Delft University of Technology

Distributed Data-Driven Decision Making in Uncertain Networked Systems with Applications in Smart Energy Systems

Rostampour, Vahab

DOI
10.4233/uuid:b620d797-9e41-4ea4-8fd2-6375aab609a9

Publication date
2018

Document Version
Final published version

Citation (APA)
Rostampour, V. (2018). *Distributed Data-Driven Decision Making in Uncertain Networked Systems with Applications in Smart Energy Systems*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:b620d797-9e41-4ea4-8fd2-6375aab609a9

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Distributed Data-Driven Decision Making in Uncertain Networked Systems

with Applications in Smart Energy Systems

Vahab Rostampour Samarin

Cover illustration: The fore- and background present uncertain networked systems such that the blur nodes show the local (private) uncertainty sources and the blur links between neighboring nodes present the concept of common uncertainty sources.
Cover design: *Vahab Rostampour Samarin*

DISTRIBUTED DATA-DRIVEN DECISION MAKING IN UNCERTAIN NETWORKED SYSTEMS

WITH APPLICATIONS IN SMART ENERGY SYSTEMS

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof.dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Monday 24 September 2018 at 10 a.m.

by

Vahab ROSTAMPOUR SAMARIN

Master of Science in Automation Engineering, Politecnico di Milano, Italy
born in Shaherey, Iran.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Dr. ir. T. Keviczky	Delft University of Technology, promotor
Prof. dr. ir. N. van de Wouw	Delft University of Technology, promotor
Independent members:	
Prof. dr. ir. B. De Schutter	Delft University of Technology
Prof. dr. ir. P. Palensky	Delft University of Technology
Prof. dr. ir. R. Scattolini	Politecnico di Milano, Italy
Prof. dr. ir. A. Teixeira	Uppsala University, Sweden

This dissertation has been completed in partial fulfillment of the requirements of the Dutch Institute of Systems and Control (DISC) for graduate studies. The work presented in this dissertation has been supported by the Uncertainty Reduction in Smart Energy Systems (URSES) research program funded by the Dutch organization for scientific research (NWO) and Shell under the project Aquifer Thermal Energy Storage Smart Grids (ATES-SG) under grant agreement number 408-13-030.

ISBN 978-94-6186-951-7

Copyright © 2018 by Vahab Rostampour Samarin.

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by an means, electronics or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission of the author.

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

Printed in the Netherlands.

To my wonderful parents,
Ahmad and Goli,

... and my beloved wife,
Zahra

To my sweet Alma.

SUMMARY

Due to recent advances in technology and irreversible societal trends many applications in control engineering become increasingly complex and interconnected. In addition to the ubiquity of large-scale systems that can be modeled as a network of interconnected systems, another challenge relates to imperfect models – for instance due to unknown parameters or environmental conditions – which are regarded in general as modeling uncertainties. Introducing uncertainties in parameters of a constrained optimal control design problem can increase the difficulty of finding an optimal solution. Ignoring uncertainties, however, can lead to results that may cause significant damages or losses in real-world applications, such as smart thermal grids and power networks.

A standard practice in robust control design is to consider a so-called worst-case (robust) approach for the uncertainties of a given system. However, this gives rise to some limitations and bottlenecks in practical applications, e.g., conservatism and computational complexity. Nowadays, it is becoming easier to collect a large number of different types of data subject to modeling uncertainty, thanks to the availability of cheaper and more pervasive sensors. These sensors are able to measure many different types of variables. This presents novel challenges and opportunities for control design engineering, and highlights the necessity of introducing a new paradigm that relies on using available historical data or the so-called scenarios.

This dissertation aims to develop a rigorous distributed approach to decision making using scenario-based techniques for large-scale networks of interconnected uncertain dynamical systems (called agents). A scenario program is a finite-dimensional optimization problem in which an objective function is minimized under constraints that are associated with finitely many, independently and identically distributed (i.i.d.), scenarios of a random parameter. Theoretical and practical interest in scenario programs originates from the fact that these problems are typically efficiently solvable while being closely related to robust and chance-constrained programs. In the former, the constraint is enforced for all admissible random parameters, whereas in the latter, the constraint is enforced up to a given level of probability. However, finding solutions of the resulting large-scale scenario optimization problem for uncertain networked systems poses several difficulties, e.g., computational cost for a central control unit.

The main contribution of this dissertation is the design of a technique to decompose a large-scale scenario program into small-scale distributed scenario programs for each agent. Building on existing results in literature, we provide novel guarantees to quantify the robustness of the resulting solutions in a distributed framework. In this setting, each agent needs to exchange some information with its neighboring agents that is necessary due to the statistical learning features of the proposed setup. However, this inter-agent communication scheme might give rise to some concerns about the agents' private information. We therefore present a novel privatized distributed framework, based on the so-called differential privacy concept, such that each agent can share requested infor-

mation while preserving its privacy. In addition, a soft communication scheme based on a set parametrization technique, along with the notion of probabilistically reliable set, is introduced to reduce the required communication burden. Such a reliability measure is incorporated into the feasibility guarantees of agent decisions in a probabilistic sense. The theoretical guarantees of the proposed distributed scenario-based decision making framework coincide with the centralized counterpart, however the scaling of the results with the number of agents remains an issue.

Motivated by an application to Smart Thermal Grids (STGs), the second contribution of this dissertation is the design of a distributed data-driven energy management framework for building climate comfort systems that are interconnected in a grid via Aquifer Thermal Energy Storage (ATES) systems. The ATES system is considered as a seasonal storage system that can be a heat source or sink, or a storage for thermal energy. In STGs, the objective is to keep the energy balance between uncertain thermal energy demand and production units of individual buildings. This requires coordination between multiple buildings and long-term planning often months in advance in the presence of two types of uncertainty, namely local (private) and common uncertainty sources. While the private uncertainty source refers to uncertain thermal energy demand of individual buildings, the common uncertainty source describes the uncertain common resource pool (ATES) between neighbors.

The third contribution of this dissertation is the application of our proposed approach to the problem of reserve scheduling for power networks with renewable generation based on an AC optimal power flow model. We first formulate such a problem using stochastic semidefinite programming (SDP) in infinite-dimensional space, which is in general computationally intractable. Using a novel affine policy, we approximate the infinite-dimensional SDP as a tractable finite-dimensional SDP, and explicitly quantify the performance of the approximation. We then use the geographical pattern of the power system to decompose the large-scale system into a multi-area power network. A consensus form of the Alternating Direction Method of Multipliers (ADMM) is provided to find a feasible solution for both local and overall multi-area network. Using the proposed distributed data-driven decision making framework, each area can have its own information about uncertainties to achieve local feasibility certificates, while conforming to the overall feasibility of the multi-area power network under mild conditions.

The fourth contribution of this dissertation is the design of a distributed anomaly detection strategy for a network of interconnected uncertain nonlinear systems. The networked system being monitored is modeled as the interconnection of overlapping agents that share some state components. For each agent, a local threshold is then designed based on the concept of probabilistic set approximation using polynomial superlevel sets. The threshold set is parametrized in a way to bound arbitrarily well the residuals produced in healthy condition by an observer based residual generator. We also formulate a second problem to maximize the sensitivity of the obtained threshold set, with respect to the possible signature of faulty events. In order to achieve this, each agent needs to communicate some information with neighboring agents to minimize their false alarm ratio. Therefore, the proposed privatized distributed framework together with the so-called soft communication scheme is applied, firstly, to preserve agents' privacy, and secondly, to reduce the required communication.

The proposed contributions aim to bridge the gap between theoretical and practical challenges in the application of large-scale scenario programs for uncertain networked systems and bring data-driven decision making strategies closer to actual implementation in distributed interconnected uncertain systems.

SAMENVATTING

Vanwege recente ontwikkelingen in technologie en maatschappij zijn veel applicaties in de regeltechniek steeds complexer geworden. Met de opkomst van grootschalige systemen, welke gemodelleerd kunnen worden als een netwerk van gekoppelde systemen, is er een uitdaging in het modeleren van imperfecties, zoals de onzekerheid in de parameters of de omgeving condities. Aan de ene kant kan het bestaan van onzekere parameters in een begrensd optimaal regelaarontwerp probleem het vinden van een optimale oplossing vermoeilijken. Aan de andere kant, het negeren van onzekerheden kan leiden tot resultaten die significante schade of verliezen aan kunnen brengen in echte applicaties, zoals bijvoorbeeld *smart thermal grids* en energienetwerken.

Een standaard methode binnen de robuuste regeltechniek is het uitgaan van het zogenaamde worst-case (robuust) aanpak van onzekerheden in een gegeven systeem. In de praktijk geeft zo'n methode een aantal limitaties, bijvoorbeeld conservatisme en het uirvoeren van complexe berekeningen. Met dank aan de beschikbaarheid van goedkopere en meer universele sensoren wordt het tegenwoordig makkelijker om grote aantallen verschillende datatypes te verkrijgen, welke onderhevig zijn aan meetonzekerheden. Deze sensoren zijn in staat om veel verschillende typen variabelen te meten. Dit resulteert in nieuwe uitdagingen en mogelijkheden voor regeltechnici en benadrukt de noodzaak van een nieuw paradigma dat gebruik maakt van beschikbare historische data, ook wel de scenario's genoemd.

Het doel van deze dissertatie is het ontwikkelen van een rigoureuze gedistribueerde aanpak voor besluitvorming door middel van scenario gebaseerde technieken voor grootschalige netwerken van gekoppelde onzekere dynamische systemen (agenten genoemd). Een scenario programma is een eindigdimensionaal optimalisatie probleem waarin een doelfunctie wordt geminimaliseerd onder randvoorwaardes die geassocieerd wordt met een eindig aantal, onafhankelijk en gelijk gedistribueerde scenario's van een aselecte parameter. Theoretische en praktische interesse in scenario programma's komt voort uit het feit dat deze problemen doorgaans efficiënt oplosbaar zijn, terwijl deze nauw verwant zijn aan de zogenaamde robuuste- en kans-gelimiteerde programma's. Bij robuust-gelimiteerde programma's dekken de randvoorwaardes alle mogelijke aselecte parameters, terwijl bij kans-gelimiteerde programma's, de randvoorwaarden worden opgelegd tot op een gegeven kans. Echter, het vinden van oplossingen van het resulterende grootschalige optimalisatie probleem voor het onzekere netwerksysteem ondervindt meerdere moeilijkheden, zoals bijvoorbeeld rekentijd voor de centrale besturingseenheid.

De hoofdbijdrage van deze dissertatie is het ontwerp van een techniek om een grootschalig scenario programma op te delen in kleinschalige gedistribueerde scenario programma's voor iedere agent. Voortwerkend op bestaande resultaten in de literatuur geven wij nieuwe garanties die de robuustheid kwantificeren van de gegenereerde oplossingen in een gedistribueerd raamwerk. In deze aanpak moet iedere agent informatie

uitwisselen met naburige agenten. Dit is nodig wegens de statistische leerfuncties van de voorgestelde aanpak. Echter, het communicatieschema tussen de agenten zou aanleiding kunnen geven tot bezorgdheid over de privacy gevoeligheid van de informatie van de agenten. Daarom presenteren wij een nieuw geprivatiseerd en gedistribueerd framework, gebaseerd op het zo genaamde differentiaal privacy concept zodat elke agent gevraagde informatie kan delenen dus de privacy wordt gegarandeerd. Daarnaast introduceren wij een soft communicatieschema om de gevraagde communicatielast te verminderen, gebaseerd op een set-parametrisatie techniek en de notie van een probabilistische betrouwbare set. Zo'n betrouwbaarheidsmaatstaf is opgenomen in de haalbaarheidsgaranties van de besluitvorming van de agent in probabilistische zin. De theoretische garanties van het voorgestelde gedistribueerde scenario gebaseerde besluitvorming raamwerk komt overeen met zijn gecentraliseerde tegenhanger, alhoewel het schalen van de resultaten met het aantal agenten een probleem blijft.

De tweede bijdrage van deze dissertatie is gemotiveerd door de toepassing in *Smart Themeral Grids* (STGs) (Slimme Warmte Netwerken). De bijdrage is het ontwerp van een gedistribueerd data-gedreven energie management raamwerk voor gebouwklimaatcomfortsystemen, die in een netwerk aan elkaar zijn gekoppeld via *Aquifer Themeral Energy Storage* (ATES) (Aquifer Warmte Energie Opslag) systemen. Het ATES systeem wordt beschouwd als een seizoensgebonden opslagsystem die als een warmtebron, koelelement, of een opslag voor thermische energie kan fungeren. Het doel in de STGs is om de energiebalans tussen de onzekere warmte-energievraag en productie van individuele gebouwen te waarborgen. Hiervoor is er coördinatie tussen meerdere gebouwen en lange termijnplanning nodig, vaak van maanden vooruit, met de aanwezigheid van twee typen onzekerheden, namelijk lokale (private) en gemeenschappelijke onzekerheidsoorzaken. De private onzekerheidsoorzaak refereert naar de onzekere warmte-energievraag van individuele gebouwen en de gemeenschappelijke onzekerheidsbron beschrijft de onzekerheid van de gemeenschappelijke energiebronnen (ATES) tussen buren.

De derde bijdrage van deze dissertatie is de toepassing van onze voorgestelde aanpak in een probleem van reserveplanning voor energienetwerken met hernieuwbare opwekking gebaseerd op een AC optimale stroomsterkte model. Allereerst formuleren we het probleem door middel van stochastische semi-definiete programmering (SDP) in oneindigdimensionale ruimte, dat in het algemeen computationeel onhandelbaar is. Door gebruik te maken van een nieuwe affiene policy, benaderen we de oneindigdimensionale SDP als een handelbare eindigdimensionale SDP en kwantificeren we expliciet de prestaties van de benadering. Daarna gebruiken we het geografische patroonen van het energiesysteem om het grootschalige systeem op te delen in een multi-gebiedsenergienetwerk. Een consensus variant van de *Alternating Direction Method of Multipliers* (ADMM) wordt gebruikt om een haalbare oplossing te vinden voor zowel de lokale als het algemene multi-gebiedsenergienetwerken. Door gebruik te maken van ons voorgestelde gedistribueerde data-gedreven besluitvorming raamwerk, kan ieder gebied zijn eigen informatie over de onzekerheden gebruiken om lokale haalbaarheidscertificaten te bewerkstelligen en tevens te voldoen aan de algemene haalbaarheid van het multi-gebiedsenergienetwerk onder generieke condities.

De vierde bijdragen van deze dissertatie is het ontwerp van een gedistribueerde ano-

malie-detectie strategie voor een netwerk van gekoppelde onzekere niet-lineaire systemen. De gemonitorde netwerksystemen worden gemodelleerd als een inter-connectie van overlappende agenten die enkele toestandscomponenten delen. Voor iedere agent wordt er een lokale drempelwaarde ontworpen, gebaseerd op het concept van benaderingen van de probabilistische set door middel van polynome drempelwaardesets. De drempelwaardesets worden geparametriseerd op een manier zodat de residuen arbitrair goed te begrenzen zijn. Deze sets worden in normale operatieomstandigheden geproduceerd door een observer-gebaseerde residugenerator. Daarnaast formuleren wij een tweede probleem om de gevoeligheid van de verkregen drempelwaardeset te maximaliseren ten opzichte van de mogelijke karakteristiek van anomalieën. Om dit te bereiken moet iedere agent enige informatie communiceren met naburige agenten om de valsealarm ratio te minimaliseren. Het voorgestelde geprivatiseerde en gedistribueerde framework, samen met het zo genoemde soft communicatieschema is toegepast om, allereerst, de privacy van agenten te behouden en, ten tweede, de benodigde communicatie te reduceren.

De voorgestelde bijdragen richten op het overbruggen van de kloof tussen theoretische en praktische uitdagingen in de toepassing van grootschalige scenario programma's voor onzekere netwerksystemen, het dichterbij brengen van data-gedreven besluitvorming strategieën en een werkelijke implementatie in gedistribueerde gekoppelde onzekere systemen.

ACKNOWLEDGEMENTS

This dissertation is the fruits of almost four years of work at the Delft Center for Systems and Control (DCSC) of Delft University of Technology. Many wonderful people have accompanied me in this unique experience. In this short note, I would like to acknowledge those special individuals whom without their support, help and encouragement the completion of this dissertation would not have been possible.

First and foremost, I would like to express my sincere gratitude to my advisor, Tamás Keviczky, for his close supervision and giving me the possibility to achieve my goals and for the trust and his endless support throughout these four years. I would like to thank him for always being positive and understanding and for giving me the time and freedom to develop my own ideas. I consider myself blessed for getting a chance to work with Nathan van de Wouw as my second advisor. I owe him a debt of gratitude for his contagious enthusiasm and sharp way of thinking that I learned to love research and get motivated and respect any aspect of it to perform the best way possible.

I have greatly appreciated all my Ph.D. exam committee members: Bart De Schutter, Peter Palensky, Riccardo Scattolini, and André Teixeira for taking their valuable time to read my dissertation and to provide constructive feedback. A special thanks to Bart De Schutter for his support during these four years at DCSC and to Peter Palensky and Riccardo Scattolini for accepting to be part of my Ph.D. exam committee. I am specially grateful to André Teixeira, not only for serving in my Ph.D. exam committee, but most importantly for all our discussions and collaborations on the privacy aspect of the results reported in Chapter 5 of this dissertation. I would also like to thank Riccardo Ferrari for his motivation and having a great practical vision in many different aspects of control engineering related to Chapter 5. Without his patient collaboration throughout these years, this dissertation would not have been the same.

I would like to extend my great appreciation to all my friends and current and former colleagues at DCSC, and of course the fantastic secretariat team, for creating a warm and productive research environment. I was honored to be part of this department for their full support and encouragement during these four years. A special thanks goes to Laura Ferranti and Eunice Herrera for all the interesting discussions that we had during the lunches and their friendship. I would also like to thank Sjoerd Boersma, Cees Verdier, Manon Kok, and Erik Steur for their help in the translation of the summary in Dutch, and Nathan van de Wouw for translating the propositions in Dutch.

There have been many people, that I have worked with and learned from them all these four years. I would like to acknowledge the amazing people involved in the ATES Smart Grids project, especially Marc Jaxa-Rozen and Martin Bloemendal for all the technical and personal discussions that we had during our weekly-based meetings, and their true friendship. Cheers to Wicak Ananduta and Ole ter Haar for being amazing master thesis students. I learned from them and hope that I was able to pass over at least some of the world-class supervision I received. Thanks to Janani Venkatasubramanian

and Valentijn van de Scheur, my current master thesis students, for their understanding about my limited time during the last year and also thanks to Lakshminarayanan Nandakumar for having a fruitful discussion on his master thesis work.

Last, but more importantly, I would like to use this opportunity to thank my parents, Ahmad and Goli, and my mother in law, Mahin, and my brothers and sisters, for all the love and for supporting me by all means for my decision of a life far from home. Words can never be enough to express my sincerest thank to my family. My deepest love and gratitude goes to my beloved wife, Zahra. What I owe her can not be expressed in words; this is just an opportunity to admit how important it is having her in my life: I love you. Alma, my little beautiful princess, you are the most wonderful gift in my life and thanks a lot for your presence in my life.

Vahab Rostampour
Delft, August 2018

Contents

Summary **vii**

Samenvatting **xi**

Acknowledgements **xv**

1 Introduction **1**
- 1.1 Research Motivation . . . 1
  - 1.1.1 Research Project . . . 2
  - 1.1.2 Research Problems . . . 5
- 1.2 Outline and Contributions . . . 10
- 1.3 Publications by the Author . . . 13

2 Probabilistic Energy Management for Building Climate Comfort in STGs with Seasonal Storage Systems **17**
- 2.1 Introduction . . . 18
  - 2.1.1 Related Works . . . 18
  - 2.1.2 Contributions . . . 19
  - 2.1.3 Structure . . . 20
  - 2.1.4 Notation . . . 20
- 2.2 System Dynamics Modeling . . . 21
  - 2.2.1 Seasonal Storage Systems . . . 21
  - 2.2.2 Thermal Energy Demand Profile . . . 23
  - 2.2.3 Building Climate Comfort Systems . . . 24
- 2.3 Energy Management Problem . . . 27
  - 2.3.1 Energy Balance in Single Agent System . . . 28
  - 2.3.2 ATES in Smart Thermal Grids . . . 30
  - 2.3.3 Problem Formulation in Multi-Agent Network . . . 32
  - 2.3.4 Move-Blocking Scheme . . . 33
- 2.4 Computationally Tractable Framework . . . 33
- 2.5 Numerical Study . . . 37
  - 2.5.1 Simulation Setup . . . 38
  - 2.5.2 Simulation Results . . . 38
- 2.6 Conclusions . . . 41

3 Distributed Stochastic MPC for Large-Scale Systems with Private and Common Uncertainty Sources **43**
- 3.1 Introduction . . . 44
  - 3.1.1 Related Works . . . 44
  - 3.1.2 Contributions . . . 45

3.1.3 Structure . . . 46
3.1.4 Notations . . . 46
3.2 Problem Statement . . . 46
3.3 Distributed Scenario MPC . . . 50
3.4 Information Exchange Scheme . . . 55
3.5 Plug-and-Play Operational Framework . . . 57
3.6 Numerical Study . . . 59
3.6.1 Three-Room Case Study . . . 59
3.6.2 Three-Building (ATES Systems) Case Study . . . 61
3.7 Conclusions . . . 64

4 Distributed Stochastic Reserve Scheduling in AC Power Systems with Uncertain Generation 65
4.1 Introduction . . . 66
4.1.1 Related Works . . . 66
4.1.2 Contributions . . . 67
4.1.3 Structure . . . 68
4.1.4 Notations . . . 69
4.2 Problem Formulation . . . 69
4.2.1 AC OPF Problem . . . 69
4.2.2 Convexified AC OPF Problem . . . 71
4.2.3 Convexified AC OPF Reserve Scheduling Problem . . . 72
4.3 Proposed Tractable Reformulation . . . 75
4.3.1 Infinite to Semi-Infinite Program: Affine Policy . . . 75
4.3.2 Semi-Infinite to Finite Program: Randomized Approach . . . 77
4.3.3 Infinite to Finite Program: Direct Approach . . . 78
4.4 Distributed Framework . . . 79
4.4.1 Multi-Area Decomposition . . . 80
4.4.2 Distributed Multi-Area AC OPF Problem via ADMM . . . 84
4.4.3 Distributed Multi-Area SP-OPF-RS Problem via ADMM . . . 86
4.5 Numerical Study . . . 92
4.5.1 Simulation Setup . . . 92
4.5.2 Simulation Results: Part One . . . 92
4.5.3 Simulation Results: Part Two . . . 95
4.6 Conclusions . . . 100

5 Privatized Distributed Anomaly Detection for Large-Scale Nonlinear Uncertain Systems 101
5.1 Introduction . . . 102
5.1.1 Related Works . . . 102
5.1.2 Contributions . . . 103
5.1.3 Structure . . . 104

5.2 Problem Statement 105
5.2.1 Large-Scale System Dynamics 105
5.2.2 Subsystem Dynamics 106
5.2.3 Residual Generator 108
5.2.4 Anomaly Detection Threshold Design Problem 110
5.3 Distributed Probabilistic Threshold Set Design 112
5.3.1 Set-Based Threshold Design 112
5.3.2 Maximization of Anomaly Detectability 114
5.3.3 Cascaded Problem Formulation Scheme 114
5.3.4 Computationally Tractable Methodology 115
5.4 Privatized Distributed Anomaly Detection 117
5.4.1 The Concept of Differential Privacy 117
5.4.2 Privacy-Preserving Framework 119
5.4.3 Privatized Inter-Agent Information Exchange Scheme 122
5.4.4 Privatized Distributed Probabilistic Threshold Set 125
5.5 Numerical Study 125
5.5.1 Privacy Preservation 126
5.5.2 Privatized Distributed Anomaly Detection 131
5.6 Conclusions 132

6 Conclusions and Recommendations 135
6.1 Conclusions 135
6.2 Recommendations for Future Research 138

A Proofs of Chapter 2 143

B Proofs of Chapter 3 149

C Proofs of Chapter 4 155

D Proofs of Chapter 5 159

Bibliography 165

Author Biography 179

List of Publications 181

1

Introduction

The aim of this dissertation is to develop a technique to distributed data-driven decision making for large-scale networks of interconnected uncertain dynamical systems, which we refer to as agents. The behavior of each agent is considered to be described by a mathematical model, that is employed for the control design (i.e., decision making) process. In the presence of uncertainties, such a model is often considered as a stochastic dynamical system, which is in general hard to control in an optimal fashion. Using available historical data or so-called scenarios of the uncertainties in each agent, we aim to provide a distributed control strategy for such a network of interconnected agents by enforcing several operational constraints up to a certain level of probability.

This introductory chapter presents the research motivation based on an application to Smart Thermal Grids (STGs) of building climate comfort systems that are interconnected via Aquifer Thermal Energy Storage (ATES). Starting with a brief overview on different types of geothermal energy systems and formulating several questions related to ATES systems, we provide a brief explanation on the developments and achievements in the modeling of such a networked system. This yields a foundation for the research problems to be addressed throughout this dissertation. We further highlight the main contributions, related to proposed solutions to these problems, along with the outline of this dissertation. Finally, we conclude this chapter with the list of publications, on which each chapter is based, together with the author's related papers and other publications that are not included in this dissertation.

1.1. Research Motivation

Motivated by an application to STGs with ATES, this research is supported by the Dutch organization for scientific research (NWO) under the project ATES STGs. The main challenge of this research is to first develop a mathematical model for the dynamics of a single ATES system integrated into the building climate comfort system, and then, to develop a mathematical model for possible mutual interactions between ATES systems in STGs, leading to a network of interconnected ATES systems. We therefore

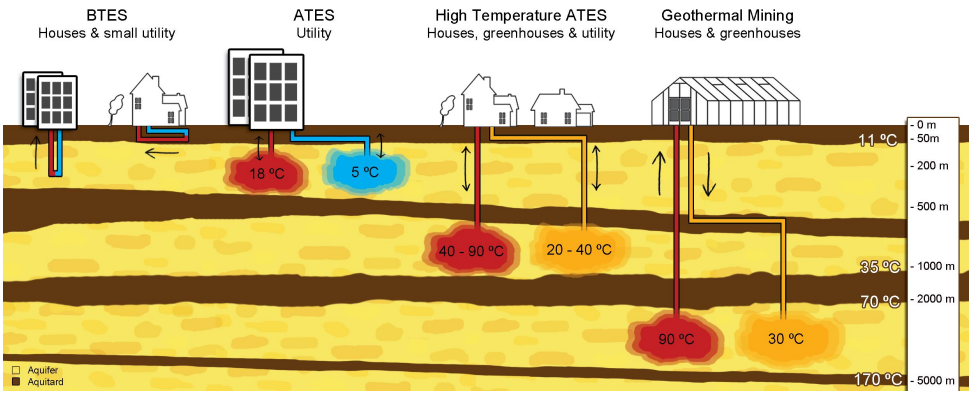


Figure 1.1: Different types of geothermal energy systems. The Figure is taken from [14].

start with a brief overview of the NWO research project, along with some questions and answers based on our developments. We then conclude with the statement of several interesting research problems inspired by this application, and that are addressed as the main research contributions throughout this dissertation.

1.1.1. RESEARCH PROJECT

Geothermal seasonal energy storage gives rise to an attractive way to reduce greenhouse gas emissions in cities with moderate climates. Figure 1.1 depicts different types of geothermal energy systems. These systems take advantage of the high storage capacity of the Earth to temporarily store thermal energy. During summer, the surplus heat of a building is stored and used to heat the building in the cold winter months, while the building can be cooled during warm summer months using the winter cold.

The NWO research project, in the scope of which the research in this dissertation, is focused on a network of buildings that are interconnected via ATES systems. Motivated by the fact that ATES systems are cost-effective,they have a high adoption potential in dense urban areas, where many buildings stand side by side on top of a suitable aquifer. The basic principle of ATES is its use of the subsurface to overcome the seasonal discrepancy between the availability and demand for thermal energy in the built environment. To proceed further, we break down our research project into the following two main steps:

- ▷ **A single ATES system:** It is considered as a heat source or sink, or as a storage for thermal energy demand of building climate comfort system. This functionality is achieved by injection and extraction of water into and from saturated underground aquifers. An ATES systems are suitable for heating and cooling of utility buildings such as offices, hospitals, universities and greenhouses.

- ▷ **A network of interconnected ATES systems:** In dense urban environments, the proximity of hot and cold wells in nearby ATES system installations may lead to unwanted mutual underground aquifer interactions leading to suboptimal operation or conservative design choices (such as excessively large, unused permits).

Such interactions are dynamically time-varying and plagued by uncertainty due to the absence of detailed underground models and cooperation between operators regarding the influence of nearby systems.

We next describe both steps in more detail and put forward some fundamental questions related to each step, and the answers we developed in our research.

SINGLE ATES SYSTEM

Developing a dynamical model for an ATES system, that is suitable in the scope optimal building climate comfort control problems, is one of the important fundamental milestones of this project. To further investigate this step, the following questions are considered:

Q_1 What kind of mathematical model can describe the slow dynamical behavior of an ATES system? What are the relevant variables that represent the behavior of the system?

Q_2 How to integrate an ATES system in the building climate comfort system in order to achieve a desired level of comfort for the building? How to overcome the difficulties that arise from the slow dynamical behavior of an ATES system compared to an hourly-based performance specification of building climate comfort system?

In order to answer the above questions, we achieved the following results:

A_1 In [142], we presented a single energy storage model for both wells of an ATES system that contains three different operating modes of an ATES system, namely: charging, discharging and storing, along with a heat pump for cold seasons. This yields a mixed logical dynamical system model. An agent-based geohydrological simulation environment (MODFLOW) is also developed.

$A_{1,2}$ Building upon our model in [142], we developed a complete and sophisticated building dynamical model integrated with a new ATES model in [137] and [135]. In particular, in [137] a detailed building climate comfort model considering all operational modes (heating, cooling, silence) in the presence of uncertain outside weather is developed. In [135], we consider the dynamics of stored thermal energy over time in each well of an ATES system to be proportional to the volume and temperature of water. In this work, we present a novel mathematical model for both the dynamical behavior of volume and the temperature of water in each well of an ATES system, together with detailed steps for estimating the model parameters. However, the overall system dynamics becomes a hybrid nonlinear (signomial) model. Although nonlinear optimization problems with constraints defined by signomials are normally harder to solve than those defined by only polynomials, signomial optimization problems often provide a much more accurate mathematical representation of real-world nonlinear optimization problems [65]. These highlight that such a dynamical model is not suitable for the control problem of large-scale systems with year-long prediction horizons, due to the computational issues.

A_2 None of the developed ATES models in [135, 137, 142] can be used for the optimal control purpose of large-scale all-season systems. Therefore, we developed a new linear model of an ATES system by keeping track of changes in the volume of the stored water and the amount of thermal energy content in each well separately in [145]. This yields a more suitable control-oriented, computationally tractable model, compared to the previous works in [135, 137, 142].

NETWORK OF INTERCONNECTED ATES SYSTEMS

Consider now a single ATES system integrated in a building climate comfort system, which we refer to as a single agent system. We now aim to extend such a single agent model into a multi-agent networked system in the presence of uncertainties due to the absence of detailed underground models. The following questions are raised to complete the modeling of the network of interconnected ATES systems:

Q_3 How to model a network of building climate comfort systems to manage the thermal energy balance between production units and demands? What are the requirements in such a setting to represent the main characteristics of STGs, and also to enable theoretical control studies?

Q_4 How to incorporate an ATES system in a network of building climate comfort systems? How to cope with the differences between slow dynamical ATES systems and the building climate comfort systems with hourly-based operations? How to deal with the local uncertain thermal energy demand of each agent? How to deal with uncertainties due to the absence of detailed underground models in the network of interconnected ATES systems? How to cooperate between agents in order to prevent the mutual interactions between nearby ATES systems?

The first set of questions are more related to the modeling of a network problem, whereas the second set of questions are about the integration of ATES systems in the network problem. The following developments are achieved in order to answer the above questions:

A_3 In [144], we developed an STG model of a network of interconnected building climate comfort systems in the presence of local uncertain thermal energy demand. This is referred to as the private uncertainty source, since such uncertainties can be due to the occupancy level of the building. We first formulated a large-scale stochastic mixed-integer (non-convex) problem to balance the local uncertain thermal energy demands and the conventional production units (boiler, micro-combined heat and power, etc.). Using available historical data (scenarios) of the local uncertainty sources and the so-called robust randomized technique in [98], we then developed a computationally tractable framework and implemented a model predictive control (MPC) paradigm. A refinement of such a problem and a distributed framework to address a multi-agent network with private uncertainty sources is presented in [134].

A_4 Building upon our works in [134, 144], we extended such a setting to integrate each building climate comfort system with an ATES system and allowed them to have

three different operating modes, namely: heating, cooling, and storing. We modeled the possible mutual interactions between neighboring ATES systems as uncertain coupling constraints. This yields a network problem with two different uncertainty sources. A local (private) uncertainty source that represents an uncertain thermal energy demand for each agent, and a common uncertainty source representing the uncertain common energy pool (ATES) between neighboring agents.

The above exposition on ATES for STGs motivates generic research objectives revolving around the modeling and control of uncertain networked systems as considered in this dissertation. We next elaborate on some concrete research problems Section 1.1.2.

1.1.2. RESEARCH PROBLEMS

Motivated by the application of STGs with ATES systems, the generic research objectives of this dissertation are related to the following four problems.

Problem 1

▷ **Modeling and controlling an uncertain networked system that represents an STG of ATES systems**
The following aspects need to be addressed:

$P_{1(a)}$ **Modeling of single agent system**
How to model an individual building with a single ATES system dynamics as a seasonal storage system by taking into consideration the hybrid nature of each building due to the different operating modes (heating, cooling, and storing)?

$P_{1(b)}$ **Modeling of a multi-agent networked system**
How to model a network of interconnected buildings via ATES systems by taking into consideration the time-varying unwanted mutual interactions between neighboring ATES systems?

$P_{1(c)}$ **Controlling a multi-agent networked system**
How to develop a control framework to achieve certain performance, e.g., a desired comfort level of each building, while satisfying some physical limitations and operational constraints and taking into consideration the time scale discrepancy between the ATES system dynamics and a single building climate comfort system?

1

$P_{1(d)}$ **Dealing with uncertainties in a networked system**
In the presence of uncertainty, (e.g., uncertain thermal energy demand for each building) and/or uncertain coupling constraints, (e.g., uncertain common energy pool (ATES)), how to achieve (state or input) constraint satisfaction? It might be of interest to achieve constraint feasibility up to a certain level of probability which gives rise to chance-constrained problems that are in general hard to solve. How to develop a computationally tractable framework using a scenario-based technique for a stochastic MPC strategy? What kind of theoretical guarantees can be provided by extending the existing results in literature, e.g., the scenario approach [27, 30, 32] and the robust randomized approach [98, 144, 181] to large-scale uncertain networked systems?

This problem is indeed an aggregation of Q_1-Q_4 in Section 1.1.1 together with some questions focused on the control aspects of an uncertain networked system from a centralized perspective. The resulting problem of centralized control leads to a large-scale scenario optimization problem. However, finding solutions of the resulting large-scale scenario optimization problem for uncertain networked systems poses several difficulties, e.g., the computational cost for a central control unit. This highlights the second problem of interest in this dissertation.

Problem 2

▷ **Developing a distributed data-driven decision making framework for uncertain networked system**
The following aspects need to be addressed:

$P_{2(a)}$ **Decomposing the large-scale scenario program**
Motivated by the computational complexity issue, such as in [149] and [138] convex and non-convex scenario programs, how to decompose such a large-scale scenario optimization problem into the small-scale distributed scenario programs? How to deal with private and common uncertainty sources? Is there a way to decompose scenarios between neighboring agents? Under which conditions can we achieve some theoretical guarantees for the feasibility of both local and network constraints? What are the limitations and bottlenecks for the extension of existing guarantees to the networked systems?

$P_{2(b)}$ **Developing a distributed framework for a network of interconnected uncertain systems**
How to deal with a network of agents that are dynamically coupled and/or have coupling operating constraints? What sort of distributed algorithms [22] can be employed?

$P_{2(c)}$ **Handling communication burden between neighboring agents**
In such a distributed framework, how should agents communicate with each other in order to achieve certain performance criteria? If agents are requested to share certain scenarios, how to handle the communication limitations in such networked systems? What are the consequences of the limited communications between neighboring agents in terms of their local constraint satisfaction?

Problem 2 is the main research problem that eventually yields a distributed data-driven decision making framework for uncertain networked systems as the main contribution of this dissertation. It will be shown that such an achievement can be used also in a general setting depending on the way that each agent interacts with its neighboring agents.

Building upon the previous problem to develop a distributed data-driven decision making framework, the next problem addresses the possibility of applying such a framework into the transmission system operators (TSOs) problem. TSOs are entities entrusted with transporting energy in the electrical power networks. Two critical issues in such a energy network, namely safety and reliability, are the most important responsibility for TSOs. Safety refers to managing any failure on their electrical generation sources, and reliability is to coordinate the balance between supply of and demand for electricity. Such issues are formulated as an optimal power flow (OPF) management problem and becoming very challenging when there are some uncertain generation units (wind power) in power networks.

Problem 3

▷ **Developing a distributed data-driven decision making framework for optimal power flow (OPF) problem in AC power systems with uncertain generation**
The following aspects need to be addressed:

1

$P_{3(a)}$ **Formulating the OPF problem using AC model of power network**
Is it possible to apply such a framework into the transmission system operators (TSOs) problem? How to make a decision about set points for power grid generators, as well as optimizing OPF steady-state set points for generators? How accurately can a power network model be considered in these problems, i.e., using an AC (nonlinear) or DC (linear) model? If using an AC model, how to deal with the non-convexity of the resulting optimization problem [82, 149]? How to cope with renewable power generation units?

$P_{3(b)}$ **Dealing with the infinite-dimensional optimization problem**
In the presence of uncertainty (e.g., uncertain generation and/or demand), the OPF management equipped with reserve scheduling (RS) problem that leads to be an infinite-dimensional optimization problem. How to approximately solve such a intractable problem? Is the a way to bridge from infinite-dimensional to finite-dimensional space and characterizing the approximation level?

$P_{3(c)}$ **Handling the multi-area AC power network problem**
Can we develop a distributed framework for OPF problem using AC modeling of power networks? How to decompose a large-scale AC power network into multi-area sub-networks? How to share the required information in such a setting to cope with renewable power generation units?

The OPF management problem is one of the most widely studied problems for the last decades starting from 1962 [34]. There are still ongoing researches to find OPF solutions for the recent power systems challenges such as a market liberalization [8] and/or a large penetration of uncertain energy sources [132]. Aiming at various application domains for our proposed distributed data-driven decision making framework, requires different considerations such as the level of detail OPF model formulation that represents power systems. These highlight the necessity of understanding the general concept application domains in order to develop a distributed framework to maintain stability of power system operations.

In the control of large-scale network of interconnected uncertain systems, it is uncommon that a control system operates continuously, uninterrupted based on optimal scheduled plans. Due to the presence of mechanical and/or electrical hardware (components), both actuators and sensors can fail in realistic situations. The task of monitoring and diagnosis involves generating a diagnostic signal sensitive to the occurrence of specific anomalies. This task is typically accomplished by designing a filter with all available

information as inputs, and an output which implements a non-zero mapping from the faults to the residual. The existing approaches based on filter design [110] or traditional robust design [44] all consider a centralized unit to process the information. Using a centralized monitoring system for large-scale networked systems is computationally expensive, and therefore it is crucial to develop a distributed framework to detect when failures happen, and to identify as soon as possible which failures have taken place for safety critical processes. This highlights our motivation to address the next problem.

Problem 4

▷ **Developing a distributed data-driven decision making framework for anomaly detection of a large-scale nonlinear uncertain system**
The following aspects need to be addressed:

$P_{4(a)}$ **Decomposing a large-scale nonlinear uncertain system**
How to increase the resiliency of uncertain networked systems to anomalies (faults)? How to endow such systems with smart architectures capable of monitoring, detecting, isolating and counteracting such anomalies and threats?

$P_{4(b)}$ **Developing a set-based probabilistic threshold set framework**
Can we extend such a classical approach to a data-driven approach using available historical scenarios of the so-called healthy residuals? Such residuals can take values in arbitrary shaped, possibly non-convex regions. How to develop a threshold set that is built with arbitrary shape to distinguish healthy residuals from anomalies?

$P_{4(c)}$ **Developing a privatized distributed detection framework**
How to develop a distributed setting to have multiple thresholds to process data and eventually yield an appropriate anomaly detection with isolation? Distributed methods require communication [17, 54, 57, 115, 128, 179, 180], which may be undesirable as it may lead to leaking privacy-sensitive information. For instance, consider a large-scale network system where neighboring diagnosis nodes are each monitoring different sub-grids with distributed energy sources and each is managed by its own grid operator. The two grid operators must exchange data about nodes on their respective boundaries in order to allow for grid balancing, but they would rather keep private the way that they are allocating energy supply to their different energy sources and satisfying their energy demand [62, 155]. How to design a privatized distributed anomaly detection framework?

1

Table 1.1: Classification of the dissertation contents based on the class of networked systems in the proposed framework for each problem.

	Dynamics				Constraints			Uncertainty		Proposed Framework		
	Discrete time											
	Linear	Nonlinear	Coupled	Hybrid	Input	State	Coupled	Worst Case	Stochastic	Centralized	Distributed	
Problem 1	√			√	√	√	√	√	√	√		Chapter 2
Problem 2	√		√		√	√	√		√		√	Chapter 3
Problem 3		(static) √			√	√	√		√		√	Chapter 4
Problem 4		√	√		√	√			√		√	Chapter 5

1.2. OUTLINE AND CONTRIBUTIONS

Based on the above research problems, this dissertation aims to develop a distributed data-driven decision making framework for uncertain networked systems. The proposed framework should be equipped with computationally efficient distributed algorithms, as the problems of interest are all related to large-scale networked systems. It should also explore the availability of data (scenarios), which can be used to first learn some statistical properties of the problem, and then to provide some theoretical guarantees on the feasibility of decisions in distributed setting. It is important to note that data-driven approaches do not necessarily require a statistical model of the uncertainties, although some model of the uncertainties may still be necessary for generating scenarios beyond the cardinality available from historical data [26].

The main contribution of this dissertation is motivated by a common issue between all the proposed problems in Section 1.1.2. From a centralized unit for all aforementioned problems, the resulting formulation will be a large-scale scenario program which is computationally demanding. We therefore propose a technique to decompose such a problem into small-scale distributed scenario programs for each agent. We then quantify the robustness of the resulting solutions for each agent in a distributed framework to guarantee a priori feasibility of each agent locally and globally under some mild conditions building on existing results in literature, see [27, 30, 32], and the references therein. The theoretical guarantees of the proposed distributed data-driven decision making coincide with the centralized counterpart. This is achieved under the assumption that neighboring agents can exchange some requested scenarios without any communication constraints. Such inter-agent communications might however give rise to privacy concerns.

We next propose a novel privatized communication framework, such that each agent can share requested information while preserving its privacy. It is important to highlight that in such a setting, each agent requests a certain number of scenarios from its neighbors, which is called a hard communication scheme. This means that agents are not flexible to decide about the number of scenarios that should be sent to their neighbors. In order to relax this restriction, we introduce a soft communication scheme using a set parametrization technique, together with the notion of probabilistically reliable sets. Such a reliability measure of the soft communication scheme is incorporated into the feasibility guarantees of agent decisions in a probabilistic sense.

Table 1.1 presents a classification of the dissertation contents based on the class of networked systems that can be formed in the proposed framework for each problem. We now explain the detailed contributions of individual chapters based on their related problem, as they are presented in Table 1.1.

CHAPTER 2

This chapter concentrates on the modeling of STGs with ATES systems and the development of an optimal control framework to address Problem 1.

ATES as a seasonal storage system has not, to the best of our knowledge, been considered in STGs. We develop a novel large-scale stochastic hybrid dynamical model to predict the dynamics of thermal energy imbalance in STGs consisting of building climate comfort systems with hourly-based operation and ATES as a seasonal energy storage system. Based on our previous work in [135] and [142], we extend an ATES system model to predict the amount of stored water and thermal energy. We first incorporate the ATES model into a building climate comfort problem, and then formulate a large-scale STG problem by taking into consideration geographical coupling constraints between ATES systems. Using an MPC paradigm to achieve a desired level of comfort for buildings, we formulate a finite-horizon mixed-integer quadratic optimization problem with multiple chance constraints at each sampling time leading to a non-convex problem, which is difficult to solve.

We next propose a move-blocking control scheme to enable our stochastic MPC to handle long prediction horizons and an hourly-based operation of the building climate comfort systems together with a seasonal variation of desired optimal operation of the ATES system in a unified framework. The time-scale discrepancy between the ATES system dynamics and building climate comfort systems are explicitly accounted for in the developed MPC-based optimization formulation. Our proposed control strategy offers a long enough prediction horizon to prevent mutual interactions between ATES systems with much less computational time compared to a fixed prediction horizon that is sampled densely (i.e., every hour).

We finally develop a computationally tractable framework to approximate a solution of our proposed MPC formulation based on our previous work in [144]. In particular, we extend the framework in [144] to cope with multiple chance constraints which provides a more flexible approximation technique compared to the so-called robust randomized approach [98, 100], which is only suitable for a single chance constraint. Our framework is closely related to, albeit different from, the approach of [158].

The contents of this chapter have been accepted to appear in the IEEE Transactions on Smart Grids [148].

CHAPTER 3

Problem 2 is addressed in this chapter by proposing a decomposition technique for a large-scale system dynamics model in order to support the distribution of the resulting centralized scenario problem at each sampling time. Moreover, a novel communication scheme is introduced to reduce communication burden between the distributed small-scale problems. The main key ingredients of this chapter are as follows.

We first provide a technique to decompose the large-scale scenario program into distributed scenario programs that exchange a certain number of scenarios with each other in order to make local decisions. We show that such a decomposition technique can be applied to large-scale linear systems with both private (local) and common uncertainty sources. This yields a flexible and practical plug-and-play distributed scenario MPC framework.

We then quantify the level of robustness of the resulting solutions using our proposed distributed scenario MPC framework, and provide two new a priori probabilistic guarantees for the desired level of constraint fulfillment under some mild conditions of both cases of private and common uncertainty sources.

We finally develop a so-called soft communication scheme between neighboring agents, based on a set parametrization technique together with the notion of a probabilistically reliable sets, in order to reduce the required communication between each subproblem. We show how to incorporate the probabilistic reliability notion into existing results, and provide new guarantees for the desired level of constraint violations.

The contents of this chapter is currently under review for the IEEE Transactions on Control of Networked Systems [147].

CHAPTER 4

This chapter addresses Problem 3 by developing a distributed framework to manage the OPF problem using an AC model of power system in the presence of uncertain generation. In this chapter, we formulate a distributed data-driven reserve scheduling (RS) problem to cope with high penetration of wind power using an AC power network model. To the best of our knowledge, such a distributed framework has not been yet addressed in the related literature and this is the first work in this direction. The summary of the contributions in this chapter are as follows.

We provide a novel reformulation of the RS problem using an AC OPF model of power systems with wind power generation, leading to an infinite-dimensional SDP which is in general computationally intractable. We propose an approximation of the infinite-dimensional semidefinite program (SDP) with tractable finite-dimensional SDPs using an affine policy inspired by practical aspects of the problem. We explicitly quantify the exactness of the approximation, and provide a priori probabilistic feasibility guarantees to optimally schedule generating units while simultaneously determining the geographical allocation of the required reserve. We also provide another formulation of the OPF-RS problem, similarly to [149] with some modifications, and compare the proposed formulations in terms of worst-case computational complexity analysis.

We develop a distributed stochastic framework to carry out multi-area RS using an AC OPF model of power networks with wind power generation. We provide a technique to decompose a large-scale finite-dimensional SDP into small-scale problems by exploring the connections between the properties of a power network and chordal graphs. A noticeable feature of our distributed setup is that each local area of the power system may have different local accuracy regarding available wind power, and receives a priori probabilistic feasibility certificates to optimally schedule local generating units together with local allocation of the required reserve. This is based on the results developed in Chapter 3 and [146, 147]. We then provide consensus ADMM algorithms for both OPF and OPF-RS problems in a similar manner to [72, 96], with some modifications to cope with stochasticity of the formulations.

The contents of this chapter are currently under review for the IEEE Transactions on Power Systems [150].

CHAPTER 5

This chapter focuses on the development of a distributed anomaly detection framework in a privatized setting to address Problem 4.

We provide a general formulation for the dynamics of a large-scale nonlinear uncertain system together with a decomposition into a number of interconnected subsystems, by extending existing results from [54]. We then introduce a novel fault detection threshold set design problem, using the concept of probabilistic set approximation through *polynomial superlevel sets* [40]. The proposed approach requires communication between a number of agents, one for each subsystem, and such communication may involve privacy sensitive measurements. For designing threshold sets, we formulate a two-stage chance-constrained optimization problem, in which the first step is aimed at fulfilling a probabilistic robustness constraint, and the second step maximizes the performance of detection with respect to a given class of faults. A computationally tractable framework is given for the solution of the chance constrained problem, through a data-driven technique, along with theoretical guarantees.

We next develop a differentially private distributed framework to preserve the privacy of the exchanged information between neighboring subsystems. This makes use of a pre-processing scheme to achieve the privacy of control input using output measurements as the database. And, finally, our proposed soft communication scheme in Chapter 2 is employed to overcome the communication bandwidth constraints, such that each agent will share a set with all its neighboring agents together with a reliability of information for the shared set. The reliability measure of neighboring subsystems is incorporated in the probabilistic guarantees for each subsystem in terms of new level of local false alarms.

The contents of this chapter are currently under review for the IEEE Transactions on Automatic Control [141].

CHAPTER 6

This chapter concludes this dissertation with some remarks on the main contributions and recommendations for directions of future work.

1.3. PUBLICATIONS BY THE AUTHOR

This dissertation is based on several results in previously published or submitted articles in international journals and conferences. An overview of the publications that each chapter is based on is provided below.

- ▷ Chapter 2 is based on
  - **V. Rostampour**, T. Keviczky, *Probabilistic Energy Management for Building Climate Comfort in STGs with Seasonal Storage Systems*, To appear in IEEE Transactions on Smart Grid. [148]

1

Other related publications

- **V. Rostampour**, T. Keviczky, *Energy Management for Building Climate Comfort in Uncertain Smart Thermal Grids with ATES*, in International Federation of Automatic Control World Congress (IFAC), pp. 13698-13705, Toulouse, France, Jul 2017. [145]
- **V. Rostampour**, T. Keviczky, *A MPC Framework of GSHP coupled with ATES System in Heating and Cooling Networks of a Building*, in Proceedings of International Energy Agency (IEA) Conference on Heat Pump, Rotterdam, The Netherlands, May 2017. [137]
- **V. Rostampour**, M. Bloemendal, T. Keviczky, *A Control-Oriented Model For Combined Building Climate Comfort and ATES System*, in Proceedings of European Geothermal Congress (EGC), Strasbourg, France, Sep 2016. [135]
- **V. Rostampour**, M. Jaxa-Rozen, M. Bloemendal, T. Keviczky, *Building Climate Energy Management in Smart Thermal Grids via ATES Systems*, Energy Procedia, Vol. 97, pp. 59-66, Jun 2016. [142]
- **V. Rostampour**, T. Keviczky, *Robust Randomized Model Predictive Control for Energy Balance in Smart Thermal Grids*, in Proceedings of European Control Conference (ECC), pp. 1201-1208, Aalborg, Denmark, Jun 2016. [144]

▷ Chapter 3 is based on

- **V. Rostampour**, T. Keviczky, *Distributed Stochastic MPC for Large-Scale Systems with Private and Common Uncertainty Sources*, Submitted to IEEE Transactions on Control of Networked Systems, Feb 2018. [147]

Other related publications

- **V. Rostampour**, M. Jaxa-Rozen, M. Bloemendal, J. Kwakkel, T. Keviczky, *Aquifer Thermal Energy Storage (ATES) Smart Grids: Large-Scale Seasonal Energy Storage as A Distributed Energy Management Solution*, Submitted to Applied Energy, Aug 2018. [143]
- **V. Rostampour**, T. Keviczky, *Distributed Stochastic Model Predictive Control Synthesis for Large-Scale Uncertain Linear Systems*, to appear in Proceedings of American Control Conference (ACC), Milwaukee, WI, US, Jun 2018. [146]
- **V. Rostampour**, W. Ananduta, T. Keviczky, *Distributed Stochastic Thermal Energy Management in Smart Thermal Grids*, Submitted to PowerWeb: Intelligent Energy Systems, Chapter 9, Springer, Mar 2018. [134]

▷ Chapter 4 is based on

- **V. Rostampour**, O. ter Haar, T. Keviczky, *Distributed Stochastic Reserve Scheduling in AC Power Systems With Uncertain Generation*, Submitted to IEEE Transactions on Power Systems, Dec 2017. [150]

Other related publications

- **V. Rostampour**, O. ter Haar, T. Keviczky, *Computationally Tractable Reserve Scheduling for AC Power Systems with Wind Power Generation*, Submitted to PowerWeb: Intelligent Energy Systems, Chapter 6, Springer, Mar 2018. [152]
- **V. Rostampour**, O. ter Haar, T. Keviczky, *Tractable Reserve Scheduling in AC Power Systems With Uncertain Wind Power Generation*, in Proceedings of Conference on Decision and Control (CDC), pp. 2647-2654, Melbourne, Australia, Dec 2017. [151]

▷ Chapter 5 is based on

- **V. Rostampour**, R. Ferrari, A. H. Teixeira, T. Keviczky, *Privatized Distributed Anomaly Detection for Large-Scale Nonlinear Uncertain Systems*, Submitted to IEEE Transactions on Automatic Control, July 2018. [141]

Other related publications

- **V. Rostampour**, R. Ferrari, A. H. Teixeira, T. Keviczky, *Differentially Private Distributed Fault Diagnosis for Large-Scale Nonlinear Uncertain Systems*, to appear in IFAC Conference on Fault Detection, Supervision and Safety (SAFEPROCESS), Warsaw, Poland, Aug 2018. [140]
- **V. Rostampour**, R. Ferrari, T. Keviczky, *A Set Based Probabilistic Approach to Threshold Design for Optimal Fault Detection*, in Proceedings of American Control Conference (ACC), pp. 5422-5429, Seattle, WA, US, May 2017. [139]

▷ Other publications not included in this dissertation

- S.Boersma, **V. Rostampour**, B.Doekemeijer, J.W. van Wingerden, T. Keviczky, *Wind Farm Active Power Tracking Using Nonlinear MPC*, Submitted to IFAC Conference on Nonlinear Model Predictive Control (NMPC), Madison, Wisconsin, US, Aug 2018. [20]
- S.Boersma, **V. Rostampour**, B.Doekemeijer, W. van Geest, J.W. van Wingerden, *A Centralized Model Predictive Wind Farm Controller in PALM Providing Power Reference Tracking: LES Study*, to appear in Journal of Physics — Conference Series (Torque), Milan, Italy, Jun 2018. [19]
- M. Bloemendal, M. Jaxa-Rozen, **V. Rostampour**, *Use it or Lose it: Adaptive Governance of Aquifers with ATES*, in Proceedings of International Energy Agency (IEA) Conference on Heat Pump, Rotterdam, The Netherlands, May 2017. [16]
- **V. Rostampour**, Dieky Adzkiya, Sadegh Soudjani, Bart De Schutter, T. Keviczky, *Chance Constrained Model Predictive Controller Synthesis for Stochastic Max-Plus Linear Systems*, in Proceedings of Systems, Man, and Cybernetics Conference (SMC), pp. 3581-3588, Budapest, Hungary, Oct 2016. [133]

- M. Jaxa-Rozen, M. Bloemendal, **V. Rostampour**, J. Kwakkel, *Assessing the Sustainable and Optimal Application of Aquifer Thermal Energy Storage*, in Proceedings of European Geothermal Congress (EGC), Strasbourg, France, Sep 2016. [71]
- **V. Rostampour**, P. M. Esfahani, T. Keviczky, *Stochastic Nonlinear MPC of an Uncertain Batch Polymerization Reactor*, in IFAC Conference on Nonlinear Model Predictive Control (NMPC), pp. 540-545, Seville, Spain, Sep 2015. [138]

2

Probabilistic Energy Management for Building Climate Comfort in STGs with Seasonal Storage Systems

This chapter presents an energy management framework for building climate comfort (BCC) systems interconnected in a grid via aquifer thermal energy storage (ATES) systems in the presence of two types of uncertainty (private and common). ATES can be used either as a heat source (hot well) or sink (cold well) depending on the season. We consider the uncertain thermal energy demand of individual buildings as private uncertainty source and the uncertain common resource pool (ATES) between neighbors as common uncertainty source. We develop a large-scale stochastic hybrid dynamical model to predict the thermal energy imbalance in a network of interconnected BCC systems together with mutual interactions between their local ATES. We formulate a finite-horizon mixed-integer quadratic optimization problem with multiple chance constraints at each sampling time, which is in general a non-convex problem and hard to solve. We then provide a computationally tractable framework by extending the so-called robust randomized approach and offering a less conservative solution for a problem with multiple chance constraints. A simulation study is provided to compare completely decoupled, centralized and move-blocking centralized solutions. We also present a numerical study using a geohydrological simulation environment (MODFLOW) to illustrate the advantages of our proposed framework.

2.1. INTRODUCTION

GLOBAL energy consumption has significantly increased due to the combined factors of increasing population and economic growth over the past few decades. This increasing consumption highlights the necessity of employing innovative energy saving technologies. Smart Thermal Grids (STGs) can play an important role in the future of the energy sector by ensuring a heating and cooling supply that is more reliable and affordable for thermal energy networks connecting various households, greenhouses and other buildings, which we refer to as agents. STGs allow for the adaptation to changing circumstances, such as daily, weekly or seasonal variations in supply and demand by facilitating each agent with smart thermal storage technologies.

Aquifer thermal energy storage (ATES) is a less well-known sustainable storage system that can be used to store large quantities of thermal energy in aquifers. Aquifers are underground porous formations containing water that are suitable for seasonal thermal energy storage. It is especially suitable for climate comfort systems of large buildings such as offices, hospitals, universities, musea and greenhouses, see [71]. Most buildings in moderate climates have a heat shortage in winter and a heat surplus in summer. Where aquifers exist, this temporal discrepancy can be overcome by seasonally storing and extracting thermal energy into and out of the subsurface, enabling the reduction of energy usage and CO_2 emissions of climate comfort systems in buildings. Figure 2.1 depicts the operating modes of an ATES system for a single building.

