Sign. Process.*, vol. 60, no. 6, pp. 2941–2951, 2012.

[125] A. A. Nasir, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, "Beamforming design for wireless information and power transfer systems: Receive power-splitting versus transmit time-switching," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 876–889, 2017.

[126] Y. Shi, H. D. Tuan, and P. Apkarian, "Nonconvex spectral optimization algorithms for reduced-order $\mathcal{H}_\infty$LPV-LFT controllers," *Int. J. Nonlinear Robust Control*, vol. 27, 2017.

[127] "Distribution system operators." https://www.swissgrid.ch/swissgrid/en/home/experts/dso.html. Accessed: 2017-11-27.

[128] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999.

[129] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1." http://cvxr.com/cvx, Mar. 2014.

[130] C. Jin, J. Tang, and P. Ghosh, "Optimizing electric vehicle charging: A customer's perspective," *IEEE Trans. Veh. Tech.*, vol. 62, pp. 2919–2927, Sept 2013.

[131] J. Zheng, X. Wang, K. Men, C. Zhu, and S. Zhu, "Aggregation model-based optimization for electric vehicle charging strategy," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 1058–1066, 2013.

[132] L. Chen, C. Y. Chung, Y. Nie, and R. Yu, "Modeling and optimization of electric vehicle charging load in a parking lot," in *2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 1–5, Dec 2013.

[133] C. S. Antúnez, J. F. Franco, M. J. Rider, and R. Romero, "A new methodology for the optimal charging coordination of electric vehicles considering vehicle-to-grid technology," *IEEE Trans. Sustainable Energy*, vol. 7, pp. 596–607, April 2016.

[134] "Tesla model s." https://en.wikipedia.org/wiki/Tesla_Model_S. Accessed: 2017-06-06.

[135] "Australian energy market operator." https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Data-dashboard#price-demand. Accessed: 2017-06-06.

[136] G. Krumpholz, K. Clements, and P. Davis, "Power system observability: a practical algorithm using network topology," *IEEE Trans. Power Apparatus and Systems*, no. 4, pp. 1534–1542, 1980.

[137] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow (P-OPF) with gaussian and gamma distributions," *IEEE Trans. Power Systems*, vol. 20, no. 2, pp. 773–781, 2005.

[138] J. Grainger and J. W. Stevenson, *Power System Analysis*. New York: McGraw-Hill, 1994.

[139] J. De La Ree, V. C. anf J. Thorp, and A. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Trans. Smart Grid*, vol. 1, pp. 20–27, Jun. 2010.

[140] H. V. Poor, *An Introduction to Signal Detection and Estimation (second edition).* New York: Springer-Verlag, 1994.

[141] H. D. Tuan, D. H. Pham, B. Vo, and T. Q. Nguyen, "Entropy of general Gaussian distributions and MIMO channel capacity maximizing precoder and decoder," in *Proc. 2007 IEEE Inter. Conf. Acous. Speech Signal Process (ICASSP 07)*, pp. III–325–III–328, May 2007.

[142] R. Kavasseri and S. K. Srinivasan, "Joint placement of phasor and power flow measurements for observability of power systems," *IEEE Trans. Power Systems*, vol. 26, no. 4, pp. 1929–1936, 2011.

[143] M. Göl and A. Abur, "Observability and criticality analyses for power systems measured by phasor measurements," *IEEE Trans. Power Systems*, vol. 28, no. 3, pp. 3319–3326, 2013.

[144] E. Che, H. D. Tuan, and H. H. Nguyen, "Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks," *IEEE Trans. Wirel. Commun.*, vol. 13, pp. 5481–5495, Oct. 2014.

[145] H. H. M. Tam, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, "Joint load balancing and interference management for small-cell heterogeneous networks with limited backhaul capacity," *IEEE Trans. Wirel. Commun.*, vol. 16, pp. 872–884, Feb. 2017.

[146] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Trans. Power Systems*, vol. 21, pp. 1608–1615, Nov 2006.

[147] H. D. Tuan, T. T. Son, H. Tuy, and H. H. Nguyen, "Optimum multi-user detection by nonsmooth optimization," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3444–3447, 2011.

[148] B. A. Cipra, "The best of the 20th century: editors name top 10 algorithms," *SIAM News*, vol. 33, pp. 1–2, Dec. 2000.

[149] "CPLEX optimizer." https://www.ibm.com/analytics/data-science/prescriptive-analytics/cplex-optimizer. Accessed: 2018-02-14.

[150] J. A. Bengua, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Joint sensor and relay power control in tracking Gaussian mixture targets by wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 492–506, 2018.

[151] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Commun.*, vol. 62, pp. 488–499, Feb. 2014.