2.1.1. RELATED WORKS

There are various studies in literature related to buildings integrated into a smart grid [125, 165]. Modeling a building heating system connected to a heat pump can be found in [104], an experimental model with a focus on heating, ventilation, and air conditioning (HVAC) systems in [162], using multi-HVAC systems in [156]. Models for building system dynamics together with HVAC controls are typically linear [173] for obvious computational purposes. For instance resistance and capacitance circuit models, that represent heat transfer and thermodynamical properties of the building, are commonly used for building control studies [45, 94, 175]. PID controllers for HVAC systems are widely used in many commercial buildings [2]. Model predictive control (MPC), on the other hand, has received a lot of attention [106, 114, 119], since it can handle large-scale dynamical systems subject to hard constraints, e.g., equipment limitations. Using demand response for smart buildings [123], MPC can be used in building climate comfort (BCC) problems [92, 116]. MPC can overcome BCC problems even in decentralized or distributed setting and it is shown that has several advantages compared to PID controllers [114, 119, 163].

STGs have been studied implicitly in the context of micro combined heat and power systems, see [166], or general smart grids, e.g., see [78] and [79]. Building heat demand with a dynamical storage tank was considered in [167], whereas in [120] an adaptive-grid model for dynamic simulation of thermocline thermal energy storage systems was developed. A deterministic view on STGs was studied by a few researchers [127], [90], [154]. STGs with uncertain thermal energy demands have been considered in [51], where a MPC strategy was employed with a heuristic Monte Carlo sampling approach to make the solution robust. A dynamical model of thermal energy imbalance in STGs with a

probabilistic view on uncertain thermal energy demands was established in [144], where a stochastic MPC with a theoretical guarantee on the feasibility of the obtained solution was developed.

2.1.2. CONTRIBUTIONS

ATES as a seasonal storage system has not, to the best of our knowledge, been considered in STGs. In [135] and [142], a dynamical model for an ATES system integrated in a BCC system has been developed. Following these studies, the first results toward developing an optimal operational framework to control ATES systems in STGs is presented here. In this framework, uncertain thermal energy demands are considered along with the possible mutual interactions between ATES systems, which may cause limited performance and reduced energy savings. The main contributions of this chapter are threefold:

a) We develop a novel large-scale stochastic hybrid dynamical model to predict the dynamics of thermal energy imbalance in STGs consisting of BCC systems with hourly-based operation and ATES as a seasonal energy storage system. Based on our previous work in [135] and [142], we extend an ATES system model to predict the amount of stored water and thermal energy. We first incorporate the ATES model into a BCC problem and then, formulate a large-scale STGs problem by taking into consideration the geographical coupling constraints between ATES systems. Using an MPC paradigm, we formulate a finite-horizon mixed-integer quadratic optimization problem with multiple chance constraints at each sampling time leading to a non-convex problem, which is difficult to solve.

b) We next propose a move-blocking control scheme to enable our stochastic MPC framework to handle long prediction horizons and an hourly-based operation of the BCC systems together with a seasonal variation of desired optimal operation of the ATES system in a unified framework. In practice, the BCC systems have an hourly-based operation and typically day-ahead planning compared to the ATES system that is based on a seasonal operation. Using a fixed prediction horizon length, e.g., least common multiple of these two systems, may turn out to be computationally prohibitive, however also necessary in order to represent ATES interaction dynamics. The time scale discrepancy between the ATES system dynamics and BCC systems are explicitly accounted for in the developed MPC-based optimization formulation. Our proposed control strategy offers a long enough prediction horizon to prevent mutual interactions between ATES systems with much less computational time compared to a fixed prediction horizon that is sampled densely (i.e., every hour).

c) We develop a computationally tractable framework to approximate a solution of our proposed MPC formulation based on our previous work in [144]. In particular, we extend the framework in [144] to cope with multiple chance constraints which provides a more flexible approximation technique compared to the so-called robust randomized approach [98, 100], which is only suitable for a single chance constraint. Our framework is closely related to, albeit different from, the approach of [158]. In [158], the problem formulation is convex and consists of an objective function with multiple chance constraints, in which the terms in objective and constraints are univariate. In contrast, our problem formulation is mixed-integer and the objective function consists of separable additive components.

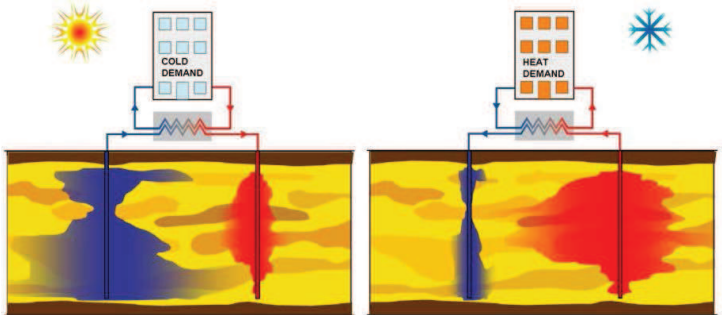


Figure 2.1: Operational modes of an ATES system during warm (left) and cold (right) seasons. Figure is taken from [142].

It is important to highlight that two major difficulties arising in stochastic hybrid MPC, namely recursive feasibility and stability, are not in the scope of this work, and they are subject of our ongoing research work. Thus, instead of analyzing the closed-loop asymptotic behavior, in this work we focus on individual stochastic hybrid MPC problem instances from the optimization point of view and derive probabilistic guarantees for multiple chance constraints fulfillment.

2.1.3. STRUCTURE

The layout of this chapter is as follows. Section 2.2 describes dynamics of an ATES system and a BCC system. In Section 2.3, we first formulate an energy management problem in a single agent, and then, extend it to a network of multiple agents. We present three different setups, namely: one with completely decoupled agents, a centralized problem, and a move-blocking centralized problem formulation. In Section 2.4, we develop a computationally tractable framework to solve these problems, and Section 2.5 provides a simulation study with a comparison between these three different settings. In addition, the numerical results obtained via a geohydrological simulation environment (MODFLOW) are shown. Section 2.6 concludes this chapter with some remarks.

2.1.4. NOTATION

The following international system of units is used throughout the chapter: Kelvin [K] and Celsius [°C] are the units of temperature, Meter [m] is the unit of length, Hour [h] is the unit of time, Kilogram [kg] is the unit of mass, Watt [W] is the unit of power, Joule [J], kiloWatt-hour [kWh], and MegaWatt-hour [MWh] are the units of energy.

$\mathbb{R}, \mathbb{R}_+$ denote the real and positive real numbers, and $\mathbb{N}, \mathbb{N}_+$ the natural and positive natural numbers, respectively. We operate within n-dimensional space $\mathbb{R}^n$ composed by column vectors $u, v \in \mathbb{R}^n$. The Cartesian product over n sets $\mathcal{X}_1, \cdots, \mathcal{X}_n$ is given by: $\prod_{i=1}^n \mathcal{X}_i = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n = \{(x_1, \cdots, x_n) : x_i \in \mathcal{X}_i\}$. The cardinality of a set $\mathcal{A}$ is shown by $|\mathcal{A}| = A$.

Given a metric space Δ, and $\mathbb{P}$ a probability measure defined over Δ, its Borel σ-algebra is denoted by $\mathfrak{B}(\Delta)$. Throughout the chapter, measurability always refers to Borel measurability. In a probability space $(\Delta, \mathfrak{B}(\Delta), \mathbb{P})$, we denote the N-Cartesian product set

of Δ by Δ^N with the respective product measure by $\mathbb{P}^N$.

2.2. SYSTEM DYNAMICS MODELING

In this section, we first develop a mathematical model for an ATES system dynamics as a single seasonal storage system. We then describe the steady-state dynamical model for a building to capture its thermal energy demand profile during heating and cooling modes based on our previous work in [137]. We finally present the BCC system where we introduce the so-called thermal energy imbalance dynamics. Using the thermal energy imbalance dynamics of a BCC system, we integrate an ATES system into the building thermal energy production unit which consists of a boiler, a chiller, a heat pump, a heat exchanger and storage tanks as the heating and cooling modes equipment.

2.2.1. SEASONAL STORAGE SYSTEMS

Consider an ATES system consisting of warm and cold wells to store warm water during warm season and cold water during cold season, respectively. Each well can be described as a single thermal energy storage where the amount of stored energy is proportional to the temperature difference between stored water and aquifer ambient water. Stored thermal energy from the last season is going to be used for the current season and so forth. Depending on the season, the operating mode (heating or cooling) of an ATES system changes, by reversing the direction of water flow between wells as it is shown in Figure 2.1.

During a cold season, for heating purposes, the direction of water is from the warm well to the cold well through a heat exchanger to extract the stored thermal energy from the water. The return water is cooled down and stored in the cold well. This procedure is opposite during a warm season for cooling purposes of the BCC system. An ATES system can be characterized by some physically meaningful parameters. The most relevant features that can describe the status of an ATES system for the purpose of optimal control is the stored volume of water together with the thermal energy content in each well. A free manipulated variable in this setting is the pump flow rate that is used to circulate water from one well to the other through a heat exchanger.

We therefore define the states that can describe the ATES system dynamics to be the volume of water, $\mathrm{V}^h_{\mathrm{a},k}$ [m^3], $\mathrm{V}^c_{\mathrm{a},k}$ [m^3], and the thermal energy content, $\mathrm{S}^h_{\mathrm{a},k}$ [Wh], $\mathrm{S}^c_{\mathrm{a},k}$ [Wh], of warm and cold wells. The superscripts "h" and "c" refer to the heating and cooling operating modes of an ATES system, respectively, and the subscript "a" denotes the ATES system variables. Consider the following first-order difference equations as ATES system model dynamics:

$$\mathrm{V}^h_{\mathrm{a},k+1} = \mathrm{V}^h_{\mathrm{a},k} - \tau(u^h_{\mathrm{a},k} - u^c_{\mathrm{a},k}), \tag{2.1a}$$

$$\mathrm{V}^c_{\mathrm{a},k+1} = \mathrm{V}^c_{\mathrm{a},k} + \tau(u^h_{\mathrm{a},k} - u^c_{\mathrm{a},k}), \tag{2.1b}$$

$$\mathrm{S}^h_{\mathrm{a},k+1} = \eta_{\mathrm{a},k}\,\mathrm{S}^h_{\mathrm{a},k} - (h^h_{\mathrm{a},k} - h^c_{\mathrm{a},k}), \tag{2.1c}$$

$$\mathrm{S}^c_{\mathrm{a},k+1} = \eta_{\mathrm{a},k}\,\mathrm{S}^c_{\mathrm{a},k} + (c^h_{\mathrm{a},k} - c^c_{\mathrm{a},k}), \tag{2.1d}$$

where $\eta_{\mathrm{a},k} \in (0,1)$ is a lumped coefficient of thermal energy losses in aquifers, $u^h_{\mathrm{a},k}$ and $u^c_{\mathrm{a},k}$ [m^{3}h^{-1}] are control variables corresponding to the pump flow rate of ATES system

during heating and cooling modes at each sampling time $k = 1, 2, \cdots$, respectively, with τ [h] as the sampling period. The variable $u^h_{\mathrm{a},k}$ circulates water from warm well to cold well, whereas $u^c_{\mathrm{a},k}$ takes water from cold well and injects into warm well of ATES system, during heating modes and cooling modes of the BCC system, respectively. The variables $h^h_{\mathrm{a},k}$ [Wh], $c^h_{\mathrm{a},k}$ [Wh] denote the thermal energy that is extracted from warm well and injected into cold well of ATES system during heating mode of BCC system, respectively. The variables $c^c_{\mathrm{a},k}$ [Wh], $h^c_{\mathrm{a},k}$ [Wh] are the thermal energy that is extracted from cold well and injected into warm well of ATES system during cooling mode of BCC system, respectively. These variable are defined by:

$$\begin{cases} h^h_{\mathrm{a},k} = \alpha_{h,k}\,\tau\,u^h_{\mathrm{a},k} \\ c^h_{\mathrm{a},k} = \alpha_{c,k}\,\tau\,u^h_{\mathrm{a},k} \end{cases}, \begin{cases} c^c_{\mathrm{a},k} = \alpha_{c,k}\,\tau\,u^c_{\mathrm{a},k} \\ h^c_{\mathrm{a},k} = \alpha_{h,k}\,\tau\,u^c_{\mathrm{a},k} \end{cases}, \tag{2.2}$$

where $\alpha_{h,k} = \rho_w\,c_{pw}\,(\mathrm{T}^h_{\mathrm{a},k} - \mathrm{T}^{\mathrm{amb}}_{\mathrm{a},k})$, and $\alpha_{c,k} = \rho_w\,c_{pw}\,(\mathrm{T}^{\mathrm{amb}}_{\mathrm{a},k} - \mathrm{T}^c_{\mathrm{a},k})$ are the thermal power coefficients of warm and cold wells, respectively. The parameters ρ_w [kgm^{-3}] and c_{pw} [Jkg^{-1}K^{-1}] are density and specific heat capacity of water, respectively. $\mathrm{T}^h_{\mathrm{a},k}$ [K], $\mathrm{T}^c_{\mathrm{a},k}$ [K], and $\mathrm{T}^{\mathrm{amb}}_{\mathrm{a},k}$ [K] denote the temperature of water inside warm well, cold well and the ambient aquifer, respectively. We also define $h_{\mathrm{a},k}$ [Wh], and $c_{\mathrm{a},k}$ [Wh] to be the amount of thermal energy that can be delivered to the building during heating and cooling modes, respectively, as follows:

$$\begin{cases} h_{\mathrm{a},k} = \alpha_k\,\tau\,u^h_{\mathrm{a},k} \\ c_{\mathrm{a},k} = \alpha_k\,\tau\,u^c_{\mathrm{a},k} \end{cases}, \tag{2.3}$$

where $\alpha_k = \alpha_{h,k} + \alpha_{c,k}$ is the total thermal power coefficient. In previous work [135], we have also developed a control-oriented model for the integrated ATES into BCC system where we consider the dynamical behavior of the volume and temperature of water in each well of ATES system.

It is important to note that the amount of cold energy content, $\mathrm{S}^c_{\mathrm{a},k}$, in the cold well is a function of the temperature difference between the temperature of the stored water in the cold well and the ambient aquifer temperature, $\mathrm{T}^c_{\mathrm{a},k} - \mathrm{T}^{\mathrm{amb}}_{\mathrm{a},k}$, which leads to have an unusual negative signed for the energy content in the cold well. Without loss of generality to avoid any further technical issue, we neglect the negative sign in the proposed model (2.1) by simply considering the temperature difference between the ambient aquifer temperature and the temperature of the stored water in the cold well, $\mathrm{T}^{\mathrm{amb}}_{\mathrm{a},k} - \mathrm{T}^c_{\mathrm{a},k}$. It is also considered to have no energy flow between wells in an ATES system through the aquifers based on the operator's experience and historical information.

Let us now discuss the dynamics of ATES system in (2.1). Equations (2.1a) to (2.1d) describe the evolution of water volume and the thermal energy content in warm and cold wells, respectively. During cold seasons, for the heating purpose of BCC system, the amount $\tau u^h_{\mathrm{a},k}$ of volume of warm water from the warm well is extracted to provide amount $h_{\mathrm{a},k}$ of thermal energy, and meanwhile, the amount $c^h_{\mathrm{a},k}$ of thermal energy is stored in the cold well of ATES system. As for the cooling purpose of BCC system during warm seasons, the amount $\tau u^c_{\mathrm{a},k}$ of volume of cold water from the cold well is extracted to provide amount $c_{\mathrm{a},k}$ of thermal energy, while the amount $h^c_{\mathrm{a},k}$ of thermal energy is

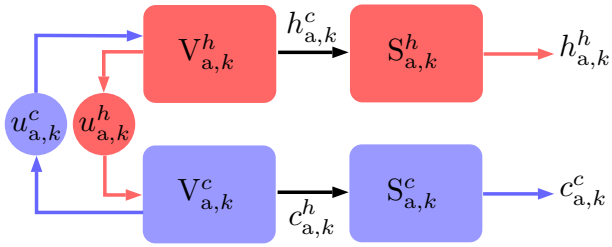


Figure 2.2: Operational block diagram of ATES system showing the relation between an ATES system variables. Operating modes during cold and warm seasons are shown via red color and blue color, respectively.

injected in the warm well of ATES system. The operations of the ATES system are visualized in Figure 2.2, which represents the relation between the variables of ATES system. Operating modes during cold and warm seasons are shown via red color and blue color, respectively. The following assumption is made due to the existing operational practice, and it is not restrictive for our proposed model.

Assumption 1 *There is either no operation or only one operating mode active in ATES systems, which leads to either both control variables being zero or only one control variable being nonzero at any time instant:*

$$u^h_{a,k}\, u^c_{a,k} = 0\,, \quad k = 1, 2, \cdots .$$

The dynamics of ATES system in (2.1) can be also written in a more compact format for each agent $i \in \{1, \cdots, N\}$:

$$x^{\mathrm{a}}_{i,k+1} = a^{\mathrm{a}}_{i,k} x^{\mathrm{a}}_{i,k} + b^{\mathrm{a}}_{i,k} u^{\mathrm{a}}_{i,k}\,, \tag{2.4}$$

where $x^{\mathrm{a}}_{i,k} = \begin{bmatrix} V^h_{a,k} & V^c_{a,k} & S^h_{a,k} & S^c_{a,k} \end{bmatrix}^\top \in \mathbb{R}^4$ and $u^{\mathrm{a}}_{i,k} = \begin{bmatrix} u^h_{a,k} & u^c_{a,k} \end{bmatrix}^\top \in \mathbb{R}^2$ denote the state and the control vectors, respectively, and $a^{\mathrm{a}}_{i,k}, b^{\mathrm{a}}_{i,k}$ can be obtained via (2.1). Note that there are some operational constraints on the ATES control variable as well,

$$u^{\min}_{\mathrm{a}} \le u^h_{a,k} \le u^{\max}_{\mathrm{a}}\,, \tag{2.5a}$$

$$u^{\min}_{\mathrm{a}} \le u^c_{a,k} \le u^{\max}_{\mathrm{a}}\,, \tag{2.5b}$$

where $u^{\min}_{\mathrm{a}}, u^{\max}_{\mathrm{a}}$ represent the minimum and maximum pump flow rate of ATES system, respectively.

The proposed model for an ATES system in (2.4) is a linear time-varying discrete-time system, due to the variation of the temperatures in both wells and the ambient aquifer (2.2). In Section 2.2.3, we will integrate (2.4) into a BCC system dynamics.

2.2.2. Thermal Energy Demand Profile

A dynamical model of building thermal energy demand was developed in our previous work [137] to determine the thermal energy demand of a building at each sampling time

k, considering the desired indoor air temperature and the outside weather conditions. We refer to the BCC system that determines the level of thermal energy demand $\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k}$ [Wh] at each sampling time k via

$$\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k} = f_{\mathrm{B}}(p^{\mathrm{B}}_{s}, \mathrm{T}^{\mathrm{B}}_{\mathrm{des},k}, \vartheta_k)\,, \tag{2.6}$$

where p^{B}_{s} corresponds to a parameter vector of the building characteristics, $\mathrm{T}^{\mathrm{B}}_{\mathrm{des},k}$ [°C] is the desired indoor air temperature of the building, and $\vartheta_k = [\mathrm{T}^{\mathrm{B}}_{o,k}, \mathrm{I}_{o,k}, \mathrm{v}_{o,k}, \mathrm{Q}_{p,k}, \mathrm{Q}_{e,k}] \in \mathbb{R}^5$ is a vector of uncertain variables that contains the outside air temperature, the solar radiation, the wind velocity, the thermal energy produced due to occupancy by people, and electrical devices, and lighting inside the building, respectively. This yields the building thermal energy demand that takes into account the overall building effects, e.g., zones, walls, humans and non-human thermal energy sources with the outside uncertain weather conditions. Since we are mainly interested in capturing the variation of thermal energy demand w.r.t. the outside air temperature $\mathrm{T}^{\mathrm{B}}_{o,k}$, the uncertain variable ϑ_k, is assigned to $\mathrm{T}^{\mathrm{B}}_{o,k}$, and the rest of the variables are fixed to their nominal (forecast) values at each sampling time k.

The operating modes (heating or cooling) of BCC system are determined based on the sign of $\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k}$ at each sampling time k. The variable $\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k}$ with positive and negative signs, represents the thermal energy demand during heating mode and the building surplus thermal energy during cooling mode, respectively. $\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k}$ is zero represents the comfort mode of building, and thus, in such a case no heating or cooling is requested. We also distinguish between the thermal energy demand of building during heating mode $h_{\mathrm{d},k}$, and cooling mode $c_{\mathrm{d},k}$, using the relation: $\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k} = h_{\mathrm{d},k} - c_{\mathrm{d},k}$.

The following technical assumption is necessary for the measurability of the uncertainty.

Assumption 2 *The mapping from the uncertain variable ϑ_k to the thermal energy demand $Q^{B}_{d,k}$ is a measurable function (2.6), so that $Q^{B}_{d,k}$ can be viewed as a random variable on the same probability space as ϑ_k. Moreover, the thermal energy demand at each sampling time k can be either zero (no thermal energy demand) or only for heating $h_{d,k}$ (cooling $c_{d,k}$) mode of the BCC system, which leads to:*

$$h_{d,k}\, c_{d,k} = 0\,,\ k = 1, 2, \cdots\,.$$

Figure 2.3 shows the thermal energy demand profile of a building for the last five years with respect to the outside registered weather data in The Netherlands. The top panel in Figure 2.3 depicts $\mathrm{Q}^{\mathrm{B}}_{\mathrm{d},k}$ as the result of (2.6), whereas the middle and bottom panels show the thermal energy demand during heating mode $h_{d,k}$ and the thermal energy surplus during cooling mode $c_{d,k}$, respectively.

2.2.3. BUILDING CLIMATE COMFORT SYSTEMS

Consider a single agent (i.e., building) $i \in \{1, \cdots, N\}$ that is facilitated with a boiler, a heat pump, a storage tank for the heating mode, and a chiller, a storage tank for the cooling mode together with an ATES system that is available for both operating modes (see Figure 2.4). We now focus on the modeling of energy balance for the BCC system.

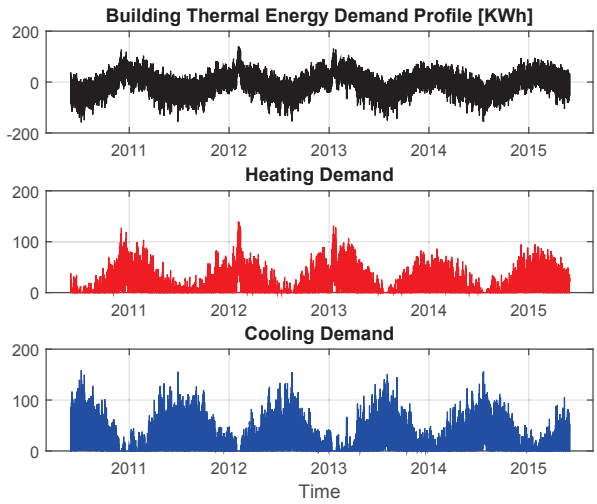


Figure 2.3: Thermal energy demand profile of a building during 2010-2015 with respect to the outside registered weather data in The Netherlands [5]. The black line shows $Q^{B}_{d,k}$, the red line is related to the thermal energy demand during heating mode $h_{d,k}$ and the blue line corresponds to the thermal energy surplus during cooling mode $c_{d,k}$.

Define two vectors of control variables during heating and cooling modes in each agent i at each sampling time k, to be

$$u^{h}_{i,k} = \begin{bmatrix} h_{\text{boi},k} & h_{\text{im},k} \end{bmatrix}^{\top} \in \mathbb{R}^2, \; u^{c}_{i,k} = \begin{bmatrix} c_{\text{chi},k} & c_{\text{im},k} \end{bmatrix}^{\top} \in \mathbb{R}^2.$$

The variables $h_{\text{boi},k}$, $c_{\text{chi},k}$, $h_{\text{im},k}$, and $c_{\text{im},k}$ denote the production of boiler, chiller, the imported energies from external parties during heating and cooling modes, respectively. We consider boiler and chiller operating limits that constrain their production within a certain bound for cost effective maintenance of such equipment. Define $v_{\text{boi},k} \in \{0,1\}$ and $v_{\text{chi},k} \in \{0,1\}$ to be two binary variables to decide about the ON/OFF status of the boiler and chiller, respectively. Consider now to the following conditional situations:

$$\text{boiler}: \begin{cases} v_{\text{boi},k} = 1 & \text{if} \quad h^{\min}_{\text{boi}} \leq h_{\text{boi},k} \leq h^{\max}_{\text{boi}} \\ v_{\text{boi},k} = 0 & \text{otherwise} \end{cases}, \tag{2.7a}$$

$$\text{chiller}: \begin{cases} v_{\text{chi},k} = 1 & \text{if} \quad c^{\min}_{\text{chi}} \leq c_{\text{chi},k} \leq c^{\max}_{\text{chi}} \\ v_{\text{chi},k} = 0 & \text{otherwise} \end{cases}, \tag{2.7b}$$

where $h^{\min}_{\text{boi}}$, $h^{\max}_{\text{boi}}$, $c^{\min}_{\text{chi}}$, $c^{\max}_{\text{chi}}$ denote the minimum and maximum capacity of thermal energy production of boiler and chiller, respectively.

We define two variables to capture the thermal energy imbalance errors during heating mode $x^{h}_{i,k} \in \mathbb{R}$, and an imbalance error of the cooling mode $x^{c}_{i,k} \in \mathbb{R}$. They are related to the difference between the level of the storage tank with the forecasted thermal energy demand, $h^{\text{f}}_{\text{d},k}$, $c^{\text{f}}_{\text{d},k}$, during heating and cooling modes, respectively, which are formally

defined using the following relations:

$$x^h_{i,k} = h_{\mathrm{s},k} - h^{\mathrm{f}}_{\mathrm{d},k}\,, \tag{2.8a}$$

$$x^c_{i,k} = c_{\mathrm{s},k} - c^{\mathrm{f}}_{\mathrm{d},k}\,. \tag{2.8b}$$

Herein, $h_{\mathrm{s},k}$, and $c_{\mathrm{s},k}$ represent the level of storage tank during heating and cooling modes, respectively, and obey the following dynamics:

$$h_{\mathrm{s},k+1} = \eta^h_{\mathrm{s},k} x^h_{i,k} + \eta^h_{\mathrm{s},k}\left(h_{\mathrm{boi},k} + h_{\mathrm{im},k} + \alpha_{\mathrm{hp},k} h_{\mathrm{a},k}\right), \tag{2.9}$$

$$c_{\mathrm{s},k+1} = \eta^c_{\mathrm{s},k} x^c_{i,k} + \eta^c_{\mathrm{s},k}\left(c_{\mathrm{chi},k} + c_{\mathrm{im},k} + c_{\mathrm{a},k}\right), \tag{2.10}$$

where $\alpha_{\mathrm{hp},k} = \mathrm{COP}_k(\mathrm{COP}_k - 1)^{-1}$ is related to the effect of the heat pump during heating mode and COP_k stands for the coefficient of performance of heat pump at each sampling time k. The parameters $\eta^h_{\mathrm{s},k}, \eta^c_{\mathrm{s},k} \in (0,1)$ denote the thermal loss coefficients due to inefficiency of storage tank during heating and cooling modes, respectively. The variables $h_{\mathrm{a},k}$ and $c_{\mathrm{a},k}$ are defined in (2.3) and are related to the ATES system model. It is important to note that $h_{\mathrm{a},k}$ and $c_{\mathrm{a},k}$ are dependent on the pump flow rates $u^h_{\mathrm{a},k}$ and $u^c_{\mathrm{a},k}$ of the ATES system during heating and cooling modes of the BCC system, respectively. We now substitute $h_{\mathrm{s},k}$, and $c_{\mathrm{s},k}$ as in (2.9) into (2.8) to derive the dynamical behavior of the thermal energy imbalance $x^h_{i,k}$ and $x^c_{i,k}$ that are given by

$$x^h_{i,k+1} = a^h_{i,k} x^h_{i,k} + b^h_i u^h_{i,k} + b^{\mathrm{a},h}_{i,k} u^{\mathrm{a}}_{i,k} + c^h_{i,k} w^h_{i,k}\,, \tag{2.11a}$$

$$x^c_{i,k+1} = a^c_{i,k} x^c_{i,k} + b^c_i u^c_{i,k} + b^{\mathrm{a},c}_{i,k} u^{\mathrm{a}}_{i,k} + c^c_{i,k} w^c_{i,k}\,, \tag{2.11b}$$

where $a^h_{i,k} = \eta^h_{\mathrm{s},k}$, $a^c_{i,k} = \eta^c_{\mathrm{s},k}$, $b^h_{i,k} = \begin{bmatrix}\eta^h_{\mathrm{s},k} & \eta^h_{\mathrm{s},k}\end{bmatrix}$, $b^c_{i,k} = \begin{bmatrix}\eta^c_{\mathrm{s},k} & \eta^c_{\mathrm{s},k}\end{bmatrix}$, $b^{\mathrm{a},h}_{i,k} = \begin{bmatrix}\tilde{\eta}^h_{\mathrm{s},k} & 0\end{bmatrix}$, $b^{\mathrm{a},c}_{i,k} = \begin{bmatrix}0 & \eta^c_{\mathrm{s},k}\alpha_k\tau\end{bmatrix}$, $c^h_i = -1$, $c^c_i = -1$, and $\tilde{\eta}^h_{\mathrm{s},k} = \eta^h_{\mathrm{s},k}\alpha_{\mathrm{hp},k}\alpha_k\tau$. The variables $w^h_{i,k} = h^{\mathrm{f}}_{\mathrm{d},k+1}$ and $w^c_{i,k} = c^{\mathrm{f}}_{\mathrm{d},k+1}$ refer to the forecast of thermal energy demand during heating and cooling modes in the next time step, respectively. The only uncertain variable in each agent i at each sampling time k is considered to be the deviation of actual thermal energy demand from its forecast value as defined in Section 2.2.2, and therefore, $w^h_{i,k}$ and $w^c_{i,k}$ represent uncertain parameters.

Consider now the system dynamics for each agent i by concatenating the thermal energy imbalance errors during heating and cooling modes (2.11) together with the state vector of the ATES system (2.4) as follows:

$$x_{i,k+1} = a_{i,k} x_{i,k} + b_{i,k} u_{i,k} + c_{i,k} w_{i,k}\,, \tag{2.12}$$

where $x_{i,k} = \begin{bmatrix} x^{h\top}_{i,k} & x^{c\top}_{i,k} & x^{\mathrm{a}\top}_{i,k}\end{bmatrix}^\top \in \mathbb{R}^6$ and $u_{i,k} = \begin{bmatrix} u^{h\top}_{i,k} & u^{c\top}_{i,k} & u^{\mathrm{a}\top}_{i,k}\end{bmatrix}^\top \in \mathbb{R}^6$ denote the state and the control vectors, respectively, and $w_{i,k} = \begin{bmatrix} w^h_{i,k} & w^c_{i,k}\end{bmatrix}^\top \in \mathcal{W}_{i,k} \subseteq \mathbb{R}^2$ is the uncertainty vector such that $\mathcal{W}_{i,k}$ is an unknown uncertainty set. The system matrices $a_{i,k}, b_{i,k}, c_{i,k}$ can be readily derived from their definitions and we omit them in the interest of space.

The proposed model for a BCC system in (2.12) is a stochastic hybrid linear time-varying discrete-time system. It is important to note that the hybrid nature of (2.12) is

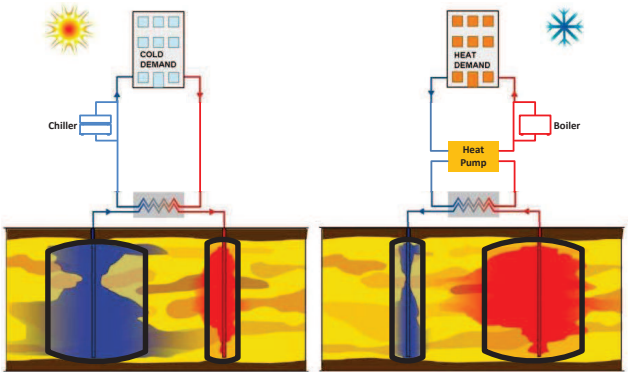


Figure 2.4: Heating and cooling operating modes of BCC system with an ATES system during warm (left) and cold (right) seasons.

due to the fact that each equipment (boiler and chiller) can be either ON or OFF as in (2.7) depending on heating and cooling modes of the building. This possibility therefore changes the proposed thermal energy imbalance error dynamics (2.11).

In order to provide a desired thermal comfort for each BCC system in the following section, we will develop a control framework based on the MPC paradigm where (2.12) is used to predict the thermal energy imbalance error dynamics together with the ATES system dynamics for each agent $i \in \{1, \cdots, N\}$, and then, extend this to a network of interconnected BCC systems. Moreover, we will provide a solution method to overcome an important challenge of the network of BCC systems due to the spatial distribution of ATES systems. An important remark is that the variations of system parameters in the proposed dynamical model (2.12) evolve on a much slower time-scale compared to the system dynamics and, therefore, we consider the system dynamics (2.12) to be time-invariant in the following parts. It is worth mentioning that our proposed control technique in this chapter can be easily extended to cope with time-varying parameters by considering them as multiplicative uncertainty sources, see e.g., [146, 147].

2.3. ENERGY MANAGEMENT PROBLEM

In this section, we formulate an optimization problem for heating and cooling modes of the BCC system integrated with ATES which we refer to as a single agent energy management problem. We then extend the single agent problem to a network with multiple agents that can be producers and consumers of thermal energy in an STG setting. In such a setting, there might be some unwanted overlap (mutual interactions) between the stored water in the wells of neighboring ATES systems (see Figure 2.5) through aquifers. Such an unwanted mutual interactions between warm and cold wells, clearly, reduce the energy efficiency of the ATES systems. The goal of the agents is to match the local consumption and production and to avoid mutual interactions between their ATES systems in the network and thereby improve energy efficiency.

2.3.1. ENERGY BALANCE IN SINGLE AGENT SYSTEM

Consider an MPC problem with a finite prediction horizon N_h for agent $i \in \{1, \cdots, N\}$, and introduce the subscript t in our notation to characterize the value of the planning quantities for a given time $t \in \mathcal{T}$, where the set of predicted time steps is denoted by $\mathcal{T} := \{k, k+1, \cdots, k+N_h-1\}$. Using the subscript $t|k$, we refer to the t time step prediction of variables at the simulation time step k.

Define $v_{i,t|k} = \begin{bmatrix} v_{\text{boi},t|k} & v_{\text{chi},t|k} \end{bmatrix}^\top \in \{0,1\}^2$ as a vector of binary variables to decide about the ON/OFF status of boiler and chiller in each agent $i \in \{1, \cdots, N\}$. We also take into account the startup cost of boiler and chiller using $c^{\text{su}}_{i,t|k} = \begin{bmatrix} c^{\text{su}}_{\text{boi},t|k} & c^{\text{su}}_{\text{chi},t|k} \end{bmatrix}^\top$ and add $c^{\text{su}}_{i,t|k}$ into the control decision variables $u_{i,t|k} = \begin{bmatrix} u^{h\top}_{i,t|k} & u^{c\top}_{i,t|k} & u^{\text{a}\top}_{i,t|k} & c^{\text{su}\top}_{i,t|k} \end{bmatrix}^\top \in \mathbb{R}^8$ for each agent i at each time step $t|k$.

The goal of each agent i is to map the local thermal energy supply of production units to the local thermal energy demand of BCC system. Our goal thus is to formulate an optimization problem to find the control input $u_{i,t|k}$ for each agent i such that the thermal energy imbalance errors stay as a small as possible at minimal production cost and to satisfy physical constraints of heating and cooling modes equipment at each sampling time k. We therefore associate a quadratic cost function with each agent i at each prediction time step k as follows:

$$J_i(x_{i,t|k}, u_{i,t|k}) = x^\top_{i,t|k} Q_i \, x_{i,t|k} + u^\top_{i,t|k} R_i \, u_{i,t|k} \,, \tag{2.13}$$

where $Q_i = \text{diag}\left(\begin{bmatrix} q^h_i & q^c_i & \mathbf{0}_{1\times 4} \end{bmatrix}\right) \in \mathbb{R}^{6\times 6}$ is a weighting matrix coefficient of thermal energy imbalance errors, $R_i = \text{diag}(r_i) \in \mathbb{R}^{8\times 8}$ indicates a diagonal matrix with the cost vector r_i on its diagonal, and r_i is defined as

$$r_i = \begin{bmatrix} r_{\text{boi}} & r^h_{\text{im}} & r_{\text{chi}} & r^c_{\text{im}} & r^h_{\text{a}} & r^c_{\text{a}} & 1 & 1 \end{bmatrix}^\top \in \mathbb{R}^8 \,,$$

where r_{boi} (r_{chi}) represents the cost of natural gas that is used by boiler (chiller), r^h_{im} (r^c_{im}) denotes the cost of imported thermal energy from an external party during heating (cooling) mode, and r^h_{a} (r^c_{a}) corresponds to the pumping electricity cost of ATES system to extract the required thermal energy during heating (cooling) modes. The other entries of r_i represent the start-up costs. The proposed cost function consists of two main parts which leads to the regulation of imbalance errors to zero at minimal production cost together with minimum energy balance error of ATES system in each agent i. The reason for introducing a cost function in this form is that from a computational perspective quadratic cost functions are motivated by convexity and differentiability arguments. Note that the cost function $J_i(\cdot)$ is a random variable due to the uncertain state variables, and thus, we consider $\mathbb{E}[J_i(\cdot)]$ to obtain a deterministic cost function.

We are now in a position to formulate a finite-horizon stochastic hybrid control problem as the local energy management problem for each agent $i \in \{1, \cdots, N\}$ using the fol-

lowing chance-constrained mixed-integer optimization problem:

$$\underset{\{u_{i,t|k},v_{i,t|k}\}_{t\in\mathcal{T}}}{\text{minimize}} \quad \sum_{t\in\mathcal{T}} \mathbb{E}\big[J_i(x_{i,t|k},u_{i,t|k})\big] \tag{2.14a}$$

$$\text{subject to} \quad c^{su}_{i,t|k} \geq \Lambda^{su}(v_{i,t|k}-v_{i,t-1|k}),\, c^{su}_{i,t|k}\geq 0, \forall t\in\mathcal{T} \tag{2.14b}$$

$$v_{\text{boi},t|k}h^{\min}_{\text{boi}} \leq h_{\text{boi},t|k} \leq h^{\max}_{\text{boi}} v_{\text{boi},t|k}, \forall t\in\mathcal{T} \tag{2.14c}$$

$$v_{\text{chi},t|k}c^{\min}_{\text{chi}} \leq c_{\text{chi},t|k} \leq c^{\max}_{\text{chi}} v_{\text{chi},t|k}, \forall t\in\mathcal{T} \tag{2.14d}$$

$$h^{\min}_{\text{im}} \leq h_{\text{im},t|k} \leq h^{\max}_{\text{im}}, \forall t\in\mathcal{T} \tag{2.14e}$$

$$c^{\min}_{\text{im}} \leq c_{\text{im},t|k} \leq c^{\max}_{\text{im}}, \forall t\in\mathcal{T} \tag{2.14f}$$

$$u^{\min}_{\text{a}} \leq u^{h}_{\text{a},t|k} \leq u^{\max}_{\text{a}}, \forall t\in\mathcal{T} \tag{2.14g}$$

$$u^{\min}_{\text{a}} \leq u^{c}_{\text{a},t|k} \leq u^{\max}_{\text{a}}, \forall t\in\mathcal{T} \tag{2.14h}$$

$$\mathbb{P}\big\{x_{i,t+1|k}\geq 0 \,\big|\, x_{i,t|k}, \forall t\in\mathcal{T}\big\} \geq 1-\varepsilon_i, \{w_{i,t|k}\}_{t\in\mathcal{T}}\in\mathcal{W}_i, \tag{2.14i}$$

where Λ^{su} is a diagonal matrix including the startup costs of boiler and chiller on the diagonal, $h^{\min}_{\text{im}}$, $h^{\max}_{\text{im}}$, $c^{\min}_{\text{im}}$, $c^{\max}_{\text{im}}$ are the minimum and maximum capacity of thermal energy production for each external party during heating and cooling modes, respectively. $\varepsilon_i \in (0,1)$ is the admissible constraint violation parameter. Note that $\mathcal{W}_i$ represents the Cartesian product of $\mathcal{W}_{i,t|k}$ for all $t\in\mathcal{T}$.

In order of appearance, the constraints have the following meaning. The constraint (2.14b) captures the status change of boiler and chiller (from OFF to ON). Note that the status change from ON to OFF never appears in the cost function due to the positivity constraint of $c^{su}_{i,t|k}\geq 0$. (2.14c), (2.14d), (2.14e), (2.14f), (2.14g), (2.14h) impose box constraints (capacity limitations) on their variables. In the given lower and upper bounds of both constraints (2.14c) and (2.14d), there are multiplications with binary variables which enforce the status change of boiler and chiller, respectively. Constraint (2.14i) ensures probabilistically feasible trajectories of the thermal energy imbalance errors for in each agent w.r.t all possible realization of the uncertain variables $w^{h}_{i,t|k}$ and $w^{c}_{i,t|k}$ for all predicted time step $t\in\mathcal{T}$.

To extend the proposed formulation (2.14) to the energy management problem of smart thermal grids, we need to introduce the notation, $\boldsymbol{x}_i := \{x_{i,t+1|k}\}_{t\in\mathcal{T}} \in \mathbb{R}^{n_x}$, $\boldsymbol{u}_i := \{u_{i,t|k}\}_{t\in\mathcal{T}} \in \mathbb{R}^{n_u}$, $\boldsymbol{v}_i := \{v_{i,t|k}\}_{t\in\mathcal{T}} \in \mathbb{R}^{n_v}$, and $\boldsymbol{w}_i := \{w_{i,t|k}\}_{t\in\mathcal{T}} \in \mathbb{R}^{n_w}$, where $n_x := 6N_h$, $n_u := 9N_h$, $n_v := 2N_h$, and $n_w := 2N_h$. Given the initial value of the state $x_{i,k}$, one can eliminate the state variables from the dynamics (2.12) of each agent i:

$$\boldsymbol{x}_i = A_i x_{i,k} + B_i \boldsymbol{u}_i + C_i \boldsymbol{w}_i, \tag{2.15}$$

where the exact form of A_i, B_i and C_i matrices are omitted in the interest of space and can be found in [21, Section 9.5]. We can now rewrite the total cost function over the prediction horizon in a more compact form as follows:

$$\mathcal{J}_i(\boldsymbol{x}_i,\boldsymbol{u}_i) = \boldsymbol{x}_i^\top \boldsymbol{Q}_i \boldsymbol{x}_i + \boldsymbol{u}_i^\top \boldsymbol{R}_i \boldsymbol{u}_i,$$

where $\boldsymbol{Q}_i$ and $\boldsymbol{R}_i$ are two block-diagonal matrices with Q_i and R_i on the diagonal for each agent i. Note that the sum $\sum(\cdot)$ and the expectation $\mathbb{E}[\cdot]$ in the cost function (2.14a) are

linear operators and thus, we can change their order without loss of generality. Consider now the reformulation of (2.14) in a more compact form as follows:

$$\min_{\boldsymbol{u}_i,\boldsymbol{v}_i} \quad \mathcal{V}_i(\boldsymbol{x}_i,\boldsymbol{u}_i) = \mathbb{E}_{\boldsymbol{w}_i}\left[\mathcal{J}_i(\boldsymbol{x}_i,\boldsymbol{u}_i)\right] \tag{2.16a}$$

$$\text{s.t.} \quad E_i\boldsymbol{u}_i + F_i\boldsymbol{v}_i + P_i \le 0, \quad \forall \boldsymbol{w}_i \in \mathcal{W}_i \tag{2.16b}$$

$$\mathbb{P}_{\boldsymbol{w}_i}\left[A_i x_{i,k} + B_i\boldsymbol{u}_i + C_i\boldsymbol{w}_i \ge 0\right] \ge 1-\varepsilon_i, \tag{2.16c}$$

where E_i, F_i, P_i are matrices that are built by concatenating all constraints in (2.14). The index of $\mathbb{E}_{\boldsymbol{w}_i}, \mathbb{P}_{\boldsymbol{w}_i}$ denotes the dependency of the state trajectory $\boldsymbol{x}_i$ on the string of random scenarios $\boldsymbol{w}_i$ for each agent i. The following technical assumption is adopted.

Assumption 3 *The variable $\boldsymbol{w}_i$ is defined on $(\mathcal{W}_i, \mathfrak{B}(\mathcal{W}_i), \mathbb{P}_{\boldsymbol{w}_i})$, where $\mathcal{W}_i \subseteq \mathbb{R}^{n_w}$, $\mathfrak{B}(\cdot)$ denotes a Borel σ-algebra, and $\mathbb{P}_{\boldsymbol{w}_i}$ is a probability measure defined over $\mathcal{W}_i$.*

It is worth to mention that for our study we only need a finite number of instances of $\boldsymbol{w}_i$, and we do not require the probability space $\mathcal{W}_i$ and the probability measure $\mathbb{P}_{\boldsymbol{w}_i}$ to be known explicitly. The availability of a number of scenarios from the sample space $\mathcal{W}_i$ is enough which will become concrete in later parts of the chapter. Such samples can be for instance obtained from historical data.

The proposed optimization problem (2.16) is a finite-horizon, chance-constrained mixed-integer quadratic program, whose stages are coupled by binary variables (2.14b), and dynamics of the imbalance error (2.14i) for each agent i at each sampling time k. It is important to note that the proposed problem (2.16) is in general a non-convex problem and hard to solve. In the following section, we will develop a tractable framework to obtain a probabilistically feasible solution for each agent i. We refer to the proposed optimization problem (2.16) as a single agent problem, and whenever all agents solve this problem separately in a receding horizon fashion without any coupling constraints, it is referred to as the *decoupled solution* (DS) in the subsequent parts. We next extend the proposed single agent optimization problem (2.16) into an STG setting.

2.3.2. ATES IN SMART THERMAL GRIDS

Consider a regional thermal grid consisting of N agents with heterogeneous parameters as it was developed in the previous part. Such an STG setting however can lead to unwanted mutual interactions between ATES systems as it is illustrated in Figure 2.5. We therefore need to introduce a proper coupling constraint between neighboring agents that makes use of the following assumption.

Assumption 4 *It is considered that each well of an ATES system to be a growing reservoir with respect to the horizontal axis (see black solid line in Figure 2.4). We therefore assume to have a cylindrical reservoir with a fixed height ℓ [m] (filter screen length) and a growing radius $r^h_{a,k}, r^c_{a,k}$ [m] (thermal radius) for each well of an ATES system.*

Using the volume of stored water in each well of ATES system, one can determine the thermal radius using

$$r^h_{\mathrm{a},k} = \left(\frac{c_{pw}\mathrm{V}^h_{\mathrm{a},k}}{c_{aq}\pi\ell}\right)^{0.5}, \quad r^c_{\mathrm{a},k} = \left(\frac{c_{pw}\mathrm{V}^c_{\mathrm{a},k}}{c_{aq}\pi\ell}\right)^{0.5}, \tag{2.17}$$

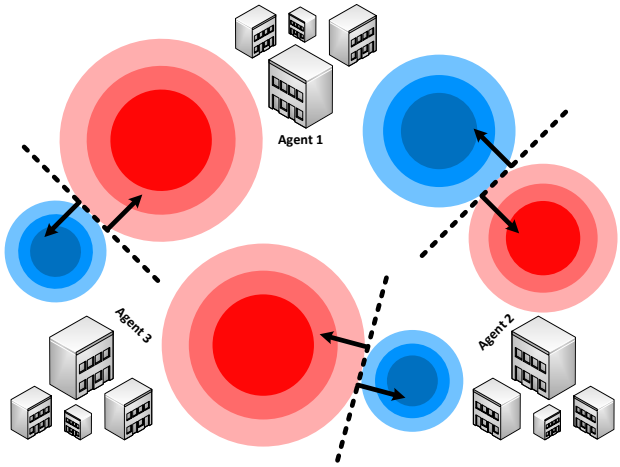


Figure 2.5: Three-agent ATES system in an STG. Each agent has a single ATES system which consists of a warm and a cold well. Horizontal cross sections of warm and cold wells are shown with red and blue circles. The black dashed lines represent the unwanted mutual interactions between ATES systems.

where $c_{aq} = (1 - n_p)c_{\text{sand}} + n_p c_{pw}$ is the aquifer heat capacity. c_{sand} [Jkg^{-1}K^{-1}] relates to the sand specific heat capacity, and n_p [−] is the porosity of aquifer. Let us now denote the set of neighbors of agent i by

$$\mathcal{N}_i \subseteq \{1, 2, \cdots, N\} \setminus \{i\} \,.$$

We impose a limitation on the thermal radius of warm well $r^h_{\text{a},k}$ and cold well $r^c_{\text{a},k}$ of ATES system in each agent i, based on the corresponding wells of its neighbor $j \in \mathcal{N}_i$:

$$(r^h_{\text{a},k})_i + (r^c_{\text{a},k})_j \leq d_{ij} \,, \; j \in \mathcal{N}_i \,, \tag{2.18}$$

where d_{ij} is a given distance between agent i and its neighbor $j \in \mathcal{N}_i$. This constraint prevents overlapping between the growing domains of warm and cold wells of ATES systems in an STG setting. Due to the nonlinear transformation in (2.17), we propose the following reformulation of this constraint to simplify the problem:

$$(\mathrm{V}^h_{\text{a},k})_i + (\mathrm{V}^c_{\text{a},k})_j \leq \mathrm{V}_{ij} - \bar{\delta}_{ij,k} \,, \tag{2.19}$$

where $\mathrm{V}_{ij} = c_{aq}\pi\ell\,(d_{ij})^2/c_{pw}$ denotes the total volume of common resource pool between agent i and its neighbor $j \in \mathcal{N}_i$. The variable $\bar{\delta}_{ij,k} = 2c_{aq}\pi\ell\,(\bar{r}^h_{\text{a},k})_i\,(\bar{r}^c_{\text{a},k})_j/c_{pw}$ represents a time-varying parameter that captures the mismatch between the linear and nonlinear constraint relations. The following corollary is a direct result of the above reformulation.

Corollary 1 *If $(\bar{r}^h_{a,k})_i$ and $(\bar{r}^c_{a,k})_j$ represent the current thermal radius of warm and cold wells of ATES system in agent i and j, respectively, then constraints (2.18) and (2.19) are equivalent.*

The proof is provided in Appendix A. □

Definition 1 *We define $\delta_{ij,k}$ to be a common uncertainty source between each agent i and its neighboring agent $j \in \mathcal{N}_i$, using the following model:*

$$\delta_{ij,k} := \bar{\delta}_{ij,k}(1 \pm 0.1\zeta), \tag{2.20}$$

where ζ is a random variable defined on some probability space, $\bar{\delta}_{ij,k}$ is constructed by using two given possible $(\bar{r}^h_{a,k})_i, (\bar{r}^c_{a,k})_j$ realizations that can be obtained using historical data in the DS framework. Since the mapping (2.20) from ζ to $\delta_{ij,k}$ is measurable, one can view $\delta_{ij,k}$ as a random variable on the same probability space as ζ.

2.3.3. PROBLEM FORMULATION IN MULTI-AGENT NETWORK

We now formulate the energy management problem for ATES systems in STGs as follows:

$$\min_{\{\boldsymbol{u}_i,\boldsymbol{v}_i\}_{i=1}^N} \quad \sum_{i=1}^N \mathcal{V}_i(\boldsymbol{x}_i, \boldsymbol{u}_i) \tag{2.21a}$$

$$\text{s.t.} \quad E_i\boldsymbol{u}_i + F_i\boldsymbol{v}_i + P_i \le 0, \tag{2.21b}$$

$$\mathbb{P}_{\boldsymbol{w}_i}\left[A_i x_{i,k} + B_i\boldsymbol{u}_i + C_i\boldsymbol{w}_i \ge 0\right] \ge 1 - \varepsilon_i, \forall \boldsymbol{w}_i \in \mathcal{W}_i, \tag{2.21c}$$

$$\mathbb{P}_{\boldsymbol{\delta}_{ij}}\left[H_i\boldsymbol{x}_i + H_j\boldsymbol{x}_j \le \bar{V}_{ij} - \boldsymbol{\delta}_{ij}\right] \ge 1 - \bar{\varepsilon}_{ij}, , \forall \boldsymbol{\delta}_{ij} \in \Delta_{ij}, \tag{2.21d}$$

$$\forall j \in \mathcal{N}_i, \forall i \in \{1, 2, \cdots, N\},$$

where H_i, H_j are coefficient matrices of appropriate dimensions, $\bar{V}_{ij} \in \mathbb{R}^{N_h}$ is the upper-bound on the total common resource pool, $\boldsymbol{\delta}_{ij}$ is a vector of common uncertainty variables, and $\bar{\varepsilon}_{ij} \in (0,1)$ denotes the level of admissible coupling constraint violation for each agent i and $\forall j \in \mathcal{N}_i$. $\bar{V}_{ij}$ can be expressed as $\bar{V}_{ij} = \mathbf{1}^{N_h} \otimes V_{ij}$, using the Kronecker product. Notice that the index of $\mathbb{P}_{\boldsymbol{\delta}_{ij}}$ denotes the dependency of the state trajectories on the string of random common scenarios $\boldsymbol{\delta}_{ij} = \{\delta_{ij,t|k}\}_{t\in\mathcal{T}} \subseteq \mathbb{R}^{N_h =: n_\delta}$.

Assumption 5 *The variable $\boldsymbol{\delta}_{ij}$ is considered to be a random vector on some probability space $(\Delta_{ij}, \mathfrak{B}(\Delta_{ij}), \mathbb{P}_{\boldsymbol{\delta}_{ij}})$, where $\Delta_{ij} \subseteq \mathbb{R}^{n_\delta}$, $\mathfrak{B}(\cdot)$ denotes a Borel σ-algebra, and $\mathbb{P}_{\boldsymbol{\delta}_{ij}}$ is a probability measure defined over Δ_{ij}.*

Assumption 6 *The variables $\boldsymbol{w}_i \in \mathbb{R}^{n_w}$ and $\boldsymbol{\delta}_{ij} \in \mathbb{R}^{n_\delta}$ are two vectors of independent random scenarios from two disjoint probability spaces $\mathcal{W}_i$ and Δ_{ij}, respectively.*

We refer to the proposed optimization problem (2.21) as a multi-agent network problem, and whenever the proposed problem (2.21) is solved in a receding horizon fashion, it is mentioned as the *centralized solution* (CS) in the following parts. The feasible set of (2.21) is in general non-convex and hard to determine explicitly due to the presence of chance constraints (2.21c), (2.21d). In what follows, we will develop a tractable framework to obtain probabilistically feasible solutions for all agents.

2.3.4. MOVE-BLOCKING SCHEME

The proposed system dynamics in (2.12) for each agent i consists of a BCC system dynamics (2.6) with typically an hourly-based operation, and an ATES system (2.4) that is based on a seasonal variation of desired optimal operation. This leads to a control problem that is sensitive w.r.t. the prediction horizon length, e.g., (2.16) and (2.21). Using a fixed prediction horizon length, e.g., least common multiple of these two systems, may turn out to be computationally prohibitive, however, also necessary in order to represent ATES interaction dynamics. We therefore aim to formulate a move-blocking strategy to reduce the number of control variables.

Consider $\mathcal{T} = \{k, k+1, \cdots, k+N_h-1\}$ to be the set of sampling time instances within the full prediction horizon, and $\mathcal{T}_u = \{\tau_1, \tau_2, \cdots, \tau_{T_u}\} \subseteq \mathcal{N}_h$ to be the set of sampling instances at which the control input is updated with $T_u = |\mathcal{T}_u|$. We introduce a new vector of multi-rate decision variables $\tilde{\boldsymbol{u}}_i \in \mathbb{R}^{N_u T_u}$ which are related to the original ones by:

$$\boldsymbol{u}_i = \boldsymbol{\Psi}\tilde{\boldsymbol{u}}_i \,, \tag{2.22}$$

where $\boldsymbol{\Psi} = \begin{bmatrix}\Psi_1 & \Psi_2 & \cdots & \Psi_{T_u}\end{bmatrix} \in \mathbb{R}^{N_u N_h \times N_u T_u}$ is a linear mapping matrix. For all $m \in \{1, \cdots, T_u\}$, we construct

$$\Psi_m = \begin{bmatrix}\psi_{1,m}^\top & \psi_{2,m}^\top & \cdots & \psi_{N_h,m}^\top\end{bmatrix}^\top \in \mathbb{R}^{N_u N_h \times N_u} \,, \tag{2.23}$$

where $\psi_{l,m} \in \mathbb{R}^{N_u \times N_u}$ for all $l \in \{1, 2, \cdots, N_h\}$ is defined as

$$\psi_{l,m} = \begin{cases} \mathbf{1} & \text{if } k+l-1 = \tau_m \\ 0 & \text{otherwise} \end{cases} \,, \tag{2.24}$$

where $\mathbf{1} \in \mathbb{R}^{N_u \times N_u}$ represents an identity matrix.

We reformulate the optimization problem (2.21) using the proposed move-blocking scheme (2.24), and whenever the reformulation of (2.21) is solved in a receding horizon fashion, it is referred to as the *move-blocking centralized solution* (MCS).

2.4. COMPUTATIONALLY TRACTABLE FRAMEWORK

In this section, we provide a framework to approximately solve the chance-constrained optimization problem (2.21), which is in general difficult to solve. To this end, we employ a data-driven approach to approximate the chance constraints using some available samples of uncertainties. We first extract at random some instances of the uncertainties (scenarios), where the scenarios are independent and identically distributed (i.i.d.) and, then, find the optimal solution of the problem with only the constraints associated with the extracted scenarios.

An important requirement of such a technique is to have a convex problem w.r.t the decision variables, which is not the case in our formulation (2.21). To tackle such a mixed-integer chance-constrained problem, one can use a worst-case mixed-integer reformulation technique as it was initially introduced in [109]. Due to the large-scale network problem (2.21), such a reformulation leads to enormous cost of computation and it is indeed an intractable approach. Following the so-called robust randomized

technique [98], the reformulation is done in a way to provide a feasible solution for all scenarios of the uncertainty realizations in a probabilistic sense.

The idea of robust randomized approach is the following. First, an auxiliary chance-constrained optimization problem is formulated to determine a probabilistic bounded set of random variables. This yields a bounded set of uncertainty that is a subset of the uncertainty space and contains a portion of the probability mass of the uncertainty with high confidence level. Then, a robust version of the initial problem subject to the uncertainty confined in the obtained set is solved. We here extend this framework in order to be able to handle a problem with multiple chance constraints based on the idea of the robust randomized approach [98].

Consider $\boldsymbol{y}_i = (\boldsymbol{u}_i, \boldsymbol{v}_i) \in \mathbb{R}^{(n_u+n_v)=n_y}$, $\boldsymbol{y} = \mathrm{col}(\boldsymbol{y}_i)_{i=1}^N$, where $\mathrm{col}(\cdot)$ is an operator to stack elements. Define $\boldsymbol{w} = \mathrm{col}(\boldsymbol{w}_i)_{i=1}^N \subseteq \mathcal{W}$ to be the private uncertainty sources for a network of agents, $\boldsymbol{\delta}_i = \mathrm{col}(\boldsymbol{\delta}_j)_{j\in\mathcal{N}_i} \subseteq \Delta_i$ to be the common uncertainty sources for each agent, and $\boldsymbol{\delta} = \mathrm{col}(\boldsymbol{\delta}_i)_{i=1}^N \subseteq \Delta$ to be the common uncertainty sources for a multi-agent network, where

$$\mathcal{W} := \prod_{i=1}^N \mathcal{W}_i \,,\ \Delta_i := \prod_{j\in\mathcal{N}_i} \Delta_{ij} \,,\ \Delta := \prod_{i=1}^N \Delta_i \,.$$

Consider now the proposed optimization problem in (2.21) in a more compact format:

$$\min_{\boldsymbol{y}} \quad \sum_{i=1}^N \mathcal{V}_i(\boldsymbol{x}_i, \boldsymbol{u}_i) \tag{2.25a}$$

$$\text{s.t.} \quad \mathbb{P}_{\boldsymbol{w}}\left[\boldsymbol{y} \in \prod_{i=1}^N \mathcal{Y}_i(\boldsymbol{w}_i)\right] \geq 1-\varepsilon \,,\ \forall \boldsymbol{w} \in \mathcal{W} \tag{2.25b}$$

$$\mathbb{P}_{\boldsymbol{\delta}}\left[\boldsymbol{y} \in \prod_{i=1}^N \bigcap_{j\in\mathcal{N}_i} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij})\right] \geq 1-\bar{\varepsilon} \,,\ \forall \boldsymbol{\delta} \in \Delta \tag{2.25c}$$

where $\varepsilon := \sum_{i=1}^N \varepsilon_i \in (0,1)$, $\bar{\varepsilon} := \sum_{i=1}^N \sum_{j\in\mathcal{N}_i} \bar{\varepsilon}_{ij} \in (0,1)$. $\mathcal{Y}_i(\boldsymbol{w}_i) \in \mathbb{R}^{n_y}$ and $\mathcal{Y}_{ij}(\boldsymbol{\delta}_{ij}) \in \mathbb{R}^{n_y}$ are defined[1] by

$$\mathcal{Y}_i(\boldsymbol{w}_i) := \left\{\boldsymbol{y}_i \in \mathbb{R}^{n_y} : E_i\boldsymbol{u}_i + F_i\boldsymbol{v}_i + P_i \leq 0 \,, , A_i x_{i,k} + B_i\boldsymbol{u}_i + C_i\boldsymbol{w}_i \geq 0\right\} ,$$
$$\mathcal{Y}_{ij}(\boldsymbol{\delta}_{ij}) := \left\{\left(\boldsymbol{y}_i, \boldsymbol{y}_j\right) \in \mathbb{R}^{2n_y} : H_i\boldsymbol{x}_i + H_j\boldsymbol{x}_j \leq \bar{V}_{ij} - \boldsymbol{\delta}_{ij}\right\}.$$

It is important to note that $\check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij}) \in \mathbb{R}^{2n_yN_i}$ represents the cylindrical extension[2] of $\mathcal{Y}_{ij}(\boldsymbol{\delta}_{ij})$. In the subsequent parts, we refer to the constraint (2.25b) as the agents' private chance constraints, and to the constraint (2.25c) as the agents' common chance constraints. The proposed formulation (2.25) is a mixed-integer quadratic optimization problem with multiple chance constraints, due to the binary variables $\{\boldsymbol{v}_i\}_{i=1}^N$ and the

[1] Both sets have a dependency on the initial value of the state $x_{i,k}$ for each agent i at each sampling time k. Given $x_{i,k}$, we here highlight the dependency of these sets on the uncertainties $\boldsymbol{w}_i$ and $\boldsymbol{\delta}_{ij}$ for each agent i at each sampling time k.

[2] Cylindrical extension replicates the membership degrees from the existing dimensions into the new dimensions [178, Section 4].

chance constraints (2.25b), (2.25c). The index of $\mathbb{P}_{\boldsymbol{w}}$ and $\mathbb{P}_{\boldsymbol{\delta}}$ denote the dependency on the string of random scenarios $\boldsymbol{w} \in \mathcal{W}$ and $\boldsymbol{\delta} \in \Delta$, respectively.

Building upon our previous work in [144], we extend the so-called robust randomized approach in [98, 100] to be able to handle a problem with multiple chance constraints. Problem (2.25) is a stochastic program with multiple chance constraints, where $\mathbb{P}_{\boldsymbol{w}}$ and $\mathbb{P}_{\boldsymbol{\delta}}$ denote two different probability measures related to the private and common uncertainty sources, respectively. In summary, one can reformulate the chance constraints in (2.25) using a worst-case chance constraint defined by

$$\max_{\eta \in \mathcal{N}_{\text{MCP}}} \mathbb{P}\left[f_\eta(\boldsymbol{y}, \cdot)\right] \geq 1 - \tilde{\varepsilon}\,, \tag{2.26}$$

where $\tilde{\varepsilon} = \min_{\eta \in \mathcal{N}_{\text{MCP}}}\{\varepsilon_\eta\}$, $f_\eta(\boldsymbol{y}, \cdot)$ denotes the η-th chance constraint function, and $\mathcal{N}_{\text{MCP}}$ is the set of indices of chance constraint functions formulated in (2.25). However, this procedure clearly leads to a considerable amount of conservatism, due to the fact that it requires the solution to satisfy all constraints with the highest probability $1 - \tilde{\varepsilon}$. We instead employ the robust randomized approach for each chance constraint function $f_k(\boldsymbol{y}, \cdot), k \in \mathcal{N}_{\text{MCP}}$, separately. Our framework is also related to, albeit different from the approach of [158], since the feasible set in (2.25) is non-convex. Moreover, the problem formulation in [158] consists of an objective function with multiple chance constraints, in which the terms in objective and constraints are univariate w.r.t. the decision variables. In contrast, the objective function in our problem formulation (2.25) consists of separable additive components and constraint functions are also separable w.r.t. (2.25b), (2.25c) between each agent $i = 1, \cdots, N$ and $\forall\, j \in \mathcal{N}_i$.

Define $\mathcal{B}_i, \bar{\mathcal{B}}_{ij}$ to be two bounded sets of private and common uncertainty sources for each agent i, respectively. $\mathcal{B}_i$, $\bar{\mathcal{B}}_{ij}$ are assumed to be axis-aligned hyper-rectangular sets [98, Proposition 1]. This is not restrictive and any convex set with convex volume could have been chosen instead as in [139]. We parametrize $\mathcal{B}_i(\boldsymbol{\gamma}) := [\overline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}]$ by $\boldsymbol{\gamma} = (\overline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}) \in \mathbb{R}^{2n_w}$, and $\bar{\mathcal{B}}_{ij}(\boldsymbol{\lambda}) := [\overline{\boldsymbol{\lambda}}, \underline{\boldsymbol{\lambda}}]$ by $\boldsymbol{\lambda} = (\overline{\boldsymbol{\lambda}}, \underline{\boldsymbol{\lambda}}) \in \mathbb{R}^{2n_\delta}$, and consider the following chance-constrained optimization problem:

$$\begin{cases} \min\limits_{\boldsymbol{\gamma}} & \left\|\overline{\boldsymbol{\gamma}} - \underline{\boldsymbol{\gamma}}\right\|_1 \\ \text{s.t.} & \mathbb{P}\left\{\boldsymbol{w}_i \in \mathcal{W}_i \;\middle|\; \boldsymbol{w}_i \in [\underline{\boldsymbol{\gamma}}, \overline{\boldsymbol{\gamma}}]\right\} \geq 1 - \varepsilon_i \end{cases}, \tag{2.27a}$$

$$\begin{cases} \min\limits_{\boldsymbol{\lambda}} & \left\|\overline{\boldsymbol{\lambda}} - \underline{\boldsymbol{\lambda}}\right\|_1 \\ \text{s.t.} & \mathbb{P}\left\{\boldsymbol{\delta}_{ij} \in \Delta_{ij} \;\middle|\; \boldsymbol{\delta}_{ij} \in [\underline{\boldsymbol{\lambda}}, \overline{\boldsymbol{\lambda}}]\right\} \geq 1 - \bar{\varepsilon}_{ij} \end{cases}. \tag{2.27b}$$

Following the so-called scenario approach in [27], one can determine the number of required uncertainty scenarios to formulate a tractable problem, using $N_s = \frac{2}{\epsilon}(\xi + \ln\frac{1}{\nu})$, where ξ is the dimension of decision vector, ϵ, $1 - \nu$ are the level of violation, and the confidence level, respectively. We determine N_{s_i} by substituting $\xi = 2n_w$, $\epsilon = \varepsilon_i$, $\nu = \beta_i$, and determine $\bar{N}_{s_{ij}}$ by substituting $\xi = 2n_\delta$, $\epsilon = \bar{\varepsilon}_{ij}$, $\nu = \bar{\beta}_{ij}$, for all agent $i \in \{1, 2, \cdots, N\}$.

We next define $\mathcal{S}_i = \{\boldsymbol{w}_i^{(1)}, \cdots, \boldsymbol{w}_i^{(N_{s_i})}\} \subset \mathcal{W}_i$, $\bar{\mathcal{S}}_{ij} = \{\boldsymbol{\delta}_{ij}^{(1)}, \cdots, \boldsymbol{\delta}_{ij}^{(\bar{N}_{s_{ij}})}\} \subset \Delta_{ij}$ and formu-

late a tractable version of (2.27a) and (2.27b) by

$$\begin{cases} \min\limits_{\boldsymbol{\gamma}} & \left\|\overline{\boldsymbol{\gamma}}-\underline{\boldsymbol{\gamma}}\right\|_1 \\ \text{s.t.} & \boldsymbol{w}_i \in [\underline{\boldsymbol{\gamma}},\overline{\boldsymbol{\gamma}}] \quad , \quad \boldsymbol{w}_i \in \mathcal{S}_i \end{cases} , \tag{2.28a}$$

$$\begin{cases} \min\limits_{\boldsymbol{\lambda}} & \left\|\overline{\boldsymbol{\lambda}}-\underline{\boldsymbol{\lambda}}\right\|_1 \\ \text{s.t.} & \boldsymbol{\delta}_{ij} \in [\underline{\boldsymbol{\lambda}},\overline{\boldsymbol{\lambda}}] \quad , \quad \boldsymbol{\delta}_{ij} \in \bar{\mathcal{S}}_{ij} \end{cases} . \tag{2.28b}$$

The optimal solutions $(\boldsymbol{\gamma}^*, \boldsymbol{\lambda}^*)$ of the proposed tractable problem are probabilistically feasible for the chance-constrained problems, [30, Theorem 1]. Moreover, $\boldsymbol{\gamma}^*$, and $\boldsymbol{\lambda}^*$ also characterize our desired probabilistic bounded sets $\mathcal{B}_i^*$ and $\bar{\mathcal{B}}_{ij}^*$, respectively. Note that $\mathcal{S}_i$ and $\bar{\mathcal{S}}_{ij}$ are two collections of random scenarios that are i.i.d.

After determining $\mathcal{B}_i^*$ and $\bar{\mathcal{B}}_{ij}^*$ for all agents $i \in \{1, \cdots, N\}$, we are now able to reformulate the robust counterpart of the original problem (2.25) via:

$$\min_{\boldsymbol{y}} \quad \sum_{i=1}^{N} \mathcal{V}_i(\boldsymbol{x}_i, \boldsymbol{u}_i) \tag{2.29a}$$

$$\text{s.t.} \quad \boldsymbol{y} \in \prod_{i=1}^{N} \bigcap_{\boldsymbol{w}_i \in (\mathcal{B}_i^* \cap \mathcal{W}_i)} \mathcal{Y}_i(\boldsymbol{w}_i) , \tag{2.29b}$$

$$\boldsymbol{y} \in \prod_{i=1}^{N} \bigcap_{j \in \mathcal{N}_i} \bigcap_{\boldsymbol{\delta}_{ij} \in (\bar{\mathcal{B}}_{ij}^* \cap \Delta_{ij})} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij}) . \tag{2.29c}$$

Note that the aforementioned problem is not a randomized program, and instead, the constraints have to be satisfied for all values of the private uncertainty in $(\mathcal{B}_i^* \cap \mathcal{W}_i)$, and common uncertainty in $(\bar{\mathcal{B}}_{ij}^* \cap \Delta_{ij})$. The proposed problem (2.29) is a robust mixed-integer quadratic program. In [12], it was shown that robust problems are tractable [144, Proposition 1], and remain in the same class as the original problems, e.g., robust mixed-integer programs remain mixed-integer programs, for a certain class of uncertainty sets, such as in our problem (2.29), where the uncertainty is bounded in a convex set. The following theorem quantifies the robustness of solution obtained by (2.29) w.r.t. the initial problem (2.25).

Theorem 1 *Let $\varepsilon_i, \bar{\varepsilon}_{ij} \in (0,1)$ and $\beta_i, \bar{\beta}_{ij} \in (0,1)$ for all $j \in \mathcal{N}_i$, for each $i \in \{1, \cdots, N\}$ be chosen such that $\varepsilon = \sum_{i=1}^{N} \varepsilon_i \in (0,1)$, $\beta = \sum_{i=1}^{N} \beta_i \in (0,1)$, $\bar{\varepsilon}_i = \sum_{j \in \mathcal{N}_i} \bar{\varepsilon}_{ij} \in (0,1)$, $\bar{\beta}_i = \sum_{j \in \mathcal{N}_i} \bar{\beta}_{ij} \in (0,1)$ and $\bar{\varepsilon} = \sum_{i=1}^{N} \bar{\varepsilon}_i \in (0,1)$, $\bar{\beta} = \sum_{i=1}^{N} \bar{\beta}_i \in (0,1)$. Determine $\mathcal{B}_i^*$ and $\bar{\mathcal{B}}_{ij}^*$ by constructing $\mathcal{S}_i, \bar{\mathcal{S}}_{ij}$ and solving (2.28) for all $j \in \mathcal{N}_i$, for each $i \in \{1, \cdots, N\}$. If $\boldsymbol{y}_s^*$ is a feasible solution of the problem (2.29), then $\boldsymbol{y}_s^*$ is also a feasible solution for the chance constraints (2.25b) and (2.25c), with the confidence levels of $1-\beta$ and $1-\bar{\beta}$, respectively.*

The proof is provided in Appendix A. □

The interpretation of Theorem 1 is as follows. The obtained solutions via (2.29) for all agents $i = 1, \cdots, N$ have feasibility guarantees with $1-\varepsilon_i$ and $1-\bar{\varepsilon}_{ij}$ probabilities for the private and common uncertain sources $\boldsymbol{w}_i$ and $\boldsymbol{\delta}_{ij}$ with high confidence levels $1-\beta_i$

and $1-\bar{\beta}_{ij}$, respectively. To keep the robustness level of the solutions for the whole network problem, these choices have to follow a certain design rule. It is important to mention that in order to maintain the violation level for the whole network the violation level of individual agent needs to decrease which may lead to very conservative results for each agent, since the number of required samples needs to increase in the proposed formulation (2.28).

Remark 1 *We also approximate the objective function empirically for each agent i following the approach in [138]. $\mathbb{E}_{\boldsymbol{w}_i}[\mathcal{J}_i(\cdot)]$ can be approximated by averaging the value of its argument for some number of different scenarios, which plays a tuning parameter role. Using $N_{s_i^0}$ as the tuning parameter, consider $N_{s_i^0}$ number of different scenarios of $\boldsymbol{w}_i$ to build $\mathcal{S}_i^0 = \{\boldsymbol{w}_i^{(1)}, \cdots, \boldsymbol{w}_i^{(N_{s_i^0})}\} \subset \mathcal{W}_i$ for each agent $i = 1, \cdots, N$. Then one can approximate the cost function empirically as follows:*

$$\sum_{i=1}^{N} \mathcal{V}_i(\boldsymbol{x}_i(\boldsymbol{w}_i), \boldsymbol{u}_i) = \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{w}_i \in \mathcal{W}_i}[\mathcal{J}_i(\boldsymbol{x}_i(\boldsymbol{w}_i), \boldsymbol{u}_i)] \approx \sum_{i=1}^{N} \frac{1}{N_{s_i^0}} \sum_{\boldsymbol{w}_i \in \mathcal{S}_i^0} \mathcal{J}_i(\boldsymbol{x}_i(\boldsymbol{w}_i), \boldsymbol{u}_i).$$

It is worth mentioning that one can employ scenario removal algorithms to improve the objective value, leading to a tradeoff between feasibility and optimality, see e.g., [31, 109].

Remark 2 *A tractable decoupled solution (DS) formulation for (2.16) can be achieved by removing the robust coupling constraint (2.29c) from (2.29). Since there is no longer a coupling constraint, each agent i can therefore solve its problem independently.*

The solution of (2.29) is $\{u^*_{i,k|k}, v^*_{i,k|k}, \cdots, u^*_{i,k+N_h-1|k}, v^*_{i,k+N_h-1|k}\}_{i=1}^N$, which is the optimal input sequence. Based on an MPC paradigm, the current input at time step k is implemented in the system dynamics (2.12) using the first element of optimal solutions as $\{u_{i,k}, v_{i,k}\}_{i=1}^N := \{u^*_{i,k|k}, v^*_{i,k|k}\}_{i=1}^N$ and we proceed in a receding horizon fashion. This means (2.29) is solved at each time step k by using the current measurement of the state $\{x_{i,k}\}_{i=1}^N$. It is important to highlight that the feasibility guarantees in Theorem 1 are independent from the sampling rate of the real continuous-time system. It is however very important to have a discrete-time system model that can predict the real system behavior as precisely as possible. Once such a suitable discrete-time system model is developed, one can use our proposed tractable frameworks (DS, CS, and MCS), and instead of analyzing the closed-loop asymptotic behavior, achieve the fulfillment of multiple chance constraints from an optimization point of view and have a-priori probabilistic feasibility guarantees via Theorem 1.

2.5. NUMERICAL STUDY

In this section, we present a simulated case study for a three-agent ATES system in an STG, as it is shown in Figure 2.5. We determine the thermal energy demands of three buildings, that had been equipped with ATES systems, modeled using realistic parameters and the actual registered weather data in the city center of Utrecht, The Netherlands, where these buildings are located. We refer interested readers to [5, Appendix A] for the complete detailed parameters of this case study.

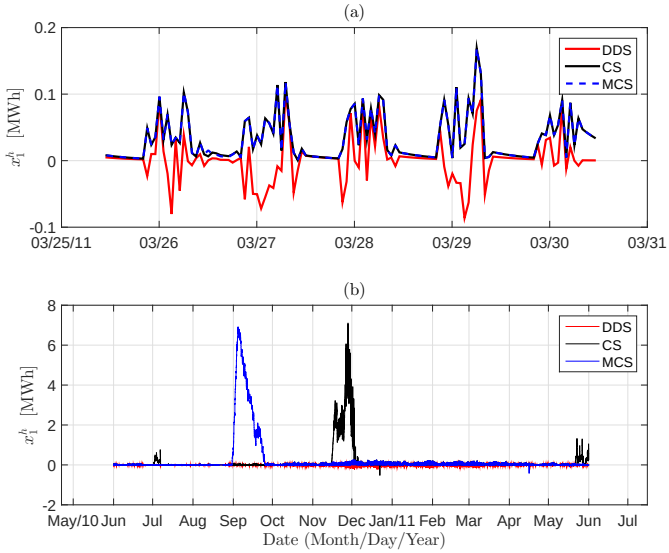


Figure 2.6: A-posteriori feasibility validation of the obtained results via DDS, CS, and MCS formulations for the imbalance error dynamics in the first building of the three-agent ATES-STG example. Figure 3(a) focuses on a randomly chosen five-day period to allow a better comparison, whereas Figure 3(b) presents the complete one year results.

2.5.1. SIMULATION SETUP

We simulate three problem formulations, namely: DS (decoupled solution), CS (centralized solution), and MCS (move-blocking centralized solution), using the proposed tractable framework (2.29). The simulation time is one year from June 2010 to June 2011 with hourly-based sampling time. The prediction horizon for DS and CS is a day-ahead (24 hours), whereas for MCS is a whole season (3 months). The multi-rate control actions in MCS are considered to be hourly-based during first day, daily-based in the first week, weekly-based within the first month, and monthly-based for the rest of the season. We also simulate a deterministic DS (DDS) for comparison purposes, where the uncertain elements ($\boldsymbol{w}_i$) are fixed to their forecast value for each agent $i = 1,2,3$. In order to generate scenarios from the private uncertainty sources, we use a discrete normal stochastic process, where the thermal energy demand of each building varies within 10% of its actual value at each sampling time. A similar technique is used for the common uncertainty sources. The simulation environment was MATLAB with YALMIP as the interface [86] and Gurobi as a solver.

2.5.2. SIMULATION RESULTS

Figure 2.6 and Figure 2.7 (a) depict a-posteriori feasibility validation of the private chance constraint of agent 1 and the common chance constraint between agent 1 and agent 2. A-posteriori feasibility validation of the results in this section refer to the single simulation study and not to the a-posteriori probability of the constraint feasibility, which can be achieved using Monte-Carlo (MC) approach. It is important to note that the results ob-

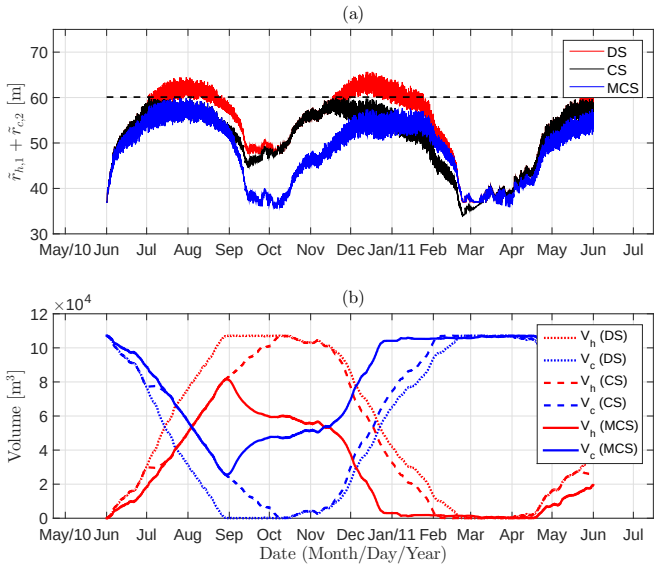


Figure 2.7: A-posteriori feasibility validation of the obtained results via DS, CS, and MCS formulations for the common coupling constraint between the first and second building of the three-agent ATES-STG example is presented in Figure 4(a). Figure 4(b) shows the ATES system state trajectories (volume of the stored water in the warm and cold wells of the first building) in the three-agent ATES-STG example.

tained for the other two buildings are very similar, and therefore we focus on the results of the first building (agent 1). To illustrate the functionality of our proposed framework to deal with the private chance constraint, in Figure 2.6, we present the a-posteriori feasibility validation of the obtained results via DDS, CS, and MCS formulations. Figure 2.6 (a) shows the obtained results for the last five days in March 2011, and Figure 2.6 (b) shows the results for one year simulation from June 2010 until June 2011. In Figure 2.6 the "red" color denotes the solution of DDS, "black" color shows the solution of CS, and "blue" presents the solution of MCS.

Figure 2.6 (a) focuses on a randomly chosen five-day period to allow a better comparison between the results of DDS, CS, and MCS. It is clearly shown that the obtained results via CS and MCS, provide a feasible (nonnegative) trajectory of the thermal energy imbalance error during heating mode, whereas the solution of DDS, leads to some violations throughout the simulation time. Notice that all three proposed approaches, namely DS, CS, and MCS, achieved the feasibility of the private chance constraint in a probabilistic sense as it is guaranteed in Theorem 1. We present the results obtained via DDS to highlight such an achievement, whereas the results obtained via DS is omitted to demonstrate the other achievements.

In Figure 2.6 (b), the complete one year results of DDS, CS, and MCS are shown. Two important observations are as follows: the obtained results of CS and MCS have very small number of violations, much less than our desired level of violations, throughout the simulation time. This yields a less conservative approach compared to the classical robust control approach (see [21, Ch.14]). As the second observation, in the results of CS

and MCS one can see some instances of a large non-zero imbalance error, which is expected: By taking into account the coupling constraints between agents, the solutions of agents are going to extract the stored thermal energy from their ATES systems to prevent the mutual interactions between their ATES systems as in Figure 2.7 (a). Interestingly, the results of MCS show that agent 1 starts to extract the stored thermal energy from its ATES system sooner due to its longer prediction horizon, compared to CS.

Figure 2.7 (a) shows the evaluation of our proposed reformulation for the coupling constraint in (2.19) together with the a-posteriori feasibility validation of the common chance constraint between agent 1 and agent 2. We plot the obtained $\tilde{r}_{h,1} + \tilde{r}_{c,2}$ using DS, CS, and MCS formulations. As it is clearly shown DS results are violating the coupling constraint which leads to overlap between the stored water in warm well of ATES system in agent 1 and the stored water in cold well of ATES system in agent 2. This is due to the fact that there are no coupling constraints in the DS framework and each agent works without any information from neighboring agents. It is important to highlight that the results obtained via DDS and DS are the same in terms of the ATES system dynamical behavior. This is due to the fact that the cost parameter associated with the ATES system pump is the same in both DDS and DS formulations, and thus ATES systems participate in the agent energy management in the same way, regardless of the private chance constraints. We also present the evolution of the stored water volume in each well of the ATES system for agent 1 using the obtained results via DS, CS, and MCS formulations in Figure 2.7 (b) to illustrate the impact of the different formulations.

It is worth to mention that Figure 2.6 and Figure 2.7 illustrate all main contributions: 1) having a probabilistically feasible solution for each agent w.r.t. the private uncertainty sources as it is encoded via (2.25b), 2) respecting the common resource pool between neighboring agents in STGs as it is formulated in (2.25c) (the first and second outcomes are the direct results of our theoretical guarantee in Theorem 1), and 3) prediction using a longer horizon yields an anticipatory control decision that improves the operation of an ATES system. This is a direct consequence of our proposed move-blocking scheme in (2.24).

Figure 2.8 summarizes the results in terms of average thermal efficiency that we obtained by integrating our control strategy, DS and CS, into Python to build a live-link with MODFLOW, a more realistic aquifer simulation environment[3] [64]. Figure 2.8 is presented to highlight the impact of considering the proposed coupling constraints, as it is formulated in (2.25c), versus the decoupled setting. The impact of our control strategy, DS (red) and CS (blue), on average thermal energy efficiency [15] in each building illustrates that we can store and retrieve the same amount of thermal energy in ATES systems, in a more efficient way due to information exchange between the agents to prevent the mutual interactions between wells using the results of MCS and CS compared to DS.

[3] MODFLOW is a modular hydrologic model, and it is considered an international standard for aquifer simulation and predicting groundwater conditions and interactions.

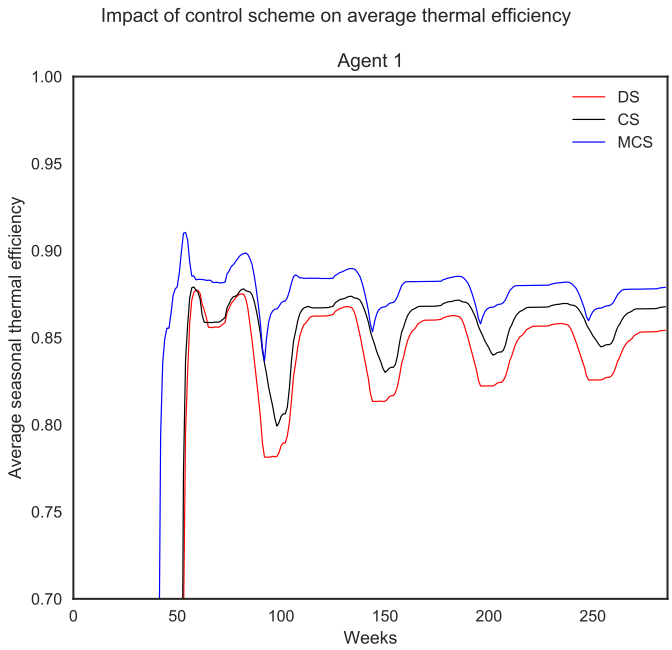


Figure 2.8: Impact of DS and CS on average thermal efficiency.

2.6. CONCLUSIONS

This chapter proposed a stochastic MPC framework for an energy management problem in STGs consisting of ATES systems integrated into BCC systems. We developed a large-scale stochastic hybrid model to capture thermal energy imbalance errors in an ATES-STG. In such a framework, we formalized two important practical concerns, namely: 1) the balance between extraction and injection of energy from and into the aquifers within a certain period of time; 2) the unwanted mutual interaction between ATES systems in STGs. Using our developed model, we formulated a finite-horizon mixed-integer quadratic optimization problem with multiple chance constraints. To solve such a problem, we proposed a tractable formulation based on the so-called robust randomized approach. In particular, we extended this approach to handle a problem with multiple chance constraints. We simulated our proposed framework using a three-agent ATES-STG example which confirmed the expected performance improvements.

3

Distributed Stochastic MPC for Large-Scale Systems with Private and Common Uncertainty Sources

In this chapter, we present a distributed stochastic model predictive control (SMPC) approach for large-scale linear systems with private and common uncertainties in a plug-and-play framework. Typical SMPC approaches for such problems involve formulating a large-scale finite-horizon chance-constrained optimization problem at each sampling time, which is in general non-convex and difficult to solve. Using an approximation, the so-called scenario approach, we formulate a large-scale scenario program and provide a theoretical guarantee to quantify the robustness of the obtained solution. However, such a reformulation leads to a computational tractability issue, due to the large number of required scenarios. We present two novel ideas in this chapter to address this issue. We first provide a technique to decompose the large-scale scenario program into distributed scenario programs that exchange a certain number of scenarios with each other in order to compute local decisions. We show the exactness of the decomposition with a-priori probabilistic guarantees for the desired level of constraint fulfillment for both the cases of private and common uncertainty sources. As our second contribution, we develop an inter-agent soft communication scheme based on a set parametrization technique together with the notion of probabilistically reliable sets to reduce the required communication between the subproblems. We show how to incorporate the probabilistic reliability notion into existing results and provide new guarantees for the desired level of constraint violations. Two different simulation studies of both private and common uncertainty sources are presented to illustrate the advantages of the proposed framework.

3.1. INTRODUCTION

STOCHASTIC model predictive control (SMPC) has attracted significant attention in the recent control literature, due to its ability to provide an alternative, often less conservative way to handle uncertain systems. SMPC takes into account the stochastic characteristics of the uncertainties and thereby the system constraints are treated in a probabilistic sense, i.e. using chance constraints [75, 103]. SMPC computes an optimal control sequence that minimizes a given objective function subject to the uncertain system dynamics model and chance constraints in a receding horizon fashion [122]. Chance constraints enable SMPC to offer an alternative approach to achieve a less conservative solution compared to robust model predictive control (MPC) [9], since it directly incorporates the tradeoff between constraint feasibility and control performance.

Distributed MPC has been an active research area in the past decades, due to its applicability in different domains such as power networks [169], chemical plants [138], process control [176], and building automation [93]. For such large-scale dynamic systems with state and input constraints, distributed MPC is an attractive control scheme. In distributed MPC one replaces large-scale optimization problems stemming from centralized MPC with several smaller-scale problems that can be solved in parallel. These problems make use of information from other subsystems to formulate finite-horizon optimal control problems. In the presence of uncertainties, however, the main challenge in the formulation of distributed MPC is how the controllers should exchange information through a communication scheme among subsystems (see, e.g., [81], and references therein). This highlights the necessity of developing distributed control strategies to cope with the uncertainties in subsystems while at the same time minimizing information exchange through a communication framework.

3.1.1. RELATED WORKS

In order to handle uncertainties in distributed MPC, some approaches are based on robust MPC [38, 126]. Assuming that the uncertainty is bounded, a robust optimization problem is solved at each sampling time, leading to a control law that satisfies the constraints for all admissible values of the uncertainty. The resulting solution using such an approach tends to be conservative in many cases. Tube-based MPC, see for example [33] and the references therein, was considered in a plug-and-play (PnP) decentralized setup in [129], and it has been recently extended to distributed control systems [41] for a collection of linear stochastic subsystems with independent dynamics. While in [41] coupled chance constraints were considered separately at each sampling time, in this chapter we consider a chance constraint on the feasibility of trajectories of dynamically coupled subsystems. Our approach is motivated by [129] to reduce the conservativeness of the control design. Other representative approaches for SMPC of a single stochastic system include affine parametrization of the control policy [68], the randomized (scenario) approach [28, 89, 121, 159], and the combined randomized and robust approach [98, 144, 181]. None of these approaches, to the best of our knowledge, have been considered in a distributed control setting.

This chapter aims to develop a systematic approach to distributed SMPC using the scenario MPC technique. Scenario MPC approximates SMPC via the so-called scenario (sample) approach [27, 30], and if the underlying optimization problem is convex with

respect to the decision variables, finite sample guarantees can be provided. Following such an approach, the computation time for a realistic large-scale system of interest becomes prohibitive, due to the fact that the number of samples to be extracted tends to be high, and consequently leads to a large number of constraints in the resulting optimization problem. To overcome the computational burden caused by the large number of constraints, in [85, 88] a heuristic sample-based approach was used in an iterative distributed fashion via dual decomposition such that all subsystems collaboratively optimize a global performance index. In another interesting work [29], a multi-agent consensus algorithm was presented to achieve consensus on a common value of the decision vector subject to random constraints such that a probabilistic bound on the tails of the consensus violation was also established. However, in most of the aforementioned references the aim to reduce communication among subsystems, which we refer to as agents, has not been addressed.

3.1.2. CONTRIBUTIONS

Our work in this chapter differs from the above references in two important aspects which have not been, to the best of our knowledge, considered in literature. A decomposition technique based on the large-scale system dynamics is employed to distribute the resulting centralized scenario optimization problem at each sampling time and a novel communication scheme is introduced to reduce the communication between the small-scale problems. We first propose a technique to decompose the large-scale (centralized) scenario optimization problem into small-scale scenario programs such that they can be solved in parallel at the cost of exchanging a certain number of scenarios with each other in order to compute local decisions. We then provide new a-priori probabilistic guarantees to quantify the robustness of the resulting solution and show the theoretical links to the existing centralized guarantees. To reduce the communications between agents required by our proposed framework, we finally develop a new set-based communication setup together with the notion of probabilistically reliable information. We then incorporate the probabilistic reliability notion into existing results and provide new guarantees for the desired level of constraint violations. The main contributions of this chapter are as follows:

- ▷ We provide a technique to decompose the large-scale scenario program into distributed scenario programs that exchange a certain number of scenarios with each other in order to compute local decisions. We show that such a decomposition technique can be applied to large-scale linear systems with both private (local) and common uncertainty sources. This yields a flexible and practical plug-and-play distributed scenario MPC framework.

- ▷ We quantify the level of robustness of resulting solutions using our proposed distributed scenario MPC framework and provide two new a-priori probabilistic guarantees for the desired level of constraint fulfillment under some mild conditions for both cases of private and common uncertainty sources.

- ▷ We develop an inter-agent soft communication scheme based on a set parametrization technique together with the notion of probabilistically reliable set to reduce

the required communication between each subproblem. We show how to incorporate the probabilistic reliability notion into existing results and provide new guarantees for the desired level of constraint violations.

It is important to highlight that two major difficulties arising in SMPC, namely recursive feasibility [87] and stability, are not in the scope of this chapter, and they are subject of our ongoing research work. Thus, instead of analyzing the closed-loop asymptotic behavior, in this chapter we focus on individual SMPC problem instances from the optimization point of view and derive probabilistic guarantees for constraint fulfillment in a distributed setting.

3.1.3. STRUCTURE

The structure of this chapter is as follows. Section 3.2 describes a mathematical model of the control system dynamics. We first formulate an SMPC problem for a large-scale linear system with uncertain parameters and additive disturbances, then provide a reformulation, namely the centralized scenario MPC. A theoretical study on the connections of these two control problems based on the existing results is provided. In Section 3.3, we propose a decomposition technique to distributed scenario MPC and show that such a decomposition technique can be applied to both cases of private and common uncertainty sources. We then analyze the robustness of the obtained solution compared to the centralized scenario MPC formulated in the previous section. Section 3.4 introduces two types of inter-agent communication schemes between each subproblem, namely hard and soft communications, and then proceeds to quantify the robustness of the proposed schemes. Section 3.5 provides a summary of our developments to establish a practical plug-and-play distributed SMPC framework that considers network changes by agents which want to join or leave the network. Section 3.6 presents two different simulation studies to illustrate the functionality of our theoretical achievements, whereas in Section 3.7, we conclude this chapter with some remarks.

3.1.4. NOTATIONS

$\mathbb{R}, \mathbb{R}_+$ denote the real and positive real numbers, and $\mathbb{N}, \mathbb{N}_+$ the natural and positive natural numbers, respectively. We operate within the n-dimensional space $\mathbb{R}^n$ composed of column vectors $u, v \in \mathbb{R}^n$. The Cartesian product over n sets $\mathcal{X}_1, \cdots, \mathcal{X}_n$ is given by: $\prod_{i=1}^n \mathcal{X}_i = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n = \{(x_1, \cdots, x_n) : x_i \in \mathcal{X}_i\}$. The cardinality of a set $\mathcal{A}$ is denoted by $|\mathcal{A}| = A$. We denote a block-diagonal matrix with blocks X_i, $i \in \{1, \cdots, n\}$, by $\mathrm{diag}_{i\in\{1,\cdots,n\}}(X_i)$, and a vector consisting of stacked sub vectors x_i, $i \in \{1, \cdots, n\}$, by $\mathrm{col}_{i\in\{1,\cdots,n\}}(x_i)$.

Given a metric space Δ, its Borel σ-algebra is denoted by $\mathfrak{B}(\Delta)$. Throughout the chapter, measurability always refers to Borel measurability. In a probability space $(\Delta, \mathfrak{B}(\Delta), \mathbb{P})$, we denote the N-Cartesian product set of Δ by Δ^N and the respective product measure by $\mathbb{P}^N$.

3.2. PROBLEM STATEMENT

This section provides an overview of the control problem statement. We first describe the dynamics of a large-scale uncertain linear system together with input and state con-

straint sets and the control objective. We then formulate a centralized SMPC for such a large-scale control system problem. Finally, a tractable reformulation based on the scenario MPC [157] together with theoretical connections are provided.

Consider a discrete-time uncertain linear system with additive disturbance in a compact form as follows:

$$x_{k+1} = A(\delta_k)x_k + B(\delta_k)u_k + C(\delta_k)w_k \,, \tag{3.1}$$

with a given initial condition $x_0 \in \mathbb{R}^n$. Here $k \in \mathcal{T} := \{0, 1, \cdots, T-1\}$ denotes the time instance, $x_k \in \mathcal{X} \subset \mathbb{R}^n$ and $u_k \in \mathcal{U} \subset \mathbb{R}^m$ correspond to the state and control input, respectively, and $w_k \in \mathbb{R}^p$ represents an additive disturbance. The system matrices $A(\delta_k) \in \mathbb{R}^{n\times n}$ and $B(\delta_k) \in \mathbb{R}^{n\times m}$ as well as $C(\delta_k) \in \mathbb{R}^{n\times p}$ are random, since they are known functions of an uncertain variable δ_k that influences the system parameters at each time step k.

Assumption 7 *Random variables $\boldsymbol{w} := \{w_k\}_{k\in\mathcal{T}}$ and $\boldsymbol{\delta} := \{\delta_k\}_{k\in\mathcal{T}}$ are defined on probability spaces $(\mathcal{W}, \mathfrak{B}(\mathcal{W}), \mathbb{P}_{\boldsymbol{w}})$ and $(\Delta, \mathfrak{B}(\Delta), \mathbb{P}_{\boldsymbol{\delta}})$, respectively. $\boldsymbol{w}$ and $\boldsymbol{\delta}$ are two independent random processes, where $\mathbb{P}_{\boldsymbol{w}}$ and $\mathbb{P}_{\boldsymbol{\delta}}$ are two different probability measures defined over $\mathcal{W}$ and Δ, respectively, and $\mathfrak{B}(\cdot)$ denotes a Borel σ-algebra. The support sets $\mathcal{W}$ and Δ of $\boldsymbol{w}$ and $\boldsymbol{\delta}$, respectively, together with their probability measures $\mathbb{P}_{\boldsymbol{w}}$ and $\mathbb{P}_{\boldsymbol{\delta}}$ are entirely generic. In fact, $\mathcal{W}$, Δ and $\mathbb{P}_{\boldsymbol{w}}$, $\mathbb{P}_{\boldsymbol{\delta}}$ do not need to be known explicitly. Instead, the only requirement is availability of a "sufficient number" of samples, which will become concrete in later parts of the chapter. Such samples can be for instance obtained by a model learned from available historical data [117].*

The system in (3.1) is subject to constraints on the system state trajectories and control input. Consider the state and control input constraint sets to be compact convex in the following form

$$\mathcal{X} := \{x \in \mathbb{R}^n : Gx \le g\} \,, \mathcal{U} := \{u \in \mathbb{R}^m : Hu \le h\} \,, \tag{3.2}$$

where $G \in \mathbb{R}^{q\times n}$, $g \in \mathbb{R}^q$, $H \in \mathbb{R}^{r\times m}$, and $h \in \mathbb{R}^r$. Keeping the state inside a feasible set $\mathcal{X} \subset \mathbb{R}^n$ for the entire prediction horizon may be too conservative and result in loss of performance. In particular, this is the case when the best performance is achieved close to the boundary of $\mathcal{X}$, and thus, constraint violations will be unavoidable due to the fact that the system parameters in (3.1) are imperfect and uncertain. To tackle such a problem, we will consider chance constraints on the state trajectories to avoid violation of the state variable constraints probabilistically even if the disturbance $\boldsymbol{w}$ or uncertainty $\boldsymbol{\delta}$ has unbounded support. Notice that a robust problem formulation [9] cannot cope with problems having an unbounded disturbance set.

In order to find a stabilizing full-information controller that leads to admissible control inputs $\boldsymbol{u} := \{u_k\}_{k\in\mathcal{T}}$ and satisfies the state constraints, we follow the traditional MPC approach. The design relies on the standard assumption of the existence of a suitable pre-stabilizing control law, see, e.g., [129, Proposition 1]. To cope with the state prediction under uncertainty and disturbance, we employ a parametrized feedback policy [68] and split the control input: $u_k = Kx_k + v_k$ with $v_k \in \mathbb{R}^m$ as a free correction input variable to compensate for disturbances.

The control objective is to minimize a cumulative quadratic stage cost of a finite horizon cost $J(\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ that is defined as follows:

$$J(\boldsymbol{x}, \boldsymbol{u}) = \mathbb{E}\left[\sum_{k=0}^{T-1}\left(x_k^\top Q x_k + u_k^\top R u_k\right) + x_T^\top P x_T\right], \tag{3.3}$$

with $Q \in \mathbb{R}^{n\times n}_{\geq 0}$, and $R \in \mathbb{R}^{m\times m}_{>0}$. Consider $\boldsymbol{x} := \{x_k\}_{k\in\mathcal{T}}$, $(A, Q^{\frac{1}{2}})$ to be detectable and P to be the solution of the discrete-time Lyapunov equation:

$$\mathbb{E}[\,A_{cl}(\delta_k)^\top P A_{cl}(\delta_k)\,] + Q + K^\top R K - P \preceq 0\,, \tag{3.4}$$

for the closed-loop system with $A_{cl}(\delta_k) = A(\delta_k) + B(\delta_k)K$. Each stage cost term is taken in expectation $\mathbb{E}[\cdot]$, since the argument x_k is a random variable. Using $\boldsymbol{v} = \{v_k\}_{k\in\mathcal{T}}$, consider now the following stochastic control problem:

$$\min_{\boldsymbol{v}\in\mathbb{R}^{Tm}} \quad J(\boldsymbol{x}, \boldsymbol{u}) \tag{3.5a}$$

$$\text{s.t.} \quad x_{k+1} = A(\delta_k)x_k + B(\delta_k)u_k + C(\delta_k)w_k\,, \tag{3.5b}$$

$$\mathbb{P}[\,x_{k+\ell} \in \mathcal{X}\,,\ \ell \in \mathbb{N}_+\,] \geq 1 - \varepsilon\,, \tag{3.5c}$$

$$u_k = Kx_k + v_k \in \mathcal{U}\,, \quad \forall k \in \mathcal{T}\,, \tag{3.5d}$$

where x_0 is initialized based on the measured current state, and $\varepsilon \in (0,1)$ is the admissible state constraint violation parameter of the large-scale system (3.1). The objective function is assumed to be a quadratic function; however, this is not a restriction and any generic convex function can be chosen instead. It is important to mention that the parameters of constraint sets, $\mathcal{X}, \mathcal{U}$, and the objective function $J(\cdot)$ can be time-varying with respect to the sampling time $k \in \mathcal{T}$. For the clarity of our problem formulation, we assume time-invariance. The state trajectory $x_{k+\ell}, \forall \ell \in \mathbb{N}_+$, has a dependency on the random variables $\boldsymbol{w}$ and $\boldsymbol{\delta}$, and thus, the chance constraint can be interpreted as follows: the probability of violating the state constraint at the future time step $\ell \in \mathbb{N}_+$ is restricted to ε, given that the state of the system in (3.1) is measurable at each time step $k \in \mathcal{T}$. It is important to highlight that in the proposed chance constraint (3.5c), the future time step index $\ell \in \mathbb{N}_+$ is bounded by the finite value of the prediction horizon T and this holds for all the following proposed formulations. Even though $\mathcal{U}$ and $\mathcal{X}$ are compact convex sets, due to the chance constraint on the state trajectory, the feasible set of the optimization problem in (3.5) is a non-convex set, in general.

Remark 3 *Instead of the chance constraint on the state trajectory of form* (3.5c)*, one can also bound the average rate of state constraint violations [157]. Moreover, one can also define the cost function* (3.5a) *as a desired quantile of the sum of discounted stage costs ("value-at-risk"), instead of the sum of expected values. Instead of a state feedback law, one can also consider a nonlinear disturbance parametrization feedback policy over the prediction horizon, similar to [144], using the scenario approximation. Such a parametrization does not affect the convexity of the resulting optimization [121].*

To handle the chance constraint (3.5c), we recall a scenario-based approximation [157]. w_k and δ_k at each sampling time $k \in \mathcal{T}$ are not necessarily independent and identically distributed (i.i.d.). In particular, they may have time-varying distributions and/or

be correlated in time. We assume that a "sufficient number" of i.i.d. samples of the disturbance $\boldsymbol{w} \in \mathcal{W}$ and $\boldsymbol{\delta} \in \Delta$ can be obtained either empirically or by a random number generator. We denote the sets of given finite samples (scenarios) of uncertain variables with $\mathcal{S}_{\boldsymbol{w}} := \{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_s)}\} \in \mathcal{W}^{N_s}$ and $\mathcal{S}_{\boldsymbol{\delta}} := \{\boldsymbol{\delta}^{(1)}, \cdots, \boldsymbol{\delta}^{(N_s)}\} \in \Delta^{N_s}$, respectively.

Following the approach in [138], we approximate the expected value of the objective function empirically by averaging the value of its argument for some number of different scenarios, which plays a tuning parameter role. Using $N_{\bar{s}}$ as the tuning parameter, consider $N_{\bar{s}}$ number of different scenarios of $\boldsymbol{w}$ and $\boldsymbol{\delta}$ to build

$$\bar{\mathcal{S}}_{\boldsymbol{w},\boldsymbol{\delta}} = \left\{ (\boldsymbol{w}^{(i)}, \boldsymbol{\delta}^{(i)}) \,:\, \boldsymbol{w}^{(i)} \in \mathcal{W} \,,\, \boldsymbol{\delta}^{(i)} \in \Delta \,,\, i = 1, \cdots, N_{\bar{s}} \right\} ,$$

which has the cardinality $|\bar{\mathcal{S}}_{\boldsymbol{w},\boldsymbol{\delta}}| = N_{\bar{s}}$. We then approximate the cost function empirically as follows:

$$J(\boldsymbol{x}, \boldsymbol{u}) = \mathbb{E}_{(\boldsymbol{w},\boldsymbol{\delta})} \left[\sum_{k=0}^{T-1} V(x_k(w_k, \delta_k), u_k) \right] \approx \frac{1}{N_{\bar{s}}} \sum_{(\boldsymbol{w}^{(i)}, \boldsymbol{\delta}^{(i)}) \in \bar{\mathcal{S}}_{\boldsymbol{w},\boldsymbol{\delta}}} \sum_{k=0}^{T-1} V(x_k(w_k^{(i)}, \delta_k^{(i)}), u_k) ,$$

where

$$V(x_k(w, \delta_k), u_k) = \left(x_k(w_k, \delta_k)^\top Q x_k(w_k, \delta_k) + u_k^\top R u_k \right) + x_T(w_k, \delta_k)^\top P x_T(w_k, \delta_k) .$$

Notice that $x_k(w_k, \delta_k)$ indicates the dependency of the state variables on the random variables.

We are now in a position to formulate an approximated version of the proposed stochastic control problem in (3.5) using the following finite horizon scenario program:

$$\min_{\boldsymbol{v} \in \mathbb{R}^{Tm}} \quad J(\boldsymbol{x}, \boldsymbol{u}) \tag{3.6a}$$

$$\text{s.t.} \quad x_{k+1}^{(i)} = A(\delta_k^{(i)}) x_k^{(i)} + B(\delta_k^{(i)}) u_k^{(i)} + C(\delta_k^{(i)}) w_k^{(i)} , \tag{3.6b}$$

$$x_{k+\ell}^{(i)} \in \mathcal{X} \,,\, \ell \in \mathbb{N}_+ \,,\, \forall \boldsymbol{w}^{(i)} \in \mathcal{S}_{\boldsymbol{w}} \,,\, \forall \boldsymbol{\delta}^{(i)} \in \mathcal{S}_{\boldsymbol{\delta}} , \tag{3.6c}$$

$$u_k^{(i)} = K x_k^{(i)} + v_k \in \mathcal{U} \,, \quad \forall k \in \mathcal{T} , \tag{3.6d}$$

where superscript (i) indicates a particular sample realization. The solution of (3.6) is the optimal control input sequence $\boldsymbol{v}^* = \{v_k^*, \cdots, v_{k+T-1}^*\}$. Based on the MPC paradigm, the current input is implemented as $u_k := K x_k + v_k^*$ and we proceed in a receding horizon fashion. This means that the problem (3.6) is solved at each time step k by using the current measurement of the state x_k. Note that new scenarios are needed to be generated at each sampling time $k \in \mathcal{T}$. It is worth to mention that while the constraint (3.6d) is probabilistic for all prediction time step, at the initial time step $k = 0$ is deterministic and the superscript (i) can be dropped as there is only one measured current state x_k. This holds for all the following proposed formulations.

The key features of the proposed optimization problem (3.6) are as follows: 1) there is no need to know the probability measures $\mathbb{P}_{\boldsymbol{w}}$ and $\mathbb{P}_{\boldsymbol{\delta}}$ explicitly, only the capability of obtaining random scenarios is enough, 2) formal results to quantify the level of approximations are available. In particular, the results follow the so-called scenario approach

[30], which allows to bound a-priori the violation probability of the obtained solution via (3.6).

In the following theorem, we restate the explicit theoretical bound of [30, Theorem 1], which measures the finite scenarios behavior of (3.6).

Theorem 2 *Let $\varepsilon, \beta \in (0,1)$ and $N_s \geq \mathsf{N}(\varepsilon, \beta, Tm)$, where*

$$\mathsf{N}(\varepsilon, \beta, Tm) := \min\left\{ N \in \mathbb{N} \,\middle|\, \sum_{i=0}^{Tm-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \leq \beta \right\}.$$

If the optimizer of problem (3.6) $\boldsymbol{v}^* \in \mathbb{R}^{Tm}$ is applied to the discrete-time dynamical system (3.1) for a finite horizon of length T, then, with at least confidence $1-\beta$, the original constraint (3.5c) is satisfied for all $k \in \mathcal{T}$ with probability more than $1-\varepsilon$.

It was shown in [30] that the above bound is tight. The interpretation of Theorem 2 is as follows: when applying $\boldsymbol{v}^*$ in a finite horizon control problem, the probability of constraint violation of the state trajectory remains below ε with confidence $1-\beta$:

$$\mathbb{P}^{N_s}\left[\mathcal{S}_{\boldsymbol{w}} \in \mathcal{W}^{N_s}, \mathcal{S}_{\boldsymbol{\delta}} \in \Delta^{N_s} : \mathrm{Vio}(\boldsymbol{v}^*) \leq \varepsilon\right] \geq 1-\beta,$$

with

$$\mathrm{Vio}(\boldsymbol{v}^*) := \mathbb{P}\left[\boldsymbol{w} \in \mathcal{W}, \boldsymbol{\delta} \in \Delta : x_{k+\ell} = A_{cl}(\delta_k)x_k + B(\delta_k)v_k^* + C(\delta_k)w_k \notin \mathcal{X}, \ell \in \mathbb{N}_+ \,\middle|\, x_k = x_0\right],$$

where $A_{cl}(\delta_k) = A(\delta_k) + B(\delta_k)K$. It is worth mentioning that the proposed constraint on the control input in (3.6d) is also met in a probabilistic sense, due to the feedback parametrization and the nature of the scenario approach that appears in the proposed optimization problem (3.6).

Remark 4 *One can obtain an explicit expression for the desired number of scenarios N_s as in [3], where it is shown that given $\varepsilon, \beta \in (0,1)$ and e the Euler constant, then $N_s \geq \frac{e}{e-1}\frac{1}{\varepsilon}\left(Tm + \ln\frac{1}{\beta}\right)$. It is important to note that N_s is used to construct the sets of scenarios, $\mathcal{S}_{\boldsymbol{w}}$, $\mathcal{S}_{\boldsymbol{\delta}}$ to obtain a probabilistic guarantee for the desired level of feasibility, while the number of scenarios $N_{\bar{s}}$ is just a tuning variable to approximate the objective function empirically.*

We formulated a large-scale SMPC (3.5) together with a tractable reformulation based on the proposed centralized scenario MPC (3.6). Figure 3.1 shows a pictorial representation of (3.6) as a large-scale network of interconnected agents to summarize this section. It is worth mentioning that such a large-scale SMPC (3.5) is initially proposed in [148] for a network of interconnected buildings in smart thermal grids. In the following section, we will provide a distributed framework to solve the proposed problem in (3.6) by decomposing the large-scale system dynamics (3.1).

3.3. DISTRIBUTED SCENARIO MPC

In this section, we describe a decomposition technique to partition the large-scale system dynamics in (3.1). By taking into consideration two possible uncertainty sources,

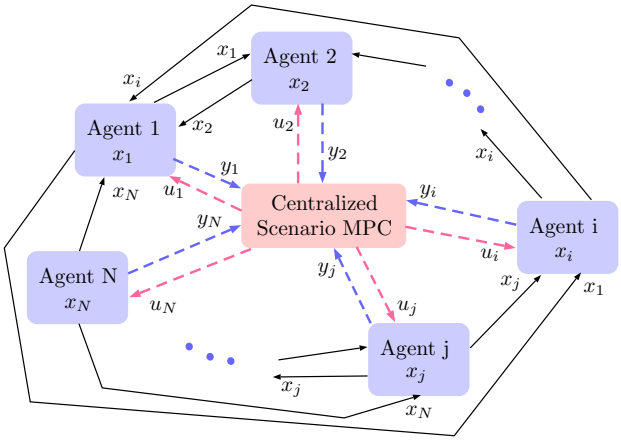


Figure 3.1: Centralized scenario MPC that corresponds to the problem (3.6). The measurement variable y_i is the full state information x_i which is sent to the centralized controllers for all agents $i \in \mathcal{N}$.

namely private (local) and common uncertainties, for the resulting network of interconnected subsystems (agents), we provide the theoretical links to the results that we provided in the previous section.

Consider a partitioning of the system dynamics (3.1) through a decomposition into N subsystems and let $\mathcal{N} = \{1, 2, \cdots, N\}$ be the set of subsystem indices. The state variables x_k, control input signals u_k and the additive disturbance w_k can be considered as $x_k = \mathrm{col}_{i\in\mathcal{N}}(x_{i,k})$, $u_k = \mathrm{col}_{i\in\mathcal{N}}(u_{i,k})$ and $w_k = \mathrm{col}_{i\in\mathcal{N}}(w_{i,k})$, respectively, where $x_{i,k} \in \mathbb{R}^{n_i}$, $u_{i,k} \in \mathbb{R}^{m_i}$, $w_{i,k} \in \mathbb{R}^{p_i}$, and $\sum_{i\in\mathcal{N}} n_i = n$, $\sum_{i\in\mathcal{N}} m_i = m$, $\sum_{i\in\mathcal{N}} p_i = p$. The following assumption is important in order to be able to partition the system parameters.

Assumption 8 *The control input and the additive disturbance of the subsystems are decoupled, e.g., $u_{i,k}$ and $w_{i,k}$ only affect subsystem $i \in \mathcal{N}$ for all $k \in \mathcal{T}$. Consider the state and control input constraint sets $\mathcal{X}$ and $\mathcal{U}$ of large-scale system dynamics (3.1) as defined in (3.2) to be $\mathcal{X} = \prod_{i\in\mathcal{N}} \mathcal{X}_i$, and $\mathcal{U} = \prod_{i\in\mathcal{N}} \mathcal{U}_i$ such that $\mathcal{X}_i$ and $\mathcal{U}_i$ for all subsystem $i \in \mathcal{N}$ are given in the following form:*

$$\mathcal{X}_i := \{x \in \mathbb{R}^{m_i} : G_i\, x \leq g_i\}\,,\ \mathcal{U}_i := \{u \in \mathbb{R}^{p_i} : H_i\, u \leq h_i\}\,, \tag{3.7}$$

where $G = diag_{i\in\mathcal{N}}(G_i)$, $H = diag_{i\in\mathcal{N}}(H_i)$, $g = col_{i\in\mathcal{N}}(g_i)$ and $h = col_{i\in\mathcal{N}}(h_i)$.

It is important to note that under Assumption 8 and the condition (3.7), there is no coupling constraints between each subsystem $i \in \mathcal{N}$ and its neighboring subsystems $j \in \mathcal{N}_i$. Instead, in this chapter, we focus on the subsystem $i \in \mathcal{N}$ that is dynamically coupled with all its neighboring subsystems $j \in \mathcal{N}_i$ as it is presented in (3.9).

We refer to the additive disturbance $w_{i,k}$ as a private (local) uncertainty source of each subsystem $i \in \mathcal{N}$, since it is assumed that it affects only the subsystem $i \in \mathcal{N}$. The uncertain variable δ_k is considered as a common uncertainty source between all subsystems $i \in \mathcal{N}$. Observe the fact that every common uncertain phenomenon can be

considered as a local uncertain variable, e.g., the outside weather condition for neighboring buildings. Therefore, we also consider to have $\delta_k = \mathrm{col}_{i\in\mathcal{N}}(\delta_{i,k})$ and refer to both random variables $w_{i,k}$ and $\delta_{i,k}$ as a local uncertainty sources.

We are now able to decompose the large-scale system matrices

$$B(\delta_k) = \mathrm{diag}_{i\in\mathcal{N}}(B_i(\delta_{i,k})) \in \mathbb{R}^{n\times m}\ ,\ C(\delta_k) = \mathrm{diag}_{i\in\mathcal{N}}(C_i(\delta_{i,k})) \in \mathbb{R}^{n\times p}\ ,$$

and consider $A(\delta_k) \in \mathbb{R}^{n\times n}$ in the following form:

$$A(\delta_k) = \begin{bmatrix} A_{11}(\delta_{1,k}) & \cdots & A_{1N}(\delta_{1,k}) \\ \vdots & \ddots & \vdots \\ A_{N1}(\delta_{N,k}) & \cdots & A_{NN}(\delta_{N,k}) \end{bmatrix},$$

where $A_{ij}(\delta_{i,k}) \in \mathbb{R}^{n_i\times n_j}$, $B_i(\delta_{i,k}) \in \mathbb{R}^{n_i\times m_i}$, and $C_i(\delta_{i,k}) \in \mathbb{R}^{n_i\times p_i}$. Define the set of neighboring subsystems of subsystem i as follows:

$$\mathcal{N}_i = \left\{ j \in \mathcal{N}\backslash\{i\} \;\middle|\; A_{ij}(\delta_{i,k}) \neq \mathbf{0} \right\}\ , \tag{3.8}$$

where $\mathbf{0}$ denotes a matrix of all zeros with proper dimension. Note that if subsystems are decoupled, they remain decoupled regardless of the uncertainties $\delta_{i,k}$ for all $i \in \mathcal{N}$. Consider now a large-scale network that consists of N interconnected subsystems, and each subsystem can be described by an uncertain discrete-time linear time-invariant system with additive disturbance of the form

$$\begin{cases} x_{i,k+1} & = A_{ii}(\delta_{i,k})x_{i,k} + B_i(\delta_{i,k})u_{i,k} + q_{i,k} \\ q_{i,k} & = \sum_{j\in\mathcal{N}_i} A_{ij}(\delta_{i,k})x_{j,k} + C_i(\delta_{i,k})w_{i,k} \end{cases}. \tag{3.9}$$

Following Assumption 8, one can consider a linear feedback gain matrix K_i for each subsystem $i \in \mathcal{N}$ such that $K = \mathrm{diag}_{i\in\mathcal{N}}(K_i)$. Using K_i in each subsystem, we assume that there exists P_i for each subsystem $i \in \mathcal{N}$ such that $P = \mathrm{diag}_{i\in\mathcal{N}}(P_i)$ preserves the condition in (3.4). Consider now the objective function of each subsystem $i \in \mathcal{N}$ in the following form:

$$J_i(\boldsymbol{x}_i, \boldsymbol{u}_i) := \mathbb{E}\left[\sum_{k=0}^{T-1} \left(x_{i,k}^\top Q_i x_{i,k} + u_k^\top R_i u_{i,k} \right) + x_{i,T}^\top P_i x_{i,T} \right],$$

where $Q_i \in \mathbb{R}_{\geq 0}^{n_i\times n_i}$, $R_i \in \mathbb{R}_{>0}^{m_i\times m_i}$ such that $Q = \mathrm{diag}_{i\in\mathcal{N}}(Q_i)$, and $R = \mathrm{diag}_{i\in\mathcal{N}}(R_i)$. Note that $\boldsymbol{x}_i = \mathrm{col}_{k\in\mathcal{T}}(x_{i,k})$ and $\boldsymbol{u}_i = \mathrm{col}_{k\in\mathcal{T}}(u_{i,k})$ such that $\boldsymbol{x} = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{x}_i)$ and $\boldsymbol{u} = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{u}_i)$.

Using $\boldsymbol{v}_i = \mathrm{col}_{k\in\mathcal{T}}(v_{i,k})$ such that $\boldsymbol{v} = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{v}_i)$, we decompose the proposed formulation in (3.6) into the following finite horizon scenario program for each subsystem $i \in \mathcal{N}$:

$$\min_{\boldsymbol{v}_i\in\mathbb{R}^{Tm_i}} \quad J_i(\boldsymbol{x}_i, \boldsymbol{u}_i) \tag{3.10a}$$

$$\text{s.t.} \quad x_{i,k+1}^{(i)} = A_{ii}(\delta_{i,k}^{(i)})x_{i,k}^{(i)} + B_i(\delta_{i,k}^{(i)})u_{i,k}^{(i)} + q_{i,k}^{(i)}, \tag{3.10b}$$

$$x_{i,k+\ell}^{(i)} \in \mathcal{X}_i\,,\ \ell \in \mathbb{N}_+\,,\ \forall \boldsymbol{w}_i^{(i)} \in \mathcal{S}_{\boldsymbol{w}_i}\,,\ \forall \boldsymbol{\delta}_i{}^{(i)} \in \mathcal{S}_{\boldsymbol{\delta}_i} \tag{3.10c}$$

$$u_{i,k}^{(i)} = K_i x_{i,k}^{(i)} + v_{i,k} \in \mathcal{U}_i\,, \quad \forall k \in \mathcal{T}, \tag{3.10d}$$

where $\boldsymbol{w}_i = \mathrm{col}_{k\in\mathcal{T}}(w_{i,k}) \in \mathcal{W}_i$ and $\boldsymbol{\delta}_i = \mathrm{col}_{k\in\mathcal{T}}(\delta_{i,k}) \in \Delta_i$ such that $\mathcal{W} = \prod_{i\in\mathcal{N}} \mathcal{W}_i$ and $\Delta = \prod_{i\in\mathcal{N}} \Delta_i$. The sets $\mathcal{S}_{\boldsymbol{w}_i} := \{\boldsymbol{w}_i^{(1)}, \cdots, \boldsymbol{w}_i^{(N_{s_i})}\} \in \mathcal{W}_i^{N_{s_i}}$ and $\mathcal{S}_{\boldsymbol{\delta}_i} := \{\boldsymbol{\delta}_i^{(1)}, \cdots, \boldsymbol{\delta}_i^{(N_{s_i})}\} \in \Delta_i^{N_{s_i}}$ denote sets of given finite samples (scenarios) of disturbance and uncertainties in each subsystem $i \in \mathcal{N}$, such that $\mathcal{S}_{\boldsymbol{w}} = \prod_{i\in\mathcal{N}} \mathcal{S}_{\boldsymbol{w}_i}$ and $\mathcal{S}_{\boldsymbol{\delta}} = \prod_{i\in\mathcal{N}} \mathcal{S}_{\boldsymbol{\delta}_i}$. Note that we use indices in parenthesis to refer to each scenario of the random variables, e.g., (i), whereas indices without parenthesis refer to each subsystem $i \in \mathcal{N}$.

Remark 5 *The proposed constraint* (3.10c) *represents an approximation of the following chance constraint on the state of each subsystem $i \in \mathcal{N}$:*

$$\mathbb{P}[\, x_{i,k+\ell} \in \mathcal{X}_i \,,\, \ell \in \mathbb{N}_+ \,] \geq 1 - \varepsilon_i \,, \tag{3.11}$$

where $\varepsilon_i \in (0,1)$ is the admissible state constraint violation parameter of each subsystem (3.9)*. One can also consider $\alpha_i = 1 - \varepsilon_i$ as the desired level of state feasibility parameter of each subsystem* (3.9)*.*

In the following proposition, we provide a connection between the proposed optimization problem in (3.10) and the optimization problem in (3.6).

Proposition 1 *Given Assumption 8 and the block-diagonal structure for the state-feedback controller $K = \mathrm{diag}_{i\in\mathcal{N}}(K_i)$ for the large-scale system dynamics* (3.1)*, the optimization problem in* (3.10) *is an exact decomposition of the optimization problem in* (3.6)*.*

The proof is provided in Appendix B. □

The following theorem can be considered as the main result of this section to quantify the robustness of the solutions obtained by (3.10).

Theorem 3 *Let $\varepsilon_i, \beta_i \in (0,1)$ be chosen such that $\varepsilon = \sum_{i\in\mathcal{N}} \varepsilon_i \in (0,1)$, $\beta = \sum_{i\in\mathcal{N}} \beta_i \in (0,1)$, and $N_{s_i} \geq \mathsf{N}(\varepsilon_i, \beta_i, Tm_i)$ for all subsystem $i \in \mathcal{N}$. If $\boldsymbol{v}^* = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{v}_i^*)$, the collection of the optimizers of problem* (3.10) *for all subsystem $i \in \mathcal{N}$, is applied to the discrete-time dynamical system* (3.1) *for a finite horizon of length T, then, with at least confidence $1-\beta$, the original constraint* (3.5c) *is satisfied for all $k \in \mathcal{T}$ with probability more than $1-\varepsilon$.*

The proof is provided in Appendix B. □

The interpretation of Theorem 3 is as follows. In the proposed distributed scenario program (3.10), each subsystem $i \in \mathcal{N}$ can have a desired level of constraint violation ε_i and a desired level of confidence level $1-\beta_i$. To keep the robustness level of the collection of solutions in a probabilistic sense (3.5c) for the discrete-time dynamical system (3.1), these choices have to follow a certain design rule, e.g. $\varepsilon = \sum_{i\in\mathcal{N}} \varepsilon_i \in (0,1)$ and $\beta = \sum_{i\in\mathcal{N}} \beta_i \in (0,1)$. This yields a fixed ε, β for the large-scale system (3.1) and the individual ε_i, β_i for each subsystem $i \in \mathcal{M}$. It is important to mention that in order to maintain the violation level for the large-scale system with many partitions, i.e. $|\mathcal{N}| = N \uparrow$, the violation level of individual agents needs to decrease, i.e. $\varepsilon_i \downarrow$, which may lead to conservative results for each subsystem, since the number of required samples needs to increase, i.e. $N_{s_i} \uparrow$. Addressing such a limitation is subject of ongoing research work.

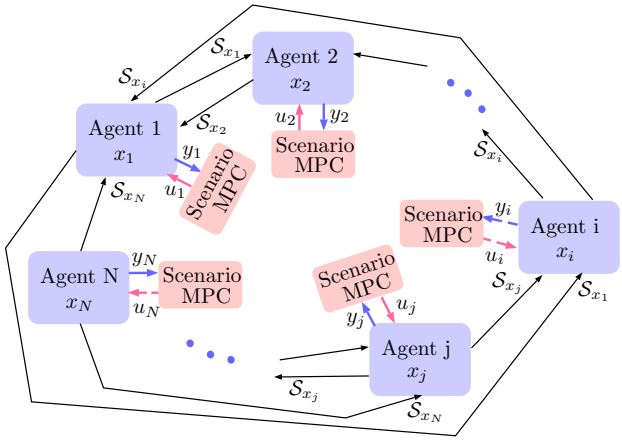


Figure 3.2: Distributed scenario MPC which corresponds to the problem (3.10). The measurement variable y_i is the full state information x_i which is sent to the centralized controllers for all agent $i \in \mathcal{N}$.

An important key feature of the proposed distributed scenario program in (3.10) compared to the optimization problem in (3.6) is as follows. Using the proposed distributed framework, we decompose a large-scale scenario program (3.6) with N_s number of scenarios into N small-scale scenario programs (3.10) with N_{s_i} number of scenarios. This yields a significant reduction in the computation time complexity of scenario programs compared to (3.6) by using the proposed distributed scenario program (3.10). Using the subsystem dynamics in (3.9), agent $i \in \mathcal{N}$ substitutes $q_{i,k}^{(i)}$ in the proposed scenario optimization problem (3.10) with the following relation:

$$q_{i,k}^{(i)} = \sum_{j \in \mathcal{N}_i} A_{ij}(\delta_{i,k}^{(i)}) x_{j,k}^{(i)} + C_i(\delta_{i,k}^{(i)}) w_{i,k}^{(i)},$$

where $\delta_{i,k}^{(i)}$ and $w_{i,k}^{(i)}$ are the local scenarios of random variables that are available in each subsystem by definition $\boldsymbol{w}_i^{(i)} \in \mathcal{S}_{\boldsymbol{w}_i}$ and $\boldsymbol{\delta}_i^{(i)} \in \mathcal{S}_{\boldsymbol{\delta}_i}$, and taking into consideration that the interaction dynamics model $A_{ij}(\delta_{i,k}^{(i)})$ by each neighboring agent $j \in \mathcal{N}_i$ is also available for agent $i \in \mathcal{N}$. Hence, the only information that subsystem $i \in \mathcal{N}$ needs is an N_{s_i} number of samples of the state variable $\boldsymbol{x}_j^{(i)} = \mathrm{col}_{k \in \mathcal{T}}(x_{j,k}^{(i)}) \in \mathbb{X}_j := \mathcal{X}_j^T$ from all its neighboring subsystems $j \in \mathcal{N}_i$ at each sampling time $k \in \mathcal{T}$.

It is important to note that even though the proposed distributed scenario optimization problem in (3.10) yields a reduction of computation time complexity, it however requires more communication between each subsystem, since at each sampling time $k \in \mathcal{T}$ all neighboring agents $j \in \mathcal{N}_i$ of the agent $i \in \mathcal{N}$ should send a set of scenarios of the state variable $\mathcal{S}_{\boldsymbol{x}_j} := \{\boldsymbol{x}_j^{(1)}, \cdots, \boldsymbol{x}_j^{(N_{s_i})}\} \in \mathbb{X}_j^{N_{s_i}}$ to the agent $i \in \mathcal{N}$.

To summarize this section, we present a network of agents that are dynamically coupled with their own local scenario MPC in Figure 3.2. In the next section, we will propose a novel inter-agent information exchange scheme to provide a more flexible and less conservative framework to exchange information between agents.

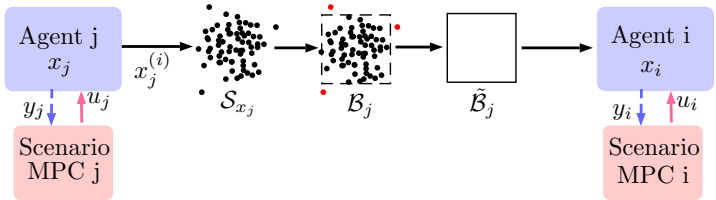


Figure 3.3: Pictorial representation of the proposed inter-agent soft communication scheme. $\mathcal{S}_{\boldsymbol{x}_j}$ is the set of $\tilde{N}_{s_i}$ scenarios, $\mathcal{B}_j$ is the parametrized set used in the optimization problem (3.12), and $\tilde{\mathcal{B}}_j$ is the solution of the optimization problem (3.12). The measurement variables y_i and y_j are the full state information x_i and x_j, respectively, which are sent to the controllers.

3.4. INFORMATION EXCHANGE SCHEME

In this section, we first describe the information exchange between agents and then propose a set-based information exchange scheme which will be referred to as a *soft communication protocol* later in this section. We finally provide the theoretical results to quantify robustness of the proposed information exchange scheme between neighboring agents.

When the proposed distributed framework (3.10) is applied to the large-scale scenario program (3.6), all neighboring agent $j \in \mathcal{N}_i$ of the agent $i \in \mathcal{N}$ should send a set of scenarios of the state variable $\mathcal{S}_{\boldsymbol{x}_j} := \{\boldsymbol{x}_j^{(1)}, \cdots, \boldsymbol{x}_j^{(N_{s_i})}\} \in \mathbb{X}_j^{N_{s_i}}$ to agent i at each sampling time $k \in \mathcal{T}$. It is of interest to address the issue of how an agent $j \in \mathcal{N}_i$ can send the contents of $\mathcal{S}_{\boldsymbol{x}_j}$ to the agent $i \in \mathcal{N}$. We propose the following two schemes:

1) Following our proposed setup in (3.10) to achieve a probabilistic guarantee for the obtained solution, agent $i \in \mathcal{N}$ requests from its neighboring agents to send the complete set of data $\mathcal{S}_{\boldsymbol{x}_j}$, element by element such that the number of required samples N_{s_i}, is chosen according to Theorem 3 in order to have a given probabilistic guarantee for the optimizer $\boldsymbol{v}_i^*$. We refer to this scheme as a *hard communication protocol* between agents. Its advantage is that it is simple and transmits exactly the contents of $\mathcal{S}_{\boldsymbol{x}_j}$, but due to possibly high values of N_{s_i}, it may turn out to be too costly in terms of required communication bandwidth.

2) To address this shortcoming, we propose another scheme, where agent $j \in \mathcal{N}_i$ sends instead a suitable parametrization of a set that contains all the possible values of data with a desired level of probability (*the level of reliability*) $\tilde{\alpha}_j$. By considering a simple family of sets, for instance boxes in $\mathbb{R}^{n_j}$, communication cost can be kept down at reasonable levels. We refer to this scheme as a *soft communication protocol* between agents (see Figure 3.3). Such a scheme may be understood as a cascaded scenario scheme similar to the one in [99], where a sufficient number of scenarios was determined in order to establish probabilistic feasibility for two cascaded scenario programs subject to a similar source of uncertainty. Our soft communication setting however differs from [99], since each agent is subject to its own uncertainty source. We aim to incorporate the reliability notion of such a soft communication scheme into the feasibility bound of each agent, in addition to determining the number of required scenarios that can be obtained as a corollary of our results presented so far.

We now describe the soft communication protocol in more detail. The neighboring agent $j \in \mathcal{N}_i$ has to first generate $\tilde{N}_{s_i}$ samples of $\boldsymbol{x}_j$ in order to build the set $\mathcal{S}_{\boldsymbol{x}_j}$. It is important to notice that in the soft communication protocol the number $\tilde{N}_{s_i}$ of samples generated by agent j may be different than the one needed by agent i, which is N_{s_i}, as will be remarked later. Let us then introduce $\mathcal{B}_j \subset \mathbb{R}^{n_j}$ as a bounded set containing all the elements of $\mathcal{S}_{\boldsymbol{x}_j}$. We assume for simplicity that $\mathcal{B}_j$ is an axis-aligned hyper-rectangular set [98]. This is not a restrictive assumption and any convex set, e.g. ellipsoids and polytopes, could have been chosen instead as described in [148]. We can define $\mathcal{B}_j := [-\boldsymbol{b}_j, \boldsymbol{b}_j]$ as an interval, where the vector $\boldsymbol{b}_j \in \mathbb{R}^{n_j}$ defines the hyper-rectangle bounds.

Consider now the following optimization problem that aims to determine the set $\mathcal{B}_j$ with minimal volume by minimizing $\|\boldsymbol{b}_j\|_1$:

$$\begin{cases} \min\limits_{\boldsymbol{b}_j \in \mathbb{R}^{n_j}} & \|\boldsymbol{b}_j\|_1 \\ \text{s.t.} & \boldsymbol{x}_j^{(l)} \in [-\boldsymbol{b}_j, \boldsymbol{b}_j]\,,\ \forall \boldsymbol{x}_j^{(l)} \in \mathcal{S}_{\boldsymbol{x}_j}\,, \\ & l = 1, \cdots, \tilde{N}_{s_i} \end{cases} \tag{3.12}$$

where $\tilde{N}_{s_i}$ is the number of samples $\boldsymbol{x}_j \in \mathcal{S}_{\boldsymbol{x}_j}$ that neighboring agent j has to take into account in order to determine $\mathcal{B}_j$. If we denote by $\tilde{\mathcal{B}}_j = [-\tilde{\boldsymbol{b}}_j, \tilde{\boldsymbol{b}}_j]$ the optimal solution of (3.12) computed by the neighbor agent j, then for implementing the soft communication protocol, agent j needs to communicate only the vector $\tilde{\boldsymbol{b}}_j$ along with the level of reliability $\tilde{\alpha}_j$ to the agent i.

Definition 2 *A set $\tilde{\mathcal{B}}_j$ is called $\tilde{\alpha}_j$−reliable if*

$$\mathbb{P}\left[\boldsymbol{x}_j \in \mathbb{X}_j : \boldsymbol{x}_j \notin [-\tilde{\boldsymbol{b}}_j, \tilde{\boldsymbol{b}}_j]\right] \leq 1 - \tilde{\alpha}_j\,, \tag{3.13}$$

and we refer to $\tilde{\alpha}_j$ as the level of reliability of the set $\tilde{\mathcal{B}}_j$.

We now provide the following theorem to determine $\tilde{\alpha}_j$ as the level of reliability of the set $\tilde{\mathcal{B}}_j$.

Theorem 4 *Fix $\tilde{\beta}_j \in (0,1)$ and let*

$$\tilde{\alpha}_j = \sqrt[\tilde{N}_{s_i} - n_j]{\frac{\tilde{\beta}_j}{\binom{\tilde{N}_{s_i}}{n_j}}}\,. \tag{3.14}$$

We then have

$$\mathbb{P}^{\tilde{N}_{s_i}}\left\{\{\boldsymbol{x}_j^1, \cdots, \boldsymbol{x}_j^{\tilde{N}_{s_i}}\} \in \mathbb{X}_j^{\tilde{N}_{s_i}} : \mathbb{P}\left[\boldsymbol{x}_j \in \mathbb{X}_j : \boldsymbol{x}_j \notin [-\tilde{\boldsymbol{b}}_j, \tilde{\boldsymbol{b}}_j]\right] \leq 1 - \tilde{\alpha}_j\right\} \geq 1 - \tilde{\beta}_j\,.$$

The proof is provided in Appendix B. □

Theorem 4 implies that given an hypothetical new sample $\boldsymbol{x}_j \in \mathbb{X}_j$, agent $j \in \mathcal{N}_i$ has a confidence of at least $1 - \tilde{\beta}_j$ that the probability of $\boldsymbol{x}_j \in \tilde{\mathcal{B}}_j = [-\tilde{\boldsymbol{b}}_j, \tilde{\boldsymbol{b}}_j]$ is at least $\tilde{\alpha}_j$. Therefore, one can *rely on $\tilde{\mathcal{B}}_j$ up to $\tilde{\alpha}_j$ probability.* The number of samples $\tilde{N}_{s_j}$ in the

proposed formulation (3.12) is a design parameter chosen by the neighboring agent $j \in \mathcal{N}_i$. In Figure 3.3 the number of red dots refers to the difference between N_{s_j} and $\tilde{N}_{s_j}$. We however remark that one can also set a given $\tilde{\alpha}_j$ as the desired level of reliability and obtain from (3.14) the required number of samples $\tilde{N}_{s_i}$.

When an agent $i \in \mathcal{N}$ and its neighbor $j \in \mathcal{N}_i$ adopt the soft communication scheme, there is an important effect on the probabilistic feasibility of agent i, following Remark 5. Such a scheme introduces some level of stochasticity on the probabilistic feasibility of agent i, due to the fact that the neighboring information is only *probabilistically reliable*. This will affect the local probabilistic robustness guarantee of feasibility as it was discussed in Theorem 3 and consequently in Theorem 2. To accommodate the level of reliability of neighboring information, we need to marginalize a joint cumulative distribution function (cdf) probability of $\boldsymbol{x}_i$ and the generic sample $\boldsymbol{x}_j \in \mathbb{X}_j$ appearing in Theorem 4. We thus have the following theorem, which can be regarded as the main theoretical result of this section.

Theorem 5 *Given $\tilde{\alpha}_j \in (0,1)$ for all $j \in \mathcal{N}_i$ and a fixed $\alpha_i = 1-\varepsilon_i \in (0,1)$, the state trajectory of a generic agent $i \in \mathcal{M}$ is probabilistically $\bar{\alpha}_i$–feasible for all $\boldsymbol{w}_i \in \mathcal{W}_i$, $\boldsymbol{\delta}_i \in \Delta_i$, i.e.,*

$$\mathbb{P}\left[x_{i,k+\ell} \in \mathcal{X}_i\,,\, \ell \in \mathbb{N}_+\right] \geq \bar{\alpha}_i\,, \tag{3.15}$$

where $\bar{\alpha}_i = 1 - \frac{1-\alpha_i}{\tilde{\alpha}_i}$ such that $\tilde{\alpha}_i = \prod_{j\in\mathcal{N}_i}(\tilde{\alpha}_j)$.

The proof is provided in Appendix B. □

Following the statement of Theorem 5, it is straightforward to observe that if for all neighboring agents $j \in \mathcal{N}_i$, $\tilde{\alpha}_j \to 1$ then $\bar{\alpha}_i \to \alpha_i$. This means that if *the level of reliability* of the neighboring information is one, i.e. $\mathbb{P}\left[\boldsymbol{x}_j \in \tilde{\mathcal{B}}_j : \forall j \in \mathcal{N}_i\right] = 1$, then, the state feasibility of agent i will have the same probabilistic level of robustness as the hard communication scheme, i.e. $\mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\right] \geq \alpha_i = 1-\varepsilon_i$. Combining this result with the statement of Theorem 3, the proposed soft communication scheme introduces some level of stochasticity on the feasibility of the large-scale system as in (3.5c). In particular, $\varepsilon_i \in (0,1)$ the level of constraint violation in each agent $i \in \mathcal{N}$ will increase, since it is proportional with the inverse of $\prod_{j\in\mathcal{N}_i}(\tilde{\alpha}_j) \in (0,1)$, and therefore, $\varepsilon = \sum_{i\in\mathcal{N}} \varepsilon_i \in (0,1)$ will also increase. After receiving the parametrization of $\tilde{\mathcal{B}}_j$ and the level of reliability $\tilde{\alpha}_j$, agent $i \in \mathcal{N}$ should immunize itself against all possible variation of $\boldsymbol{x}_j \in \tilde{\mathcal{B}}_j$ by taking the worst-case of $\tilde{\mathcal{B}}_j$, similar to the worst-case reformulation proposed in [144, Proposition 1]. It is important to notice that in this way, we decoupled the sample generation of agent $j \in \mathcal{N}_i$ from agent $i \in \mathcal{N}$.

We summarize this section by mentioning that Figure 3.3 depicts a conceptual representation of the proposed soft communication scheme between two neighboring agents. In the next section, we provide an operational framework that uses our developments in preceding sections in a more practical framework namely plug-and-play (PnP) operations.

3.5. PLUG-AND-PLAY OPERATIONAL FRAMEWORK

We summarize our proposed distributed scenario MPC in Algorithm 1 such that agents communicate with each other by using our proposed soft inter-agent communication

3

Algorithm 1 Distributed Scenario MPC

1: **Decompose** the large-scale dynamical system (3.1) into N agents as the proposed form in (3.9)
2: **Determine** the index set of neighboring agents $\mathcal{N}_i$ for each agent $i \in \mathcal{N}$
3: **For** each agent $i \in \mathcal{N}$ **do**
4: **fix** initial state $x_{i,0} \in \mathcal{X}_i$, $\varepsilon_i \in (0,1)$, and $\beta_i \in (0,1)$ such that

$$\varepsilon = \sum_{i\in\mathcal{N}} \varepsilon_i \in (0,1)\,,\ \beta = \sum_{i\in\mathcal{N}} \beta_i \in (0,1)$$

5: **initialize** $\tilde{\mathcal{B}}_j$ for all neighboring agents $j \in \mathcal{N}_i$
6: **determine** $N_{\bar{s}_i} \in (0,+\infty)$ to approximate the objective function, and N_{s_i} following Theorem 3 to approximate the chance constraint (3.11) in an equivalent sense
7: **generate** $N_{\bar{s}_i}$, N_{s_i} scenarios of $\boldsymbol{w}_i$, $\boldsymbol{\delta}_i$ to determine the sets of $\bar{\mathcal{S}}_{(\boldsymbol{w}_i,\boldsymbol{\delta}_i)}$ and $\mathcal{S}_{\boldsymbol{w}_i}$, $\mathcal{S}_{\boldsymbol{\delta}_i}$
8: **solve** the proposed optimization problem in (3.10) by taking into account the worst-case of $\tilde{\mathcal{B}}_j$ and determine $\boldsymbol{v}_i^*$
9: **generate** $\tilde{N}_{s_i}$ scenarios of $\boldsymbol{x}_i$ using the dynamical system of agent i in form of (3.9) and $\boldsymbol{v}_i^*$ together with $\mathcal{S}_{\boldsymbol{w}_i}$, $\mathcal{S}_{\boldsymbol{\delta}_i}$
10: **determine** set $\tilde{\mathcal{B}}_i$ by solving the optimization problem (3.12)
11: **send** the set $\tilde{\mathcal{B}}_i$ to all neighboring agents $j \in \mathcal{N}_i$
12: **receive** the sets $\tilde{\mathcal{B}}_j$ from all neighboring agents $j \in \mathcal{N}_i$
13: **apply** the first input of solution $u_{i,k}^* = K_i x_{i,k} + v_{i,k}^*$ into the uncertain subsystem (3.9)
14: **measure** the state and substitute it as the initial state of the next step $x_{i,0}$
15: **set** $k \leftarrow k+1$
16: **goto** Step (7)
17: **End for**

Algorithm 2 Plug-and-Play Operation

Plug-in Operation
1: **Add** the number of new subsystems into the previous number of subsystems, e.g. one additional agent $N \leftarrow N+1$ such that $|\mathcal{N}| = N+1$
2: **Update** the index set of neighboring agents $\mathcal{N}_i$ for each agent $i \in \mathcal{N}$
3: **Goto** Step (3) of Algorithm 1
Plug-out Operation
4: **Remove** the number of new subsystems from the previous number of subsystems, e.g. one exclusion agent $N \leftarrow N-1$ such that $|\mathcal{N}| = N-1$
5: **Update** the index set of neighboring agents $\mathcal{N}_i$ for each agent $i \in \mathcal{N}$
6: **Goto** Step (3) of Algorithm 1

scheme. Note that in case of the hard communication scheme, each agent needs to generate N_{s_i} scenarios and send exactly all of them to all its neighboring agents $j \in \mathcal{N}_i$. In other words, the following changes have to be made in Algorithm 1. $\tilde{N}_{s_i}$ will be substituted by N_{s_i} in Step 9 and Step 10 will be removed. Steps 11 and 12 will send and receive exactly N_{s_i} samples, respectively.

In Algorithm 1 it is assumed that the feedback control gain matrices K_i for all agent $i \in \mathcal{N}$ are given (3.4), and the coupling terms $A_{ij}(\delta_{i,k})$ are known between each agent $i \in \mathcal{N}$ and its neighboring agents $j \in \mathcal{N}_i$. It is important to note that Step 5 of Algorithm 1, initializes $\tilde{\mathcal{B}}_j$ for all neighboring agents $j \in \mathcal{N}_i$ to be used for the initial iteration in Step 8, and then, at each iteration all agent $i \in \mathcal{N}$ will send and receive $\tilde{\mathcal{B}}_j$ from all its neighboring agents $j \in \mathcal{N}_i$ as in Steps 11 and 12, respectively.

We also summarize the steps that are needed for plug-in and plug-out operations of each agent $i \in \mathcal{N}$ in Algorithm 2. Note that in a plugged-in or plugged-out operation all agents $i \in \mathcal{N}$ have to update their ε_i with β_i to respect the condition in Theorem 3 as in (4) to achieve the desired level of constraint feasibility for the large-scale system (3.1). One can also redesign K_i to potentially improve the local control performance of each agent $i \in \mathcal{N}$.

3.6. NUMERICAL STUDY

This section presents two case studies to illustrate the functionality of our proposed framework to deal with private and common uncertainty sources in networked control problems. The simulation environment for both cases was MATLAB with YALMIP as the interface [86] and Gurobi as the solver.

We simulate four problem formulations, a centralized SMPC (CSMPC) using (3.6), a distributed SMPC (DSMPC) via the distributed scenario program in (3.10), and DSMPC with the proposed soft communication scheme with 0.85−reliability (DSMPCS−0.85) as described in Definition 2 and DSMPCS−0.50, both following Algorithm 1 in a closed-loop control system framework. For comparison purposes, we also present the results obtained via decoupled SMPC (DeSMPC), where the impact of coupling neighboring dynamics in (3.9) are relaxed.

In Figure 3.5, Figure 3.6 and Figure 3.7, we evaluate our proposed framework in terms of a-posteriori feasibility validation of the obtained results in both case studies. The "red" line represents the results obtained via DeSMPC, the "blue" line shows the results obtained via CSMPC, the "magenta" presents the results obtained by using DSMPC, the "dark green" and "light green" lines show the results obtained via DSMPCS−0.85 and DSMPCS−0.50, respectively. The "black" lines indicate the bounds of the three dynamically coupled systems.

3.6.1. THREE-ROOM CASE STUDY

We simulate a three-room building climate comfort system shown in Figure 3.4 such that the temperature of rooms are dynamically coupled without any common constraints. The outside weather temperature and the related looses, e.g. through windows, is considered as the private uncertainty.

Consider now a three-room building system dynamics: $x_{k+1} = A(\delta_k)x_k + B(\delta_k)u_k +$

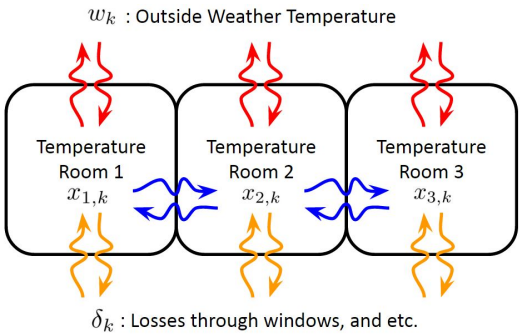


Figure 3.4: An example of three-room building climate comfort system. $x_{1,k}$, $x_{2,k}$ and $x_{3,k}$ are states of the building that are related to the temperature dynamics of rooms. The temperature dynamics are influenced by the outside weather temperature (w_k) shown by 'red' arrows and the possible losses (δ_k) through windows, and etc., of the rooms which are represented via 'orange' arrows. Moreover, the 'blue' arrows denotes the impact of the neighboring rooms on each other.

$C(\delta_k)w_k$, where

$$A = \begin{bmatrix} 0.2 & 0.3 & 0 \\ 0.2 & 0.1 & 0.1 \\ 0.2 & 0 & 0.4 \end{bmatrix}, \; B = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}, \; C = \begin{bmatrix} 0.02 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.02 \end{bmatrix},$$

such that $A(\delta_k) = A + \delta_k$ and $B(\delta_k) = B + \delta_k$ as well as $C(\delta_k) = C + \delta_k$. The system matrices are a simplified model of a three-room building such that the states $x_{i,k}$ for $i = 1,2,3$, denote the temperature of rooms. The uncertain variable $\delta_k \in \mathbb{R}$ represents the modeling errors, losses through windows, and $w_k \in \mathbb{R}$ can be realized as the outside weather temperatures.

To generate random scenarios from the additive disturbance, we built a discrete normal process such that one day hourly-based forecasted (nominal) outside weather temperature is used which varies within 10% of its nominal scenario at each sampling time. As for the uncertainty δ_k, we generate a random variable from a normal distribution with a mean value 0, variance 1 and a maximal magnitude of 0.01 at each sampling time.

The initial state variables are $[21\ 19\ 23]^\top$ and the objective is to keep the temperature of rooms within our desired lower $[20.5\ 18.5\ 22.5]^\top$ and $[21.5\ 19.5\ 23.5]^\top$ upper bounds at the minimum control input u_k. The control input u_k is also constrained to be within -1.5 [kWh] and 1.5 [kWh] for all three rooms, due to actuator saturation. The initialization of the $\tilde{\mathcal{B}}_j$ for all neighboring agents $j \in \mathcal{N}_\rangle$ as in Step 5 in Algorithm 1 can be done for instance by assuming the initial temperature of the neighboring rooms are known for all rooms.

In Figure 3.5, the dynamically coupled state trajectories for all three rooms are shown. One can clearly see in Figure 3.5 that the dynamically coupled state trajectories are feasible in a probabilistic sense, since the agent operations are within the lower and upper bounds compared to DeSMPC that violates the constraints completely; the obtained solution via DeSMPC is completely outside of the feasible areas after the first sampling time and thus, we just keep the other results for our discussions. This is a direct result

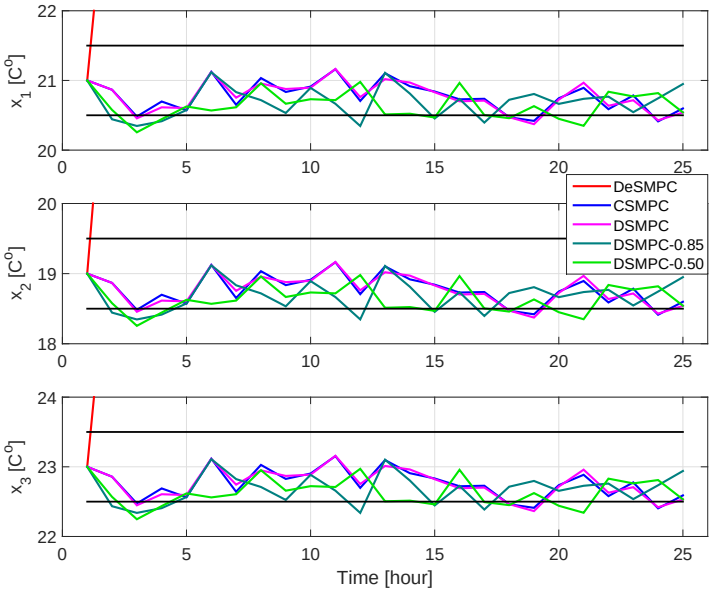


Figure 3.5: A-posteriori feasibility validation of the obtained results. The "red", "blue", "magenta", "dark green", and "light green" lines are related to the obtained results via DeSMPC, CSMPC, DSMPC, DSMPCS−0.85, and DSMPCS−0.50, respectively. The "black" lines are related to the upper bound values. The top, middle, and down plots are related to the temperatures of room 1, 2, and 3, respectively.

of Theorem 3 such that the obtained solutions via our proposed formulations have to be probabilistically feasible, that can be clearly seen in Figure 3.5, since the trajectories are on the lower bounds.

3.6.2. THREE-BUILDING (ATES SYSTEMS) CASE STUDY

We next simulate the thermal energy demands of three buildings modeled using realistic parameters and the actual registered weather data in the city center of Utrecht, The Netherlands, where these buildings are located and these had been equipped with aquifer thermal energy storage (ATES) systems [148]. An ATES system consists of two wells (warm and cold water wells) and it is considered as a heat source or sink, or as a storage for thermal energy that operates in a seasonal mode. A large-scale network of interconnected buildings, that are constrained via the state variables of ATES systems, which are the volume of water and the thermal energy content of each well, was modeled in [148]. To prevent overlap of nearby systems there are constraints on a growing thermal radius, $r^h_{i,k}$ [m], $r^c_{i,k}$ [m], of each well of an ATES system with its neighboring agents, e.g. $r^h_{i,k} + r^c_{j,k} \leq d_{ij}$ for each agent $i \in \mathcal{N}$ and its neighboring agent $j \in \mathcal{N}_i$. In [148, Corollary 1], it was shown that such a constraint on the growing thermal radius can be replaced with a constraint on the state variables (volume of water) of the ATES systems. We refer interested readers to [148] for the detailed explanations on this case study.

It is important to note that based on the decomposition technique in Section 3.3, the

subsystem $i \in \mathcal{N}$ is assumed to be dynamically coupled with all its neighboring subsystems $j \in \mathcal{N}_i$, and the feasible set are assumed to be decoupled as defined in (3.7). This leads to an issue for applying Algorithm 1 to this case study due to the aforementioned coupling constraints. To overcome this difficulty, it is assumed that variations in the predicted actions of each agent $i \in \mathcal{N}$ at each sampling time $k \in \mathcal{T}$ are bounded within the samples that are sent to the neighboring agents $j \in \mathcal{N}_i$. Such an assumption is inspired by practical reasons, since the ATES system dynamics are very slow as reported in [148]. In this way, the coupling constraints can be considered as the local (decoupled) uncertain constraints such that the neighboring actions are given as the uncertainty source. We are then able to apply Algorithm 1 to this case study as well.

A simulation study is carried out for one year based on actual weather conditions from March 2011-March 2012 in a weekly-based time steps with three months prediction horizon to meet the coupling constraints between ATES systems. We introduce additive disturbance and uncertainty sources into the deterministic first-order dynamical model of [142, 148]. It has been reported in [135] that the ambient temperature of water in aquifers is changing over the long life-time usage of ATES systems. We capture such a behavior by introducing an additive unknown (disturbance) source of energy which yields a time correlation function via the initial value of energy content variable of an ATES system. In addition to this, the system parameters of an ATES system are a function of the stored water temperature in each well, e.g. see [148, Figure 4]. We therefore introduce a small level of perturbation as an uncertainty in the parameters of the ATES system dynamics.

To generate random scenarios from the additive disturbance, we built a discrete normal process such that the nominal scenario is 10% of the amplitude of the energy content in a deterministic ATES system model and varies within 5% of its nominal scenario at each sampling time. As for the uncertainty δ_k, we generate a random variable from a Gaussian distribution with a mean value 0, variance 0.3 and a maximal magnitude of 0.03 at each sampling time.

In Figure 3.6, we show a-posteriori feasibility validation of the coupling constraints between each agent $i = 1,2,3$, and neighboring agents, e.g. $r^h_{1,k} + r^c_{2,k} \leq 65$, $r^h_{2,k} + r^c_{3,k} \leq 70$, and $r^h_{3,k} + r^c_{1,k} \leq 70$. Figure 3.7 focuses on the critical time periods in Figure 3.6, where neighboring agents are injecting thermal energies with different pump flow rates. Figure 3.7(a) shows the results from mid-May to mid-August 2011, Figure 3.7(b) presents the results of December 2011 to February 2012, and Figure 3.7(c) depicts the results of mid-April to mid-July 2011.

As a first desired achievement, one can clearly see in Figure 3.6 that the constraints are feasible in a probabilistic sense, since the agent operations are quite close to the upper bounds in the critical time periods compared to DeSMPC that violates the constraints. Strictly speaking, using our proposed framework one can achieve the maximum usage of the aquifer (subsurface) to store thermal energy without affecting the neighboring thermal storage. This is a direct result of Theorem 3 such that the obtained solutions via our proposed formulations have to be probabilistically feasible, that can be clearly seen in Figure 3.7, since the trajectories are on the upper bounds.

It is worth to mention that both Figure 3.5 and Figure 3.7 illustrate our other two main contributions more precisely: 1) the obtained results via CSMPC (blue line) and DSMPC

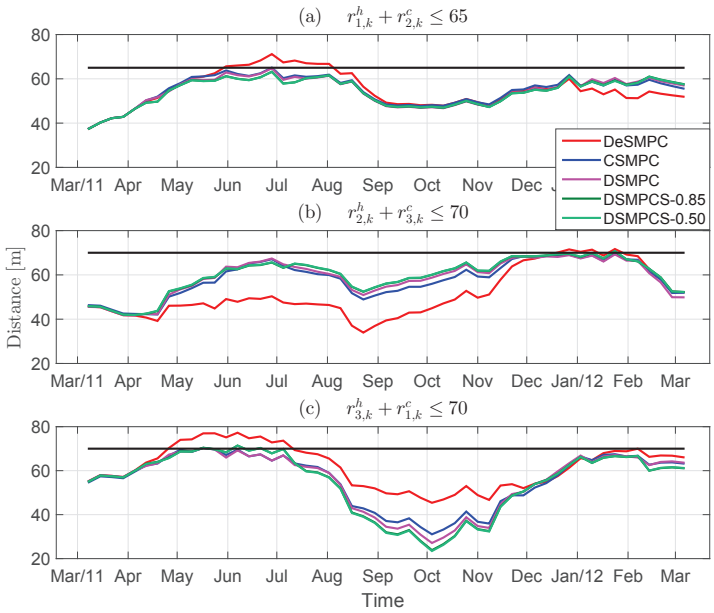


Figure 3.6: A-posteriori validation of the obtained results. The "red", "blue", "magenta", "solid green", and "dashed green" lines are related to the obtained results via DeSMPC, CSMPC, DSMPC, DSMPCS−0.85, and DSMPCS−0.50, respectively. The "black" lines are related to the upper bound values.

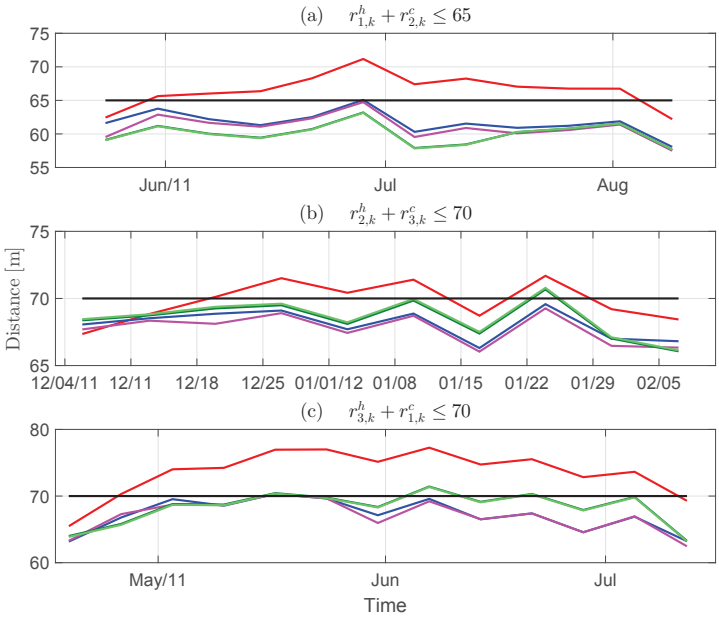


Figure 3.7: Zoom-in of the critical time periods in Figure 3.6.

(magenta line) are practically equivalent throughout the simulation studies; this is due to Proposition 1 and Theorem 3. Actually, the solutions via DSMPC are slightly more conservative compared to the results via CSMPC, and this is a direct consequence of Theorem 3. In fact the level of violation in CSMPC is considered to be $\varepsilon = 0.05$ and leading to $\varepsilon_i = 0.0167$ for all agents, which is more restrictive. 2) the proposed soft communication scheme yields less conservative solutions as explicitly derived in Theorem 5, and can be clearly seen in Figure 3.5 and Figure 3.7 with the obtained results via DSMPCS−0.85 (dark green) and DSMPCS−0.50 (light green). Following Theorem 5 the new violation level using DSMPCS−0.85 is $\bar{\varepsilon}_i = 0.0231$, and using DSMPCS−0.50 is $\bar{\varepsilon}_i = 0.0668$. It is important to notice that the violation level of global chance constraint will increase to $\bar{\varepsilon} = 0.0702$ and $\bar{\varepsilon} = 0.2004$ using DSMPCS−0.85 and DSMPCS−0.50, respectively.

3.7. CONCLUSIONS

In this chapter, we presented a rigorous approach to distributed stochastic model predictive control (SMPC) using the scenario-based approximation for large-scale linear systems with private and common uncertainty sources. We extended the existing results to quantify the robustness of the resulting solutions for both cases of private and common uncertainties in a distributed framework. We then provided a novel inter-agent soft communication scheme to minimize the amount of information exchange between each subsystem. Using a set-based parametrization technique, we introduced a reliability notion and quantified the level of feasibility of the obtained solutions via the distributed SMPC integrated with the so-called soft communication scheme in a probabilistic sense. The theoretical guarantees of the proposed distributed SMPC framework coincide with its centralized counterpart.

4

Distributed Stochastic Reserve Scheduling in AC Power Systems with Uncertain Generation

This chapter presents a framework to carry out multi-area stochastic reserve scheduling (RS) based on an AC optimal power flow (OPF) model with high penetration of wind power using distributed consensus and the alternating direction method of multipliers (ADMM). We first formulate the OPF-RS problem using semidefinite programming (SDP) in infinite-dimensional space, which is in general computationally intractable. Using a novel affine policy, we develop an approximation of the infinite-dimensional SDP as a tractable finite-dimensional SDP, and explicitly quantify the performance of the approximation. To this end, we adopt the recent developments in randomized optimization that allow a priori probabilistic feasibility guarantees to optimally schedule power generating units while simultaneously determining the geographical allocation of the required reserve. We then use the geographical patterns of the power system to decompose the large-scale system into a multi-area power network, and provide a consensus ADMM algorithm to find a feasible solution for both local and overall multi-area network. Using our distributed stochastic framework, each area can use its own wind information to achieve local feasibility certificates, while ensuring overall feasibility of the multi-area power network under mild conditions. We provide numerical comparisons with a new benchmark formulation, the so-called converted DC (CDC) power flow model, using Monte Carlo simulations for two different IEEE case studies.

4.1. INTRODUCTION

POWER transmission system operators (TSOs) aim to find an economic operating point to satisfy the power demand and network constraints by solving an optimal power flow (OPF) problem. TSOs have to deal with increasing degrees of uncertainty due to high penetration of wind power generation. While wind power has clear environmental advantages, it is a highly variable and not fully controllable resource. This imposes novel challenges and tasks for TSOs to avoid blackouts, outages, etc., and highlights the necessity of introducing a new paradigm to existing TSO functionalities.

The reserve scheduling (RS) task of TSOs deals with day-ahead scheduling of the reserve power to accommodate possible mismatches between forecasted and actual wind power. Stochastic variants of the RS problem, where violations are allowed with a small probability to achieve better performance, have received a lot of attention in the past few years, see [100, 118, 171, 174] and the references therein. A stochastic RS problem is typically formulated using a lossless DC model based on the assumption of constant voltage magnitudes and small voltage angles, while ignoring the active power losses [7]. It is worth mentioning that these assumptions do not hold in general and may lead to suboptimality or even infeasibility when implemented on real-world systems, especially for networks under a high degree of stress [164]. Using an AC representation of the power network enables the stochastic RS formulation to accurately model the effect of large deviations of wind power from its forecasted value, and to offer a-priori suitable reserves such that both active and reactive power, and complex-valued voltage are globally optimal. Due to the non-convexity of the OPF problem, identifying such an optimal operating point of a power system may not be straightforward. In [82], different reformulations and relaxations of the AC OPF problem were presented and their connections were discussed. By means of semidefinite programming (SDP), in [82] a convex relaxation was provided under the existence of a rank-one SDP solution to guarantee the recovery of a globally optimal solution of the power network.

4.1.1. RELATED WORKS

The RS problem incorporating an OPF formulation has been introduced in [149], where a chance-constrained OPF problem was formulated. With some modifications, motivated by practical aspects of the problem, the authors in [149] provided a theoretical guarantee that the OPF-RS problem yields a rank-one feasible solution. Using a heuristic Monte Carlo sampling approach, they showed that the resulting optimization problem involves an OPF problem for each wind power profile. Our work in this chapter is motivated by [149] to first rigorously provide theoretical guarantees for the feasibility of physical constraints, and then, develop a distributed reserve scheduling framework for the AC model of power network, which is, to best of our knowledge, has not been addressed in the related works.

Dealing with a chance-constrained AC OPF problem has recently received a lot of attention in literature, see e.g., [35, 46, 56, 130, 170, 172]. In [172] a Monte Carlo sampling approach from both the set of decision variables and uncertain parameters has been employed such that for each random decision variable an AC OPF problem including all random uncertain parameters is solved to determine the feasibility of the decision variable. This procedure is repeated for all random decision variables such that the final

decision is the one that is feasible and has a better cost value. Choosing an appropriate distribution for decision variables together with large computational time are drawbacks of such an algorithm.

Another interesting approach based on partial linearization methods, see e.g. [46], [56] and the references therein, has been adopted in [130]. In [130], the full AC power flow model for the forecasted operating point is formulated, and then, a linearization around this point is done to model the impact of uncertainty by assuming that the forecast errors are small. The analytical chance constraint reformulation is derived by considering a specific distribution for uncertainty, e.g. Gaussian. We note that such an approach fails whenever the distribution of uncertainty is unknown (see, e.g., [131] to handle partially unknown distributions).

While preparing the final version of this work, [35] and [170] independently gave an approach to solve the RS problem using a chance-constrained AC OPF problem in each single hour, based on the results in [171]. The OPF-RS formulation in [35] is similar to [149] with some modifications, whereas in [170] the formulation is weaker compared to [149], since the condition to distribute reserves among generators is relaxed. Even though the authors in [149] presented a complete day-ahead OPF-RS formulation with up- and down spinning reserves, the results in the aforementioned references are limited either to be heuristic or to a single hourly-based RS with the relaxed conditions. The major barrier for representing OPF-RS problems as an SDP is the necessity of defining a square SDP matrix variable, which makes the cardinality of scalar variables of the OPF-RS problem quadratic with respect to the number of buses in the power network. This may yield a very large-scale SDP problem for realistic large-scale power networks of interest.

4.1.2. CONTRIBUTIONS

Our work in this chapter differs from the aforesaid references in two important aspects. We first formulate the AC OPF problem by considering the uncertain wind power generation. Using a similar convexification to [82], we incorporate the stochastic RS into the convexified AC OPF formulation. This results in a large-scale SDP in infinite-dimensional space. We then provide a systematical approach to move from infinite to semi-infinite and then to finite-dimensional space using a novel affine policy. Such a policy differs from the existing one in the literature, see, e.g., [35, 149, 170, 171] and references therein, in technical and practical aspects, and significantly reduces the worst-case computational complexity (WCCC). Theoretical results for the exactness of the approximation and a priori feasibility certificates are provided by adopting the recent developments in randomized optimization that do not require any prior knowledge of uncertainty bounds or distributions. The proposed policy enjoys the property of operational rule of the so-called automatic generation regulator (AGR) concept in power system. Affine policies have been also used for the OPF problem using DC model of power system, e.g., [70], and most recently, in [177] for the AGR actions (up- and downspinning reserves).

We next introduce a decomposition technique by leveraging the geographical patterns in power systems to decompose the high-dimensional SDP into small-scale SDPs related to each area of a multi-area power network. We employ recent results in graph theory to break down the large-scale positive-semidefinite (PSD) constraints into small-

sized constraints. Such a technique has been also considered in [182] for state estimation problem in power systems. We then provide a distributed consensus framework using the alternating direction method of multipliers (ADMM), similarly to [72, 77, 96, 111], with some modifications for the AC OPF problem in power system. We extend such a distributed framework to handle the stochastic RS problem using the AC OPF model of a multi-area network which has not, to the best of our knowledge, been considered in literature. We highlight that such an extension is possible using the proposed affine policy which overcomes the difficulty of having defined a square SDP matrix variable for all possible wind trajectories, see, e.g., [35, 132, 149, 170, 171]. We also note that in our proposed distributed stochastic framework, each area of the power system can have its own measurements of wind power, while having feasibility guarantees for both local and overall multi-area power network under mild conditions. Two different simulation results using IEEE benchmark case studies are provided to illustrate the functionality of our developments. We also provide a new benchmark formulation for stochastic RS using DC model of power system, namely converted DC (CDC) to demonstrate more realistic comparisons. The main contributions of this work are twofold:

a) We provide a novel reformulation of the RS problem using the AC OPF model of power systems with wind power generation, leading to an infinite-dimensional SDP which is in general computationally intractable. We propose an approximation bridge from infinite-dimensional SDP to tractable finite-dimensional SDP using an affine policy inspired by practical aspects of the problem. We explicitly quantify the exactness of the approximation and provide a priori probabilistic feasibility guarantees to optimally schedule generating units while simultaneously determining the geographical allocation of the required reserve. We also provide another formulation of the OPF-RS problem, similarly to [149] with some modifications, and compare them in term of WCCC analysis.

b) We develop a distributed stochastic framework to carry out a multi-area RS using an the AC OPF model of power networks with wind power generation. We provide a technique to decompose a large-scale finite-dimensional SDP into small-scale problems by exploring the connections between the properties of a power network and the so-called chordal graphs. A noticeable feature of our distributed setup is that each local area of power system can have a freedom on their local accuracy of the wind power information and receive a priori probabilistic feasibility certificates to optimally schedule local generating units together with local allocation of the required reserve, and under certain mild conditions for overall power network. This is based on the recent results in [146, 147]. We then provide consensus ADMM algorithms to both OPF and OPF-RS problems in similar manner to [72, 96] with some modifications to cope with stochasticity of the formulations.

4.1.3. STRUCTURE

The layout of this chapter is as follows. Section 4.2 formulates the RS problem using the AC OPF model of power systems by including the uncertain wind power generation, whereas, in Section 4.3, we provide a computationally tractable reformulation to solve the resulting large-scale SDP in infinite-dimensional space. In Section 4.4, we extend the results to a multi-area power network and propose a distributed framework to solve

such problems. Section 4.5 provides two different simulation results using IEEE case studies, whereas Section 4.6 concludes this chapter with remarks.

4.1.4. NOTATIONS

$\mathbb{R}$, $\mathbb{R}_+$ denote the set of real and positive real numbers, $\mathbb{S}$, $\mathbb{S}_+$ denote the set of symmetric and positive-semidefinite matrices, respectively. $\mathbb{C}$ denotes the set of complex numbers. Vectors are denoted by lowercase-bold letters $\mathbf{a} \in \mathbb{R}^n$, and uppercase letters are reserved for matrices $A \in \mathbb{R}^{n\times n}$. $A^\top$, A^*, and A^{H} are used for the transpose, complex conjugate and conjugate transpose of a matrix, respectively. The notations $\underline{a}$ and $\overline{a}$ are used to denote the minimum and maximum allowed values, respectively. The cardinality of a set $\mathcal{A}$ is denoted by $|\mathcal{A}|$. The Frobenius norm of a matrix A is equal to the square root of the matrix trace of AA^{H}, i.e., $\|A\|_F = \sqrt{\mathrm{Tr}\left(AA^{\mathrm{H}}\right)}$.

4

4.2. PROBLEM FORMULATION

This section describes the AC OPF problem that aims to find a feasible operating point of the power system by minimizing the cost of the generation units over the optimization horizon including wind power generation. We apply a relaxation technique to convexify this problem, and then, extend the resulting formulation to include the stochastic RS problem.

4.2.1. AC OPF PROBLEM

Consider a power system with a set of buses $\mathcal{N}$, a set of lines $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$ and a set of generator buses $\mathcal{G} \subseteq \mathcal{N}$ such that $|\mathcal{N}| = N_b$ and $|\mathcal{G}| = N_G$. The set of wind power generation buses is denoted by $\mathcal{F} \subseteq \mathcal{N}$ such that $|\mathcal{F}| = N_w$, and $\mathcal{G} \bigcap \mathcal{F} = \emptyset$ which means there is no wind power at generator buses[1]. The set $\mathcal{T}$ forms a day-ahead hourly-based horizon of the RS optimization problem and in this work $|\mathcal{T}| = 24$. The vectors $\mathbf{p} \in \mathbb{R}^{N_b}, \mathbf{q} \in \mathbb{R}^{N_b}$ and $\mathbf{s} \in \mathbb{C}^{N_b}$ denote real, reactive and apparent power, respectively. Superscripts G, D, w are also used to indicate generated, demanded and wind power, respectively.

Define the decision variables to be the generator dispatch $\mathbf{p}_t^G, \mathbf{q}_t^G \in \mathbb{R}^{N_G}$ and the complex bus voltages $\mathbf{v}_t \in \mathbb{C}^{N_b}$ for each time step $t \in \mathcal{T}$. For the sake of brevity, a tilde denotes a set of variables over all time steps, e.g., $\tilde{\mathbf{a}} := \{\mathbf{a}_t\}_{t\in\mathcal{T}}$. Using the rectangular voltage notation: $\mathbf{x}_t := [\mathrm{Re}\,(\mathbf{v}_t)^\top \,\mathrm{Im}\,(\mathbf{v}_t)^\top]^\top \in \mathbb{R}^{2N_b}$, we follow [82, Lemma 1] to determine the data-matrices $Y_k, Y_k^*, Y_{lm}, Y_{lm}^*, M_k$. The cost function is the cost of real power generation, expressed as a second-order polynomial [113], see (4.1a) below, where the coefficient vectors $\mathbf{c}^{\mathrm{qu}}, \mathbf{c}^{\mathrm{li}} \in \mathbb{R}_+^{N_G}$ correspond to the quadratic and linear cost coefficients, respectively, and $[\mathbf{c}^{\mathrm{qu}}]$ represents a diagonal matrix with entries $\mathbf{c}^{\mathrm{qu}}$. We now formulate the AC OPF problem by taking into account the effect of wind power generation as follows:

$$\underset{\tilde{\mathbf{x}},\tilde{\mathbf{p}}^G,\tilde{\mathbf{q}}^G}{\text{minimize}} \qquad \sum_{t\in\mathcal{T}} (\mathbf{c}^{\mathrm{li}})^\top \mathbf{p}_t^G + (\mathbf{p}_t^G)^\top [\mathbf{c}^{\mathrm{qu}}] \mathbf{p}_t^G \tag{4.1a}$$

subject to:

[1] This is considered to streamline the presentation and it is not restrictive for our proposed framework.

a) Power generation limits $\forall k \in \mathcal{G}, \forall t \in \mathcal{T}$:

$$\begin{aligned} \underline{p_k^G} &\leq p_{k,t}^G \leq \overline{p_k^G}, \\ \underline{q_k^G} &\leq q_{k,t}^G \leq \overline{q_k^G}. \end{aligned} \tag{4.1b}$$

b) Power balance at every bus $\forall k \in \mathcal{N}, \forall t \in \mathcal{T}$:

$$\begin{aligned} \mathbf{x}_t^\top Y_k \mathbf{x}_t &= p_{k,t}^G - p_{k,t}^D + p_{k,t}^w, \\ \mathbf{x}_t^\top Y_k^* \mathbf{x}_t &= q_{k,t}^G - q_{k,t}^D, \end{aligned} \tag{4.1c}$$

where $\mathbf{p}_t^w := \{p_{k,t}^w\}_{k\in\mathcal{F}}$ is the wind power, and $\mathbf{s}_t^D := \{s_{k,t}^D\}_{k\in\mathcal{N}}$ is the demanded power such that $s_{k,t}^D = p_{k,t}^D + j q_{k,t}^D$.

c) Bus voltage limits $\forall k \in \mathcal{N}, \forall t \in \mathcal{T}$:

$$|\underline{v_k}|^2 \leq \mathbf{x}_t^\top M_k \mathbf{x}_t \leq |\overline{v_k}|^2. \tag{4.1d}$$

d) Lineflow limits $\forall (l,m) \in \mathcal{L}, \forall t \in \mathcal{T}$:

$$\left(\mathbf{x}_t^\top Y_{lm} \mathbf{x}_t\right)^2 + \left(\mathbf{x}_t^\top Y_{lm}^* \mathbf{x}_t\right)^2 \leq |\overline{s_{lm}}|^2,$$

which can be reformulated using the Schur-complement [23] to form a linear matrix inequality constraint, such that the fourth-order dependence on the voltage vector is reduced to quadratic terms:

$$\begin{bmatrix} -|\overline{s_{lm}}|^2 & \mathbf{x}_t^\top Y_{lm} \mathbf{x}_t & \mathbf{x}_t^\top Y_{lm}^* \mathbf{x}_t \\ \mathbf{x}_t^\top Y_{lm} \mathbf{x}_t & -1 & 0 \\ \mathbf{x}_t^\top Y_{lm}^* \mathbf{x}_t & 0 & -1 \end{bmatrix} \preceq 0. \tag{4.1e}$$

e) Reference bus constraint $\forall t \in \mathcal{T}$:

$$\mathbf{x}_t^\top E_{\text{ref}} \mathbf{x}_t = 0, \tag{4.1f}$$

where $E_{\text{ref}} \in \mathbb{R}^{2N_b \times 2N_b}$ is a diagonal matrix from the standard basis vector $e_{N_b + i_{\text{ref}}} \in \mathbb{R}^{2N_b}$, and i_{ref} denotes the reference bus.

Remark 6 *The power balance constraints (4.1c) can be used to reformulate the real and reactive generator dispatch in terms of the voltage vector as follows $\forall k \in \mathcal{N}, \forall t \in \mathcal{T}$:*

$$p_{k,t}^G = \mathbf{x}_t^\top Y_k \mathbf{x}_t + p_{k,t}^D - p_{k,t}^w, \tag{4.2a}$$

$$q_{k,t}^G = \mathbf{x}_t^\top Y_k^* \mathbf{x}_t + q_{k,t}^D. \tag{4.2b}$$

Using this reformulation, one can substitute for $p_{k,t}^G$ and $q_{k,t}^G$ in (4.1b) to have $\forall k \in \mathcal{N}, \forall t \in \mathcal{T}$:

$$\underline{p_k^G} \leq \mathbf{x}_t^\top Y_k \mathbf{x}_t + p_{k,t}^D - p_{k,t}^w \leq \overline{p_k^G}, \tag{4.3a}$$

$$\underline{q_k^G} \leq \mathbf{x}_t^\top Y_k^* \mathbf{x}_t + q_{k,t}^D \leq \overline{q_k^G}, \tag{4.3b}$$

where the lower and upper generation limits have been also extended to $\mathcal{N}$ using $\underline{p_k^G} = \overline{p_k^G} = 0, \forall k \in \mathcal{N} \setminus \mathcal{G}$.

Remark 7 *Following Remark 6, one can reformulate the cost function (4.1a) using the voltage vector $\mathbf{x}_t$:*

$$f_G^x(\mathbf{x}_t, \mathbf{p}_t^w, \mathbf{p}_t^D) := \sum_{k \in \mathcal{G}} c_k^{li}(\mathbf{x}_t^\top Y_k \mathbf{x}_t + p_{k,t}^D - p_{k,t}^w) + c_k^{qu}\left((\mathbf{x}_t^\top Y_k \mathbf{x}_t + p_{k,t}^D - p_{k,t}^w)\right)^2. \tag{4.4}$$

It is important to note that this function is of order four with respect to $\mathbf{x}_t$, but it can be also made quadratic[2]. To streamline the presentation, these steps are skipped.

Using $\tilde{\mathbf{x}}$, we reformulate the problem (4.1) in a more compact form:

$$\text{OPF}(\tilde{\mathbf{p}}^w) : \begin{cases} \underset{\tilde{\mathbf{x}}}{\text{minimize}} & \sum_{t \in \mathcal{T}} f_G^x(\mathbf{x}_t, \mathbf{p}_t^w, \mathbf{p}_t^D) \\ \text{subject to} & (4.1d), (4.1e), (4.1f), (4.3) \end{cases}.$$

OPF($\tilde{\mathbf{p}}^w$) is a quadratically constrained quadratic program (QCQP) in $\tilde{\mathbf{x}}$, and a non-convex optimization problem, since the data matrices $Y_k, Y_k^*, Y_{lm}, Y_{lm}^*$ are indefinite [82], which is in fact an NP-hard problem [83] and hard to solve.

4.2.2. CONVEXIFIED AC OPF PROBLEM

We can reformulate OPF($\tilde{\mathbf{p}}^w$) as an equivalent problem in a matrix variable $W_t := \mathbf{x}_t \mathbf{x}_t^\top \in \mathbb{S}^{2N_b}$ using a semi-definite reformulation (SDR) technique, see, e.g., [82, 91] and the references therein. W_t represents the operating state of the network at time step t, and therefore, it is called the state matrix. We define $\mathcal{W} \subset \mathbb{S}^{2N_b}$ as the set of feasible operating states constraints, such that $W_t \in \mathcal{W}$, using the following characteristics:

$$\begin{aligned} \mathcal{W}(\mathbf{p}^w, \mathbf{s}^D) := \Bigg\{ W \in \mathbb{S}^{2N_b} \;\Bigg|\; & \text{Tr}(E_{\text{ref}} W) = 0, \\ & \underline{p_k^G} \le \text{Tr}(Y_k W) + p_k^D - p_k^w \le \overline{p_k^G}, \quad \forall k \in \mathcal{N}, \\ & \underline{q_k^G} \le \text{Tr}\left(Y_k^* W\right) + q_k^D \le \overline{q_k^G}, \quad \forall k \in \mathcal{N}, \\ & |\underline{v_k}|^2 \le \text{Tr}(M_k W) \le |\overline{v_k}|^2, \quad \forall k \in \mathcal{N}, \forall (l,m) \in \mathcal{L}, \\ & \begin{bmatrix} -|\overline{s_{lm}}|^2 & \text{Tr}(Y_{lm} W) & \text{Tr}\left(Y_{lm}^* W\right) \\ \text{Tr}(Y_{lm} W) & -1 & 0 \\ \text{Tr}\left(Y_{lm}^* W\right) & 0 & -1 \end{bmatrix} \preceq 0 \quad \Bigg\}. \end{aligned} \tag{4.5}$$

[2] The cost function can be made linear with the use of the epigraph notation (see also [23, Section 4.1.3]). The resulting inequality constraint can be converted to an LMI using the Schur complement (see also [23, Section A.5.5]), which yields a quadratic function of $\mathbf{x}$.

Consider now the following formulation as an equivalent optimization problem to OPF($\tilde{\mathbf{p}}^w$):

$$\begin{aligned}
\underset{\tilde{W}}{\text{minimize}} \quad & \sum_{t\in\mathcal{T}} f_G(W_t, \mathbf{p}_t^w, \mathbf{p}_t^D) && && \text{(4.6a)}\\
\text{subject to} \quad & W_t \in \mathcal{W}(\mathbf{p}_t^w, \mathbf{s}_t^D), && \forall t \in \mathcal{T}, && \text{(4.6b)}\\
& W_t \succeq 0, && \forall t \in \mathcal{T}, && \text{(4.6c)}\\
& \operatorname{rank}(W_t) = 1, && \forall t \in \mathcal{T}, && \text{(4.6d)}
\end{aligned}$$

where f_G is defined by (4.4) and $W_t = \mathbf{x}_t\mathbf{x}_t^\top$. Constraints (4.6c) and (4.6d) are introduced to guarantee the exactness of SDR, and consequently, OPF($\tilde{\mathbf{p}}^w$) and (4.6) are equivalent.

The optimization problem (4.6) is non-convex, due to the presence of rank-one constraint (4.6d). Removing (4.6d) relaxes the problem to an SDP. It has been shown in [82] and later in [97] that the rank-one constraint can be dropped without affecting the solution for most power networks. In [149, Proposition 1], the authors showed that when the convex relaxation of the AC OPF problem has solutions with rank at most two, then, forcing any arbitrary selected entry of the diagonal of the matrix W_t to be zero results in a rank-one optimal solution. This condition is practically motivated since the voltage angle of one of the buses (the reference bus) is often fixed at zero in practice. We denote by C-OPF($\tilde{\mathbf{p}}^w$) the convexified version of OPF($\tilde{\mathbf{p}}^w$) problem in $\tilde{W}$, i.e., Problem (4.6) with the rank-one constraint (4.6d) removed.

In the following proposition, we restate the results in [149, Proposition 1] to highlight that our developments, in the following parts, buildup on this result and therefore the obtained solutions for the state variables using our proposed approach is the rank-one feasible solution.

Proposition 2 *If the C-OPF($\tilde{\mathbf{p}}^w$) problem has solutions with rank at most two, then, forcing any arbitrary selected entry of the diagonal of matrix W_t to be zero results in a rank-one solution W_t^*. Moreover, the corresponding value of the objective function of the proposed optimization is identical to that of the original OPF($\tilde{\mathbf{p}}^w$) problem.*

It is important to note that the optimality of the obtained solutions using our proposed approach is not in the scope of this chapter and it is subject of our ongoing research work. Instead, in this chapter we focus on feasibility and derive probabilistic guarantees for constraint fulfillment in a distributed setting.

4.2.3. CONVEXIFIED AC OPF RESERVE SCHEDULING PROBLEM

Consider a power network where a TSO aims to solve a day-ahead AC OPF problem to determine an optimal generator dispatch for the forecasted wind power trajectory such that: 1) the equipment in the power system remains safe and 2) the power balance (4.1c) in the power network is achieved. As a novel feature in our proposed formulation C-OPF($\tilde{\mathbf{p}}^w$) has a dependency on $\tilde{\mathbf{p}}^w$, and thus, it solves the AC OPF problem while taking into account the actual wind power trajectory $\tilde{\mathbf{p}}^w$. We here define the difference between a generic actual wind power realization and the forecasted wind power as the mismatch wind power at each time step, e.g. $\mathbf{p}_t^m = \mathbf{p}_t^w - \mathbf{p}_t^{w,f}$. Due to the fact that $\tilde{\mathbf{p}}^m := \{\mathbf{p}_t^m\}_{t\in\mathcal{T}}$ is a random variable, the following technical assumption is necessary in order to proceed to the next steps.

Assumption 9 *The variable $\tilde{\mathbf{p}}^m$ is defined on some probability space $(\mathcal{P}, \mathfrak{B}(\mathcal{P}), \mathbb{P})$, where $\mathfrak{B}(\cdot)$ denotes a Borel σ-algebra, and $\mathbb{P}$ is a probability measure defined over $\mathcal{P}$.*

As a top priority of TSOs is to ensure the feasibility and validity of the power network, we formulate the following problem:

$$\begin{aligned}
\underset{\tilde{W}^f, \tilde{W}}{\text{minimize}} \quad & \sum_{t \in \mathcal{T}} f_G(W_t^f, \mathbf{p}_t^{w,f}, \mathbf{p}_t^D) && && \text{(4.7a)}\\
\text{subject to} \quad & W_t^f \in \mathcal{W}(\mathbf{p}_t^{w,f}, \mathbf{s}_t^D), && \forall t \in \mathcal{T}, && \text{(4.7b)}\\
& W_t \in \mathcal{W}(\mathbf{p}_t^{w}, \mathbf{s}_t^D),\ \forall \tilde{\mathbf{p}}^m \in \mathcal{P}, && \forall t \in \mathcal{T}, && \text{(4.7c)}\\
& W_t^f \succeq 0,\ W_t \succeq 0 && \forall t \in \mathcal{T}, && \text{(4.7d)}
\end{aligned}$$

where $\tilde{\mathbf{p}}^{w,f}$ denotes the forecasted wind power trajectory, $\tilde{\mathbf{p}}^w$ is a generic wind power trajectory, $\tilde{W}^f$ is related to the state of the network in the case of forecasted wind power, and $\tilde{W}$ is a generic network state for a generic wind power trajectory. Constraints (4.7b) and (4.7c) ensure feasibility for every network state, while constraints (4.7d) enforce PSDness of the forecasted network state and the generic network state for all possible wind power trajectories.

As a second task of the TSO, the power balance of the power network has to be achieved to ensure demand satisfaction even in the presence of uncertain wind power generation. To address this issue, the TSO employs *reserve power scheduling*, using the fact that a mismatch between actual wind power and forecasted wind power can be mitigated by the controllable generators [118]. We can thus express

$$r_{k,t} := p_{k,t}^G - p_{k,t}^{G,f}, \tag{4.8}$$

where $\mathbf{r}_t = \{r_{k,t}\}_{k \in \mathcal{G}} \in \mathbb{R}^{N_G}$ denotes the amount of reserve requirement in the power network. Following Remark 6, we have:

$$\begin{aligned}
p_{k,t}^G &= \mathrm{Tr}\left(Y_k W_t\right) + p_{k,t}^D - p_{k,t}^w,\\
p_{k,t}^{G,f} &= \mathrm{Tr}\left(Y_k W_t^f\right) + p_{k,t}^D - p_{k,t}^{w,f},
\end{aligned}$$

and one can substitute these equations in (4.8) to obtain the reserve power in terms of the network states W_t and W_t^f as follows:

$$\begin{aligned}
r_{k,t} &:= \mathrm{Tr}\left(Y_k\left(W_t - W_t^f\right)\right) - (p_{k,t}^w - p_{k,t}^{w,f})\\
&= \mathrm{Tr}\left(Y_k\left(W_t - W_t^f\right)\right) - p_{k,t}^m,\\
&= \mathrm{Tr}\left(Y_k\left(W_t - W_t^f\right)\right),
\end{aligned} \tag{4.9}$$

where the term $p_{k,t}^m$ is dropped since it is assumed that there is no wind power at generator buses. The elements of $\mathbf{r}_t = \{r_{k,t}\}_{k \in \mathcal{N}_\mathcal{G}}$ can be positive and negative to represent the up- and downspinning reserve powers, respectively, such that they are deployed for a power deficit and surplus to bring balance to the network and satisfy the demand [113].

The mismatches between the demanded power and uncertain generation (wind) induce frequency deviations and activate the primary and secondary frequency controller (Automatic Generation Regulation (AGR)). The existence of an ideal primary frequency control functionality compensating for any fast time scale power deviation is assumed. Here, we focus only on the steady state behavior of the AGR reserve mechanism [171], we also define two vectors $\mathbf{d}_t^{\text{us}}, \mathbf{d}_t^{\text{ds}} \in \mathbb{R}^{N_G}$ to distribute the amount of up- or downspinning reserve powers among the available generators for each hour $t \in \mathcal{T}$. To obtain the optimal control strategies for AGR, we consider the following equality constraint, $\forall \mathbf{p}_t^m \in \mathcal{P}$, $\forall k \in \mathcal{G}$ and $\forall t \in \mathcal{T}$:

$$\begin{aligned} r_{k,t} &= \operatorname{Tr}\left(Y_k\left(W_t - W_t^f\right)\right) \\ &= -d_{k,t}^{\text{us}} \min\left(\mathbf{p}_t^m, 0\right) - d_{k,t}^{\text{ds}} \max\left(\mathbf{p}_t^m, 0\right). \end{aligned} \tag{4.10}$$

4

In order to always negate the wind power mismatch using the reserve power and bring balance to the power network, we enforce the sum of the distribution vectors to be equal to one using the following constraint $\forall t \in \mathcal{T}$:

$$\sum_{k \in \mathcal{G}} d_{k,t}^{\text{us}} = 1\,, \ \sum_{k \in \mathcal{G}} d_{k,t}^{\text{ds}} = 1\,. \tag{4.11}$$

We also distinguish between up- and downspinning reserve powers using $\mathbf{r}_t^{\text{us}}, \mathbf{r}_t^{\text{ds}} \in \mathbb{R}^{N_G}$ such that $\forall t \in \mathcal{T}$:

$$-\mathbf{r}_t^{\text{ds}} \leq \mathbf{r}_t \leq \mathbf{r}_t^{\text{us}}\,, \tag{4.12a}$$

$$0 \leq \mathbf{r}_t^{\text{us}}\,, \ 0 \leq \mathbf{r}_t^{\text{ds}}\,, \tag{4.12b}$$

where $\mathbf{r}_t^{\text{us}}$ and $\mathbf{r}_t^{\text{ds}}$ are enforced to be positive as they only appear in the reserve cost function. We now consider corresponding linear up- and downspinning cost coefficients $\mathbf{c}^{\text{us}}, \mathbf{c}^{\text{ds}} \in \mathbb{R}_+^{N_G}$ yielding the total reserve cost:

$$f_R(\mathbf{r}_t^{\text{us}}, \mathbf{r}_t^{\text{ds}}) := (c^{\text{us}})^\top \mathbf{r}_t^{\text{us}} + (c^{\text{ds}})^\top \mathbf{r}_t^{\text{ds}}.$$

Using $\Xi := \left\{\tilde{W}^f, \tilde{W}, \tilde{\mathbf{d}}^{\text{us}}, \tilde{\mathbf{d}}^{\text{ds}}, \tilde{\mathbf{r}}^{\text{us}}, \tilde{\mathbf{r}}^{\text{ds}}\right\}$ as the set of decision variables, and combining our previous discussions with the optimization problem (4.7), we are now in the position to formulate C-OPF($\tilde{\mathbf{p}}^w$) with RS problem as follows:

$$\text{C-OPF-RS}: \begin{cases} \min\limits_{\Xi} & \sum_{t \in \mathcal{T}} \left(f_G(W_t^f, \mathbf{p}_t^{w,f}) + f_R(\mathbf{r}_t^{\text{us}}, \mathbf{r}_t^{\text{ds}})\right) \\ \text{s.t.} & (4.7\text{b}), (4.7\text{c}), (4.7\text{d}), (4.10), (4.11), (4.12) \end{cases}.$$

Notice that one needs to substitute $\mathbf{r}_t$ in (4.12a) with (4.9).

C-OPF-RS is an uncertain infinite-dimensional SDP in the sense that the dimension of the decision spaces as well as the number of constraints are both infinite, due to the unknown and unbounded set $\mathcal{P}$. It is therefore computationally intractable and in general difficult to solve. In the following section, we develop an approximation technique to provide a tractable finite-dimensional SDP optimization problem.

4.3. PROPOSED TRACTABLE REFORMULATION

In this section, we propose two different reformulations for C-OPF-RS problem, and study conditions under which one can provide a finite approximation. We approach this goal in two different ways. We first develop a novel affine policy to translate the problem into semi-infinite-dimensional space, and then provide a finite approximation to the semi-infinite SDP using randomization technique. As a second approach, we directly employ a randomized technique to provide a finite approximation for the C-OPF-RS problem. The solution of each of these methods comes with a priori probabilistic performance certificates. We finally provide the WCCC analysis for both proposed methods.

4.3.1. INFINITE TO SEMI-INFINITE PROGRAM: AFFINE POLICY

Consider the proposed equality constraint (4.10), $\forall \mathbf{p}_t^m \in \mathcal{P}$, $\forall k \in \mathcal{G}$ and $\forall t \in \mathcal{T}$:

$$\mathrm{Tr}\left(Y_k\left(W_t - W_t^f\right)\right) = -d_{k,t}^{\mathrm{us}} \min\left(\mathbf{p}_t^m, 0\right) - d_{k,t}^{\mathrm{ds}} \max\left(\mathbf{p}_t^m, 0\right).$$

Since the uncertain set $\mathcal{P}$ is unknown and unbounded, it is straightforward to see that the dimension of the decision spaces for the variable W_t as well as the number of constraints are both infinite. To overcome this difficulty, we propose an *affine policy* to restrict the decision space using a *conic combination* of the generic network state W_t in C-OPF-RS. Observe that in (4.10), the network state variable W_t for every realization of the uncertainty in each time step, can be represented as a linear combination of the network forecasted state W_t^f and the up- and downspinning reserve distributions. This is also practically inspired by the AGR mechanism.

Motivated by this observation, we propose a novel parametrization of the generic network state that encodes this restriction implicitly. Define $\hat{W}_t(\mathbf{p}_t^m) \in \mathbb{S}^{2N_b}$ to be the parametrized generic network state $\tilde{\mathbf{p}}^m \in \mathcal{P}, \forall t \in \mathcal{T}$ in the following form:

$$\hat{W}_t(\mathbf{p}_t^m) := W_t^f + W_t^{\mathrm{us}} \max(-\mathbf{p}_t^m, 0) + W_t^{\mathrm{ds}} \max(\mathbf{p}_t^m, 0), \tag{4.13}$$

where $W_t^{\mathrm{us}}, W_t^{\mathrm{ds}} \in \mathbb{S}^{2N_b}$ are the coefficient matrix variables for every $t \in \mathcal{T}$. The parametrization of the network state as a conic combination of PSD matrices is, to the best of our knowledge, a novel way to make the problem more manageable.

Using the proposed conic parametrization, the generic network state is decomposed into a deterministic component and two components that scale with the positive or negative uncertainty. It is worth mentioning that both $\max(\mathbf{p}_t^m, 0)$ and $\max(-\mathbf{p}_t^m, 0)$ are always non-negative and never simultaneously non-zero such that the change in network state is determined by either W_t^{ds} or W_t^{us} in case of either a wind power surplus or deficit, respectively. The following theorem can be considered as the main result of this section to approximate C-OPF-RS in semi-infinite-dimensional space.

Theorem 6 *Given the proposed affine policy in* (4.13)*, an exact approximation of the C-OPF-RS problem can be formulated using the following optimization problem by defining*

$\hat{\Xi} := \{\tilde{W}^f, \tilde{W}^{us}, \tilde{W}^{ds}, \tilde{\mathbf{r}}^{us}, \tilde{\mathbf{r}}^{ds}\}$ *as the new set of decision variables:*

$$P\text{-}OPF\text{-}RS: \begin{cases} \min\limits_{\hat{\Xi}} & \sum_{t\in\mathcal{T}} \left(f_G(W_t^f, \mathbf{p}_t^{w,f}) + f_R(\mathbf{r}_t^{us}, \mathbf{r}_t^{ds}) \right) \\ s.t. & W_t^f \in \mathcal{W}(\mathbf{p}_t^{w,f}, \mathbf{s}_t^D) \\ & \hat{W}_t(\mathbf{p}_t^m) \in \mathcal{W}(\mathbf{p}_t^{w,f} + \mathbf{p}_t^m, \mathbf{s}_t^D) \\ & W_t^f \succeq 0, \quad W_t^{us} \succeq 0, \quad W_t^{ds} \succeq 0 \\ & -\mathbf{r}_t^{ds} \le \mathbf{r}_t \le \mathbf{r}_t^{us}, 0 \le \mathbf{r}_t^{us}, 0 \le \mathbf{r}_t^{ds} \\ & \sum_{k\in\mathcal{G}} \mathrm{Tr}\left(Y_k W_t^{us}\right) = 1 \\ & \sum_{k\in\mathcal{G}} \mathrm{Tr}\left(Y_k W_t^{ds}\right) = -1 \\ & \forall \tilde{\mathbf{p}}^m \in \mathcal{P}, \forall t \in \mathcal{T} \end{cases},$$

where $\mathbf{r}_t = \{r_{k,t}\}_{k\in\mathcal{G}}$ *should be replaced with*

$$r_{k,t} = -\mathrm{Tr}\left(Y_k W_t^{us}\right) \min(\mathbf{p}_t^m, 0) + \mathrm{Tr}\left(Y_k W_t^{ds}\right) \max(\mathbf{p}_t^m, 0).$$

Proof. The proof is provided in Appendix C. □

The interpretation of Theorem 6 is as follows. Due to the fact that the formulation in C-OPF-RS follows a linear decision rule concept as a consequence of the AGR mechanism, C-OPF-RS and P-OPF-RS are equivalent formulations using the proposed affine policy (4.13). It is however very important to realize that P-OPF-RS is always a restricted (approximation) version of a general formulation of C-OPF-RS, where, e.g., the consideration of AGR mechanism as in (4.10) has been included in a more general manner. In such a general setup formulation, one can quantify the gap between the objective values as a function of the difference between optimizers of the both problems (for more details on such a bound in a general problem formulation see [108, Theorem 3.3]).

Remark 8 *The distribution vectors of reserve power,* $\tilde{\mathbf{d}}^{us}, \tilde{\mathbf{d}}^{ds}$ *are encoded in* $\tilde{W}^{us}, \tilde{W}^{ds}$*, respectively, through the equality constraints*[3]*, and therefore, these can be determined a-posteriori using the following relations* $\forall k \in \mathcal{G}, \forall t \in \mathcal{T}$*:*

$$d_{k,t}^{us} = \mathrm{Tr}\left(Y_k W_t^{us}\right), \; d_{k,t}^{ds} = -\mathrm{Tr}\left(Y_k W_t^{ds}\right).$$

One can also define $\tilde{\mathbf{d}}^{us}, \tilde{\mathbf{d}}^{ds}$ *as two additional decision vector and replace the two equality constraints in P-OPF-RS with the following constraints:*

$$d_{k,t}^{us} = \mathrm{Tr}\left(Y_k W_t^{us}\right), \; d_{k,t}^{ds} = \mathrm{Tr}\left(Y_k W_t^{ds}\right), \tag{4.14}$$

$$\sum_{k\in\mathcal{G}} d_{k,t}^{us} = 1, \; \sum_{k\in\mathcal{G}} d_{k,t}^{ds} = -1. \tag{4.15}$$

[3] These equality constraints cause numerical issues in the solver. Using a common practice in optimization (see, e.g., [1]), we implement these constraints by introducing a slack variable $u \in \mathbb{R}_+^2$ and rewriting each equality constraint $f(a) = b$ in the form of $b - u_1 \le f(x) \le b + u_2$. By adding L1-norm of slack variables as a penalty function into the objective function, u is minimized, essentially pushing $f(a)$ to be equal to b. It is important to highlight that using such a practical way to implement these equality constraints does not lead to any kind of relaxations, since they have been evaluated afterwords by checking the optimized value of the slack variables to confirm that they are sufficiently small.

The P-OPF-RS problem is a semi-infinite-dimensional SDP due to the uncountable number of constraints corresponding to the uncertainty set $\mathcal{P}$, which is indeed an unbounded and unknown uncertainty set. In the following section, we adopt a randomization technique to approximate the P-OPF-RS problem and obtain a finite-dimensional SDP. We will also provide a performance guarantee for the feasibility of constraints in a probabilistic sense with a high level of confidence.

4.3.2. SEMI-INFINITE TO FINITE PROGRAM: RANDOMIZED APPROACH

We develop a finite approximation to the semi-infinite SDP in P-OPF-RS problem that is in general hard to solve and known to be computationally intractable [10]. To overcome this difficulty, we employ the recent developments in the area of randomized optimization, leading to a priori probabilistic guarantees for the feasibility of the obtained solutions.

Recall the randomization technique by assuming to have a 'sufficient number' of independent and identically distributed (i.i.d.) samples of the string of wind power mismatch realizations $\tilde{\mathbf{p}}^m = \{\mathbf{p}_t^m\}_{t\in\mathcal{T}}$, which can be obtained either empirically or by a random scenario generator. Notice that $\mathbf{p}_t^m$ at each sampling time t is not necessarily i.i.d., and in particular, it may have time-varying distributions and/or be correlated in time. We denote $\mathcal{S} := \{\tilde{\mathbf{p}}^{m,1}, \cdots, \tilde{\mathbf{p}}^{m,N_s}\} \in \mathcal{P}^{N_s}$ as a set of given finite multi-extraction samples (scenarios) from $\mathcal{P}$. Consider now a tractable version of P-OPF-RS using the following finite-dimensional SDP optimization problem:

$$\text{SP-OPF-RS}: \begin{cases} \underset{\hat{\Xi}\in\hat{\mathcal{X}}}{\text{minimize}} & f(\hat{\Xi}) \\ \text{subject to} & \hat{g}(\hat{\Xi}, \tilde{\mathbf{p}}^{m,i}) \leq 0\,,\ \forall \tilde{\mathbf{p}}^{m,i} \in \mathcal{S} \end{cases},$$

where $\hat{g}(\cdot)$ is the uncertain constraint function of P-OPF-RS, and all other constraints for P-OPF-RS are used to construct $\hat{\mathcal{X}}$, a deterministic feasible set for the P-OPF-RS problem.

The key features of the proposed tractable optimization problem (SP-OPF-RS) are as follows:

- there is no need to know the probability measure $\mathbb{P}$ explicitly, only the capability of obtaining random scenarios is enough.
- formal results to quantify the robustness of the obtained approximations are available. In particular the results follow the so-called scenario approach [30], which allow to bound a-priori the violation probability of the obtained solution via SP-OPF-RS.

In the following theorem, we restate the explicit theoretical bound of [30, Theorem 1] which measures the finite scenario behavior of SCP.

Theorem 7 *Let $\varepsilon \in (0,1)$, $\beta \in (0,1)$ and $N_s \geq \mathsf{N}(\varepsilon,\beta,d)$, such that*

$$\mathsf{N}(\varepsilon,\beta,d) := \min\left\{ N \in \mathbb{N} \,\middle|\, \sum_{i=0}^{d-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \leq \beta \right\}.$$

where d is the number of decision variables in P-OPF-RS problem. If the optimizer ($\hat{\Xi}^*$) of SP-OPF-RS is applied to P-OPF-RS, then the original uncertain constraint function $\hat{g}(\cdot)$ in P-OPF-RS is satisfied with probability $1-\varepsilon$ and with confidence level higher than $1-\beta$.

It was shown in [30] that the above bound is tight. The interpretation of Theorem 7 is as follows: when applying $\hat{\Xi}^*$ in P-OPF-RS problem, the probability of constraint violation remains below ε with confidence $1-\beta$:

$$\mathbb{P}^{N_s}\left[\mathcal{S}\in\mathcal{P}^{N_s} : \mathrm{Vio}(\hat{\Xi}^*)\le\varepsilon\right]\ge 1-\beta,$$

with

$$\mathrm{Vio}(\hat{\Xi}^*) := \mathbb{P}\left[\tilde{\mathbf{p}}^m\in\mathcal{P} : \hat{g}(\hat{\Xi}^*,\tilde{\mathbf{p}}^m)>0\right].$$

4

Remark 9 *Following the discussion in [35] about the practical controllable variables in power systems, one can determine the dimension of the control variable for the C-OPF-RS as well as the P-OPF-RS problem using the real generator power, voltage magnitudes and the up- and downspinning vectors for every time step. This leads to d in Theorem 7 to be given by $4TN_G$ for both problem formulations in C-OPF-RS and P-OPF-RS.*

Remark 10 *One can obtain an explicit expression for the desired number of scenarios N_s as in [3], where it is shown that given $\varepsilon,\beta\in(0,1)$ then $N_s\ge\frac{e}{e-1}\frac{1}{\varepsilon}\left(d+\ln\frac{1}{\beta}\right)$. It is important to mention that N_s is used to construct the sets of scenarios $\mathcal{S}$.*

Using the randomization technique explained above, we next provide an approximation technique to directly translate the infinite-dimensional SDP of the C-OPF-RS problem into a finite-dimensional SDP.

4.3.3. INFINITE TO FINITE PROGRAM: DIRECT APPROACH

Rather than using the affine policy (4.13) proposed in (Section 4.3.1) to obtain P-OPF-RS, and then, approximating $\mathcal{P}$ as in (Section 4.3.2) to obtain SP-OPF-RS, one can also apply directly to the C-OPF-RS problem the randomization technique explained in (Section 4.3.2) to formulate a finite-dimensional SDP. A slightly different idea has been also considered in [149] using an ad-hoc manner. To this end, we recall $\mathcal{S}:=\{\tilde{\mathbf{p}}^{m,1},\cdots,\tilde{\mathbf{p}}^{m,N_s}\}\in\mathcal{P}^{N_s}$ as a set of given finite scenarios and reformulate an approximated version of C-OPF-RS using the following optimization problem:

$$\text{SC-OPF-RS}:\begin{cases}\underset{\bar{\Xi}\in\mathcal{X}}{\text{minimize}} & f(\bar{\Xi})\\ \text{subject to} & g(\bar{\Xi},\tilde{\mathbf{p}}^{m,i})\le 0\,,\ \forall\tilde{\mathbf{p}}^{m,i}\in\mathcal{S}\end{cases},$$

where $\bar{\Xi}$ is the set of decision variables of SC-OPF-RS. It is important to note that $\bar{\Xi}$ is not the same set of optimizers (Ξ) of C-OPF-RS problem. The difference arises due to the equality constraints in (4.10) that lead to a new network state variable $\tilde{W}$ for each $\tilde{\mathbf{p}}^{m,i}\in\mathcal{S}$, and therefore,

$$\bar{\Xi}:=\left\{\tilde{W}^f,\tilde{W}^1,\tilde{W}^2,\cdots,\tilde{W}^{N_s},\tilde{\mathbf{d}}^{\mathrm{us}},\tilde{\mathbf{d}}^{\mathrm{ds}},\tilde{\mathbf{r}}^{\mathrm{us}},\tilde{\mathbf{r}}^{\mathrm{ds}}\right\}.$$

Table 4.1: WCCC Analysis.

SC-OPF-RS	$\mathcal{O}\big(T^{9.5} N_b^{6.5} N_G^{6.5} \log(1/\alpha)\big)$
SP-OPF-RS	$\mathcal{O}\big(T^{6} N_b^{4} N_G^{4} \log(1/\alpha)\big)$

The SC-OPF-RS problem is a finite-dimensional large-scale SDP, due to the fact that instead of a forecast and a generic network state, the set of decision variables now has extra N_s number of matrix variables (network states for all scenarios) in every time step, and each of the state matrices is also subject to PSD constraints. The minimum number of scenarios N_s to guarantee reasonable violation is typically quite large (see Remark 10). The resulting problem has therefore a high number of computationally expensive PSD constraints such that it is indeed computationally demanding. It is worth mentioning that the SP-OPF-RS problem contains only a forecast and a generic up- and down spinning network state subject to PSD constraints due to the proposed affine policy, and therefore, is computationally completely tractable.

To illustrate the advantages of SP-OPF-RS compared with SC-OPF-RS, we now analyze the WCCC for both problem formulations. If the resulting SDP is solved using an interior point method, the analysis in [24] states that the number of iterations needed to approximate the optimal solution with a given accuracy α is $\mathcal{O}\big(n^{1/2}\log(1/\alpha)\big)$, where n is the dimension of the decision variable. In [112], the number of floating point operations needed per iterations is characterized[4] using $\mathcal{O}\big(m^2n^2 + mn^3 + m^3)\big)$, where m is the number of constraints. The total WCCC to find an α-solution (i.e. the objective value of the α-solution is at most $\alpha > 0$ above the optimum) is then given by the worst-case number of iterations multiplied by the computational cost per one iteration:

$$\mathcal{O}\big((m^2n^2 + mn^3 + m^3)n^{1/2}\log(1/\alpha)\big).$$

In Table 4.1 the problem dimensions and the resulting WCCCs are given, where N_s is assumed to be $\mathcal{O}(TN_G)$ for both formulations. It can be seen that the SP-OPF-RS problem has much lower order of WCCC. This increase in WCCC can be explained by the N_s repetitions of the state matrices. It is important to note that in the SP-OPF-RS problem formulation, we only have three PSD constraints per time step, whereas in the SC-OPF-RS problem formulation, there are $N_s + 1$ number of PSD constraints per time step.

4.4. DISTRIBUTED FRAMEWORK

This section provides a distributed framework to solve stochastic RS using the AC OPF model of power systems. We first describe the procedures that yield a decomposition of power systems into the multi-area power networks in terms of the objective function together with the feasible sets. Next, a distributed consensus algorithm to solve multi-

[4] This concerns the WCCC for a problem with dense data-matrices, i.e. no assumption on the structure of the problem. Modern solvers, such as MOSEK [6], can achieve lower complexities by using sparsity in the problem structure. Whether or not this is the case for our problem, is outside the scope of this work. For all IPM solvers, computational complexity scales logarithmically with $1/\alpha$, and polynomially with m and n, regardless of whether they make use of the problem structure or not.

area AC OPF problem is introduced. We finally present an equivalent multi-area version of SP-OPF-RS (MASP-OPF-RS) together with a distributed consensus setup using the ADMM algorithm.

4.4.1. MULTI-AREA DECOMPOSITION

We explain the proposed decomposition procedures using the C-OPF($\tilde{\mathbf{p}}^w$) problem in (Section 4.2.2) for each time step $t \in \mathcal{T}$ to simplify the notations:

$$\underset{W \in \mathbb{S}^{2N_b}}{\text{minimize}} \quad f_G(W, \mathbf{p}^w, \mathbf{p}^D) \tag{4.16a}$$

$$\text{subject to} \quad W \in \mathcal{W}(\mathbf{p}^w, \mathbf{s}^D), \tag{4.16b}$$

$$W \succeq 0. \tag{4.16c}$$

Problem (4.16) cannot be directly decomposed, since the objective function, feasible set, and the PSD constraint are defined for a single matrix variable W that comprises for the whole power system. To approach the multi-area decomposition goal, we need to describe the division of a power network into sub-networks which we refer to as control areas.

Divide a power network into several control areas, and collect the indices in $\mathcal{A} := \{1, \ldots, N_a\}$. Define $\mathcal{N}_a \subset \mathcal{N}$ to be the subset of buses corresponding to a control area $a \in \mathcal{A}$. Every bus belongs to exactly one control area, such that $\mathcal{N}_a \cap \mathcal{N}_b = \emptyset$ for all $a, b \in \mathcal{A}, a \neq b$, and $\bigcup_{a \in \mathcal{A}} \mathcal{N}_a = \mathcal{N}$. Consider now $\mathcal{B}_a$ as the set of neighboring control area indices that are connected to area a, such that for all $a \in \mathcal{A}$:

$$\mathcal{B}_a := \left\{ b \in \mathcal{A} \mid \exists i \in \mathcal{N}_a, \exists j \in \mathcal{N}_b, (i,j) \in \mathcal{L}, \right\}.$$

The lines that interconnect the areas are called tie-lines. These lines are collected into a tie-line set $\mathcal{T}_a \subset \mathcal{L}$ for all $a \in \mathcal{A}$:

$$\mathcal{T}_a := \left\{ (i,j) \in \mathcal{L} \mid i \in \mathcal{N}_a, j \notin \mathcal{N}_a \right\}.$$

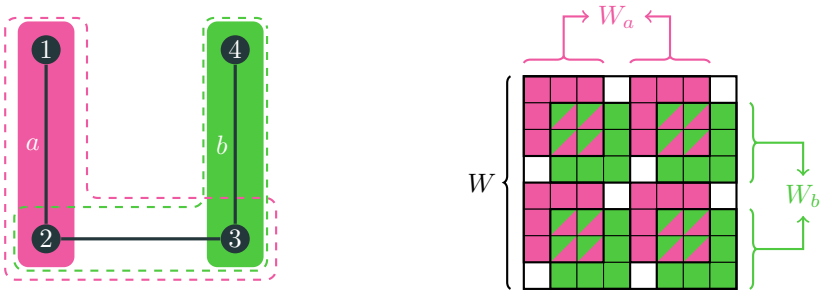
We now define the extended bus set $\mathcal{N}_a^+$ to expand the bus sets in each control area by including the endpoints of the tie-lines connected to that area for every $a \in \mathcal{A}$ as follows:

$$\mathcal{N}_a^+ := \mathcal{N}_a \cup \left\{ j \in \mathcal{N} \mid \exists i \in \mathcal{N}_a, (i,j) \in \mathcal{T}_a \right\}.$$

We now provide an illustrative example for a multi-area 4-bus power network to clarify the above notations.

Example. Two control areas are considered in Figure 4.1a for a 4-bus power network that are denoted by $\mathcal{A} = \{a, b\}$. Their corresponding bus sets are given by $\mathcal{N}_a = \{1, 2\}$ and $\mathcal{N}_b = \{3, 4\}$, shown by the shaded regions. They are connected through a single tie-line $(2, 3)$, and therefore, $\mathcal{T}_a = \{(2, 3)\}$, $\mathcal{T}_b = \{(3, 2)\}$. This leads to the extended bus sets as $\mathcal{N}_a^+ = \{1, 2, 3\}$ and $\mathcal{N}_b^+ = \{2, 3, 4\}$, depicted by the dashed lines. Finally the set of neighboring control areas are $\mathcal{B}_a = \{b\}$ and $\mathcal{B}_b = \{a\}$.

We are now ready to decompose the matrix variable W into sub-matrices corresponding to the extended control areas. Consider $W_a \in \mathbb{S}^{2|\mathcal{N}_a^+|}$ as the network state of area a which is constructed by extracting a sub-matrix from W using only the rows and



(a) The layout of 4-bus network partitioned in two areas, where dashed lines indicate the extended areas.

(b) Matrix variable and sub-matrices. The entries that overlap $[W_a]_{ab}$ and $[W_b]_{ba}$ are half-shaded.

Figure 4.1: A 4-bus example of a multi-area power network with two control areas.

columns that correspond to the buses in $\mathcal{N}_a^+$. Denote the intersection of the extended bus sets by $\mathcal{E}_{ab} := \mathcal{N}_a^+ \cap \mathcal{N}_b^+$ for every neighboring area $b \in \mathcal{B}_a$. Consider now $[W_a]_{ab}$ as the sub-matrix extracted from W_a with its rows and columns corresponding to the buses in $\mathcal{E}_{ab}$, and likewise $[W_b]_{ab}$ the extraction from W_b that corresponds to the same buses. Note that the order of the subscript does not change the shared bus set between extended control areas, and therefore $[W_a]_{ab}$ and $[W_a]_{ba}$ refer to the same extraction from W_a.

Example (cont'd). Figure 4.1b depicts the original 8×8 *state matrix for a 4-bus power network. The sub-matrices corresponding to the areas are two* 6×6 *matrices, shown by the magenta and green shaded and half-shaded entries. The sub-matrices have overlap on 16 entries, which are half-shaded. These entries correspond to the buses that are the endpoints of tie-lines (e.g., bus 2 and 3).*

We can now define the local feasibility set denoted with $\mathcal{W}_a(\mathbf{p}^w, \mathbf{s}^D)$ for all $a \in \mathcal{A}$:

$$\mathcal{W}_a(\mathbf{p}^w, \mathbf{s}^D) := \Bigg\{ W_a \in \mathbb{S}^{2|\mathcal{N}_a^+|} \,\Bigg|\, \operatorname{Tr}([E_{\text{ref}}]_a W_a) = 0, \tag{4.17a}$$

$$\underline{p_k^G} \le \operatorname{Tr}([Y_k]_a W_a) + p_k^D - C_k^w \mathbf{p}_k^w \le \overline{p_k^G}, \forall k \in \mathcal{N}_a, \tag{4.17b}$$

$$\underline{q_k^G} \le \operatorname{Tr}\big(([Y_k^*]_a W_a) + q_k^D \le \overline{q_k^G}, \forall k \in \mathcal{N}_a, \tag{4.17c}$$

$$|\underline{v_k}|^2 \le \operatorname{Tr}([M_k]_a W_a) \le |\overline{v_k}|^2, \forall k \in \mathcal{N}_a^+, \tag{4.17d}$$

$$\begin{bmatrix} -|\overline{s_{lm}}|^2 & \operatorname{Tr}([Y_{lm}]_a W_a) & \operatorname{Tr}\big([Y_{lm}^*]_a W_a\big) \\ \operatorname{Tr}([Y_{lm}]_a W_a) & -1 & 0 \\ \operatorname{Tr}\big([Y_{lm}^*]_a W_a\big) & 0 & -1 \end{bmatrix} \preceq 0, \tag{4.17e}$$

$$\forall (l,m) \in (\mathcal{N}_a^+ \times \mathcal{N}_a^+) \cap \mathcal{L} \Bigg\},$$

where the data-matrices, $Y_k, Y_k^*, Y_{lm}, Y_{lm}^*, M_k$, for each area $a \in \mathcal{A}$ are obtained via extracting the columns and rows corresponding to the buses in $\mathcal{N}_a^+$. These partitions of

the data-matrices are denoted with $[\cdot]_a$ in $\mathcal{W}_a(\mathbf{p}^w, \mathbf{s}^D)$ for all $a \in \mathcal{A}$.

Remark 11 *The power injection constraints,* (4.17b) *and* (4.17c)*, are only imposed on the buses in the area, and not on the extended area, since the power injection limits cannot be enforced on the neighboring buses $\mathcal{N}_a^+ \setminus \mathcal{N}_a$. This is due to the fact that the neighboring buses have lines that connect to buses in $\mathcal{N}_a$, the tie-line(s) in $\mathcal{T}_a$, and also have lines that connect to buses in $\mathcal{N}_b$. It is therefore impossible to define the injected power in the neighboring bus using the state matrix of area a. In other words, there is no limit on the power injection at the other end of a tie-line for each control area, allowing power to flow from and to the neighboring control areas. The line flow limits* (4.17e) *are enforced for the intersection of buses in the extended area and the set of all lines, which are equivalent to the set of lines in an area including the tie-lines.*

4

An important constraint that still requires a proper decomposition to allow a distributed solution of the problem is the PSD constraint of the full state matrix W in (4.16c). Our goal is to decompose the centralized PSD constraint (4.16c) by only imposing PSDness on the local state matrix W_a for each area $a \in \mathcal{A}$. Such a solution will be a partially filled matrix $\bar{W}$, with only those entries filled that correspond to at least one of the sub-matrices, and all other entries will be undetermined, e.g., the white cells of the matrix variables in Figure 4.1b. Various algorithms are available for *matrix completion*, the a-posteriori filling of the undetermined entries.

We approach this goal by use of the *chordal theorem* [58] that guarantees the completed matrix to be PSD if and only if specific sub-matrices are PSD. The chordal theorem states that one can reconstruct the PSD Hermitian[5] matrix using only the entries that correspond to the nodes in the maximal cliques[6] of a chordal graph[7]. The chordal theorem can thus be used to prove the equivalence between the PSDness of a matrix and the PSDness of its sub-matrices.

We now explain how one can decompose such a PSD constraint of the full state matrix W in (4.16c). Consider a graph G over $\mathcal{N}$, with its edges corresponding to the set of extended buses $\{\mathcal{N}_a^+\}$ for all areas $a \in \mathcal{A}$. This means that every bus in $\mathcal{N}_a^+$ is connected to all other buses in $\mathcal{N}_a^+$ with a single edge to form a maximal clique. This graph is then chordal [182, Proposition 1] and all maximal cliques of G are captured by the subsets $\{\mathcal{N}_a^+\}$, under the following assumptions.

Assumption 10 *Graph G with all control areas as its nodes and the tie-lines between the areas as its edges is a tree, i.e., an acyclic connected graph.*

Assumption 11 *Every control area $a \in \mathcal{A}$ has at least one bus that does not have overlap, i.e., does not have a tie-line connected to it.*

[5] A symmetric matrix is a real-valued Hermitian matrix, i.e. $\mathbb{S} \subset \mathbb{H}$, thus the chordal theorem also holds for symmetric matrices.

[6] A *clique* is a subset of nodes that together form a complete graph, i.e. the number of edges between any two nodes in a clique is equal to one. A clique is *maximal* if it is not a subset of any other cliques in the graph [168].

[7] A graph is *chordal* if every cycle of length greater than three has a chord (an edge between non-consecutive vertices in the cycle) [168].

Assumption 10 prevents multi-area power networks to have cyclic interconnections between their control areas. This can be achieved due to the fact that power networks are usually spread out geographically and mostly not intertwined. The typical number of tie-lines between power networks tends to be low. Assumption 11 will hold for almost every real-world power system, as there tend to be much more buses than tie-lines in multi-area power networks.

The decomposition is valid if and only if these two assumptions hold, thus imposing PSD constraints on the sub-matrices corresponding to the extended areas, the original matrix will also be PSD, and can be completed from the local results. This enables one to split the single PSD constraint (4.16c) on W into $|\mathcal{A}|$ smaller PSD constraints.

Remark 12 *The proposed decomposition of PSD constraint is based on the geographical (spatial) layout of the power network, rather than the sparsity in the data matrices. In the latter case, the smallest possible sub-matrices can be found, whereas in the former case, the sub-matrices correspond to the bus sets of the extended control areas used in this decomposition, and can be quite substantial.*

To decompose the problem (4.16), we define the local cost function $f_G^a(W_a, \mathbf{p}^w, \mathbf{p}^D)$ as follows:

$$f_G^a(W_a, \mathbf{p}^w, \mathbf{p}^D) := \sum_{k \in \mathcal{G}_a} c_k^{\text{li}}\big(\text{Tr}([Y_k]_a W_a) + p_k^D - p_k^w\big) + c_k^{\text{qu}}\left(\big(\text{Tr}([Y_k]_a W_a) + p_k^D - p_k^w\big)\right)^2,$$

where $\mathcal{G}_a$ corresponds to the generator buses in the control area $a \in \mathcal{A}$. We are now ready to formulate a multi-area AC OPF problem as follows:

$$\underset{\substack{\{W_a\}_{a\in\mathcal{A}},\\ \{\bar{W}_{ab}\}_{a\in\mathcal{A}, b\in\mathcal{B}_a}}}{\text{minimize}} \quad \sum_{a\in\mathcal{A}} f_G^a(W_a, \mathbf{p}^w, \mathbf{p}^D) \tag{4.18a}$$

$$\text{subject to} \quad W_a \in \mathcal{W}_a(\mathbf{p}^w, \mathbf{s}^D), \qquad \forall a \in \mathcal{A}, \tag{4.18b}$$

$$[W_a]_{ab} = \bar{W}_{ab}, \quad \forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a, \tag{4.18c}$$

$$W_a \succeq 0, \qquad \forall a \in \mathcal{A}. \tag{4.18d}$$

where $\bar{W}_{ab} \in \mathbb{S}^{2|\mathcal{E}_{ab}|}$ is an auxiliary matrix variable introduced for every pair of neighboring areas a and b. The following proposition is a direct result of [95, Theorem 1].

Proposition 3 *The optimal objective value of the decomposed Problem* (4.18) *is equal to the optimal objective of the centralized Problem* (4.16).

The proposed multi-area AC OPF problem formulation is a general consensus problem with coupling only through the equality constraints (4.18c). It has a set of local variables and constraints, which are separable between the control areas, and shared variables between two areas via the auxiliary variables $\bar{W}_{ab}, \forall a \in \mathcal{A}, b \in \mathcal{B}_a$. The local variables are related to the shared variables implicitly through the coupling constraints (4.18c).

4.4.2. DISTRIBUTED MULTI-AREA AC OPF PROBLEM VIA ADMM

Consider the multi-area AC OPF problem (4.18) as a general consensus problem. To solve the general consensus problem in a distributed setting, we employ the ADMM algorithm [22]. It has been proven that ADMM for this type of problem converges linearly [160]. We follow a similar approach as in [182] to solve the multi-area AC OPF problem (4.18) in a distributed manner.

Define the augmented Lagrangian function as follows:

$$L(\{W_a\},\{\bar{W}_{ab}\},\{\Lambda_{ab}\}) := \sum_{a\in\mathcal{A}}\Bigg(f_G^a(W_a,\mathbf{p}^w,\mathbf{p}^D) + I_{W_a}(W_a) + \sum_{b\in\mathcal{B}_a}\left(\frac{\mu}{2}\left\|[W_a]_{ab} - \bar{W}_{ab} + \frac{\Lambda_{ab}}{\mu}\right\|_F^2 + \frac{1}{2\mu}\left\|\Lambda_{ab}\right\|_F^2\right)\Bigg),$$

4

where $I_{W_a}(W_a) : \mathbb{S}^{2|\mathcal{N}_a^+|} \to \{0,+\infty\}$ is a convex indicator function for constraints (4.18b) and (4.18d) that maps to infinity if one of the constraints is violated, and to zero otherwise. Step size μ is a fixed constant, and multipliers $\Lambda_{ab} \in \mathbb{S}^{2|\mathcal{E}_{ab}|}$ are introduced for every coupling constraint. Note that unlike $\bar{W}_{ab}$ and $\bar{W}_{ba}$, the multipliers Λ_{ab} and Λ_{ba} cannot be used interchangeably, since they correspond to different constraints. The indicator function makes sure all constraints (4.18b) and (4.18d) are satisfied, and the squared norm penalty term forces constraint (4.18c) to be satisfied. We now describe the steps of the ADMM algorithm as follows:

(1) UPDATE PRIMAL VARIABLES

The multipliers and auxiliary variables are fixed at their value of the previous iteration. Consider the following minimization problem for all $a \in \mathcal{A}$:

$$W_a^{(k+1)} = \arg\min_{W_a}\left\{f_G^a(W_a,\mathbf{p}^w,\mathbf{p}^D) + I_{W_a}(W_a) + \sum_{b\in\mathcal{B}_a}\left(\frac{\mu}{2}\left\|[W_a]_{ab} - \bar{W}_{ab}^{(k)} + \frac{\Lambda_{ab}^{(k)}}{\mu}\right\|_F^2\right)\right\}, \tag{4.19}$$

Since the minimization is only in W_a, all terms of Λ_{ab} drop out. This results in $|\mathcal{A}|$ separate SDPs.

(2) UPDATE AUXILIARY VARIABLES

The resulting sub-matrices $W_a^{(k+1)}$ for all $a \in \mathcal{A}$ are used to update the auxiliary variables. The multipliers again are fixed at their previous value. Note that each area only needs to communicate the part of its local state matrix that has overlap with its neighboring area to update the auxiliary variables. If the multipliers Λ_{ab} are initialized with zero $\forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a$, [22, Section 7.1], the update of the auxiliary variable simplifies to taking the average $\forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a$:

$$\bar{W}_{ab}^{(k+1)} = \frac{1}{2}\left([W_a^{(k+1)}]_{ab} + [W_b^{(k+1)}]_{ba}\right). \tag{4.20}$$

Algorithm 3 Distributed Multi-Area AC OPF Algorithm

```
Initialize: k = 0, Λ_ab^(0) = 0, W̄_ab^(0) = 0, ∀b ∈ B_a, ∀a ∈ A
while not converged do
    for all a ∈ A do
        Update W_a^(k+1) using (4.19)
        Broadcast [W_a^(k+1)]_ab to all b ∈ B_a
        Receive [W_b^(k+1)]_ba from all b ∈ B_a
        Update W̄_ab^(k+1) using (4.20) for all b ∈ B_a
        Update Λ_ab^(k+1) using (4.21) for all b ∈ B_a
        k = k + 1
    end for
end while
```

4

(3) UPDATE MULTIPLIER VARIABLES

The multipliers are updated as follows $\forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a$:

$$\Lambda_{ab}^{(k+1)} = \Lambda_{ab}^{(k)} + \mu\left([W_a^{(k+1)}]_{ab} - \bar{W}_{ab}^{(k+1)}\right) = \Lambda_{ab}^{(k)} + \frac{\mu}{2}\left([W_a^{(k+1)}]_{ab} - [W_b^{(k+1)}]_{ba}\right). \tag{4.21}$$

Notice that no information exchange is needed for the update of the multiplier, since the parts of the state matrix of neighboring areas have already been communicated in the update of the auxiliary variables.

In Algorithm 3, we summarize the proposed distributed consensus framework using ADMM algorithm to illustrate the calculation and communication steps in each control area $a \in \mathcal{A}$. A single control area needs to solve a small-scale SDP in Algorithm 3 at each iteration that can be considered as the highest computational cost in the proposed algorithm. In Algorithms 3 and 3, only the parts of the local state matrix that have overlap with the neighbors are shared, using only simple operations, e.g., matrix addition, subtraction and scaling. As an advantage of our proposed framework, all control areas in a power network are able to reach consensus by exchanging only the relevant part of the local state matrix with their neighboring areas.

We define the consensus on the shared variables $\bar{W}_{ab}$ (as the convergence criteria) of the proposed ADMM algorithm as follows:

$$\eta^{(k)} = \sum_{a\in\mathcal{A}} \sum_{b\in\mathcal{B}_a} \|[W_a^{(k)}]_{ab} - \bar{W}_{ab}^{(k)}\|_F^2 .$$

If the residual sequence $\{\eta^{(k)}\}_{k=1}^{+\infty}$ is sufficiently small, all control areas of the power network have reached consensus on the shared variables $\bar{W}_{ab}$. We assume that the Slater's constraint qualification [161] holds for the proposed problem formulations meaning that their feasible set has a non-empty interior, and it thus admits at least one strictly feasible solution. We can now provide convergence of Algorithm 3. A similar convergence analysis of ADMM for sparse SDPs and for decomposable SDPs can be found in [96] and [72].

Theorem 8 *Assume that Slater's condition [161] holds for the multi-area AC OPF problem (4.18), and consider the iterative steps given in Algorithm 3. Then the following statements hold:*

- *The residual sequence* $\{\eta^{(k)}\}_{k=0}^{+\infty}$ *tends to* 0 *in a non-increasing way as* k *goes to* $+\infty$*, and consequently, for all* $b \in \mathcal{B}_a$*, and each* $a \in \mathcal{A}$*:*

$$[W_a^{(+\infty)}]_{ab} = [W_b^{(+\infty)}]_{ba} = \bar{W}_{ab}^{(+\infty)}.$$

- *The sequence* $\{W_a^{(k)}\}_{\forall a \in \mathcal{A}}$ *generated by Algorithm 3 converges to an optimal solution* $\{W_a^*\}_{\forall a \in \mathcal{A}}$ *of the multi-area AC OPF problem (4.18) as* k *tends to* $+\infty$*.*

Proof. The theorem follows from [66] that studies the convergence of a standard ADMM problem. The details are omitted for brevity. □

Remark 13 *The proposed algorithm uses a Gauss-Siedel update on the primal variables and the auxiliary variables, after which the multiplier variables are updated [22]. Since either the primal or the auxiliary variables are fixed in the Gauss-Siedel steps, the problem can be distributed for all the areas. The main advantage to distribute a large-scale AC OPF problem is the ability of finding local solutions for each control area based on the information received in the previous iteration. Such calculations can therefore be carried out in parallel. Although an actual parallel implementation is outside the scope of this work, it is important to mention that the proposed algorithm is amenable to such an implementation. ADMM algorithms typically need a large number of iterations to converge to high accuracy, so the local problems need to be solved many times before finding a good enough solution. Thus, the ADMM approach without parallelization might not be the quickest method to solve the AC OPF problem. Such a distributed framework is advantageous especially when the global AC OPF problem is too large and cannot be solved within polynomial time due to the curse of dimensionality or memory limitations and computational constraints.*

In the following part, we extend the proposed distributed multi-area AC OPF algorithm to solve the stochastic RS problem.

4.4.3. DISTRIBUTED MULTI-AREA SP-OPF-RS PROBLEM VIA ADMM

It is now of interest to include wind power realizations into the multi-area AC OPF problem (4.18), i.e., extend the deterministic multi-area AC OPF problem to a stochastic RS problem. It is very important to highlight that this extension is possible only using the proposed formulation (SP-OPF-RS) in Section 4.3.2, whereas this extension is not possible using the other formulation (SC-OPF-RS) in Section 4.3.3 due to the curse of dimensionality and computational burden.

Our goal now is to decompose the SP-OPF-RS problem. Given Assumption 10 and Assumption 11, we approach this goal by imposing the following technical assumption.

Assumption 12 *The set of scenarios of wind trajectories* $\mathcal{S}$*, as defined in Section 4.3.2, is given to all control areas of the power network.*

The condition in Assumption 12 is enforced due to the fact that the uncertainty source $\mathcal{P}$ is a common uncertainty source between all control areas, and therefore, the set of scenarios $\mathcal{S}$ has to be common between all control areas. We relax this condition later in this section.

Consider the following affine policy for the local network state of each area $\forall a \in \mathcal{A}$, $\forall \tilde{\mathbf{p}}^m \in \mathcal{S}$, and $\forall t \in \mathcal{T}$:

$$\hat{W}_{a,t}(\mathbf{p}_t^m) := W_{a,t}^f + \max(-\mathbf{p}_t^m, 0) W_{a,t}^{\text{us}} + \max(\mathbf{p}_t^m, 0) W_{a,t}^{\text{ds}},$$

where $W_{a,t}^f, W_{a,t}^{\text{us}}, W_{a,t}^{\text{ds}} \in \mathbb{S}^{2|\mathcal{N}_a^+|}$ are related to the sub-matrices from $W_t^f, W_t^{\text{us}}, W_t^{\text{ds}}$ using only the rows and columns corresponding to the buses in $\mathcal{N}_a^+$. Define the local reserve cost per each time step $t \in \mathcal{T}$ for all $a \in \mathcal{A}$ as follows:

$$f_R^a(\mathbf{r}_{a,t}^{\text{us}}, \mathbf{r}_{a,t}^{\text{ds}}) := \sum_{k \in \mathcal{G}_a} \mathbf{c}^{\text{us}} C_k^G \mathbf{r}_{a,t}^{\text{us}} + \mathbf{c}^{\text{ds}} C_k^G \mathbf{r}_{a,t}^{\text{ds}},$$

where C_k^G is a connection matrix for the generators such that the (i, j)-th entry is one if generator j is located at the bus i and zero otherwise. Consider $\Xi_{\text{ma}} := \{\Xi_a\}_{\forall a \in \mathcal{A}}$ where $\Xi_a = \{\tilde{W}_a^f, \tilde{W}_a^{\text{us}}, \tilde{W}_a^{\text{ds}}, \tilde{\mathbf{r}}_a^{\text{us}}, \tilde{\mathbf{r}}_a^{\text{ds}}, \tilde{\mathbf{d}}_a^{\text{us}}, \tilde{\mathbf{d}}_a^{\text{ds}}\}$ is the set of local decision variables for each control area $a \in \mathcal{A}$, and define $\Theta := \{\tilde{\tilde{W}}_{ab}^f, \tilde{\tilde{W}}_{ab}^{\text{us}}, \tilde{\tilde{W}}_{ab}^{\text{ds}}, \tilde{\tilde{\mathbf{d}}}_a^{\text{us}}, \tilde{\tilde{\mathbf{d}}}_a^{\text{ds}}\}_{\forall b \in \mathcal{B}_a, \forall a \in \mathcal{A}}$ to be the set of auxiliary variables. We are now in the position to formulate a multi-area SP-OPF-RS problem (MASP-OPF-RS) as follows:

$$\min_{\Xi_{\text{ma}}, \Theta} \quad \sum_{a \in \mathcal{A}} \sum_{t \in \mathcal{T}} \left(f_G^a(W_{a,t}^f, \mathbf{p}_t^{w,f}) + f_R^a(\mathbf{r}_{a,t}^{\text{us}}, \mathbf{r}_{a,t}^{\text{ds}}) \right)$$

$$\text{s.t.} \quad W_{a,t}^f \in \mathcal{W}_a(\mathbf{p}_t^{w,f}, \mathbf{s}_t^D) \tag{4.22a}$$

$$\hat{W}_{a,t}(\mathbf{p}_t^m) \in \mathcal{W}_a(\mathbf{p}_t^{w,f} + \mathbf{p}_t^m, \mathbf{s}_t^D) \tag{4.22b}$$

$$W_{a,t}^f \succeq 0, \quad W_{a,t}^{\text{us}} \succeq 0, \quad W_{a,t}^{\text{ds}} \succeq 0 \tag{4.22c}$$

$$-\mathbf{r}_{a,t}^{\text{ds}} \leq \mathbf{r}_{a,t} \leq \mathbf{r}_{a,t}^{\text{us}}, \; 0 \leq \mathbf{r}_{a,t}^{\text{us}}, \; 0 \leq \mathbf{r}_{a,t}^{\text{ds}} \tag{4.22d}$$

$$\text{Tr}\left([Y_k]_a W_{a,t}^{\text{us}}\right) = C_k^G \mathbf{d}_{a,t}^{\text{us}}, \; \forall k \in \mathcal{G}_a \tag{4.22e}$$

$$\text{Tr}\left([Y_k]_a W_{a,t}^{\text{ds}}\right) = C_k^G \mathbf{d}_{a,t}^{\text{ds}}, \; \forall k \in \mathcal{G}_a \tag{4.22f}$$

$$\mathbf{1}^\top \mathbf{d}_{a,t}^{\text{us}} = 1, \; \mathbf{1}^\top \mathbf{d}_{a,t}^{\text{ds}} = -1 \tag{4.22g}$$

$$[W_{a,t}^f]_{ab} = \bar{W}_{ab,t}^f, \; \forall b \in \mathcal{B}_a \tag{4.22h}$$

$$[W_{a,t}^{\text{us}}]_{ab} = \bar{W}_{ab,t}^{\text{us}}, \; \forall b \in \mathcal{B}_a \tag{4.22i}$$

$$[W_{a,t}^{\text{ds}}]_{ab} = \bar{W}_{ab,t}^{\text{ds}}, \; \forall b \in \mathcal{B}_a \tag{4.22j}$$

$$\mathbf{d}_{a,t}^{\text{us}} = \bar{\mathbf{d}}_t^{\text{us}}, \; \mathbf{d}_{a,t}^{\text{ds}} = \bar{\mathbf{d}}_t^{\text{ds}} \tag{4.22k}$$

$$\forall \tilde{\mathbf{p}}^m \in \mathcal{S}, \; \forall t \in \mathcal{T}, \; \forall a \in \mathcal{A} \tag{4.22l}$$

where $\mathbf{r}_{a,t}$ should be replaced with

$$C_k^G \mathbf{r}_{a,t} = -\text{Tr}\left([Y_k]_a W_{a,t}^{\text{us}}\right) \min(\mathbf{p}_t^m, 0) + \text{Tr}\left([Y_k]_a W_{a,t}^{\text{ds}}\right) \max(\mathbf{p}_t^m, 0), \; \forall k \in \mathcal{G}_a.$$

As a direct consequence of Proposition 3, we have the following.

Proposition 4 *The optimal objective function value of the proposed MASP-OPF-RS problem is equal to the optimal objective function value of the SP-OPF-RS problem.*

Remark 14 *The distribution of up- and down spinning reserve power in the power system is enforced through two equality constraints in the SP-OPF-RS problem formulation. One can reformulate them in terms of each control area $a \in \mathcal{A}$ at each time step $t \in \mathcal{T}$ as follows:*

$$\sum_{a\in\mathcal{A}}\sum_{k\in\mathcal{G}_a} \mathrm{Tr}\left([Y_k]_a W^{us}_{a,t}\right) = 1, \sum_{a\in\mathcal{A}}\sum_{k\in\mathcal{G}_a} \mathrm{Tr}\left([Y_k]_a W^{ds}_{a,t}\right) = -1.$$

The decomposition of such equality constraints however is not straightforward. We therefore follow the proposed reformulation (4.14) in Remark 8.

4

We now tackle the condition in Assumption 12 to provide a more flexible multi-area formulation of the MASP-OPF-RS problem. Define $\varepsilon_a \in (0,1)$ and $1-\beta_a \in (0,1)$ to be the local level of constraint violation and the local level of confidence for each control area $a \in \mathcal{A}$, respectively. Each control area can now build its own set of scenarios of wind power trajectories $\mathcal{S}_a := \{\tilde{\mathbf{p}}^{m,1}, \cdots, \tilde{\mathbf{p}}^{m,N_{s_a}}\}$ such that $N_{s_a} \geq \mathsf{N}(\varepsilon_a, \beta_a, d_a)$, where d_a is the number of decision variables in the control area $a \in \mathcal{A}$, as it is defined in Theorem 7. Using $\forall \tilde{\mathbf{p}}^m \in \mathcal{S}_a$ instead of $\forall \tilde{\mathbf{p}}^m \in \mathcal{S}$ in Equation (4.22l), we develop a more flexible multi-area formulation of the MASP-OPF-RS problem and relax Assumption 12. To quantify the robustness of the obtained solution via MASP-OPF-RS with $\mathcal{S}_a, \forall a \in \mathcal{A}$, consider the following theorem which is the main result of this section.

Theorem 9 *Let $\varepsilon_a, \beta_a \in (0,1), \forall a \in \mathcal{A}$ be chosen such that $\varepsilon = \sum_{a\in\mathcal{A}} \varepsilon_a \in (0,1)$ and $\beta = \sum_{a\in\mathcal{A}} \beta_a \in (0,1)$. If $\Xi^*_{ma} := \{\Xi^*_a\}_{\forall a\in\mathcal{A}}$ is a feasible solution of the MASP-OPF-RS problem with scenario set $\mathcal{S}_a$ for each $a \in \mathcal{A}$, then Ξ^*_{ma} is also a feasible solution of the P-OPF-RS problem with probability higher than $1-\varepsilon$ and with confidence level of at least $1-\beta$.*

Proof. Based on an important observation that each control area $a \in \mathcal{A}$ can consider a common uncertainty source $\mathcal{S}$ as a local (private) source of uncertainty and build its own (local) set of scenarios $\mathcal{S}_a$, the proof follows the similar steps as [146, 147, Theorem 2], that studies the quantification of the feasibility error with private and common uncertainty sources in a distributed setup using randomization technique, with some minor modifications. We provide a complete proof in Appendix C. □

The following corollary is a direct result of Theorem 9. Decompose the P-OPF-RS problem using the proposed approach in Section 4.4.1 into the multi-area P-OPF-RS problem (MAP-OPF-RS). This reformulation is straightforward and therefore it is omitted for the sake of brevity.

Corollary 2 *The local optimal solution Ξ^*_a for all $a \in \mathcal{A}$ obtained via the MASP-OPF-RS problem is a feasible solution for the MAP-OPF-RS problem with probability higher than $1-\varepsilon_a$ and with confidence level of $1-\beta_a$.*

We now continue by developing a distributed framework for the proposed formulation of the MASP-OPF-RS problem in (4.22). Define $\Lambda^f_{ab,t}, \Lambda^{\mathrm{us}}_{ab,t}, \Lambda^{\mathrm{ds}}_{ab} \in \mathbb{S}^{2|\mathcal{E}_{ab}|}$ for all

$b \in \mathcal{B}_a$, for each $a \in \mathcal{A}$ at each time step $t \in \mathcal{T}$ as the multipliers for the first three consensus constraints, (4.22h), (4.22i), (4.22j), and $\lambda^{\text{us}}_{a,t}, \lambda^{\text{ds}}_{a,t} \in \mathbb{R}^{N_G}$ for all $a \in \mathcal{A}$ at each time step $t \in \mathcal{T}$ for the last two consensus constraints (4.22k). We then collect all the multipliers in Γ. One can denote the local objective function for each control area $a \in \mathcal{A}$ as:

$$f^a(\Xi_{\text{a}}) := f^a_G\big(W^f_{a,t}, \mathbf{p}^{w,f}_t\big) + f^a_R(\mathbf{r}^{\text{us}}_{a,t}, \mathbf{r}^{\text{ds}}_{a,t}) + I_{W_a}(\Xi_{\text{a}}),$$

where $I_{W_a}(\Xi_{\text{a}})$ is a convex indicator function for all constraints except the consensus constraints. Consider now the augmented Lagrangian of the MASP-OPF-RS problem as follows:

$$\begin{aligned} L(\Xi_{\text{ma}}, \Theta, \Gamma) = \sum_{a\in\mathcal{A}} \sum_{t\in\mathcal{T}} \Bigg\{ & f^a(\Xi_{\text{a}}) + \frac{\mu}{2}\left\| \mathbf{d}^{\text{us}}_{a,t} - \bar{\mathbf{d}}^{\text{us}}_t + \frac{\lambda^{\text{us}}_{a,t}}{\mu} \right\|_2^2 + \frac{\mu}{2}\left\| \mathbf{d}^{\text{ds}}_{a,t} - \bar{\mathbf{d}}^{\text{ds}}_t + \frac{\lambda^{\text{ds}}_{a,t}}{\mu} \right\|_2^2 \\ & + \sum_{b\in\mathcal{B}_a} \Bigg(\frac{\mu}{2}\left\| [W^f_{a,t}]_{ab} - \bar{W}^f_{ab,t} + \frac{\Lambda^f_{ab,t}}{\mu} \right\|_F^2 \\ & \quad + \frac{\mu}{2}\left\| [W^{\text{us}}_{a,t}]_{ab} - \bar{W}^{\text{us}}_{ab,t} + \frac{\Lambda^{\text{us}}_{ab,t}}{\mu} \right\|_F^2 \\ & \quad + \frac{\mu}{2}\left\| [W^{\text{ds}}_{a,t}]_{ab} - \bar{W}^{\text{ds}}_{ab,t} + \frac{\Lambda^{\text{ds}}_{ab,t}}{\mu} \right\|_F^2 + f(\Gamma) \Bigg) \Bigg\}, \end{aligned}$$

where $f(\Gamma)$ indicates terms that are related to the multipliers Γ. We can now use ADMM algorithm to solve MASP-OPF-RS. The steps are comparable with the steps described in Section 4.4.2, i.e. update primal, update auxiliary variables and update multipliers. We now describe the steps of the ADMM algorithm as follows:

(1) UPDATE PRIMAL VARIABLES

The multipliers and auxiliary variables are fixed at their value of the previous iteration. Consider the following minimization problem for all $a \in \mathcal{A}$:

$$\begin{aligned} \Xi^{(k+1)}_{\text{a}} = \arg\min_{\Xi_{\text{a}}} \sum_{t\in\mathcal{T}} \Bigg\{ & f^a(\Xi_{\text{a}}) + \frac{\mu}{2}\left\| \mathbf{d}^{\text{us}}_{a,t} - \bar{\mathbf{d}}^{\text{us}}_t + \frac{\lambda^{\text{us}}_{a,t}}{\mu} \right\|_2^2 + \frac{\mu}{2}\left\| \mathbf{d}^{\text{ds}}_{a,t} - \bar{\mathbf{d}}^{\text{ds}}_t + \frac{\lambda^{\text{ds}}_{a,t}}{\mu} \right\|_2^2 \\ & + \sum_{b\in\mathcal{B}_a} \Bigg(\frac{\mu}{2}\left\| [W^f_{a,t}]_{ab} - \bar{W}^f_{ab,t} + \frac{\Lambda^f_{ab,t}}{\mu} \right\|_F^2 \\ & \quad + \frac{\mu}{2}\left\| [W^{\text{us}}_{a,t}]_{ab} - \bar{W}^{\text{us}}_{ab,t} + \frac{\Lambda^{\text{us}}_{ab,t}}{\mu} \right\|_F^2 \\ & \quad + \frac{\mu}{2}\left\| [W^{\text{ds}}_{a,t}]_{ab} - \bar{W}^{\text{ds}}_{ab,t} + \frac{\Lambda^{\text{ds}}_{ab,t}}{\mu} \right\|_F^2 \Bigg) \Bigg\}, \end{aligned} \tag{4.23}$$

where $f(\Gamma)$ is omitted, since the minimization is only in Ξ_{a}. This results in $|\mathcal{A}|$ number of small-scale SDPs in parallel, which can be considered the most computationally expensive step.

(2) UPDATE AUXILIARY VARIABLES

The resulting $\Xi_{\mathrm{a}}^{(k+1)}$ for all $a \in \mathcal{A}$ are used to update the auxiliary variables. The multipliers again are fixed at their previous value. Note that each area only needs to communicate the part of its local variables that have overlap with its neighboring area to update the auxiliary variables. If the multipliers are initialized with zero $\forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a$, [22, Section 7.1], the update of the auxiliary variable simplifies to taking the average $\forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a$:

$$\begin{aligned}
\bar{W}_{ab,t}^{f,(k+1)} &= \frac{1}{2}\left([W_{a,t}^{f,(k+1)}]_{ab} + [W_{b,t}^{f,(k+1)}]_{ba}\right), \\
\bar{W}_{ab,t}^{\mathrm{us},(k+1)} &= \frac{1}{2}\left([W_{a,t}^{\mathrm{us},(k+1)}]_{ab} + [W_{b,t}^{\mathrm{us},(k+1)}]_{ba}\right), \\
\bar{W}_{ab,t}^{\mathrm{ds},(k+1)} &= \frac{1}{2}\left([W_{a,t}^{\mathrm{ds},(k+1)}]_{ab} + [W_{b,t}^{\mathrm{ds},(k+1)}]_{ba}\right), \\
\bar{\mathbf{d}}_t^{\mathrm{us},(k+1)} &= \frac{1}{|\mathcal{A}|}\sum\nolimits_{a\in\mathcal{A}} \mathbf{d}_{a,t}^{\mathrm{us},(k+1)}, \\
\bar{\mathbf{d}}_t^{\mathrm{ds},(k+1)} &= \frac{1}{|\mathcal{A}|}\sum\nolimits_{a\in\mathcal{A}} \mathbf{d}_{a,t}^{\mathrm{ds},(k+1)}.
\end{aligned} \tag{4.24}$$

(3) UPDATE MULTIPLIER VARIABLES

The multipliers are updated as follows $\forall a \in \mathcal{A}, \forall b \in \mathcal{B}_a$:

$$\begin{aligned}
\Lambda_{ab,t}^{f,(k+1)} &= \Lambda_{ab,t}^{f,(k)} + \mu\left([W_{a,t}^{f,(k+1)}]_{ab} - \bar{W}_{ab,t}^{f,(k+1)}\right), \\
\Lambda_{ab,t}^{\mathrm{us},(k+1)} &= \Lambda_{ab,t}^{\mathrm{us},(k)} + \mu\left([W_{a,t}^{\mathrm{us},(k+1)}]_{ab} - \bar{W}_{ab,t}^{f,(k+1)}\right), \\
\Lambda_{ab,t}^{\mathrm{ds},(k+1)} &= \Lambda_{ab,t}^{\mathrm{ds},(k)} + \mu\left([W_{a,t}^{\mathrm{ds},(k+1)}]_{ab} - \bar{W}_{ab,t}^{f,(k+1)}\right), \\
\lambda_{a,t}^{\mathrm{us},(k+1)} &= \lambda_{a,t}^{\mathrm{us},(k)} + \mu\left(\mathbf{d}_{a,t}^{\mathrm{us},(k+1)} - \bar{\mathbf{d}}_t^{\mathrm{us},(k+1)}\right), \\
\lambda_{a,t}^{\mathrm{ds},(k+1)} &= \lambda_{a,t}^{\mathrm{ds},(k)} + \mu\left(\mathbf{d}_{a,t}^{\mathrm{ds},(k+1)} - \bar{\mathbf{d}}_t^{\mathrm{us},(k+1)}\right).
\end{aligned} \tag{4.25}$$

Notice that no information exchange is needed for the update of the multiplier, since the parts of the state matrix of neighboring areas have already been communicated in the update of the auxiliary variables.

Algorithm 4 summarizes the proposed distributed stochastic framework using a consensus ADMM algorithm to solve the MASP-OPF-RS problem. Consider the energy sequence $\{\xi^{(k)}\}_{k=1}^{+\infty}$ as a measure for convergence of Algorithm 4 as follows:

$$\begin{aligned}
\xi^{(k)} = \sum_{a\in\mathcal{A}}\sum_{t\in\mathcal{T}} \Bigg\{ &\left\|\mathbf{d}_{a,t}^{\mathrm{us}} - \bar{\mathbf{d}}_t^{\mathrm{us}}\right\|_2^2 + \left\|\mathbf{d}_{a,t}^{\mathrm{ds}} - \bar{\mathbf{d}}_t^{\mathrm{ds}}\right\|_2^2 + \\
&\sum_{b\in\mathcal{B}_a}\Big(\|[W_{a,t}^{f,(k)}]_{ab} - \bar{W}_{ab,t}^{f,(k)}\|_F^2 + \\
&\|[W_{a,t}^{\mathrm{us},(k)}]_{ab} - \bar{W}_{ab,t}^{\mathrm{us},(k)}\|_F^2 + \\
&\|[W_{a,t}^{\mathrm{ds},(k)}]_{ab} - \bar{W}_{ab,t}^{\mathrm{ds},(k)}\|_F^2\Big)\Bigg\}.
\end{aligned} \tag{4.26}$$

Algorithm 4 Distributed Stochastic MASP-OPF-RS

Initialize: $k = 0, \Gamma^{(0)} = 0, \Theta^{(0)} = 0, \forall b \in \mathcal{B}_a, \forall a \in \mathcal{A}$
Fix $\varepsilon_a \in (0,1)$ and $\beta_a \in (0,1)$, $\forall a \in \mathcal{A}$ such that

$$\varepsilon = \sum_{a\in\mathcal{A}} \varepsilon_a \in (0,1), \beta = \sum_{a\in\mathcal{A}} \beta_a \in (0,1)$$

Build the set of local scenarios $\mathcal{S}_a$, $\forall a \in \mathcal{A}$
while not converged **do**
    for all $a \in \mathcal{A}$ **do**
        Update $\Xi_a^{(k+1)}$ using (4.23)
        Broadcast $[\Xi_a^{(k+1)}]_{ab}$ to all $b \in \mathcal{B}_a$
        Receive $[\Xi_a^{(k+1)}]_{ba}$ from all $b \in \mathcal{B}_a$
        Update $\Theta^{(k+1)}$ using (4.24) for all $b \in \mathcal{B}_a$
        Update $\Gamma^{(k+1)}$ using (4.25) for all $b \in \mathcal{B}_a$
        $k = k + 1$
    end for
end while

If $\xi^{(k)}$ is sufficiently small, all control areas of the power network have reached consensus on Θ.

The following theorem is a direct result of Theorem 8 to provide the convergence property of Algorithm 4.

Theorem 10 *Assume that Slater's condition [161] holds for the MASP-OPF-RS problem (4.22), and consider the iterative steps given in Algorithm 4. Then the following statements hold:*

- *The residual sequence $\{\xi^{(k)}\}_{k=0}^{+\infty}$ tends to 0 in a non-increasing way as k goes to $+\infty$, and therefore, we have $\forall b \in \mathcal{B}_a$ and for each $a \in \mathcal{A}$ at each time step $t \in \mathcal{T}$:*

$$[W_{a,t}^{f,(+\infty)}]_{ab} = [W_{b,t}^{f,(+\infty)}]_{ba} = \bar{W}_{ab}^{f,(+\infty)},$$
$$[W_{a,t}^{us,(+\infty)}]_{ab} = [W_{b,t}^{us,(+\infty)}]_{ba} = \bar{W}_{ab}^{us,(+\infty)},$$
$$[W_{a,t}^{ds,(+\infty)}]_{ab} = [W_{b,t}^{ds,(+\infty)}]_{ba} = \bar{W}_{ab}^{ds,(+\infty)},$$

  and $\forall a \in \mathcal{A}$ at each time step $t \in \mathcal{T}$:

$$\mathbf{d}_{a,t}^{us,+\infty} = \bar{\mathbf{d}}_t^{us,+\infty}, \mathbf{d}_{a,t}^{ds,+\infty} = \bar{\mathbf{d}}_t^{ds,+\infty}.$$

- *The sequence $\{\Xi_a^{(k)}\}_{\forall a\in\mathcal{A}}$ generated by Algorithm 4 converges to an optimal solution $\{\Xi_a^*\}_{\forall a\in\mathcal{A}}$ of the MASP-OPF-RS problem (4.22) as k tends to $+\infty$.*

Proof. The theorem follows from [66] that studies the convergence of a standard ADMM problem. The details are omitted for brevity. □

4.5. NUMERICAL STUDY

In this section, we carry out numerical simulation studies to illustrate the performance of the proposed formulations and distributed framework. After a short description of the simulation setup, we present the simulation results in two different parts. We first provide a simulation study for the IEEE 30-bus power system using the proposed formulation in the SP-OPF-RS problem and compare it with the stochastic RS problem using a DC model of power network. We also develop a new benchmark formulation, namely a converted DC (CDC) approach to have a more sophisticated comparison [151]. As the second part of simulation results, we construct a realistic multi-area case study, and then, solve the MASP-OPF-RS problem using the proposed distributed consensus framework in Section 4.4. We also provide a comparison using the SP-OPF-RS problem.

4.5.1. SIMULATION SETUP

We fix $\varepsilon = 10^{-2}$ and $\beta = 10^{-5}$ to obtain the number of required scenarios of wind power trajectories at each hour $N_s = 541$ as in Remark 10. To generate trajectories for the wind power, we follow the approach of [117] together with a data-set corresponding to the hourly aggregated wind power production of Germany over the period 2006-2011. The nominal load power is obtained from MATPOWER [8] [183], and multiplied with a time-varying load profile similar to [151].

We perform Monte Carlo simulations to check the violation probability of the solutions a posteriori for both parts of simulation results. Power flows of the network are simulated for 10000 new wind power trajectories using MATPOWER, where the power and voltage magnitude of generators and all the loads are fixed without imposing any constraints. The wind power is implemented as a negative load on the wind-bus. Afterward, the resulting power flows and voltage magnitudes are evaluated by means of counting the number of violated constraints.

To solve all proposed formulations, we use Matlab together with YALMIP [86] as an interface and MOSEK [6] as a solver. All optimizations are run on a MacBook Pro with a 2,4 GHz Intel Core i5 processor and 8 GB of RAM.

4.5.2. SIMULATION RESULTS: PART ONE

We carried out a simulation study using the 30-bus IEEE benchmark power system [37], where only a single wind-bus infeed at bus 10 is considered. After obtaining a solution, the scheduled generator power (the generator power based on the forecast wind trajectory) and the voltage magnitudes are extracted from $\tilde{W}^f = \{W_t^f\}_{t\in\mathcal{T}}$ for all time steps using the following relations $\forall k \in \mathcal{G}, \forall t \in \mathcal{T}$:

$$p_{k,t}^G = \mathrm{Tr}\left(Y_k W_t^f\right) + p_{k,t}^D - p_{k,t}^w, \tag{4.27a}$$

$$q_{k,t}^G = \mathrm{Tr}\left(Y_k^* W_t^f\right) + q_{k,t}^D, \tag{4.27b}$$

$$|v_{k,t}| = \sqrt{W_t^f(k,k) + W_t^f(N_b+k, N_b+k)}\,. \tag{4.27c}$$

[8] MATPOWER is a non-commercial software for solving power flow problems using successive quadratic programming.

A DC model of the power network is used to solve the OPF-RS problem as a benchmark approach for comparison purposes. A detailed description of the DC model can be obtained from [100] and [171]. The solution of the benchmark program is the real generator power and distribution vectors for every hour, $\left\{\mathbf{p}_t^{G,\mathrm{dc}}, \mathbf{d}_t^{\mathrm{us,dc}}, \mathbf{d}_t^{\mathrm{ds,dc}}\right\}$. One also needs the reactive generator power and generator voltage magnitudes in order to have a more realistic comparison. In [149], the nominal value of such variables were extracted from the MATPOWER test case for all time steps and scenarios. This is called the nominal DC solution. This will result in large violations, since the reactive generator power is not adapted to the time-varying demand.

We here develop a novel benchmark approach, namely converted DC (CDC), to have a more sophisticated comparison by solving the following program:

$$\begin{aligned} \min_{\tilde{W}} \quad & \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{G}} \left(p_{k,t}^{G,\mathrm{dc}} - \left(\mathrm{Tr}\,(Y_k W_t) + p_{k,t}^{D} - p_{k,t}^{w,f}\right)\right)^2 \\ \text{s.t.} \quad & W_t \in \mathcal{W}(\mathbf{p}_t^{w,f}, \mathbf{s}_t^{D}), \qquad \forall t \in \mathcal{T}, \\ & W_t \succeq 0, \qquad \forall t \in \mathcal{T}. \end{aligned} \tag{4.28}$$

The solution to this program is a feasible (AC) network state $\tilde{W} = \{W_t\}_{t\in\mathcal{T}}$ where the real generator power is as close as possible to the obtained real generator power from the DC solution. The distribution vectors used in simulation will be equal to those obtained from the original solution of the DC framework. A schematic overview of the optimization and simulation process to obtain and validate both the benchmark and proposed formulations is given in Figure 4.2.

The relative line loadings for all hours and scenarios are shown as box plots per line in Figure 4.3 for DC, CDC, and SP-OPF-RS solutions. The relative line loading is defined as the apparent power flow over a line divided by the line rating $\forall (l,m) \in \mathcal{L}$:

$$|s_{lm}^{\mathrm{rel}}| := \frac{|s_{lm}|}{|\overline{s_{lm}}|},$$

such that a loading higher than 100% corresponds to a violation of the lineflow limit. The DC benchmark results (Figure 4.3a) shows the biggest violations, followed by the CDC benchmark results (Figure 4.3b). For both benchmark results, line 10, 30, 31, and 35 are violated since the line loadings are overloaded as it is clearly shown in Figure 4.3a-b. The SP-OPF-RS solutions (Figure 4.3c) shows almost no violations for all hours and scenarios.

To further assess the performance of these results, the number of violating[9] network states is counted for each hour, and divided by the total number of scenarios to be an empirical measure of the probability on constraint violation per hour (see Figure 4.4). As expected, the DC solution shows a very high level of violation during the peak hours, $t \geq 8$. Although the CDC solution improves the chance of lineflow limit violation, the theoretical limit at the peak hours is still not respected. It is important to notice that the empirical chance of constraint violation for the SP-OPF-RS results are much below the theoretical limit (5%) with 0.05% at $t = 13$ being the highest empirical probability.

[9] A network state is violating if at least one of the line limits is not satisfied.

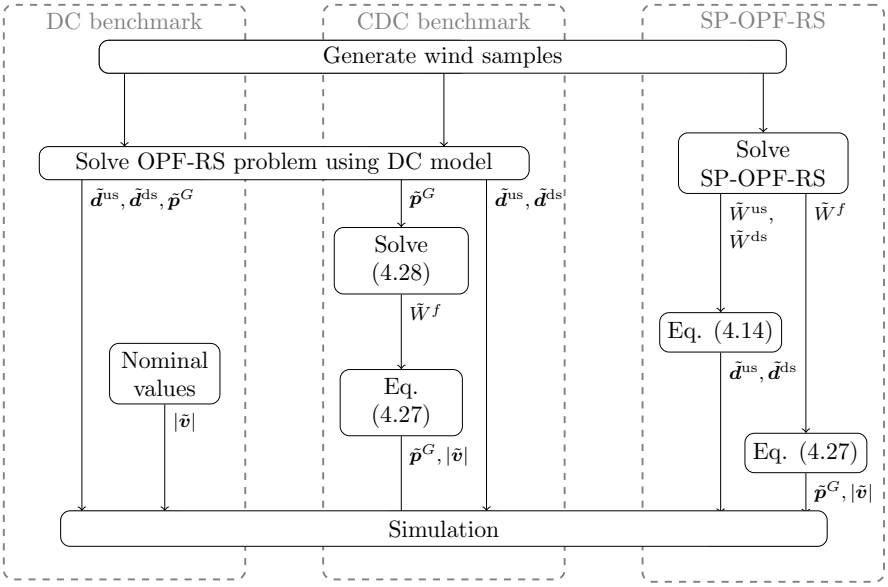


Figure 4.2: Schematic overview of optimization and simulation process for the DC, CDC benchmarks, and the SP-OPF-RS.

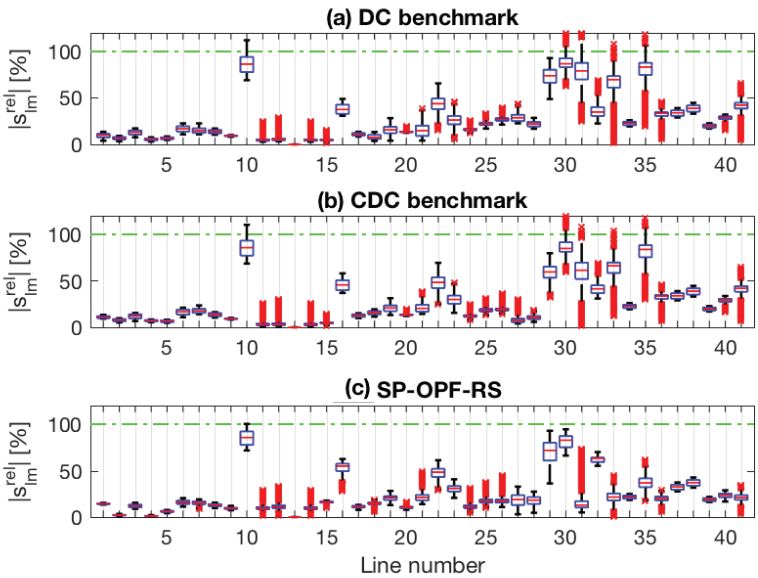


Figure 4.3: Relative line loading for all hours and scenarios per line for the IEEE 30-bus benchmark case study. The red line represents the median value, edges of each box correspond to the 25th and 75th percentiles, the whiskers extend to 99% coverage, and the red marks denote the data outliers. The upper plots (a) and (b) show the Benchmark results, and the lower plot (c) shows the SP-OPF-RS solutions.

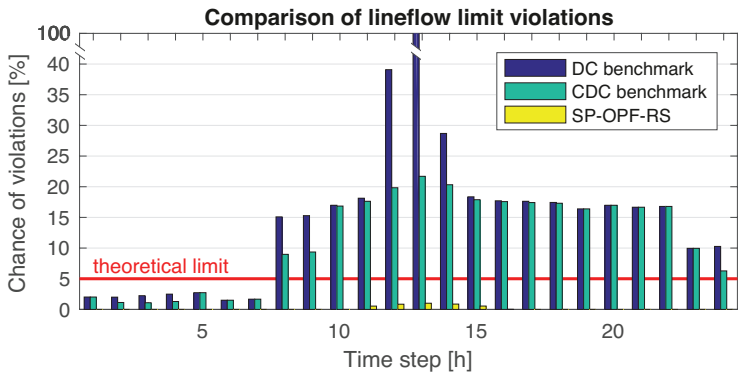


Figure 4.4: Empirical violation level of lineflow limit for different formulations for the IEEE 30-bus benchmark case study.

We next examine the bus voltage magnitudes. It is observed that the DC, and the SP-OPF-RS solutions are always within the limits for all hours and scenarios. However, for the CDC formulation the bus voltage limits show a violation of 100% for all hours. This can be explained by the fact that in the DC framework, the bus voltages are assumed to be constant at nominal value. When we implement the obtained solution in the AC framework, it can be seen that this assumption does not hold. We can thus conclude that for both the DC formulations, the empirical chance of constraint violation is much above the theoretical limits once the solution is implemented in the AC power flow simulations. The a-priori probabilistic guarantees are deemed valid for the SP-OPF-RS solutions.

4.5.3. SIMULATION RESULTS: PART TWO

We construct a two-area power network using two identical IEEE 14-bus power networks [37], and then connect a tie-line between bus 5 of the first network and bus 10 of the second network to create a realistic case study, resulting in an overall 28-bus power network with two control areas. The extended control areas are obtained by adding the endpoints of the tie-lines to the areas such that the buses are grouped in two overlapping sets as shown in Figure 4.5. We formulate the MASP-OPF-RS problem (4.22) and then use Algorithm 4 to solve the problem in a distributed consensus framework using ADMM algorithm and coordinate the local solutions of the control areas towards convergence. Algorithm 4 is run until the residual sequence $\xi^{(k)}$ is sufficiently small, i.e. below 10^{-2} for each hour, and then, the solutions of current iterates (k) are used as the optimal solutions. This happens after 158 iterations. The step size μ is selected using a heuristic approach and it is fixed to 100 which results in good performance, while convergence is fast enough (see Figure 4.6). We also formulate the SP-OPF-RS problem for the 28-bus power network and solve it in a centralized fashion for comparison purposes.

The resulting dispatch and distribution vectors are extracted from the final iterates. The generator dispatch is compared with the centralized solution in Figure 4.7. The distributed solution is almost the same as the centralized solution. The distribution vectors are displayed in Figure 4.8. It can be seen that the results are quite similar for the cen-

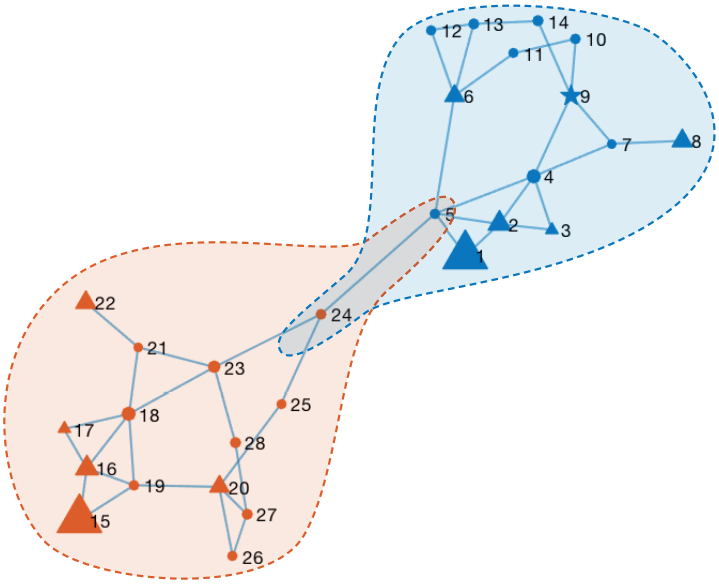


Figure 4.5: Pictorial overview of the decomposed IEEE 28-bus power system case study. Triangles, circles, and a star indicate the generator buses, load buses and the wind bus, respectively. The size of each symbol indicates the respective generator capacity or load power. The different colors of the buses indicate the different control areas of the power network, and the shaded areas depict the extended control areas.

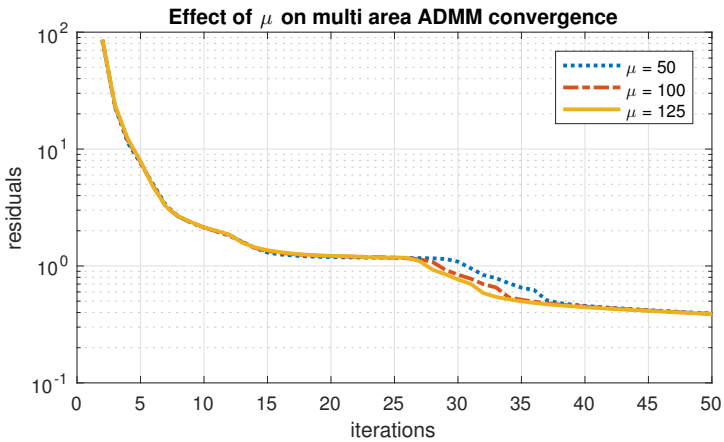
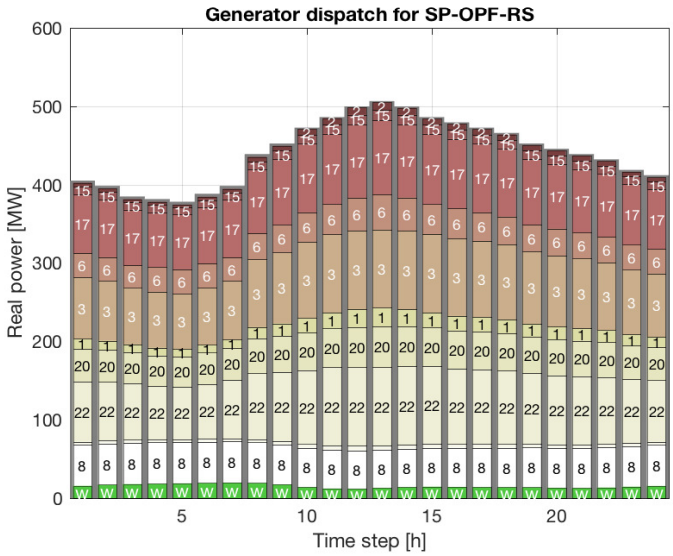
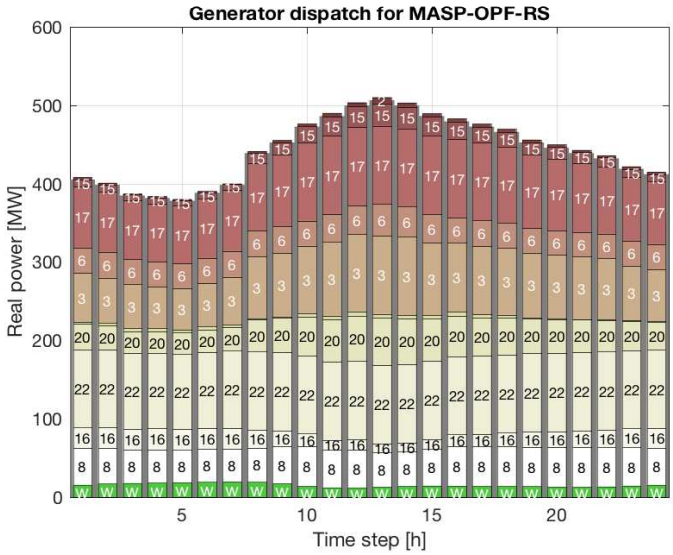


Figure 4.6: Effect of varying μ on the convergence of ADMM algorithm. The algorithm converges for large enough step-size, the convergence is very similar.

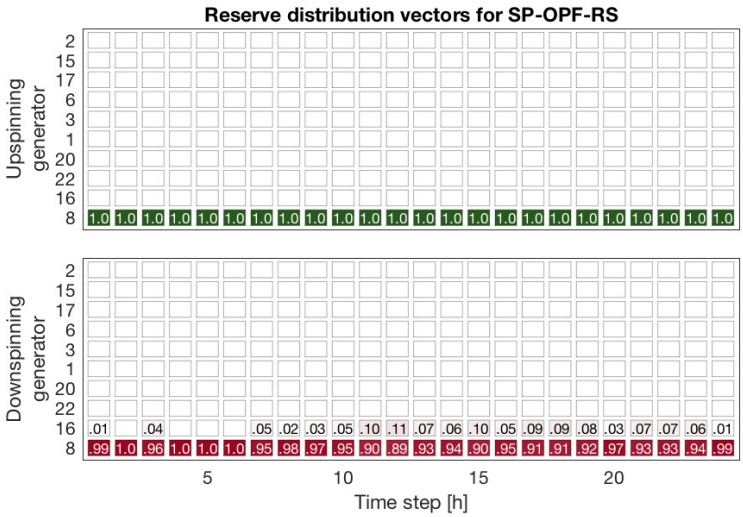


(a) centralized

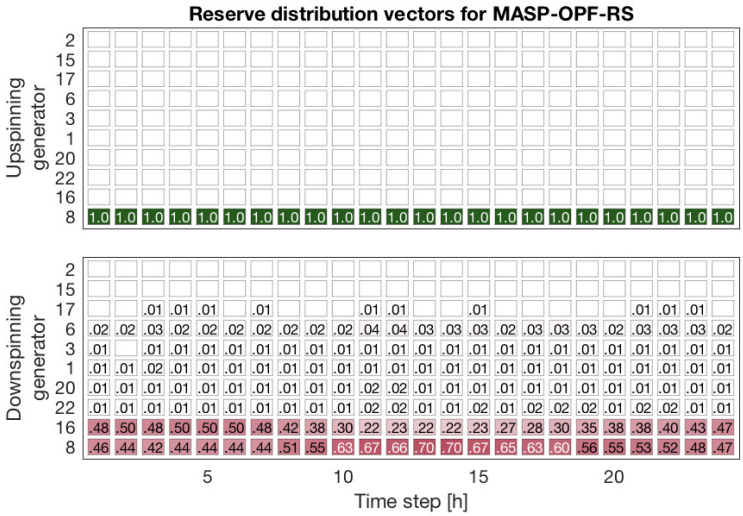


(b) distributed

Figure 4.7: Generator dispatch per hour for centralized and distributed solutions for the 28-bus test case. The grey shaded area corresponds to the total demand per hour. The numbers correspond to the generator buses, and the lowest part of each bar (green) indicates the wind power per hour.

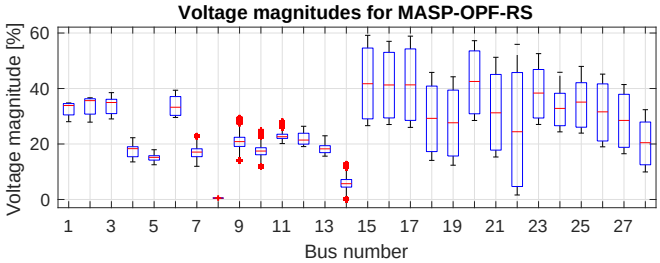


(a) centralized

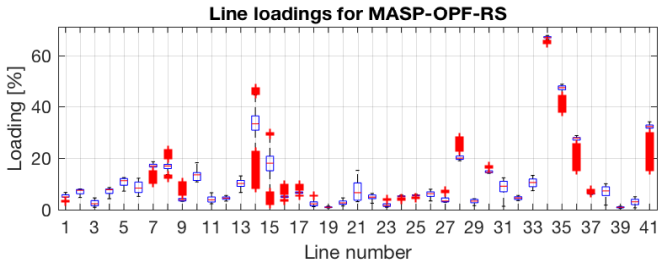


(b) distributed

Figure 4.8: Graphical display of up- and downspinning reserve distribution vectors per generator and hour for centralized and distributed solutions for the 28-bus test case. Darker cells correspond to higher contribution to the reserve power.



(a) Bus voltages



(b) Line loadings

Figure 4.9: Relative bus voltages and line loading for all hours and scenarios per line for the 28-bus test case. The red line represents the median value, edges of each box correspond to the 25th and 75th percentiles, the whiskers extend to 99% coverage, and the red marks denote the data outliers.

tralized and distributed results. The upspinning reserve is completely provided by the generator with the lowest cost which is generator number 8. The downspinning reserve is distributed over the first and second generator for the distributed solution, but in the centralized approach it is mostly provided by generator 8. For both the centralized and distributed solutions, the reserve is distributed over more generators during the peak hours, because the dispatch of the generators is higher in those hours, so less reserve power is available per generator.

We next simulate the resulting solutions with a new set of 10000 wind trajectories to compare the violation levels. The relative voltage magnitudes per bus, defined as $\forall k \in \mathcal{N}$:

$$|v_k^{\text{rel}}| := \frac{|v_k| - |\underline{v_k}|}{|\overline{v_k}| - |\underline{v_k}|},$$

such that a relative voltage magnitude below 0% corresponds to a bus voltage which is below the lower limit, and a relative voltage magnitude greater than 100% indicates a violation of the upper limit. There is no violation of the voltage magnitude limits for any of the results for all time steps and scenarios as shown in Figure 4.9a as predicted by our analysis of the developed methods. The relative line loadings for all hours and trajectories are shown as box plots per line in Figure 4.9b. The number of violating network states divided by the total number of simulations returns an empirical measure of the violation probability. Both the centralized and distributed solutions have a very low violations probability: at most 0.02% and 0.06% at the peak hours, respectively. We can conclude that the probabilistic guarantees are valid for the distributed solution.

4.6. CONCLUSIONS

We developed a framework to carry out a multi-area RS using an AC OPF model with wind power generation by distributed consensus using ADMM. The OPF-RS problem is formulated as a large-scale SDP in infinite-dimensional space, and then a novel affine policy is proposed to provide an approximation for the infinite-dimensional SDP by a finite-dimensional SDP together with explicitly quantified performance of the approximation. The proposed methodology bridges the gap between the DC and AC OPF model of power systems for RS and furnishes the TSOs with a tuning knob associated with the level of affordable probabilistic security.

Using the geographical patterns of the power system, a technique to decompose the large-scale system into a multi-area power network is provided. The consensus ADMM algorithm is then proposed to find a feasible solution for both local and overall multi-area network such that at every iteration, each area of power network solves a small-scale SDP problem, and then communicates some information to its neighbors to achieve consensus. By deriving a Lyapunov-type non-increasing function, it is shown that the proposed algorithm converges as long as Slater's condition holds. Using our distributed stochastic framework, each area can have its own wind information to achieve local feasibility certificates, while preserving overall feasibility of the multi-area power network under mild conditions.

Our theoretical developments have been demonstrated in simulation studies using IEEE benchmark power systems. The violation levels for the decomposed and centralized solutions are checked using power flow simulations to validate our decomposition method which allows for distributed solving of OPF-RS type problems.

5

Privatized Distributed Anomaly Detection for Large-Scale Nonlinear Uncertain Systems

In this chapter we design a privatized distributed anomaly detection framework for large-scale uncertain nonlinear systems. By decomposing such a system into a number of interconnected subsystems, we first design a Local Detector (LD) according to a proven model-based residual approach using a polynomial levelset approximation technique in a probabilistic sense. The resulting LD is equipped with a threshold set that is guaranteed probabilistically to contain the smallest volume of healthy residuals in an arbitrary shape. In addition, the obtained threshold is designed to be highly sensitive to the given signature of anomalies. In this setup, neighboring subsystems should exchange some information about overlapping variables, due to the statistical learning feature of the proposed approach. This might give rise to local privacy concerns of each subsystem. We next present a privatized soft communication scheme between neighboring subsystems, firstly, to preserve their privacy, and secondly to reduce the communication burden. We provide theoretical guarantees to achieve a desired level of privacy, along with a reliability measure of the received information to incorporate into a false alarm ratio of the threshold set design. Simulation studies are provided to illustrate the functionality of our theoretical developments.

5.1. INTRODUCTION

FAULT diagnosis and security for large-scale nonlinear systems, such as critical infrastructures or interconnected Cyber Physical Systems (CPS) has received increasing attention in recent years [76]. Indeed, one way to increase the resiliency of such systems to faults or deliberate cyber attacks is to endow them with architectures capable of monitoring, detecting, isolating and counteracting such anomalies and threats. Advanced model-based fault diagnosis methods have emerged in important industrial sectors, such as aerospace, as fundamental tools for guaranteeing high operational readiness levels and reducing unneeded maintenance costs [44]. A key problem to be solved for widespread industrial adoption is the development of robust methods providing satisfactory, and easy-to-tune performance in terms of the so-called *false alarm ratio* (FAR) and *missed detection ratio* (MDR).

In traditional model-based approaches a time-varying residual is produced, along with a threshold for anomaly detection. Typical threshold designs for model-based detectors are known to be plagued by high conservativeness, which leads to poor detection performance. Ideally, a model-based anomaly detection solution should be *robust* with respect to the unavoidable model and measurement uncertainties, thus having a zero or low FAR. At the same time, it should have good *fault detection* properties, which translates into a negligible MDR. For general nonlinear systems, it is customary to assume the existence of a known, static or dynamic deterministic upper bound on the uncertainties' magnitude, thus allowing to obtain a zero FAR by design [44]. Such a powerful property often comes at the cost of conservative thresholds, which lead to high MDR. Two key reasons stand behind this: the inability of traditional robust deterministic thresholds, such as norm-based or limit-checking, to tightly bound the arbitrarily shaped, possibly non-convex regions to which healthy residuals belong; and the need to account also for large, but possibly rare, values taken by the uncertainties.

5.1.1. RELATED WORKS

For large-scale systems, centralized monitoring and diagnosing architectures are rarely feasible, in contrast with distributed or decentralized ones. Decentralized solutions do not require communication between Local Detectors (LDs) which may lead to unacceptable performance. Distributed methods, which instead do require communication, are then preferable [17, 54, 57, 115, 128, 179, 180]. An unexplored issue in the distributed setting indeed arises from the need of communication between neighboring nodes. In the case where such LDs may be operated by different, possibly competing entities, mutual communication may be opposed as it may lead to leaking privacy-sensitive information. We may consider as an example a smart grid where neighboring diagnosis nodes are each monitoring different subgrids with distributed energy sources and each is managed by its own grid operator. The two grid operators must exchange data about nodes on their respective boundaries in order to allow for grid balancing, but they would rather keep private the way that they are allocating energy supply to their different energy sources and satisfying their energy demand [62, 155]. This highlights the necessity of developing techniques for privatized distributed anomaly detection in large-scale systems.

A powerful and mathematically rigorous concept for dealing with privacy problems is

differential privacy. This concept emerged in the computer science community [48, 49], but recently found applications in control systems as well, see for instance [61–63]. It assumes that each piece of *user* data whose privacy must be protected is contained in a separate record in a database. A trusted party, called *curator*, maintains such database and answers queries posed by possibly adversarial, external parties. Differential privacy aims at modifying the query output to guarantee that no adversary can guess whether a single record is present or has been altered, either by combining the results from several queries, or using side-channel information. In the previous smart grid example, the role of user data is taken by the local input applied to a subgrid, while the query corresponds to the physical phenomena of translating a given input to the boundary nodes' values that are then communicated to adversarial neighbors.

A first goal of this work is to introduce a class of novel, adaptive, parameterized threshold *sets* to reduce the conservatism of the existing thresholds [44, 54, 128, 180] by relaxing the deterministic robust zero-FAR condition, in favor of a more flexible, *probabilistic* one. Through a set-based approach to threshold design, the probability of false alarms will be defined as a user-tunable design parameter, and the detection with respect to a given class of faults will be simultaneously maximized. The use of probabilistic thresholds in model-based fault diagnosis has been investigated previously in the literature (see [44] and the references cited therein), and recently the important case of nonlinear uncertain systems has been considered [18, 107]. The use of sets in fault diagnosis has been inspired by the corpus of works on set-membership system identification [105], which initially addressed the inverse problem of finding, at each time step, the set of system parameters that could be able to explain current measurements, and compare it to a nominal one [13, 69]. Other works considered instead the direct problem of describing the admissible values of the residual in healthy condition using a set [50], with [102] being a notable example in the field of active fault diagnosis.

The second goal of this work is to present the application of a differential privacy mechanism to the proposed distributed anomaly detection framework. In particular, differential privacy will be employed to pre-process data before transmission. Under certain conditions, this leads to privatization of the local control input, by considering the output measurements as an available database to neighboring LDs. This goal will also aim at reducing communication among neighboring LDs, by proposing a soft communication scheme based on a set parametrization technique together with the notion of probabilistically reliable sets. New guarantees for the robustness level of LD in a probabilistic sense will be provided, by taking into consideration the level of probabilistic reliability of neighboring LDs.

5.1.2. CONTRIBUTIONS

The novelty of this work is three-fold and, to the best of the authors knowledge, not addressed in other related works: 1) a set-based threshold design for the LD threshold problem in nonlinear uncertain systems, with the goal of simultaneously guaranteeing robustness to uncertainties in a probabilistic sense, and maximizing detection of a given class of faults; 2) a differentially private communication scheme in a distributed anomaly detection setting, leading to a privacy guarantee for the control input of each subsystem; 3) a soft communication scheme based on a set parametrization technique, together

with a reliability notion of the neighboring LDs expressed in a probabilistic sense. In particular, the main contributions of the present chapter can be further detailed as follows:

a) The introduction of a general formulation for the dynamics of a large-scale nonlinear uncertain system, and its decomposition into a number of interconnected subsystems, by extending existing results from [54].

b) A formal definition of a novel fault detection threshold set design problem, using the concept of probabilistic set approximation through *polynomial superlevel sets* [40]; The proposed approach will require communication between a number of agents, one for each subsystem, and such communication may involve privacy sensitive measurements.

c) The formulation of a cascaded framework for designing threshold sets, through a two-stage chance-constrained optimization problem, in which the first step is aimed at fulfilling a probabilistic robustness constraint, and the second step maximizes the performance of detection with respect to a given class of faults;

d) The introduction of a computationally tractable framework for the solution of the chance-constrained cascade problem, through a randomization technique where the results of the so-called scenario approach are extended to the cascade setting, and a theoretical guarantee for the desired level of approximation is given.

e) A differentially private distributed framework to preserve the privacy of the exchanged information between neighboring subsystems. A pre-processing scheme is proposed to achieve the privacy of the local control inputs using output measurements as the database.

f) A soft communication scheme between neighboring subsystems to overcome the communication bandwidth constraints using a set parametrization technique. Each subsystem shares a set with all its neighboring subsystems together with a reliability of information for the shared set. The reliability measure of neighboring subsystems is incorporated in the probabilistic guarantees for each subsystem in terms of new levels of local false alarms.

5.1.3. STRUCTURE

The structure of this chapter is as follows. Section 5.2 describes an arbitrary nonlinear uncertain system, and its decomposition into a number of interconnected subsystems together with its faults, and provides the formal definition of the proposed threshold set design problem. The proposed framework, a set-based probabilistic threshold design, is described in Section 5.3, where a cascaded chance-constrained optimization problem is formulated, and then the theoretical results for approximately solving such a problem based on the randomization technique is provided. Section 5.4 presents a privacy preserving mechanism drawing on the concept of differential privacy, and a new communication scheme will be introduced. Two different simulation studies are provided in Section 5.5 to illustrate the effectiveness of our proposed privatized setting and distributed fault detection framework. Some final remarks will be given in Section 5.6.

5.2. PROBLEM STATEMENT

Consider a large-scale nonlinear dynamical system $\mathcal{S}$, which can be thought as originating from the interconnection of N smaller subsystems $\mathcal{S}_I$, $I = 1, \ldots, N$. The rationale for this setting is to allow each subsystem to be monitored by a dedicated agent $\mathcal{L}_I$, called *Local Detector* (LD). This leads to a distributed fault detection architecture that will avoid the limitations of centralized solutions, as anticipated in the previous section. Based on the existing approach for model-based distributed fault diagnosis literature, see e.g., [54], such a framework requires each agent to use locally available information, coming from measurements on its subsystem, and information from neighboring agents to compute a dynamical residual. Each agent will compare such local residual to a corresponding local detection threshold, the crossing of which will lead to local detection of a fault or a cyber-attack (anomaly). In the following parts, the detailed problem formulation will be presented.

Remark 15 *It is assumed that the system $\mathcal{S}$, and thus all the subsystems $\mathcal{S}_I$ into which it can be decomposed, are cyber-physical entities that can be subjected to faults or cyber-attacks. This drives the need for monitoring. In contrast, the agents $\mathcal{L}_I$ which are in charge of such monitoring are assumed to be cyber-only entities, that will not be themselves subjected to cyber-attacks*[1]*, except for the indirect effect of attacks directed at their corresponding subsystem $\mathcal{S}_I$.*

5

5.2.1. LARGE-SCALE SYSTEM DYNAMICS

Consider the nonlinear uncertain discrete-time system $\mathcal{S}$ to have the following form

$$\begin{cases} x_{k+1} &= g(x_k, u_k, w_k, f_k) \\ y_k &= x_k + v_k \end{cases}, \tag{5.1}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^n$ are the state, the input and the output of the system at discrete time index k, respectively. The full state information at each time step k is available, albeit corrupted by a measurement uncertainty $v_k \in \mathbb{R}^n$. This can be easily extended to input-output systems similarly to [53]. The variable $w_k \in \mathbb{R}^p$, instead, represents unavoidable modeling uncertainties affecting Eq. (5.1), while $f_k \in \mathcal{F} \subseteq \mathbb{R}^q$ represents a parametrization of anomalies, such as faults and malicious cyber-attacks. Such formulation is purposely as general as possible, and comprises the cases where w_k and f_k affect the overall dynamics $g : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^n$ as additive or multiplicative terms, or where these affect one or more parameters that appear in the definition of g. For instance, if $\mathcal{S}$ models an electrical circuit, w_k could correspond to parametric uncertainties in the electrical resistance of individual conductors or components, f_k could describe eventual open circuit or ground faults, and v_k represents the uncertainty in the measurements provided by a number of voltmeters connected to the circuit. In this respect, the functional dependence of g on f_k, together with its domain set $\mathcal{F}$, describe the class of the possible dynamic anomalies occurring in $\mathcal{S}$. The only assumption we

[1] While this assumption may be criticized with an argument similar to Giovenale's "Quis custodiet ipsos custodes?" (literally, "Who will guard the guards themselves?"), we believe that the case where the LD themselves may fail or be attacked can be addressed as a straightforward extension of the approach presented in this chapter, but its inclusion will impair clarity.

will adopt on the functional dependence of g on w_k and f_k is that $w_k = 0$ and $f_k = 0$ corresponds to the nominal and normal behavior of $\mathcal{S}$, that is in the absence of uncertainties and anomalies. The following two technical assumptions are also needed for the upcoming analysis.

Assumption 13 *No anomaly acts on the system before the anomaly occurrence time k_f, that is $f_k = 0$, for $0 \leq k < k_f$. Moreover, the variables x_k and u_k remain bounded before and after the occurrence of an anomaly, i.e., there exist some stability regions $\mathbb{S} := \mathbb{S}^x \times \mathbb{S}^u \subset \mathbb{R}^n \times \mathbb{R}^m$ such that $(x_k, u_k) \in \mathbb{S}$, for all k.*

Assumption 14 *The variables w_k and v_k are random variables defined on some probability spaces $(\mathcal{W}, \mathfrak{B}(\mathcal{W}), \mathbb{P}_{\mathcal{W}})$, and $(\mathcal{V}, \mathfrak{B}(\mathcal{V}), \mathbb{P}_{\mathcal{V}})$, respectively, where $\mathcal{W} \subseteq \mathbb{R}^p$, $\mathcal{V} \subseteq \mathbb{R}^n$, $\mathfrak{B}(\cdot)$ denotes a Borel σ-algebra, and $\mathbb{P}_{\mathcal{W}}, \mathbb{P}_{\mathcal{V}}$ are a probability measure defined over $\mathcal{W}, \mathcal{V}$, respectively. Furthermore, w_k and v_k are not correlated and are independent from x_k, u_k and f_k, for all k.*

Furthermore, g will be assumed to be differentiable and Lipschitz with respect to u. It is important to note that, as in [40], we do not require the sample spaces $\mathcal{W}, \mathcal{V}$ and the probability measures $\mathbb{P}_{\mathcal{W}}, \mathbb{P}_{\mathcal{V}}$ to be known explicitly, as it will be explained in Section 5.3. A distinctive, practical advantage of the proposed framework is that its implementation requires only the availability of a finite number of samples from the uncertain variables, which can for instance be obtained from historical data. Finally, we introduce the digraph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ to represent the structure of $\mathcal{S}$ similarly to [54], where $\mathcal{N} := \{x_k^{(1)}, \ldots, x_k^{(n)}\} \cup \{u_k^{(1)}, \ldots, u_k^{(m)}\} \cup \{w_k^{(1)}, \ldots, w_k^{(p)}\} \cup \{f_k^{(1)}, \ldots, f_k^{(q)}\}$ is the node set and $\mathcal{E}$ is the edge set. The fact that the edge $e = (a, b) \in \mathcal{E}$ means that the variable a influences the dynamics of the variable b, with $a, b \in \mathcal{N}$. In general, we do not expect that the dynamics of every state component $x_k^{(i)}$ depends on all the components of all the other variables. This means that the graph $\mathcal{G}$ will not be complete in general.

5.2.2. SUBSYSTEM DYNAMICS

Consider each subsystem $\mathcal{S}_I$ for all $I \in \{1, \ldots, N\}$ to be characterized by a local state $x_{I,k} \in \mathbb{R}^{n_I}$, a local input $u_{I,k} \in \mathbb{R}^{m_I}$ and a local output $y_{I,k} \in \mathbb{R}^{n_I}$ such that each of them is composed of a suitable subset of components from x_k, u_k and y_k. It is assumed that each I-th subsystem is defined by means of an n_I-tuple $\mathcal{I}_I$, called *extraction index tuple* [54], containing indices to the components of x_k making up $x_{I,k}$. Formally, this can be written as $x_{I,k} := \mathrm{col}(x_k^{(i)} : i = \mathcal{I}_I^{(j)}, j = 1, \ldots, n_I)$, and similarly for $y_{I,k}$ and $v_{I,k}$. The local input $u_{I,k}$ is instead built with all the components of u_k affecting at least one component of $x_{I,k}$, that is $u_{I,k} := \mathrm{col}(u_k^{(j)} : (u_k^{(j)}, x_k^{(i)}) \in \mathcal{E}, i \in \mathcal{I}_I, j \in \{1, \ldots, m\})$. Similarly, we can build $w_{I,k} := \mathrm{col}(w_k^{(j)} : (w_k^{(j)}, x_k^{(i)}) \in \mathcal{E}, i \in \mathcal{I}_I, j \in \{1, \ldots, p\})$ and $f_{I,k} := \mathrm{col}(f_k^{(j)} : (f_k^{(j)}, x_k^{(i)}) \in \mathcal{E}, i \in \mathcal{I}_I, j \in \{1, \ldots, q\})$. The local input $u_{I,k}$ is built with all the components of u_k that *structurally affect* at least one component of $x_{I,k+1}$, and similarly for building the local $w_{I,k}$ and $f_{I,k}$.

Definition 3 *A variable c* structurally affects *a variable $a = b(c, d)$ through a multi-input function b, and is written $c \xrightarrow{b} a$, if there exists at least a pair of distinct values $\bar{c}$ and $\bar{c}'$ and a value $\bar{d}$ such that $\bar{a} = b(\bar{c}, \bar{d})$ is distinct from $\bar{a}' = b(\bar{c}', \bar{d})$.*

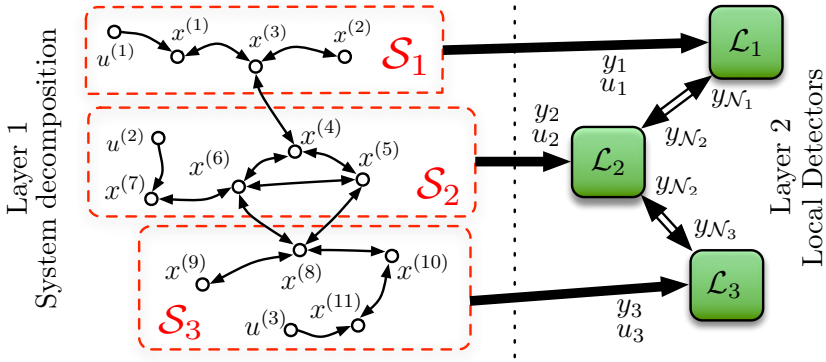


Figure 5.1: The proposed distributed anomaly detection architecture. On the left side, the decomposition of the original system $\mathcal{S}_I$ is shown, where $I = 1,2,3$: thin black lines represent causal dependency between variables. On the right, the communication and the acquisition of measurements by the agents $\mathcal{L}_I$ is depicted, where $I \in \{1,2,3\}$.

5

It is important to stress that here we have a structural knowledge of the effect of w_k and f_k on each component of g. This does not preclude the capability for our problem formulation to capture the case where the uncertainty, or the anomaly, are non parametric and arbitrary. For instance, we could assume in this case the dynamics to be decomposable as $g(x_k, u_k, w_k, f_k) = g^\star(x_k, u_k) + w_k + f_k$, where $g^\star$ represents the nominal and normal dynamics, and w_k and f_k are arbitrarily varying signals, but respecting Assumption 13 and Assumption 14. We now formally define a system decomposition $\mathcal{D}$ similarly to [54].

Definition 4 *A* decomposition $\mathcal{D}$ *of dimension* N *of the large-scale system* $\mathcal{S}$ *is a set* $\mathcal{D} \triangleq \{\mathcal{S}_1, ..., \mathcal{S}_N\}$ *made of* N *subsystems, defined through a set* $\{\mathcal{I}_1, \dots, \mathcal{I}_N\}$ *of extraction index tuples, such that for each* $I \in \{1, \dots, N\}$ *the following conditions hold: (a) the subdigraph of* $\mathcal{G}$ *induced by* $\mathcal{I}_I$ *must be weakly connected, that is, each component of* $x_{I,k}$ *must act on or must be acted on by at least another component of* $x_{I,k}$*, (b)* $\mathcal{I}_I \neq \emptyset$*, (c)* $1 \leq \mathcal{I}_I^{(j)} \leq n$*, for each* $j \in \{1, \dots, n_I\}$*, (d)* $\mathcal{I}_I \cap \mathcal{I}_J = \emptyset, \forall I \neq J : I, J \in \{1, \dots, N\}$*, and (e)* $\bigcup_{I=1}^{N} \mathcal{I}_I = \{1, ..., n\}$*.*

The set $\mathcal{D}$ contains *non-overlapping* subsystems in contrast to [54]. This is imposed via conditions (d) and (e) in Definition 4 that require the decomposition of the state x_k into local states $x_{I,k}$ to be disjoint and cover all the components of x_k. We defined the output y_k in (5.1) such that the same disjointness and covering condition translates to the local output variables $y_{I,k}$ and their associated measurement uncertainties $v_{I,k}$. Due to the generality of the g dynamics function, that may functionally depend on w_k and f_k, it may not have a disjointness property for all the local $w_{I,k}$ and $f_{I,k}$, although still the covering condition will hold. This means that the same component of the uncertainty w_k or of the anomaly parameter f_k may influence more than one subsystem. We can

proceed further and describe the dynamics of the generic subsystem $\mathcal{S}_I$ as

$$\begin{cases} x_{I,k+1} &= g_I(x_{I,k}, u_{I,k}, x_{\mathcal{N}_I,k}, w_{I,k}, f_{I,k}) \\ y_{I,k} &= x_{I,k} + v_{I,k} \end{cases}, \tag{5.2}$$

where the *local dynamical function* $g_I : \mathbb{R}^{n_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^{n_{\mathcal{N}_I}} \times \mathbb{R}^{p_I} \times \mathbb{R}^{q_I} \to \mathbb{R}^{n_I}$ can be simply obtained by taking in the right order the components of g that are contained in the index tuple $\mathcal{I}_I$. Since, as in general we cannot assume that all the resulting subsystems $\mathcal{S}_I$ are decentralized, i.e., their dynamics depend not only on the local state x_I, we introduced the *interconnection variable* $x_{\mathcal{N}_I,k}$ similarly to [54].

Definition 5 *The* interconnection variable $x_{\mathcal{N}_I,k} \in \mathbb{R}^{n_{\mathcal{N}_I}}$ *of the subsystem* $\mathcal{S}_I$ *is the vector* $x_{\mathcal{N}_I,k} := col(x_k^{(j)} : (x_k^{(j)}, x_{I,k}^{(i)}) \in \mathcal{E},\ i \in \{1,\dots,n_I\},\ j \in \{1,\dots,n\})$.

The role of $x_{\mathcal{N}_I,k}$ is to describe the functional dependence of the local dynamics g_I on state components from other subsystems, which we will call *neighboring subsystems* or simply *neighbors*. The set of all the neighbors of $\mathcal{S}_I$ will be denoted by $\mathcal{N}_I$. Note that since Assumption 13 holds for the original system $\mathcal{S}$, it will also continue to do so for every subsystem. We can introduce a stability region $\mathcal{S}_I$ for each one, where the local state x_I and input u_I are assumed to always belong. We can also easily build the sets $\mathcal{V}_I$, $\mathcal{W}_I$, $\mathcal{F}_I$ and $\mathcal{V}_{\mathcal{N}_I}$ of, respectively: the local measurement and modeling uncertainties, the local fault parameters, and the measurement uncertainties for the interconnection variable.

5.2.3. RESIDUAL GENERATOR

In the considered distributed setting, an agent $\mathcal{L}_I$ shall compute a local dynamic residual $r_{I,k}$ and compare it to a local dynamic detection threshold for the purpose of detecting anomalies. The residual shall be defined as $r_{I,k} := y_{I,k} - \hat{y}_{I,k}$, and can be computed as the output estimation error of the following nonlinear observer:

$$\begin{cases} \hat{x}_{I,k+1} &= g_I(y_{I,k}, u_{I,k}, y_{\mathcal{N}_I,k}, 0, 0) + \Lambda(\hat{y}_{I,k} - y_{I,k}) \\ \hat{y}_{I,k} &= \hat{x}_{I,k} \end{cases} \tag{5.3}$$

where $\hat{x}_I$, $\hat{y}_I \in \mathbb{R}^{n_I}$ are, respectively, the local state and output estimates, $y_{\mathcal{N}_I,k} \in \mathbb{R}^{n_{\mathcal{N}_I}}$ are the measurements of the interconnection variables $x_{\mathcal{N}_I,k}$, $\Lambda \triangleq \mathrm{diag}(\lambda^i,\ i = 1\dots n_I)$ is a diagonal matrix, and λ^i with $|\lambda^i| < 1$ denotes a filtering parameter chosen to guarantee the stability of the estimator. It can be seen that the estimator in (5.3) is defined using only quantities that are supposed to be available at run-time to the agent $\mathcal{L}_I$, namely: the nominal local dynamics g_I, the local outputs $y_{I,k}$ and inputs $u_{I,k}$, and the measurements $y_{\mathcal{N}_I,k}$ of the interconnection variables of the neighboring agents $\mathcal{L}_J$, with $J \in \mathcal{N}_I$. The local use of $y_{\mathcal{N}_I,k}$ requires some form of regular communication between neighboring agents, and the resulting privacy implications will be addressed in Section 5.4.

By using Equations (5.1) and (5.3), we can write the residual dynamics as

$$r_{I,k+1} = \Lambda\, r_{I,k} + \delta_{I,k}\,, \tag{5.4}$$

where we introduced the *total uncertainty* $\delta_{I,k}$, which is a stochastic process representing the uncertain part of the residual dynamics:

$$\begin{aligned}\delta_{I,k} &:= g_I(x_{I,k}, u_{I,k}, x_{\mathcal{N}_I,k}, w_{I,k}, f_{I,k}) - g(y_{I,k}, u_{I,k}, y_{\mathcal{N}_I,k}, 0, 0) + v_{I,k+1} \\ &= g_I(y_{I,k} - v_{I,k}, u_{I,k}, y_{\mathcal{N}_I,k} - v_{\mathcal{N}_I,k}, w_{I,k}, f_{I,k}) - g(y_{I,k}, u_{I,k}, y_{\mathcal{N}_I,k}, 0, 0) + v_{I,k+1}\,.\end{aligned} \tag{5.5}$$

Thanks to Assumption 13 and Assumption 14, it follows that $\delta_{I,k}$ is a random variable on a probability space $(\Delta_{I,k}, \mathfrak{B}(\Delta_{I,k}), \mathbb{P}_{\Delta_{I,k}})$, where $\Delta_{I,k}$ is a time-varying set defined as follows.

Definition 6 *The time-varying* total uncertainty set $\Delta_{I,k} \subset \mathbb{R}^{n_I}$ *at time index k is defined as*

$$\Delta_{I,k} := \{\delta_{I,k} \,|\, w_{I,k} \in \mathcal{W}_I,\, f_{I,k} \in \mathcal{F}_I,\, v_{I,k} \in \mathcal{V}_I,\, v_{I,k+1} \in \mathcal{V}_I,\, v_{\mathcal{N}_I,k} \in \mathcal{V}_{\mathcal{N}_I}\},$$

where $\delta_{I,k}$ *is computed according to* (5.5).

We will introduce also the following definition, as a special case of Definition 6 which is of interest for the time instants $k < k_f$ during which the subsystem $\mathcal{S}_I$ is in normal operation.

Definition 7 *The time-varying* normal total uncertainty set $\Delta^0_{I,k} \subset \mathbb{R}^{n_I}$ *at time index k is defined as*

$$\Delta^0_{I,k} := \{\delta_{I,k} \,|\, w_{I,k} \in \mathcal{W}_I,\, f_{I,k} \in \{0\},\, v_{I,k} \in \mathcal{V}_I,\, v_{I,k+1} \in \mathcal{V}_I,\, v_{\mathcal{N}_I,k} \in \mathcal{V}_{\mathcal{N}_I}\},$$

where $\delta_{I,k}$ *is computed according to* (5.5).

The role of $\Delta_{I,k}$ and $\Delta^0_{I,k}$ is to quantify the range of possible values that the total uncertainty $\delta_{I,k}$ can take, respectively, in an arbitrary condition during which an anomaly *may* be present, and a normal condition where an anomaly is not. As it will be shown in the next subsection and the remainder of the present one, $\Delta^0_{I,k}$ will have a central role in deriving a probabilistically robust detection threshold, while $\Delta_{I,k}$ will be instrumental in improving detectability.

We can now introduce a compact notation for the *residual generator* described by Equations (5.3) to (5.5), through a mapping function $\Sigma_I : \mathbb{R}^{n_I} \times \mathbb{R}^{n_I} \to \mathbb{R}^{n_I}$ defined as

$$r_{I,k+1} := \Sigma_I(r_{I,k}, \delta_{I,k})\,. \tag{5.6}$$

The mapping from the uncertain variable $\delta_{I,k} \in \Delta_{I,k}$ to the residual variables $r_{I,k+1}$ is measurable, so that the residual signal $r_{I,k+1}$ can be viewed as a random variable on the same probability space as $\delta_{I,k}$.

Given these preliminaries, it is now possible to write the following two fundamental definitions (see Figure 5.2).

Definition 8 *The time-varying* residual set $\mathcal{R}_{I,k+1}$ *at time index* $k+1$ *is defined as the image of the set* $\Delta_{I,k}$ *through* Σ_I*, that is*

$$\mathcal{R}_{I,k+1} := \Sigma_I(r_{I,k}, \Delta_{I,k}) = \{r_{I,k+1} \,|\, r_{I,k+1} = \Sigma_I(r_{I,k}, \delta_I),\, \delta_I \in \Delta_{I,k}\}.$$

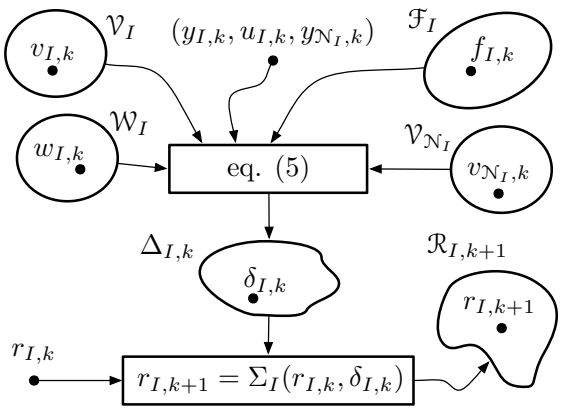


Figure 5.2: The residual set $\mathcal{R}_{I,k+1}$ can be thought of as the set obtained by computing the output Σ_I while letting $\delta_{I,k}$ vary over its domain $\Delta_{I,k}$ and fixing the residual $r_{I,k}$ to its actual value. The domain $\Delta_{I,k}$ in turn is computed through (5.5) by letting $v_{I,k}$, $w_{I,k}$, $f_{I,k}$ and $v_{\mathcal{N}_I,k}$ vary over their respective domains, and fixing the local output and input $y_{I,k}$ and $u_{I,k}$, as well as the interconnection variables measurement $y_{\mathcal{N}_I,k}$, to their actual values. The healthy residual set $\mathcal{R}^0_{I,k+1}$ can be obtained similarly, but by fixing the value $f_{I,k} = 0$.

Definition 9 *The time-varying* normal residual set $\mathcal{R}^0_{I,k+1}$ *at time index* $k+1$ *is defined as the image of the set* $\Delta^0_{I,k}$ *through* Σ_I*, that is*

$$\mathcal{R}^0_{I,k+1} := \Sigma_I(r_{I,k}, \Delta^0_{I,k}) = \{r_{I,k+1} \,|\, r_{I,k+1} = \Sigma_I(r_{I,k}, \delta_I), \delta_I \in \Delta^0_{I,k}\}.$$

For ease of notation, when there is no ambiguity, in the rest of the chapter we will drop the index I to denote that a quantity refers to the generic subsystem $\mathcal{S}_I$ or the generic agent $\mathcal{L}_I$. The index $\mathcal{N}$ will be retained to indicate the neighbor set of the generic subsystem or agent.

5.2.4. ANOMALY DETECTION THRESHOLD DESIGN PROBLEM

After building a residual generator, an important problem in anomaly detection is to design a threshold with suitable robustness and detection performance guarantees. Traditional solutions to the deterministic robust threshold design problem (see [55] for a survey) seek a threshold that bounds all possible values of the normal residual $r_{k+1} \in \mathcal{R}^0_{k+1}$, thus guaranteeing zero false alarm ratio (FAR) by design. In the norm-based approach, a scalar threshold τ bounding $\|r_{k+1}\|$ is sought, whereas in the limit checking approach a vector is found such that its j–th component $\tau^{(j)}$ bounds $|r^{(j)}_{k+1}|$. In order to minimize the missed detection ratio (MDR), such thresholds should be made as small as possible, a goal which we may express as:

$$\text{(I)} \begin{cases} \min\limits_{\tau \in \mathbb{R}} & \tau \\ \text{s.t.} & \|r\| \leq \tau \end{cases}, \quad \text{(II)} \begin{cases} \min\limits_{\tau^{(j)} \in \mathbb{R}} & \tau^{(j)} \\ \text{s.t.} & |r^{(j)}| \leq \tau^{(j)} \end{cases}, \tag{5.7}$$

where the constraints should hold for all $r \in \mathcal{R}^0_{k+1}$ and problem (II) should be solved for each $j = 1, \cdots, n_I$ independently. If we interpret the thresholds resulting from (I) and (II)

in a set-theoretic framework, it is easy to see that they lead, respectively, to the smallest ball and axis-aligned box in $\mathbb{R}^n$ containing the healthy residual set $\mathcal{R}^0_{k+1}$, see Figure 5.3(a) and Figure 5.3(b). Such solutions are clearly conservative, for two reasons. First, they use simple and convex manifolds to bound the set $\mathcal{R}^0_{k+1}$, which in general can have an arbitrary shape and be non–convex, because of the assumed nonlinearity of the system dynamics g. Secondly, bounding the entire set $\mathcal{R}^0_{k+1}$ does indeed lead to a deterministic guarantee on the FAR, but ignores the fact that in real applications some values of r_{k+1} may have a negligible probability of being produced, and as such they could be excluded in the threshold design procedure, without practical consequences.

It is the stated objective of the present chapter to address both aforesaid sources of conservatism. First, we will introduce an adaptive, parametrized set-based threshold, which could *approximate* arbitrarily well the shape of the set $\mathcal{R}^0_{k+1}$. Then, we will relax the deterministic, hard constraints of problems (I) and (II) with a probabilistic guarantee, thus reaching a desired level of FAR. We finally propose a threshold design framework, which at the same time aims to reduce the MDR.

In order to formalize the above ideas, we first define $\mathcal{T}_k \subseteq \mathbb{R}^n$ as an adaptive *threshold set* at time index k for anomaly detection, and then introduce the following novel concept.

Definition 10 An anomaly set $\mathcal{F}' \subseteq \mathcal{F}$ is said to be detectable by an adaptive threshold set $\mathcal{T}_{k+1}$ and a residual generator Σ *if* $\forall\, f_k \in \mathcal{F}'$, $\exists\, r_{k+1} \in \mathcal{R}_{k+1}$ *and* $\delta_k \in \Delta_k$ *such that* $r_{k+1} \notin \mathcal{T}_{k+1}$, *with* $r_{k+1} = \Sigma(r_k, \delta_k)$.

Definition 11 *Given the residual generator function* Σ *in* (5.6) *and a fixed* $\alpha \in [0,1]$, *an adaptive threshold set* $\mathcal{T}_k$ *is said to be* probabilistically α–robust *with respect to the random total uncertainty* $\delta_k \in \Delta^0_k$, *if*

$$\mathcal{V}(\mathcal{T}_{k+1}) := \mathbb{P}\left[r_{k+1} \notin \mathcal{T}_{k+1} \,\middle|\, r_{k+1} \in \mathcal{R}^0_{k+1}\right] < 1 - \alpha\,, \tag{5.8}$$

where $\mathcal{V}(\mathcal{T}_{k+1})$ *is the* violation probability (FAR) *of the normal residuals* $r_{k+1} \in \mathcal{R}^0_{k+1}$.

We now describe the adaptive threshold set $\mathcal{T}_k$, using a *generalized indicator function* $\mathbf{1}_{\mathcal{T}}(r, \theta_k) : \mathbb{R}^{n_I} \times \mathbb{R}^{t_I} \mapsto \mathbb{R}$ parametrized by a time varying vector $\theta_k \in \mathbb{R}^{t_I}$, as follows:

$$\mathcal{T}_k := \{r \in \mathbb{R}^{n_I} \,|\, \mathbf{1}_{\mathcal{T}}(r, \theta_k) \geq c\}\,. \tag{5.9}$$

This yields a *c–superlevel set* [23] of $\mathbf{1}_{\mathcal{T}}$, for any value of c, as the adaptive threshold set $\mathcal{T}_k$, while Definition 11 leads to an expected FAR better then $1 - \alpha$.

Remark 16 *A fundamental point in Definition 11 is that the probabilistic condition* (5.8) *is expressed in terms of the* future *normal residual belonging to the* future *threshold set. While at a given time the actual residual* r_k *is a computable deterministic quantity, its future value* r_{k+1} *is a random variable due to the fact that it linearly depends on the random variable* δ_k. *It thus makes sense to consider the probability, measured with respect to the probability space on which* δ_k *is defined, that in normal conditions* r_{k+1} *will belong to the set* $\mathcal{T}_{k+1}$. *The latter is a deterministic set that shall be computed at the current time, as will be highlighted in the next sections. We stress again the fact that Definition 11 does not require* $\mathcal{T}_{k+1}$ *to be a (proper) subset of* $\mathcal{R}^0_{k+1}$, *but only that* $\mathcal{T}_{k+1}$ *approximates it in the given*

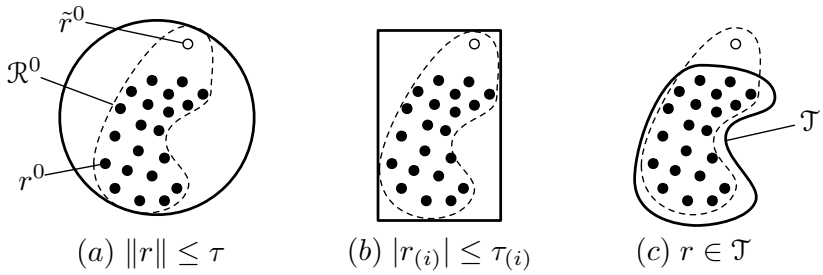


Figure 5.3: A pictorial, intuitive comparison of different robust threshold and residual evaluation approaches. Representative normal values r^0 of the residual are drawn as filled black circles, while rare ones $\tilde{r}^0$ are drawn as empty circles. For convenience, in all cases the evaluation condition is represented as membership in a set drawn with a tick line. a) Norm based. b) Limit checking. c) The proposed, probabilistic set-based approach.

probabilistic sense. This distinction will be the key in designing the solution proposed in Section 5.3.

5

5.3. DISTRIBUTED PROBABILISTIC THRESHOLD SET DESIGN

This section presents a probabilistic framework that will allow to design a local threshold set for each agent $\mathcal{L}_I$, in a distributed way. First, an optimization problem will be formulated in order to minimize the FAR in the probabilistic sense of Definition 11. Then, a second optimization problem will be introduced to maximize detectability, that is to minimize the MDR with respect to a given set of anomalies. Finally, a unified approach will be proposed to handle minimization of both FAR and of MDR by means of a two-stage program in a cascaded setting. Such a problem formulation leads to a two-stage chance-constrained optimization problem, that is in general non-convex and hard to solve. Therefore, a computationally tractable methodology will be provided, which approximates such a problem via a randomization technique. It is important to highlight that existing theoretical results on such an randomization technique will be extended in order to cope with the proposed cascaded formulation.

5.3.1. SET-BASED THRESHOLD DESIGN

In the proposed approach, we will assume the indicator function $\mathbf{1}_{\mathcal{T}}(r,\theta_k)$ to be a polynomial function of given degree d, with θ containing the polynomial coefficients in a arbitrary order. Denoting by $\pi_\xi(r)$ a vector of monomials of degree up to $\xi := \lceil d/2 \rceil$[2], we can conveniently define $\mathbf{1}_{\mathcal{T}}(r,\theta_k) := \pi_\xi(r)^\top G(\theta_k)\pi_\xi(r)$, where $G(\theta_k)$ is a matrix depending on the coefficients contained in θ_k, which is the quantity to be determined during the proposed design procedure.

We first formulate a chance-constrained optimization problem to obtain the minimum volume threshold set $\mathcal{T}_k$ that fulfills Definition 11 for a user-specified α:

$$\begin{cases} \min\limits_{\theta} & \mathrm{vol}(\mathcal{T}_k) \\ \text{s.t.} & \mathcal{V}(\mathcal{T}_k) < 1-\alpha\,, \end{cases} \tag{5.10}$$

[2] $\lceil\cdot\rceil$ is the ceiling operator which returns the smallest integer greater than or equal to its argument.

where $\mathrm{vol}(\mathcal{T}_k) := \int_{\mathcal{T}_k} dr$ is the volume or Lebesgue measure of $\mathcal{T}_k$. The rationale for seeking the minimum volume set is to minimize the threshold conservativeness, that is to maximize detectability. Note that the inequality defined in (5.10) can be simply also considered as less than equal without affecting our developments in the following parts. The following proposition and remark provide a formal description of such intuitive motivation.

Proposition 5 (Detectability) *A necessary condition for the anomaly $\mathcal{F}' \subset \mathcal{F}$ to be detectable is that $\exists\, k \geq k_f$ such that $\mathcal{R}_{k+1}^{\mathcal{F}'} \cap \overline{\mathcal{T}_{k+1}} \neq \emptyset$, where $\overline{\mathcal{T}_{k+1}}$ denotes the complement of set $\mathcal{T}_{k+1}$.*

Proof. The proof is straightforward and is based on Assumption 13, Definition 8 and Definition 10. The details are omitted for brevity. □

Remark 17 *The condition in Proposition 5 is necessary but is not sufficient in general, as it still allows for $\mathcal{R}_{k+1}^{\mathcal{F}'} \cap \mathcal{T}_{k+1} \neq \emptyset$. This means that at least one of the possible realizations of the random variable δ_k can cause the faulty residual r_{k+1} to be inside $\mathcal{T}_{k+1}$, thus preventing detection. Since at time k the next residual r_{k+1} is a random variable, the probability of such an event can in theory be computed by integrating the probability density function (pdf) of the faulty r_{k+1} over the threshold set $\mathcal{T}_{k+1}$. As we do not require in this work any specific assumption on δ_k and thus on the pdf of r_{k+1}, then minimizing the volume of $\mathcal{T}_{k+1}$ turns out to be a reasonable heuristic in order to minimize the integral of the pdf of r_{k+1} on it. A more rigorous investigation of this problem will be presented in the next subsection.*

The proposed optimization problem (5.10) is in general non-convex and hard to solve, due to the numerical complexity arising from the minimum volume objective, and the probabilistic constraint. Following [59] and [39], we restrict the range of our indicator function to be non-negative which yields $\mathbf{1}_{\mathcal{T}}$ to be a polynomial sum-of-squares (SoS) and $G(\theta_k)$ to be a symmetric Gram matrix. We are now able to bound the objective function using the relation:

$$\mathrm{vol}(\mathcal{T}_k) = \int_{\mathcal{T}_k} dr = \int_{\mathcal{B}} \mathbf{1}_{\mathcal{T}}(r,\theta_k)\, dr \leq \mathrm{trace}(G(\theta_k)M),$$

where $\mathcal{B} \in \mathbb{R}^n$ is an arbitrary compact set so that $\mathcal{T}_k \subseteq \mathcal{B}$ and $M := c^{-1}\int_{\mathcal{B}} \pi_\xi(r)\pi_\xi(r)^\top dr$ denotes the matrix of moments of the Lebesgue measure on $\mathcal{B}$ in basis $\pi_\xi(r)$. We can reformulate (5.10) as follows:

$$\min_{\theta,\gamma} \quad \gamma \tag{5.11a}$$

$$\text{s.t.} \quad G(\theta) \succeq 0\,, \tag{5.11b}$$

$$\mathrm{trace}(G(\theta)M) \leq \gamma\,, \tag{5.11c}$$

$$\mathbb{P}\left[\mathbf{1}_{\mathcal{T}}(r_{k+1}^0,\theta) \geq c\right] \geq \alpha\,. \tag{5.11d}$$

Constraint (5.11b) imposes the positive semidefiniteness of $G(\theta_k)$ in order to constrain $\mathbf{1}_{\mathcal{T}_k}$ to be SoS. We also introduced the auxiliary variable γ to allow us to upper bound the objective function, using epigraphical reformulation [11], as in constraint (5.11c). As it is explained in Remark 16, the probabilistic constraint in (5.11d) is measured with respect to the underlying random variable $\delta_k \in \Delta_k^0$.

5

5.3.2. MAXIMIZATION OF ANOMALY DETECTABILITY

To maximize the probability of a successful detection at each time step k, we propose to design $\mathcal{T}_k$ to be not only as small as possible, but also as much distant as possible from $\mathcal{R}_k^{\mathcal{F}'}$, in the sense that is described next. We assume the availability[3] of a description of the set $\mathcal{R}_k^{\mathcal{F}'}$ through a polynomial SoS generalized indicator function $\mathbf{1}_{\mathcal{R}^{\mathcal{F}'}}(r,\psi_k) := \pi_\xi(r)^\top G(\psi_k)\pi_\xi(r)$ with the same degree d and the same monomial basis $\pi_\xi(r)$ as $\mathbf{1}_{\mathcal{T}}$. By denoting with ψ^* the value of ψ_k so that $\mathbf{1}_{\mathcal{R}^{\mathcal{F}'}}(r,\psi^*) \geq c$ for all $r \in \mathcal{R}_k^{\mathcal{F}'}$, we can formulate an optimization problem for maximizing the distance between the threshold set $\mathcal{T}$ and the abnormal residual set $\mathcal{R}_k^{\mathcal{F}'}$:

$$\begin{cases} \max\limits_{\theta} & \left\| \mathbf{1}_{\mathcal{T}}(r,\theta) - \mathbf{1}_{\mathcal{R}^{\mathcal{F}'}}(r,\psi^*) \right\|_\infty \\ \text{s.t.} & \|\theta\|_\infty \leq \bar{c}\,, \end{cases} \tag{5.12}$$

where $\bar{c}$ is a given constant parameter. The objective function aims at maximizing the Chebyshev distance between $\mathbf{1}_{\mathcal{T}_k}$ and $\mathbf{1}_{\mathcal{R}_k^{\mathcal{F}'}}$, which is also known as the polynomial height [184]. Since both of them share the same monomial basis vector $\pi_\xi(r)$, this leads to the maximization of the distance $\|G(\theta) - G(\psi^*)\|_\infty$ between their Gram matrices [184]. Notice that the second constraint in (5.12) is added to ensure that the solutions remain bounded.

We next propose a cascade of the two optimization problems (5.11) and (5.12), which is in general hard to solve due to the chance constraint (5.11d).

5.3.3. CASCADED PROBLEM FORMULATION SCHEME

In order to attain our stated goal of obtaining a probabilistically α–robust threshold set, that also maximizes detection of the anomaly $f \in \mathcal{F}'$ in the sense of Definition 10 as formulated in (5.12), we propose a cascade of two chance-constrained optimization problems as follows:

$$\begin{cases} \min\limits_{\theta,\gamma} & \gamma \\ \text{s.t.} & G(\theta) \succeq 0\,, \quad \operatorname{trace}(G(\theta)M) \leq \gamma\,, \\ & \mathbb{P}\left[\mathbf{1}_{\mathcal{T}}(r_{k+1}^0,\theta) \geq c\right] \geq \alpha \end{cases} \tag{5.13a}$$

$$\begin{cases} \max\limits_{\theta} & \|G(\theta) - G(\psi^*)\|_\infty \\ \text{s.t.} & G(\theta) \succeq 0\,, \quad \operatorname{trace}(G(\theta)M) \leq \gamma^*\,, \\ & \mathbb{P}\left[\mathbf{1}_{\mathcal{T}}(r_{k+1}^0,\theta) \geq c\right] \geq \alpha \end{cases} \tag{5.13b}$$

where the quantity γ^* is the optimal cost obtained by solving the first stage (5.13a), while (5.13) has to be solved sequentially in a *lexicographic* (multi-objective) sense [101]. Note that the unnecessary constraint in (5.12) is dropped due to the introduced bound in (5.13b).

[3] A simple way to obtain such a set is to solve a problem analogous to (5.11), where the constraint (5.11d) is imposed deterministically on all elements of $\mathcal{R}_k^{\mathcal{F}'}$.

Remark 18 *The first problem (5.13a) aims at determining the minimum volume threshold set $\mathcal{T}_{k+1}$ subject to the probabilistic α–robust constraint, but in doing so is ignoring any information on the abnormal residual set $\mathcal{R}_{k+1}^{\mathcal{F}'}$. This could possibly lead to unsatisfactory detection properties due to a large intersection $\mathcal{T}_{k+1} \cap \mathcal{R}_{k+1}^{\mathcal{F}'}$. The goal of the second stage problem (5.13b) is then to find a new parameter θ, leading to a new threshold set $\mathcal{T}_{k+1}$ with the same robustness guarantee and a volume which is not worse than the one resulting from the solution of problem (5.13a), but which is as distant as possible from the set $\mathcal{R}_{k+1}^{\mathcal{F}'}$.*

The proposed optimization problem (5.13) is however non-convex and hard to solve due to chance constraints being in general difficult to enforce. In the following part, we provide a computationally tractable approach based on randomization technique together with a rigorous theoretical analysis of its properties.

5.3.4. COMPUTATIONALLY TRACTABLE METHODOLOGY

Chance-constrained optimization problems are known to be non-convex and hard to solve. However, these problems received increasing attention due to recent developments towards computationally tractable approaches. In particular, randomization techniques allow to approximate chance constraints in an equivalent sense without imposing any restriction on the probability distribution and geometric information of uncertain variables. The basic idea is very simple: chance constraints are substituted with finitely many hard constraints that correspond to samples from the uncertainty realizations. Using this approach, we are now able to formulate the following tractable optimization problem as a counterpart of the one in (5.13):

$$\begin{cases} \min\limits_{\theta,\gamma} & \gamma \\ \text{s.t.} & G(\theta) \succeq 0\,, \quad \text{trace}(G(\theta)M) \leq \gamma\,, \\ & \mathbf{1}_{\mathcal{T}}(r_{k+1}^{0,i},\theta) \geq c\,, \quad i = 1,\cdots,N_s\,, \end{cases} \tag{5.14a}$$

$$\begin{cases} \max\limits_{\theta} & \|G(\theta) - G(\psi^*)\|_\infty \\ \text{s.t.} & \theta \succeq 0\,, \quad \text{trace}(G(\theta)M) \leq \gamma^*\,, \\ & \mathbf{1}_{\mathcal{T}}(r_{k+1}^{0,i},\theta) \geq c\,, \quad i = 1,\cdots,N_s\,, \end{cases} \tag{5.14b}$$

where $r_{k+1}^{0,i} = \Sigma(r_k, \delta_k^i)$, and $\delta_k^i \in \Delta_k^0$ are samples of the random variable δ_k. We assume to be able to generate N_s samples based on the knowledge of $\Sigma(\cdot)$, and availability of the uncertainty samples from Δ_k^0, following Definition 7. It is important to mention that the same samples have to be used in both problems (5.14a) and (5.14b). Otherwise, the second problem (5.14b) may lead to infeasible results, since the optimal solution γ^* of the first problem (5.14a), that is used in the second problem (5.14b), depends on different samples.

Remark 19 *The requirement that each agent $\mathcal{L}_I$ is able to generate samples of its own uncertainties, which occur in the definition of δ_I according to (5.5), may be fulfilled by assuming a local knowledge of their probability density function (pdf). Should this knowl-*

edge be not available, samples can still be obtained using historical data recorded in normal conditions from system (5.2).

The link between the chance-constrained program and the quality of its approximation is the number of samples N_s that should be considered in order to reach a given level of confidence. This has been rigorously investigated in the *scenario approach*, a powerful randomized method developed recently (see [30] and the references therein). The crucial requirement to invoke these results is the convexity of the optimization problem in the decision variables, but unfortunately in the present case this does not hold due to the use of the Chebyshev distance in the objective (5.14b). It is, however, easy to show that (5.14b) can be transformed into a number of different convex programs. As it has been shown in [107, Lemma 4.3], the set of the solutions of (5.14b) is equivalent to the union of the solution sets of ξ different convex programs, where we recall that 2ξ is the degree of $\mathbf{1}_{\mathcal{T}}(r,\theta)$.

Due to the cascaded structure of the optimization formulation (5.14), it is not straightforward to use the theoretical results in [30]. In [30], the existence and uniqueness of the tractable program solution is assumed. This was later relaxed by applying a tie-break rule (e.g., lexicographic rule) and selecting among the optimal solutions the one with the best Euclidean distance [30, Section 2.1.5]. This is, however, not true in general for differently structured problems, such as the cascade formulation in (5.14), since in [30] a single tractable optimization program was considered. More specifically, a tie-break rule can be employed if the non–unique optimal solutions are obtained regardless of the number N_s of samples of the uncertain variable. As it is explained in the above remark, this cannot be guaranteed here in (5.14) due to the fact that the optimal solution γ^* is a random variable and depends on N_s. The following theorem extends the result obtained in [30] to the present setting.

Theorem 11 *Given d the degree of polynomial function of $\mathbf{1}_{\mathcal{T}}(r,\theta_k)$, consider $v := [\theta,\gamma]^\top \in \mathbb{R}^\ell$ to be the augmented vector of all the decision variables of* (5.14). *Let $\beta \in [0,1]$ and $N_s \geq \mathsf{N}_s(\alpha,\beta,\ell)$, where*

$$\mathsf{N}_s(\alpha,\beta,\ell) = \min\left\{ N_s \in \mathbb{N} \,\middle|\, d\sum_{i=0}^{\ell-1}\binom{N_s}{i}(1-\alpha)^i\alpha^{N_s-i} \leq \beta \right\},$$

Then, the optimizer $v^* := [\theta_b^*,\gamma^*]^\top$ of the randomized cascade convex program (5.14) *is a feasible solution of the chance-constrained cascaded optimization problem* (5.13) *with confidence level* $(1-\beta)$, *in the average.*

Proof. The proof is provided in Appendix D. □

The interpretation of Theorem 11 is as follows. Applying the results in [30, Theorem 1] leads to computing the number N_s of samples as a function of the total degrees of freedom of problem (5.14a) and of the confidence level with which it is desired to approximate (5.13a). Solving (5.14a) then yields an optimal solution (θ_a^*,γ^*), the last term of which is used as a fixed constraint for solving (5.14b). Reference [30, Theorem 1] yields a theoretical guarantee for feasibility of solution (θ_a^*,γ^*), however here we compute θ_b^* which might not be feasible for (5.13a) together with γ^*. It is important that the same N_s is used for both (5.14a) and (5.14b): otherwise, there are no guarantees that the program

(5.14b), which is based on the solution of (5.14a), is feasible. This is due to γ_a^* being a random variable and depending on the specific value of N_s.

Following the proposed optimization problem in (5.14), for the generic agent $\mathcal{L}_I$, generating samples of the normal residual r_I^0 requires the availability of samples of δ_I. According to Definition 7 together with (5.5), it emerges that for computing samples δ_I^i the agent $\mathcal{L}_I$ must know the current measured value $y_{\mathcal{N}_I,k}$ of the interconnection variables, as well as samples of candidate values for the true interconnection variables that are compatible with their measurement uncertainties. Such samples will be denoted as $x^i_{\mathcal{N}_I,k} = y_{\mathcal{N}_I,k} - v^i_{\mathcal{N}_I,k}$, with i indicating the i–th sample, and their generation necessitates in turn the ability to generate samples of the uncertainties $v_{\mathcal{N}_I,k}$. This requires that neighboring agents $\mathcal{L}_J$, with $J \in \mathcal{N}_I$, compute such samples and communicate them to $\mathcal{L}_I$ along with the measurement of their part of the interconnection variable $y_{\mathcal{N}_I,k}$. Such communication may expose private information of neighbors, such as their actual measurement or a possibly high number of samples of their measurement uncertainty, which could be used to estimate the *probability density function* (pdf) of such uncertainty. In the following section, we propose a solution to this issue.

5

5.4. PRIVATIZED DISTRIBUTED ANOMALY DETECTION

The proposed distributed fault detection scheme outlined in the previous section requires neighboring agents to communicate to each other a number of samples of their interconnection variables. While such communication may seem perfectly acceptable, there can be practical situations where it could lead to privacy concerns. We may consider as an example a smart grid where $\mathcal{S}_I$ and $\mathcal{S}_\mathcal{N}$ represent different subgrids with distributed energy sources and each managed by its own grid operator. The two grid operators must exchange data about nodes on their respective boundaries in order to allow for grid balancing, but they would rather keep private the way that they are allocating energy supply to their different energy sources and satisfying their energy demand [62, 155]. That is, they prefer to keep their local input u private.

In this section we develop a mechanism for each agent to preserve its privacy, drawing on the concept of $\epsilon-$differential privacy. We first introduce the concept of differential privacy and then present some existing results from the literature. Afterwards we accommodate such a mechanism into our proposed problem setting of Section 5.2, which yields a new framework that we refer to as differentially private distributed anomaly detection. We finally provide a so-called soft communication scheme to cope with network communication bandwidth limitation together with a novel theoretical guarantee to accommodate the privatized inter-agent information exchange scheme for our proposed probabilistic threshold design technique of Section 5.3.

5.4.1. THE CONCEPT OF DIFFERENTIAL PRIVACY

Differential privacy (DP) "addresses the paradox of learning nothing about an individual while learning useful information about a population" [49]. The initial concern that drove its development is in fact protecting the privacy of human individuals, for instance when personal health data is collected and used in medical studies [36].

As a preliminary notion, we need to introduce the concepts of *database* and of *query*.

Definition 12 *A* database *D of length n is a set $D = \{d_1, d_2, \cdots, d_n\}$ taking values in $\mathcal{D}$, where $\mathcal{D}$ is the* universe *of all possible databases.*

Definition 13 *A* query *q is a mapping $q : \mathcal{D} \to \mathbb{R}^{n_q}$, where n_q is the size of the result provided by the query.*

In DP, it is assumed that data contained in a database D can be accessed only through the results of queries, which are answered by the subject holding D, called *curator*. Protecting the privacy of an element d_i in D can thus be obtained by making the results of any query run on D insensitive enough to the single d_i. This can also be expressed by ensuring that two adjacent databases [63] are nearly indistinguishable from the answers to a query.

Definition 14 *Two databases $D = \{d_1, \cdots, d_n\}$ and $D' = \{d'_1, \cdots, d'_n\}$ are said to be* adjacent*, and it is written as $adj(D, D')$, if there exists $i \in \{1, \cdots, n\}$ such that $d_j = d'_j$ for all $j \neq i$.*

5

This is enforced by introducing so-called *mechanisms*, which are randomized mappings from the universe $\mathcal{D}$ to some subset in $\mathbb{R}^{n_q}$, and letting the curator use the mechanism in lieu of the query. A mechanism that acts on a database is said to be *differentially private* if it complies with the following definition from [48].

Definition 15 *Given $\epsilon \geq 0$ as the desired level of privacy, a randomized mechanism M preserves $\epsilon-$differential privacy if for all $\mathcal{R} \subset range(M)$ and all adjacent databases D and D' in $\mathcal{D}$, it holds that*

$$\mathbb{P}\left[M(D) \in \mathcal{R}\right] \leq e^{\epsilon}\, \mathbb{P}\left[M(D') \in \mathcal{R}\right] . \tag{5.15}$$

Remark 20 *A smaller ϵ implies higher level of privacy. By using differential privacy, one can hide information at the individual level, no matter what side information others may have. Definition 15 shows that DP is based on randomization, but is independent on the contents of databases, as long as they belong to $\mathcal{D}$ and are adjacent.*

A popular mechanism in the DP literature is the so-called Laplace mechanism, that introduces a Laplacian additive noise dependent on the query ℓ_p-sensitivity, the latter being defined as follows, similarly to [63, Definition 10].

Definition 16 *For any query $q : \mathcal{D} \to \mathbb{R}^{n_q}$, the ℓ_p-sensitivity of q under the adjacency relation, $adj(\cdot,\cdot)$, is defined as*

$$\sigma := \max\{\|q(D) - q(D')\|_p : D, D' \in \mathcal{D} \text{ s.t. } adj(D, D')\}.$$

It is worth to mention that ℓ_p-sensitivity of q does not depend on a specific database D. We now recall the following results from [61, Theorem 9] on the Laplacian mechanism.

Proposition 6 *Consider a query $q : \mathcal{D} \to \mathbb{R}^{n_q}$ whose ℓ_2-sensitivity is σ. Define the randomized mechanism M as $M(D) = q(D) + \nu$, where $\nu \in \mathbb{R}^{n_q}$ is a random vector whose probability density function is given by $p_\nu(\nu) \propto \exp(-\epsilon\|\nu\|/\sigma)$. Then the randomized mechanism M preserves $\epsilon-$differential privacy.*

5.4.2. PRIVACY-PRESERVING FRAMEWORK

The proposed privacy-preserving framework for distributed fault detection will be now presented. To simplify the notation and formulation, we will assume without loss of generality the case of a given agent $\mathcal{L}$ having a single neighbor $\mathcal{L}_{\mathcal{N}}$, connected through an interconnection variable $x_{\mathcal{N}} \in \mathbb{R}^{n_{\mathcal{N}}}$. We will also drop the time indices to simplify our notation whenever possible. Following the scheme developed in Section 5.3, the neighbor $\mathcal{L}_{\mathcal{N}}$ should send to $\mathcal{L}$ at each time index its last interconnection variable measurement $y_{\mathcal{N}}$, along with a set of samples $\mathcal{V}_{\mathcal{N}} = \{v_{\mathcal{N}}^1, \ldots, v_{\mathcal{N}}^{N_s}\}$ of its measurement uncertainty. With such data $\mathcal{L}$ can build the following set of candidate values for the variable $x_{\mathcal{N}}$ as follows:

$$\mathcal{X}_{\mathcal{N}} = \{x_{\mathcal{N}}^1, \ldots, x_{\mathcal{N}}^{N_s}\} := \{y_{\mathcal{N}}\} \ominus \mathcal{V}_{\mathcal{N}} \,, \tag{5.16}$$

where $\ominus$ is defined as $\mathcal{A} \ominus \mathcal{B} := \{a - b \,|\, a \in \mathcal{A}, b \in \mathcal{B}\}$.

From the DP viewpoint, agent $\mathcal{L}_{\mathcal{N}}$ is the curator of a database that contains the last local input $u_{\mathcal{N},k-1}$, and at time k is answering a query from $\mathcal{L}$ by providing the following set

$$D_{\mathcal{N}} := \{y_{\mathcal{N}}\} \cup \mathcal{X}_{\mathcal{N}} \,. \tag{5.17}$$

A desired goal of $\mathcal{L}_{\mathcal{N}}$ is to replace such an answer with a mechanism that guarantees the privacy of $u_{\mathcal{N}}$. It is important to mention that all elements of $D_{\mathcal{N}}$ are related to $y_{\mathcal{N}}$ through (5.16), and therefore, $D_{\mathcal{N}}$ contains no more useful information than $y_{\mathcal{N}}$ itself. Due to this fact, for the privacy analysis, we will consider the query to be only the output signal $y_{\mathcal{N}}$ by taking into account that the DP mechanism preserves privacy after any kind of functional composition [49, Proposition 2.1]. Consider now that the query submitted by $\mathcal{L}$ to $\mathcal{L}_{\mathcal{N}}$ yields the output signal $y_{\mathcal{N}}$, and let us design a DP mechanism for it.

In the remainder of this section, we will first design a DP mechanism where the control input variables $u_{\mathcal{N}}$ represent the database that $\mathcal{L}_{\mathcal{N}}$ wants to privatize. The output signal will thus be considered as the query submitted by the neighboring agents $\mathcal{L}$ (see Figure 5.4a). We then develop another DP mechanism where the output signal $y_{\mathcal{N}}$ of $\mathcal{L}_{\mathcal{N}}$ is considered to be the database that we want to privatize, which leads to the query submitted by the neighboring agents $\mathcal{L}$ (see Figure 5.4b) being an identity. To this end, we provide a theoretical link between the two proposed DP mechanisms such that under some mild conditions, one can use the second DP mechanism where the database is the output signal $y_{\mathcal{N}}$ of $\mathcal{L}_{\mathcal{N}}$ with identity query to achieve the same desired goal of privatizing the control input variables $u_{\mathcal{N}}$.

CONTROL INPUT AS DATABASE

Before proceeding further, we need an extended definition of adjacency.

Definition 17 *Two control actions $u_{\mathcal{N}}, u'_{\mathcal{N}} \in \mathcal{U} \subset \mathbb{R}^{m_{\mathcal{N}}}$ are two adjacent control inputs at time step $k-1$ if and only if $\|u_{\mathcal{N}} - u'_{\mathcal{N}}\|_0 \leq 1$, and it is written $adj(u_{\mathcal{N}}, u'_{\mathcal{N}})$. Such a distance between databases is referred to as the Hamming distance, i.e., the number of rows on which they differ. The set $\mathcal{U}$ is a compact set over which the input sequence $\{u_{\mathcal{N},k}\}_{k=0}^{\infty}$ can take values.*

Consider now that two adjacent control inputs belong to a bounded set $\mathcal{U}$ such that:

$$\max_{i\in\{1,\dots,m_{\mathcal{N}}\}} |(u_{\mathcal{N}})^{(i)} - (u'_{\mathcal{N}})^{(i)}| \leq 2\zeta\,, \tag{5.18}$$

where $\zeta \geq 0$ is a positive constant number which depends on the set $\mathcal{U}$. Since the query $q(\cdot)$ answered by $\mathcal{L}_{\mathcal{N}}$ is actually the output of the generic subsystem $\mathcal{S}_{\mathcal{N}}$, the constant σ that appears in Definition 16 can be computed as

$$\sigma_{\mathcal{N}} = \max_{\substack{u_{\mathcal{N}},u'_{\mathcal{N}}\in\mathcal{U}\\ \mathrm{adj}(u_{\mathcal{N}},u'_{\mathcal{N}})\\ \psi_{\mathcal{N}}\in\Psi}} \|g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}}) - g_{\mathcal{N}}(\psi_{\mathcal{N}},u'_{\mathcal{N}})\|_p\,, \tag{5.19}$$

where $g_{\mathcal{N}}(\psi_{\mathcal{N},k-1},u_{\mathcal{N},k-1}) := y_{\mathcal{N},k}$ represents a compact notation for $\mathcal{S}_{\mathcal{N}}$ dynamics in (5.2). The new quantity $\psi_{\mathcal{N}} \in \Psi$ represents the other variables, apart from the input $u_{\mathcal{N}}$, which influence $\mathcal{S}_{\mathcal{N}}$, and is defined as $\psi_{\mathcal{N}} := \mathrm{col}(x, x_{\mathcal{N}}, w, f)$, with $\Psi := \mathcal{S}^x \times \mathcal{S}^{x_{\mathcal{N}}} \times \mathcal{W} \times \mathcal{F}$. The bound $\sigma_{\mathcal{N}}$ can be seen as a bound on the global ℓ_p−sensitivity of the mapping function $g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}})$ with respect to the control input $u_{\mathcal{N}}$ at each time step k for all $p \geq 1$. The following assumption is needed to compute $\sigma_{\mathcal{N}}$.

5

Assumption 15 *The nonlinear function $g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}})$ of the generic subsystem $\mathcal{S}_{\mathcal{N}}$ is measurable and differentiable in $u_{\mathcal{N}}$ such that at each sampling time k*

$$\frac{\partial g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}})}{\partial u_{\mathcal{N}}} \neq 0\,, \forall u_{\mathcal{N}} \in \mathcal{U}\,, \psi_{\mathcal{N}} \in \Psi\,,$$

and there exists a constant L for all time step k, $u_{\mathcal{N}}$, $u'_{\mathcal{N}} \in \mathcal{U}$ and $\psi_{\mathcal{N}} \in \Psi$ such that:

$$\|g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}}) - g_{\mathcal{N}}(\psi_{\mathcal{N}},u'_{\mathcal{N}})\| \leq L\|\varphi_{\mathcal{N}} - \varphi'_{\mathcal{N}}\| = L\|u_{\mathcal{N}} - u'_{\mathcal{N}}\|\,, \tag{5.20}$$

where $\varphi_{\mathcal{N}}$ and $\varphi'_{\mathcal{N}}$ are two vectors obtained by concatenating $\psi_{\mathcal{N}}$ with $u_{\mathcal{N}}$ and $u'_{\mathcal{N}}$, respectively. We refer to L as the Lipschitz constant of the nonlinear function $g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}})$ of the generic subsystem $\mathcal{S}_{\mathcal{N}}$.

Remark 21 *An essential factor is the differentiability of $g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}})$ in order to derive the sensitivity of the output signal with respect to the small variations (adjacent relations) of input control signals. The key assumption is the Lipschitz condition (5.20). An approximation of the Lipschitz constant L at time step k can be calculated from (5.2) using the available values of $\psi_{\mathcal{N}} \in \Psi$ with drawing a sufficiently high number of samples of the uncertainties $v_{\mathcal{N}}$ and $w_{\mathcal{N}}$, following a Monte Carlo approach.*

Proposition 7 *The global ℓ_2−sensitivity of the output of the generic subsystem $\mathcal{S}_{\mathcal{N}}$ is bounded by $\sigma_{\mathcal{N}} \leq 2\zeta L$.*

Proof. Based on Definition 17 and (5.18) together with Assumption 15, the proof is straightforward by making use of the following relations (5.19) and (5.20), from which we can derive the inequality

$$\begin{aligned}\sigma_{\mathcal{N}} &\leq \max_{\substack{u_{\mathcal{N}},u'_{\mathcal{N}}\in\mathcal{U}\\ \mathrm{adj}(u_{\mathcal{N}},u'_{\mathcal{N}})}} L\|u_{\mathcal{N}} - u'_{\mathcal{N}}\| \\ &= \max_{\substack{u_{\mathcal{N}},u'_{\mathcal{N}}\in\mathcal{U}\\ \mathrm{adj}(u_{\mathcal{N}},u'_{\mathcal{N}})}} L \max_{i\in\{1,\dots,m_{\mathcal{N}}\}} |(u_{\mathcal{N}})^{(i)} - (u'_{\mathcal{N}})^{(i)}| \leq 2\zeta L\,.\end{aligned}$$

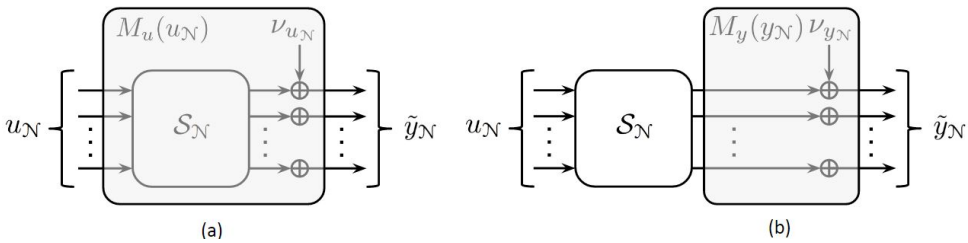


Figure 5.4: Two DP mechanisms with different databases and queries. (a) is a randomized mechanism $M_u(u_\mathcal{N})$ to preserve ϵ_u−differential privacy of input control signals. (b) is a randomized mechanism $M_y(y_\mathcal{N})$ to preserve ϵ_y−differential privacy of output control signals. With (b) the ϵ_u−differential privacy of input control signals in (a) is guaranteed under certain mild conditions (see Theorem 12).

The proof is completed. □

We are now ready to state the problem to be addressed in the present section.

Problem 1 *Find a randomized mechanism M_u such that it preserves ϵ_u−differential privacy for the neighboring agent $\mathcal{L}_\mathcal{N}$ under the adjacency relation described in Definition 17.*

Proposition 8 *The mechanism $M_u(u_\mathcal{N}) = g_\mathcal{N}(\psi_\mathcal{N}, u_\mathcal{N}) + \nu_{u_\mathcal{N}}$, where $u_\mathcal{N}$ is the control input signal and $\nu_{u_\mathcal{N}} \in \mathbb{R}^{n_\mathcal{N}}$ is a noisy vector drawn from a probability density function that is proportional to $\exp(-\epsilon_u \|\nu_{u_\mathcal{N}}\|/2\zeta L)$, is ϵ_u−differentially private.*

Proof. The proof is the direct result of combining Proposition 7 with Proposition 6. □

Remark 22 *Since the query is the output signal of $\mathcal{S}_\mathcal{N}$, following Proposition 8, the random mechanism is introduced by adding a random noise into the output signal, see Figure 5.4(a). It is worth mentioning that one can also add such a random variable into the input signal of $\mathcal{S}_\mathcal{N}$ by considering the query to be an identity mapping and consequently determining σ. We however note that such an option is not considered here, because perturbation of control input signals may lead to instability of the generic nonlinear subsystem $\mathcal{S}_\mathcal{N}$.*

OUTPUT SIGNAL AS DATABASE

We now investigate the possibility of privatizing the input control signals by just considering the output signals as a database with an identical mapping query, see Figure 5.4(b). To approach this goal we introduce the following definition to establish an adjacency relationship between two output signals.

Definition 18 *Two output signals $y_\mathcal{N}, y'_\mathcal{N} \in \mathcal{Y} \subset \mathbb{R}^{n_\mathcal{N}}$ are adjacent if and only if $\|y_\mathcal{N} - y'_\mathcal{N}\|_0 \leq 1$, and it is written as $adj(y_\mathcal{N}, y'_\mathcal{N})$. The set $\mathcal{Y}$ is a compact set over which the output sequence $\{y_{\mathcal{N},k}\}_{k=0}^{\infty}$ can take values, and since two output signals belong to $\mathcal{Y}$, we can have:*

$$\max_{i\in\{1,\ldots,n_\mathcal{N}\}} |(y_\mathcal{N})^{(i)} - (y'_\mathcal{N})^{(i)}| \leq 2\xi,$$

where $\xi \geq 0$ is a positive constant number which depends on the set $\mathcal{Y}$.

Since the query is an identity mapping, a bound $\sigma_{\mathcal{N}}$ on the global ℓ_2- sensitivity of such a query can be obtained from:

$$\sigma_{\mathcal{N}} = \max_{\substack{y_{\mathcal{N}}, y'_{\mathcal{N}} \in \mathcal{Y} \\ \text{adj}(y_{\mathcal{N}}, y'_{\mathcal{N}})}} \|y_{\mathcal{N}} - y'_{\mathcal{N}}\| \leq 2\xi\,. \tag{5.21}$$

The following proposition provides a randomized mechanism M_y such that it preserves ϵ_y-differential privacy for the agent $\mathcal{L}_{\mathcal{N}}$ under the adjacency relation of Definition 18.

Proposition 9 *The mechanism $M_y(y_{\mathcal{N}}) = y_{\mathcal{N}} + v_{y_{\mathcal{N}}}$, where $y_{\mathcal{N}}$ is the output signal and $v_{y_{\mathcal{N}}} \in \mathbb{R}^{n_{\mathcal{N}}}$ is a noisy vector drawn from a probability density function that is proportional to $\exp(-\epsilon_y \|v_{y_{\mathcal{N}}}\|/2\xi)$, is ϵ_y-differentially private.*

Proof. The proof is the direct result of combining Proposition 7 with Proposition 6. □

Theorem 12 *Let $M_u(u_{\mathcal{N}})$ and $M_y(y_{\mathcal{N}})$ be the two randomized mechanisms introduced in Propositions 8 and Proposition 9 for a generic nonlinear system dynamics $\mathcal{S}_{\mathcal{N}}$ such that they preserve ϵ_u and ϵ_y level of differential privacy with $\epsilon_y = \epsilon_u \frac{\xi}{\zeta L}$, respectively. Given ζ in Remark 5.18 and ξ in Definition 18 with L in Assumption 15, if $\xi \leq \zeta L$, then,*

$$\frac{\mathbb{P}[M_u(u_{\mathcal{N}}) \in \mathcal{R}_u]}{\mathbb{P}[M_u(u'_{\mathcal{N}}) \in \mathcal{R}_u]} = \frac{\mathbb{P}[M_y(y_{\mathcal{N}}) \in \mathcal{R}_y]}{\mathbb{P}[M_y(y'_{\mathcal{N}}) \in \mathcal{R}_y]} \leq e^{\epsilon_y} \leq e^{\epsilon_u}\,.$$

Proof. The proof is provided in Appendix D. □

It is important to highlight that Theorem 12 is the first result, to the best of our knowledge, towards privatizing a desired database, e.g. the control input actions, using another database, e.g. the output signals of a generic nonlinear system dynamics $\mathcal{S}_{\mathcal{N}}$. Theorem 12 provides a theoretical link between two randomized mechanisms $M_u(u_{\mathcal{N}})$ in Proposition 8 and $M_y(y_{\mathcal{N}})$ in Proposition 9. Strictly speaking, one can consider the output signals of a generic dynamical system $\mathcal{S}_{\mathcal{N}}$ as a database to develop a randomized mechanism $M_y(y_{\mathcal{N}})$ such that it preserves ϵ_y-differential privacy together with achieving the ϵ_u-differential privacy of the input control signals as the main desired privacy goal by considering that $\epsilon_y = \frac{\xi \epsilon_u}{\zeta L}$ and $\xi \leq \zeta L$.

5.4.3. PRIVATIZED INTER-AGENT INFORMATION EXCHANGE SCHEME

In the previous section, we have shown how to preserve the privacy of the agent $\mathcal{L}_{\mathcal{N}}$ input, assuming the query it answers to the agent $\mathcal{L}$ is just its output $y_{\mathcal{N}}$. We can now easily show that the same approach trivially extends to the original problem where the query result is the set $D_{\mathcal{N}}$, which is built by taking the actual output and a number of candidate samples for the agent state, see (5.16) and (5.17). A mechanism can be built such that the query output is modified to $\tilde{D}_{\mathcal{N}} := \{\tilde{y}_{\mathcal{N}}\} \cup \tilde{\mathcal{X}}_{\mathcal{N}}$, as it is shown in Figure 5.5, where

$$\tilde{\mathcal{X}}_{\mathcal{N}} := \{\tilde{y}_{\mathcal{N}}\} \ominus \mathcal{V}_{\mathcal{N}} = \{\tilde{x}^1_{\mathcal{N}}, \ldots, \tilde{x}^{N_s}_{\mathcal{N}}\}\,. \tag{5.22}$$

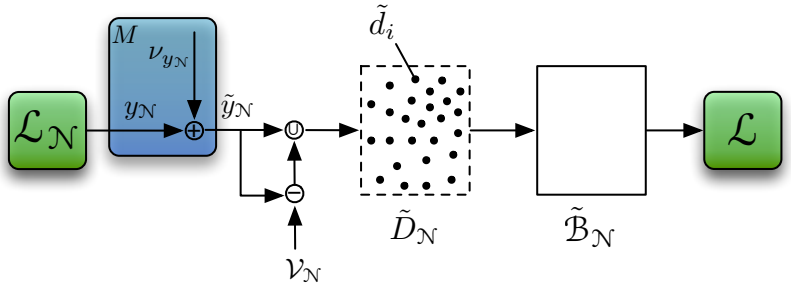


Figure 5.5: A pictorial representation of the proposed privacy-preserving inter-agent soft communication scheme. The application of the privacy mechanism M to the output $y_\mathcal{N}$ yields the privatized output $\tilde{y}_\mathcal{N}$, from which then the privatized set $\tilde{D}_\mathcal{N}$ is obtained by subtracting the measurement noise sample set $\mathcal{V}_\mathcal{N}$. Furthermore, the soft communication protocol enables that only a description of a set $\tilde{\mathcal{B}}_\mathcal{N}$ is sent to agent $\mathcal{L}$, instead of all the privatized samples $\tilde{d}_i$.

Now, it is of interest to address the issue of how agent $\mathcal{L}_\mathcal{N}$ is going to send to $\mathcal{L}$ the contents of $\tilde{D}_\mathcal{N}$. We propose the following two schemes:

a) For our proposed probabilistic threshold set design $\mathcal{T}_k$, agent $\mathcal{L}$ requests from neighboring agents to send the complete privatized inter-agent data, element by element. It is important to mention that the number of required samples of $x_\mathcal{N}$, N_s, is chosen according to Theorem 11 in order to have a given probabilistic guarantee for $\mathcal{T}_k$. We refer to this scheme as a *hard communication protocol* between agents. Its advantage is that it is simple and transmits exactly the contents of $\tilde{D}_\mathcal{N}$, but due to possibly high values of N_s it may turn out to be too costly in terms of the required communication bandwidth.

b) To address this shortcoming, we propose an alternative scheme, where instead $\mathcal{L}_\mathcal{N}$ sends a suitable parametrization of a set that contains all the possible values of its privatized data with a desired level of probability $\tilde{\alpha}_\mathcal{N}$. By considering a simple family of sets, such as for instance boxes in $\mathbb{R}^{n_\mathcal{N}}$, communication cost can be kept at reasonable levels. We refer to this scheme as a *soft communication protocol* between agents (see Figure 5.5).

We are now going to detail the soft communication protocol. The neighboring agent $\mathcal{L}_\mathcal{N}$ has to first generate $\tilde{N}_s$ samples of $x_\mathcal{N}$ and take its last privatized measurement $\tilde{y}_\mathcal{N}$ in order to build the database $\tilde{D}_\mathcal{N}$. It is important to notice that, in the soft communication protocol, the number $\tilde{N}_s$ of samples generated by $\mathcal{L}_\mathcal{N}$ may be different from that needed by $\mathcal{L}$, which is N_s, as will be explained later. More precisely, $\mathcal{L}_\mathcal{N}$ must apply the mechanism $M(y_\mathcal{N})$ to $y_\mathcal{N}$ to obtain the privatized output signals $\tilde{y}_\mathcal{N}$, then generate $\tilde{N}_s$ samples of $\tilde{x}_\mathcal{N}$ to build the privatized database $\tilde{D}_\mathcal{N}$, and send to neighboring agent $\mathcal{L}$. For sake of simplicity, we denote $\tilde{d}_i \in \tilde{D}_\mathcal{N}$ as an element of the privatized database $\tilde{D}_\mathcal{N}$.

Let us then introduce $\mathcal{B}_\mathcal{N} \subset \mathbb{R}^{n_\mathcal{N}}$ as a bounded set containing all the elements of $\tilde{D}_\mathcal{N}$. We assume for simplicity that $\mathcal{B}_\mathcal{N}$ is an axis-aligned hyper-rectangular set. This is not a restrictive assumption and any convex set could have been chosen instead as in [148]. We can so define $\mathcal{B}_\mathcal{N}$ as the Cartesian product of $n_\mathcal{N}$ intervals of the type $[-b_\mathcal{N}^{(i)}, b_\mathcal{N}^{(i)}]$, where $i = 1, \ldots, n_\mathcal{N}$ and the vector $b_\mathcal{N} \in \mathbb{R}^{n_\mathcal{N}}$ defines the hyper-rectangle bounds. For convenience, we will introduce the shorthand notation $\mathcal{B}_\mathcal{N} = [-b_\mathcal{N}, b_\mathcal{N}]$.

Consider now the following optimization problem that aims to determine the set $\mathcal{B}_{\mathcal{N}}$ with minimal volume:

$$\begin{cases} \min\limits_{b_{\mathcal{N}}} & \|b_{\mathcal{N}}\|_1 \\ \text{s.t.} & \tilde{d}_i \in [-b_{\mathcal{N}}, b_{\mathcal{N}}]\,,\ \forall \tilde{d}_i \in \tilde{D}_{\mathcal{N}}\ , \\ & i = 1, \cdots, \tilde{N}_s + 1 \end{cases} \tag{5.23}$$

where $\tilde{N}_s + 1$ is the number of privatized data elements[4] $\tilde{d}_i \in \tilde{D}_{\mathcal{N}} \subseteq \mathbb{R}^{n_{\mathcal{N}}}$ that neighboring agent $\mathcal{L}_{\mathcal{N}}$ has to take in account in order to determine $\mathcal{B}_{\mathcal{N}}$. If we denote by $\tilde{\mathcal{B}}_{\mathcal{N}} = [-\tilde{b}_{\mathcal{N}}, \tilde{b}_{\mathcal{N}}]$ the optimal solution of (5.23) computed by the neighbor agent $\mathcal{L}_{\mathcal{N}}$, then for implementing the soft communication protocol the latter needs to communicate to agent $\mathcal{L}$ only the vector $\tilde{b}_{\mathcal{N}}$ along with the probability of violation (*the level of reliability*) $\tilde{\alpha}_{\mathcal{N}}$. The level of reliability $\tilde{\alpha}_{\mathcal{N}}$ can be determined as a direct application of the scenario approach theory in [27], leading to the following result.

Theorem 13 *Fix $\tilde{\beta}_{\mathcal{N}} \in (0,1)$ and let*

$$\tilde{\alpha}_{\mathcal{N}} = \sqrt[\tilde{N}_s+1-n_{\mathcal{N}}]{\frac{\tilde{\beta}_{\mathcal{N}}}{\binom{\tilde{N}_s+1}{n_{\mathcal{N}}}}}\,. \tag{5.24}$$

We then have

$$\mathbb{P}^{\tilde{N}_s+1}\left\{\{\tilde{d}_1, \cdots, \tilde{d}_{\tilde{N}_s+1}\} \in \tilde{D}_{\mathcal{N}}^{\tilde{N}_s+1} : \mathbb{P}\{\tilde{d} \in \tilde{D}_{\mathcal{N}} : \tilde{d} \notin [-\tilde{b}_{\mathcal{N}}, \tilde{b}_{\mathcal{N}}]\} \geq 1 - \tilde{\alpha}_{\mathcal{N}}\right\} \leq \tilde{\beta}_{\mathcal{N}}. \tag{5.25}$$

Proof. Equation (5.24) is a direct result of the scenario approach theory in [27], if $\tilde{\beta}_{\mathcal{N}}$ is chosen such that

$$\binom{\tilde{N}_s+1}{n_{\mathcal{N}}} \tilde{\alpha}_{\mathcal{N}}^{\tilde{N}_s+1-n_{\mathcal{N}}} \leq \tilde{\beta}_{\mathcal{N}}\,.$$

Considering the worst-case equality in the above relation and some algebraic manipulations, one can obtain the above assertion. □

Theorem 13 implies that given an hypothetical new privatized sample $\tilde{d}$, we have a confidence of at least $1 - \tilde{\beta}_{\mathcal{N}}$ that the probability of it belonging to $\tilde{\mathcal{B}}_{\mathcal{N}} = [-\tilde{b}_{\mathcal{N}}, \tilde{b}_{\mathcal{N}}]$ is at least $\tilde{\alpha}_{\mathcal{N}}$. In other words, the optimal set $\tilde{\mathcal{B}}_{\mathcal{N}}$ is an $\tilde{\alpha}_{\mathcal{N}}$*−probabilistic approximation of the set* $\tilde{D}_{\mathcal{N}}$. Therefore, one can *rely on* $\tilde{\mathcal{B}}_{\mathcal{N}}$ *up to* $\tilde{\alpha}_{\mathcal{N}}$ *probability.*

Remark 23 *The number of samples $\tilde{N}_s$ in the proposed formulation (5.23) is a design parameter chosen by the neighboring agent $\mathcal{L}_{\mathcal{N}}$. We however remark that one can also set a given $\tilde{\alpha}_{\mathcal{N}}$ as the desired level of reliability and obtain from (5.24) the required number of samples $\tilde{N}_s$.*

Remark 24 *After receiving the parametrization of $\tilde{\mathcal{B}}_{\mathcal{N}}$ and the level of reliability $\tilde{\alpha}_{\mathcal{N}}$, agent $\mathcal{L}$ can then obtain the samples needed for computing its threshold by locally generating $N_s + 1$ samples, drawing them uniformly from inside $\tilde{\mathcal{B}}_{\mathcal{N}}$. Then it should designate,*

[4] We should remember that $\tilde{D}_{\mathcal{N}}$, contains $\tilde{N}_s$ elements corresponding to the generated samples of $\tilde{x}_{\mathcal{N}}$ plus one element corresponding to the privatized measurement $\tilde{y}_{\mathcal{N}}$.

using an arbitrary policy, one of them as the value to use for the interconnection variable measurement $y_{\mathcal{N}}$. The remaining N_s ones would be used as values of the samples for $x_{\mathcal{N}}$. In this way, we decoupled the sample generation of $\mathcal{L}_{\mathcal{N}}$ from the one of $\mathcal{L}$, preserving also the privacy of the former. We however note that the proposed soft communication protocol introduces some level of stochasticity on the probabilistic threshold design of agent $\mathcal{L}$, due to the fact that the neighboring information is probabilistically reliable. *In the following section, we characterize the threshold set probabilistic robustness as in Definition 11, and provide a new level of probability for the threshold design in order to accommodate this new situation.*

5.4.4. PRIVATIZED DISTRIBUTED PROBABILISTIC THRESHOLD SET

When an agent $\mathcal{L}$ and its neighbor $\mathcal{L}_{\mathcal{N}}$ adopt the soft communication scheme we proposed in the previous section, there is an important effect on the local probabilistic threshold set $\mathcal{T}$ computed by agent $\mathcal{L}$. Such a scheme introduces an additional level of stochasticity, as the set $\tilde{\mathcal{B}}_{\mathcal{N}}$ which is a probabilistic approximation, is communicated instead of the N_s samples that would have been sent in the hard communication scheme. This will affect the local threshold set probabilistic robustness guarantee, as explained in the following theorem. To accommodate the level of reliability of neighboring information, we need to marginalize the joint cumulative distribution function probability of the residual value at time step $k+1$ and the generic privatized sample $\tilde{d}$ appearing in Theorem 13.

Theorem 14 *Given $\tilde{\alpha}_{\mathcal{N}} \in (0,1]$ and a fixed $\alpha \in (0,1]$, then following Definition 11, the adaptive threshold set $\mathcal{T}_k$ is probabilistically $\bar{\alpha}$–robust with respect to the random total uncertainty $\delta_k \in \Delta_k^0$, i.e.,*

$$\mathbb{P}\left[r_{k+1} \in \mathcal{T}_{k+1}\right] \geq \bar{\alpha}\,, \tag{5.26}$$

where $\bar{\alpha} = 1 - \frac{1-\alpha}{\tilde{\alpha}_{\mathcal{N}}}$, and for all $r_{k+1} \in \mathcal{R}_{k+1}^0$.

Proof. The proof is provided in Appendix D. □

Theorem 14 provides a new level of the robustness for the threshold set $\mathcal{T}_k$ for each agent $\mathcal{L}$. It is straightforward to observe that if $\tilde{\alpha}_{\mathcal{N}} \to 1$ then $\bar{\alpha} \to \alpha$. This means that if *the level of reliability* of the neighboring information is one, $\mathbb{P}\left[\tilde{d} \in \tilde{\mathcal{B}}_{\mathcal{N}}\right] = 1$, then, the designed threshold set will have the same level of probabilistic robustness as the hard communication scheme, $\mathbb{P}\left[r_{k+1} \in \mathcal{T}_{k+1}\right] \geq \alpha$. It is important to note that the proposed steps for the probabilistic threshold set design that we presented in Section 5.3 are directly applicable to the results that we obtained in Theorem 14. This is due to the fact that the proposed scheme in Section 5.3 is independent from whether the hard or the soft communication scheme is used between neighboring agents.

5.5. NUMERICAL STUDY

In this section, we present the results of two numerical studies to illustrate the effectiveness of our proposed approach. First, we will introduce a minimal example with two subsystems, where in the absence of any privacy mechanism, one LD is able to detect

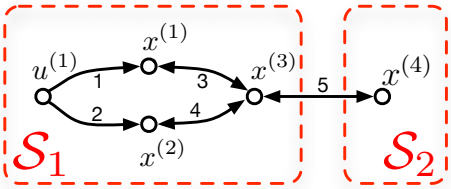


Figure 5.6: The structural graph of the 4-tank system chosen for this study, which is decomposed into two subsystems.

changes in the input of the other subsystem, while the same is not possible in the presence of a mechanism based on Proposition 8 and Proposition 9. Second, we will illustrate an implementation of the proposed distributed probabilistic anomaly detection scheme in the presence of a privacy-preserving communication mechanism.

5

5.5.1. PRIVACY PRESERVATION

Consider a water distribution network, whose structure is depicted in Figure 5.6, comprised of: 4 tanks, where levels are represented by state variables $x^{(1)}$ to $x^{(4)}$; 5 pipes, named according to the structural graph edges labels; one source, not depicted, that can deliver water to either tank 1 or 2 via pipes 1 or 2, according to the position of a valve controlled by the input $u^{(1)}$; and one point of delivery connected to tank 4.

The tank network example in Figure 5.6 is decomposed into two subsystems: $\mathcal{S}_1$, which can be thought of as a water resupply subnetwork, and $\mathcal{S}_2$, which is a customer of $\mathcal{S}_1$ and acts as a water distribution subnetwork. The operator of $\mathcal{S}_1$ can switch the valve commanded by $u^{(1)}$ in order to provide water to $\mathcal{S}_2$ either through the route $1 \to 3 \to 4$, or $2 \to 3 \to 4$. The two routes are supposed to lead to different operating costs and hence to different pricing policies that $\mathcal{S}_1$ charges to $\mathcal{S}_2$.

It is important to note that in this example, we do not consider anomalies, and the information on which route $\mathcal{S}_1$ is operating at a given moment is considered a private information, and finally, the role of $\mathcal{S}_2$ is only to distribute water to end users, which are assumed to tap into tank 4. The subsystems dynamics can then be described via the following equations:

$$\mathcal{S}_1 : \begin{cases} x_+^{(1)} = g_1^{(1)}(x_1, u_1, x_{\mathcal{N}_1}) = x^{(1)} + \frac{T}{A_1}\left[(1-u_1)\phi_{s,1} - \phi_{1,3}\right] \\ x_+^{(2)} = g_1^{(2)}(x_1, u_1, x_{\mathcal{N}_1}) = x^{(2)} + \frac{T}{A_2}\left[u_1\phi_{s,2} - \phi_{2,3}\right] \\ x_+^{(3)} = g_1^{(3)}(x_1, u_1, x_{\mathcal{N}_1}) = x^{(3)} + \frac{T}{A_3}\left[\phi_{1,3} + \phi_{2,3} - \phi_{3,4}\right] \end{cases},$$

$$\mathcal{S}_2 : \left\{ x_+^{(4)} = g_2^{(1)}(x_2, x_{\mathcal{N}_2}, w_2) = x^{(4)} + \frac{T}{A_4}\left[\phi_{3,4} - w_2\right] \right. ,$$

where g_I denotes the local dynamics of $\mathcal{S}_I$ (see Eq. (5.2)), A_i denotes the cross-section of the i–th tank, T the sampling interval, $u_1 = u^{(1)} \in \{0, 1\}$ denotes the position of a valve that can either connect tank 1, when $u^{(1)} = 0$, or tank 2, when $u^{(1)} = 1$, to a constant pressure water source serving $\mathcal{S}_1$, which is equivalent to an infinite tank at a constant level x_s. The symbol $\phi_{a,b}$ denotes the flow from tank a to tank b along the pipe connecting

them defined as:

$$\phi_{s,1} = \max\left(0, \text{sign}(x_s - x^{(1)})\sqrt{2g|x_s - x^{(1)}|}c_1\right),$$
$$\phi_{s,2} = \max\left(0, \text{sign}(x_s - x^{(2)})\sqrt{2g|x_s - x^{(2)}|}c_2\right),$$
$$\phi_{1,3} = \text{sign}(x^{(1)} - x^{(3)})\sqrt{2g|x^{(1)} - x^{(3)}|}c_3,$$
$$\phi_{2,3} = \text{sign}(x^{(2)} - x^{(3)})\sqrt{2g|x^{(2)} - x^{(3)}|}c_4,$$
$$\phi_{3,4} = \text{sign}(x^{(3)} - x^{(4)})\sqrt{2g|x^{(3)} - x^{(4)}|}c_5,$$

where c_j is the cross section of the j–th pipe and g the gravitational acceleration. Finally, the variable $w_2(t)$ is an unknown external signal representing the time-varying demand of users of the water network. It is assumed that each subsystem is endowed with a local detector (LD) $\mathcal{L}_I$, which can measure the local states: $x_1 = \left[x^{(1)}\, x^{(2)}\, x^{(3)}\right]^\top$ for $\mathcal{S}_1$ and $x_2 = \left[x^{(4)}\right]$ for $\mathcal{S}_2$. Furthermore, $\mathcal{L}_1$ shall send to $\mathcal{L}_2$ measurements of the interconnection variable $y_{\mathcal{N}_2} = \left[x^{(3)}\right]$, while $\mathcal{L}_2$ shall send to $\mathcal{L}_1$ measurements of $y_{\mathcal{N}_1} = \left[x^{(4)}\right]$.

We now propose an approach through which $\mathcal{L}_2$ can breach the privacy of $\mathcal{S}_1$ by reconstructing the current value of its local input $u^{(1)}$, and show how our proposed mechanism can be used to preserve privacy by adding a Laplacian noise to $y_{\mathcal{N}_2}$ measurements before their communication. It is assumed that $\mathcal{L}_2$ knows a model of $\mathcal{S}_1$ dynamics, which in general can be different from the actual one due to its uncertain knowledge of $\mathcal{S}_1$ parameters and/or structure. In the present case we will assume that the uncertainties are only in the parameters. Having said this, $\mathcal{L}_2$ can therefore breach $\mathcal{S}_1$ privacy by using such model and $y_{\mathcal{N}_2}$ measurements to test the hypothesis "$\mathcal{H}: u^{(1)}$ is equal to 0" against the hypothesis "$\mathcal{H}': u^{(1)}$ is equal to 1". This means that $\mathcal{L}_2$ can reconstruct through which route $\mathcal{S}_1$ is supplying water and possibly check pricing fairness of $\mathcal{S}_1$. To this purpose, $\mathcal{L}_2$ will implement the following two different estimators:

$$\hat{\mathcal{S}}_1 : \begin{cases} \hat{x}_{1,+} = \hat{g}_1(\hat{x}_1, 0, x_{\mathcal{N}_1}) + \Lambda(\hat{y}_{\mathcal{N}_2} - y_{\mathcal{N}_2}) \\ \hat{y}_{\mathcal{N}_2} = \hat{x}_1^{(3)} \end{cases}, \tag{5.27}$$

$$\hat{\mathcal{S}}_1' : \begin{cases} \hat{x}_{1,+}' = \hat{g}_1(\hat{x}_1', 1, x_{\mathcal{N}_1}) + \Lambda(\hat{y}_{\mathcal{N}_2}' - y_{\mathcal{N}_2}) \\ \hat{y}_{\mathcal{N}_2}' = \hat{x}_1'^{(3)} \end{cases}, \tag{5.28}$$

where $\hat{g}_1$ represents the uncertain $\mathcal{S}_1$ dynamics model employed by $\mathcal{L}_2$ and $y_{\mathcal{N}_2}$ is the measurement of the interconnection variable $x^{(3)}$ sent by $\mathcal{L}_1$ to $\mathcal{L}_2$. The variables $\hat{x}_1$ and $\hat{x}_{1,+}'$ and, respectively, $\hat{y}_{\mathcal{N}_2}$ and $\hat{y}_{\mathcal{N}_2}'$ are estimates of $\mathcal{S}_1$ states and of the interconnection variable computed by $\mathcal{L}_2$ under the two hypotheses. By comparing the absolute value of the scalar residuals $r := y_{\mathcal{N}_2} - \hat{y}_{\mathcal{N}_2}$ and $r' := y_{\mathcal{N}_2} - \hat{y}_{\mathcal{N}_2}'$ to a fixed scalar threshold τ, the two hypotheses can be tested. Although the estimators in Eq. (5.27) and the use of residuals and thresholds resembles the anomaly detection approach introduced in Section 5.2, here we are not implementing a distributed anomaly detection scheme but just proposing similar techniques for breaching the privacy of one subsystem.

Given the tank cross sections as 1, 1, 5 and 2 m^2, and the pipe cross sections equal to 0.25, 0.25, 0.2, 0.5 and 0.2 m^2, it can be seen that the two water supply routes on which

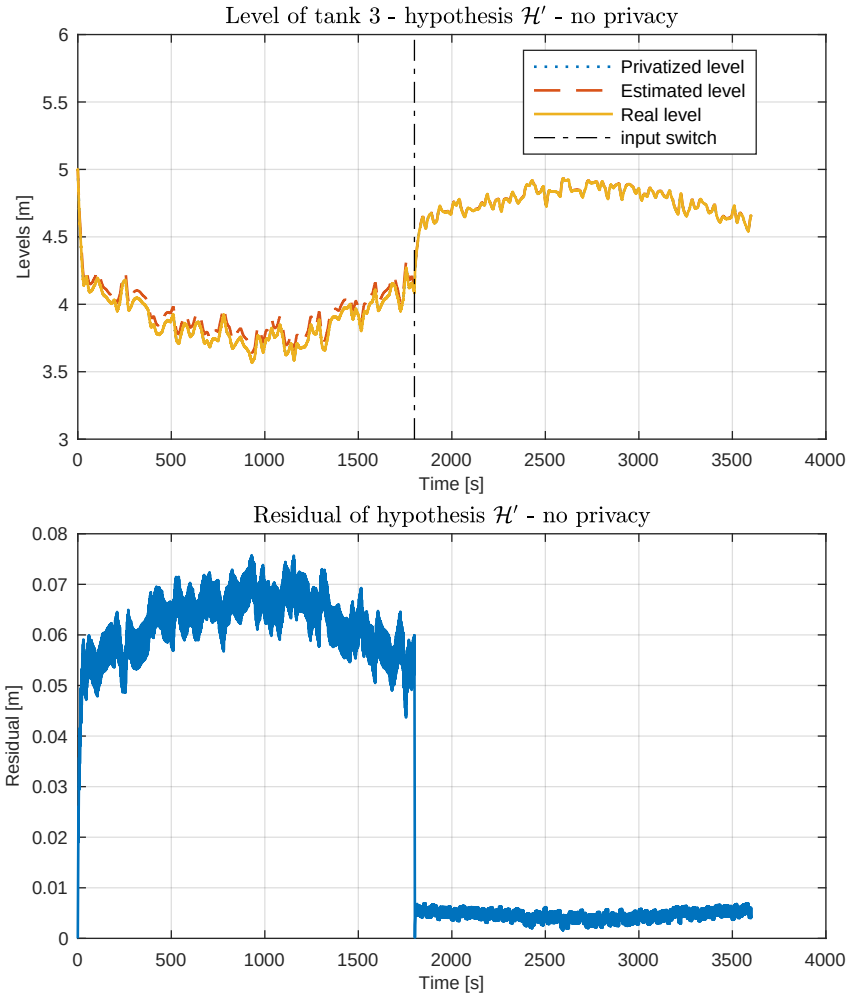


Figure 5.7: Third tank estimated level and residual computed by $\mathcal{L}_2$ for testing hypothesis $\mathcal{H}'$, when no privacy mechanism is applied to communication from $\mathcal{L}_1$ to $\mathcal{L}_2$.

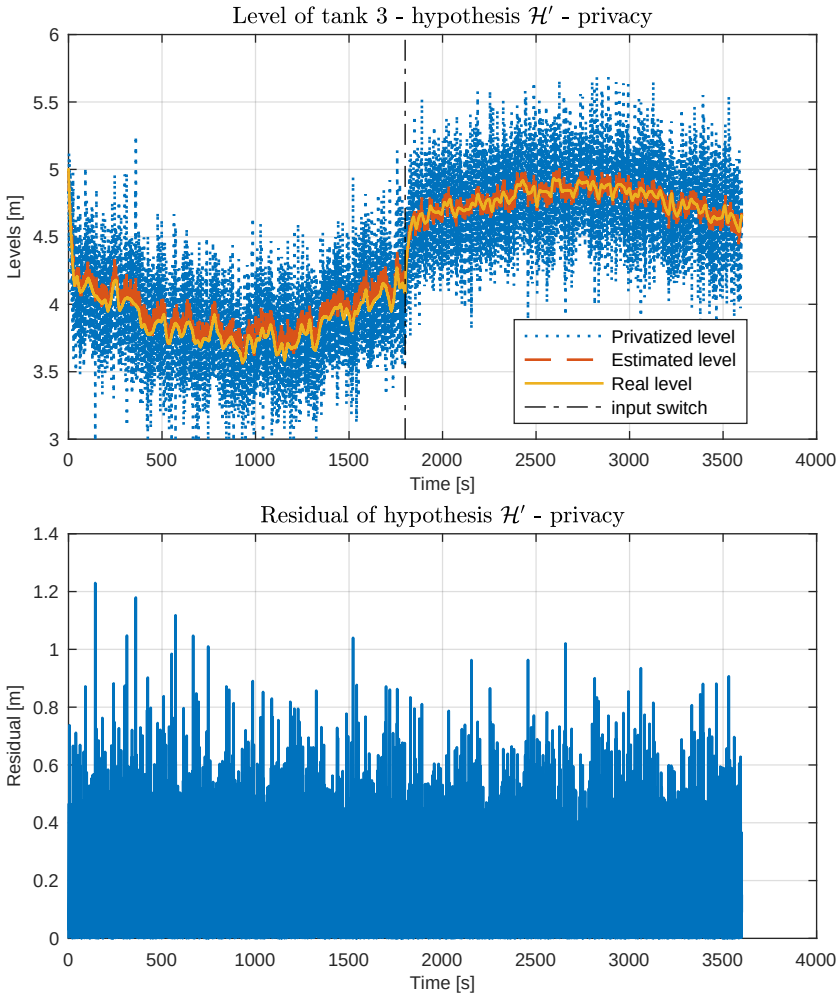


Figure 5.8: Third tank estimated level and residual computed by $\mathcal{L}_2$ for testing hypothesis $\mathcal{H}'$, when Laplacian privacy mechanism is applied to communication from $\mathcal{L}_1$ to $\mathcal{L}_2$.

	$\tau/\max(M(D))$							
case	0.01	0.05	0.1	0.25	0.5	0.75	0.95	e^{ϵ}
no privacy	**3**	**1701**	**1627**	**621**	**161**	**3.53**	1.04	
privacy	1.15	1.08	1.03	1.01	1.00	1.00	1.00	1.65

Table 5.1: Numerical evaluation of Privacy Probability Ratio. Numerical estimation of the PPR is computed using samples from the simulation, for several test sets $\mathcal{R}$. Each set corresponds to one threshold τ, that here is reported as normalized on the maximum value of the range of the mechanism M. Bold values denote cases where the differential privacy condition PPR $\leq e^{\epsilon}$ does not hold.

$\mathcal{S}_1$ can operate differ only in the cross section of the pipe feeding the third tank, being 0.2 in one case and 0.5 m^2 in the other. The unknown user demand w_2 is assumed to follow the expression $0.6 + 0.25 * \sin(2\pi/T_d\, t) + w_d$, where $T_d = 1$ hour is the demand periodicity and w_d is a random number obtained by sampling every 15 s from a normal distribution with zero mean and variance equal to 0.25. The sampling time is $T = 0.1$ s and $\mathcal{S}_1$ water supply policy is to use the first route during the first half of every hour, and the second in the second half. Finally, the model used by $\mathcal{L}_2$ is affected by a random parametric uncertainty with a maximum magnitude equal to 2% of the nominal values.

Figure 5.7 shows the estimated level for tank 3 and the residual for the hypothesis $\mathcal{H}'$, when no privacy mechanism is in place. As it can be noticed, any static threshold τ in the range 0.01 to 0.03 will lead to a successful testing of this hypothesis. Similar results can be obtained for the hypothesis $\mathcal{H}$. We can cast this in the DP setting by defining the adjacent databases $D := \{0\}$ and $D' := \{1\}$ such that u_1 will belong at any time to only one of them, and considering the query result to be the residual r' produced under hypothesis $\mathcal{H}'$. The DP condition in Definition 15 can then be checked by defining the test set $\mathcal{R} := [0\ \tau]$.

In order to introduce the mechanism described in Proposition 8, we first need to compute the query sensitivity introduced in Definition 16. For the present case, this is equal to $\sigma = \max(|r(t) - r'(t)|) = 0.075$, where the max is evaluated over the steps of the numerical simulation. Figure 5.8 shows the simulation results when a Laplacian noise corresponding to a privacy level of $\epsilon = 0.5$ is added by $\mathcal{L}_1$ to the interconnection variable $y_{\mathcal{N}_2}$ before transmission to $\mathcal{L}_2$.

Table 5.1 reports numerical results obtained by checking condition (5.15) in the case of no privacy mechanism, which is equivalent to an identity mechanism M, and in the case where we introduce a Laplacian noise mechanism. To make evaluation of the condition more clear, we tabulate here the numerical evaluation on simulation data of the *Privacy Probability Ratio* (PPR), which is defined to be PPR $:= \mathbb{P}[M(D)\in\mathcal{R}]/\mathbb{P}[M(D')\in\mathcal{R}]$, and accordingly the privacy is preserved if PPR $\leq e^{\epsilon}$. As it can be clearly seen from Table 5.1, the privacy condition (5.15) is satisfied for all of the considered test sets $\mathcal{R}$. This means that $\mathcal{L}_2$ is not able, by comparing the residual r' to any static threshold τ, to decide whether the hypothesis $\mathcal{H}'$ is true or false. Indeed, both hypotheses values have similar probabilities, as the PPR is close to 1.

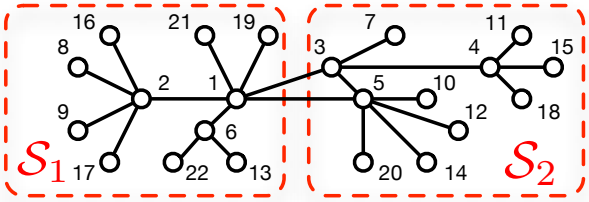


Figure 5.9: The structural graph of the 22-tank system chosen for the numerical study. It will be decomposed into two subsystems: one comprising node 1 and all the nodes to its left; the other comprising the remaining nodes on its right. The interconnection between the two susbsytems is represented by the two edges (1,3) and (1,5), corresponding to two pipes.

5.5.2. PRIVATIZED DISTRIBUTED ANOMALY DETECTION

The system under study will be a multi-tank system (see [54] for details on modeling such a system), whose structural graph contains 22 nodes, each representing a state variable corresponding to the level of a tank, while edges represent pipes interconnecting such tanks (Figure 5.9). The graph has been obtained by application of the Barabási-Albert model [4], which leads to scale-free networks. After labeling the nodes according to their degree, in descending order, the following two subsystems have been obtained by defining the extraction index tuples:

$$\begin{cases} \mathcal{I}_1 = (1,2,6,8,9,13,16,17,19,21,22) & \in \mathbb{R}^{11} \\ \mathcal{I}_2 = (3,4,5,7,10,11,12,14,15,18,20) & \in \mathbb{R}^{11} \end{cases}.$$

Finally, in order to make the interconnection between the two resulting subsystems asymmetric and thus more interesting, an edge between nodes 1 and 3 has been added, on top of the edges produced by the Barabási-Albert algorithm.

The actual tank cross sections have been chosen equal to $1\,\mathrm{m}^2$, while pipe cross sections are equal to $0.2\,\mathrm{m}^2$. Drains with the same section as interconnecting pipes have been assumed to be connected to terminal nodes (i.e. nodes with unitary degree). A single source pump, with a sinusoidal time profile with a frequency of 0.1 Hz, has been instead connected to tank no. 1. All tank levels are assumed to be measured, with a Gaussian measurement uncertainty with zero mean and a standard deviation equal to 0.05 m. When building the LD estimators, a Gaussian parametric uncertainty is introduced, having zero mean and a variance equal to 5% and 7.5%, respectively, of the tanks and pipe cross sections. The privacy mechanism M has been generated using the following values: $\epsilon = 0.1$, $\zeta = 0.01$, $\tilde{N}_s = 16$. Finally, each LD will generate $N_s = 512$ samples for computing their threshold sets, using a fourth-order polynomial as indicator function.

The fault that is presented in the current study represents a clogging in the pipe between tanks 1 and 3, reducing its flow to 50% of its nominal value. The reason we have chosen this kind of fault is that it affects exactly each subsystem interconnection variables, and as such may be hidden, that is made undetectable, by the introduction of the privacy mechanism. The following figures will present the results obtained by simulating such fault occurring at time $T_f = 250\,\mathrm{s}$. In order to make it possible to represent graphically the 11-dimensional residuals and threshold sets for the two LDs, we have chosen to

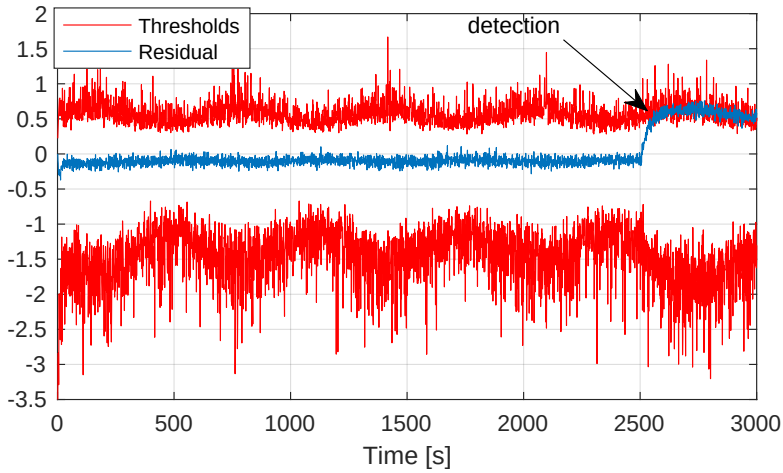


Figure 5.10: Projection of the residual and the threshold set of agent $\mathcal{L}_1$ on component no. 1, as a function of time.

consider only their projection on the multitank components no. 1, and on no. 3 and no. 5, respectively. As only the components no. 1 and 3 will be affected by the fault, this is not going to hide any information, with component no. 5 presented only for reference.

In Figure 5.10, the residual and the threshold set of the first agent, projected on the component corresponding to tank no. 1, are depicted. As in this case we are considering only one dimension, the residual can be plotted as a curve, and the threshold set projection, being a time-varying interval, can be represented by plotting two more curves corresponding to its bounds. Detection is successfully achieved shortly after the fault time.

In Figure 5.11, we instead depict the residual and threshold set for the second agent. As in this case we want to present their behaviour along the components corresponding to tanks no. 3 and 5, a time-sequence of two-dimensional plots are given. Here we can notice that occasionally the residual value can fall outside the threshold set, e.g., in sub-figures (d)-(e), as we may expect given the current probabilistic approach in designing the threshold set. After the fault time the residual is consistently outside the threshold set, see sub-figures (f)-(g)-(d), thus allowing for detection.

5.6. CONCLUSIONS

In this chapter, we developed a rigorous technique for designing a distributed anomaly detection framework for a large-scale uncertain nonlinear system. After decomposing such a system into a network of interconnected uncertain subsystems, our developments consist of three novel contributions. For each subsystem, we first provided a threshold design technique using the polynomial levelset approximation to determine a set that contains minimum volume of healthy residuals in a probabilistic sense. Such a threshold set is then reshaped to be highly sensitive to the class of given anomalies. We then equipped each subsystem with a privatized setup based on the differential pri-

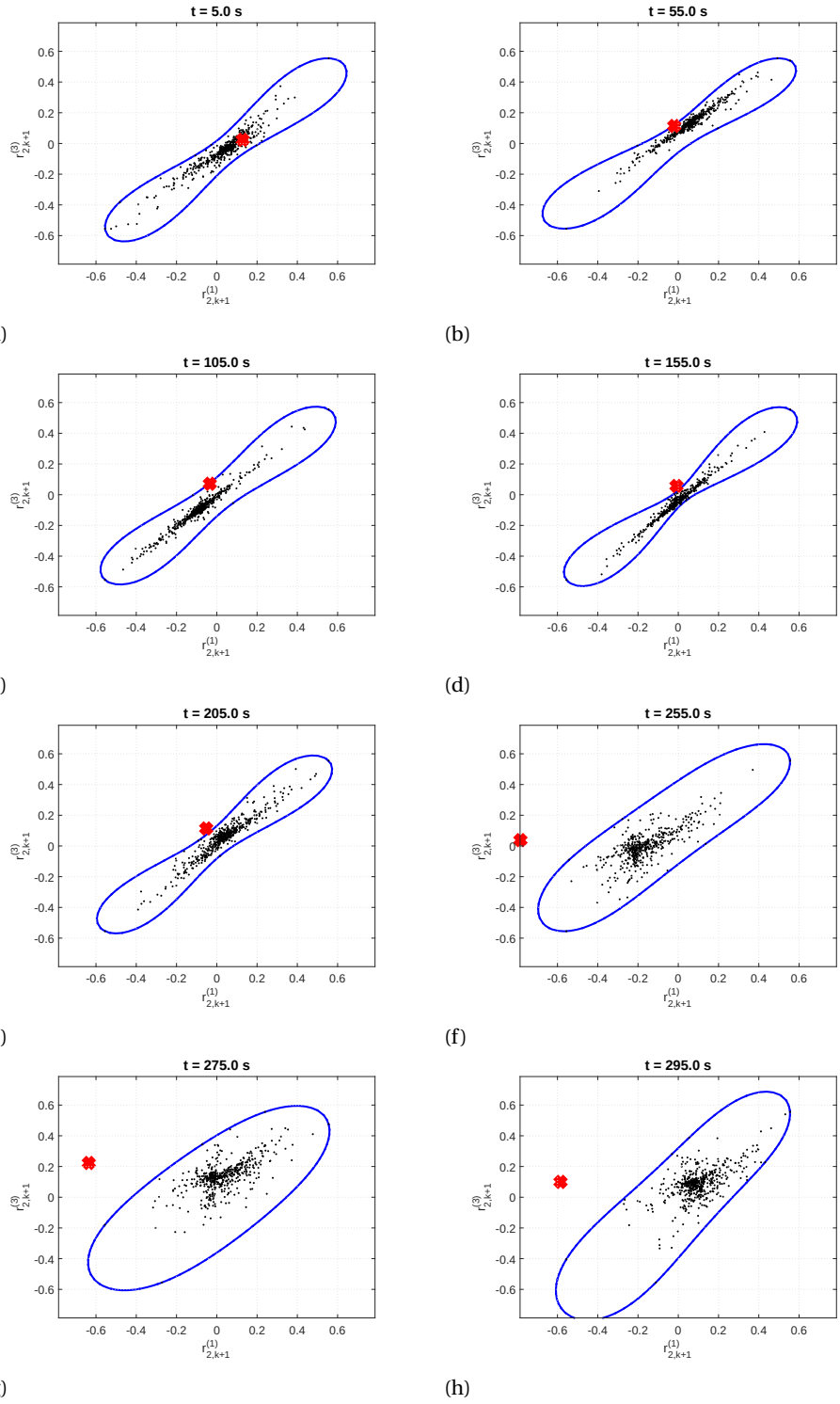


Figure 5.11: Projection of the residual and the threshold set of agent $\mathcal{L}_2$ on component no. 3 and 5, at various instants in time. At each sampling period, the healthy residuals and the threshold set are shown with block dots and blue line, respectively, and the actual residual is presented via red cross symbol.

vacy concept to pre-process the overlapping variables before sending to its neighboring subsystems. We finally developed a soft communication scheme and furnished each subsystem with a tuning knob associated with the reliability level of sending neighbors information to reduce the required communication burden.

6

Conclusions and Recommendations

This chapter provides a summary of the scientific and technical achievements for each chapter, along with several suggestions and recommendations on some open problems as future research directions.

6.1. Conclusions

The main contribution of the work presented in this dissertation is a distributed framework for decision making using scenario-based techniques in large-scale networks of interconnected uncertain dynamical systems. To achieve this goal, the following results were developed:

- ▷ **Decomposition of a large-scale scenario program into small-scale distributed scenario programs for each agent**
  We addressed two different decomposition techniques in terms of the way in which agents interact with other: 1) Dynamically coupled uncertain systems with only local constraints, and 2) Uncertain systems linked only by coupling constraints. The former refers to the case where agent dynamics influence each other with a causal dependency, while their feasible sets are completely decoupled. In the latter case, the agents' dynamics are independent and their feasible sets are considered to be coupled. Such dependencies lead to different distributed techniques, and therefore different developments, e.g., plug-and-play framework and consensus ADMM algorithm.

- ▷ **Theoretical guarantees to quantify the robustness of the resulting solutions in a distributed framework**
  We extended the existing results in literature to cope with multi-agent networked systems subject to local (private) and common uncertainty sources. The proposed

setting furnishes each agent with a tuning knob associated with the level of constraint satisfaction in a probabilistic sense. The theoretical guarantees of the proposed distributed data-driven decision making approach coincide with the centralized counterpart under some mid conditions. i.e., a certain number of scenarios from the neighboring agent should be received.

- ▷ **Soft communication scheme between neighboring agents to reduce the communication burden**
  In the proposed distributed setting, each agent requests a certain number of scenarios from its neighbors, which is called a hard communication scheme. This means that agents are not flexible to decide about the number of scenarios that should be sent to their neighbors. We relaxed such a restriction by introducing a soft communication scheme using a set parametrization technique, together with the notion of probabilistically reliable sets. Such a reliability measure of the soft communication scheme is incorporated into the feasibility guarantees of agent decisions in a probabilistic sense.

- ▷ **Distributed framework to share requested information while preserving privacy of individual agent**
  In the proposed distributed setup, the inter-agent soft communication scheme might give rise to some concerns about the agents' private information, e.g., local control inputs. We therefore present a novel privatized distributed framework, based on the so-called differential privacy concept, such that each agent can share requested information while preserving its privacy. We then equipped each agent with a privatized setup to pre-process the requested information (scenarios) before sending to its neighboring agents.

In the following, we briefly summarize the highlights of each chapter.

CHAPTER 2

This chapter proposed an energy management framework in STGs consisting of ATES systems integrated into BCC systems using a stochastic MPC paradigm. We developed a large-scale stochastic hybrid model to capture thermal energy imbalance errors in an ATES-STG. We formalized two important practical concerns by developing dynamical models to capture: a) the balance between extraction and injection of energy from and into the aquifers within a certain period of time, and b) the unwanted mutual interaction between ATES systems in STGs. Using our developed model, we formulated a finite-horizon mixed-integer quadratic optimization problem with multiple chance constraints. To solve such a difficult problem, we proposed a tractable formulation based on the so-called robust randomized approach. In particular, we extended this approach to handle a problem with multiple chance constraints.

CHAPTER 3

This chapter presented an approach to distributed stochastic MPC using the scenario-based approximation for large-scale linear systems with private and common uncertainty sources. We extended the existing results to quantify the robustness of the resulting solutions for both cases of private and common uncertainties in a distributed

framework. We then provided a novel inter-agent soft communication scheme to minimize the amount of information exchange between each subsystem. Using a set-based parametrization technique, we introduced a reliability notion and quantified the level of feasibility of the obtained solutions via the distributed stochastic MPC integrated with the so-called soft communication scheme in a probabilistic sense. The theoretical guarantees of the proposed distributed stochastic MPC framework coincide with its centralized counterpart.

CHAPTER 4

In this chapter, we developed a framework to solve a multi-area reserve scheduling (RS) problem using an AC optimal power flow (OPF) model with wind power generation by distributed consensus using ADMM. The OPF-RS problem is formulated as a large-scale semidefinite program (SDP) in infinite-dimensional space, and then a novel affine policy is proposed to provide an approximation for the infinite-dimensional SDP by a finite-dimensional SDP together with explicitly quantified performance of the approximation. The proposed methodology bridges the gap between the DC and AC OPF model of power systems for RS and furnishes the TSOs with a tuning knob associated with the level of affordable probabilistic security.

Using the geographical patterns of the power system, a technique to decompose the large-scale system into a multi-area power network is provided. We then proposed the consensus ADMM algorithm to find a feasible solution for both local and overall multi-area network such that at every iteration, each area of the power network solves a small-scale SDP problem, and then communicates some information to its neighbors to achieve consensus. By deriving a Lyapunov-type non-increasing function, it is shown that the proposed algorithm converges as long as Slater's condition holds. Using our distributed stochastic framework, each area can have its own wind information to achieve local feasibility certificates, while preserving overall feasibility of the multi-area power network under mild conditions.

CHAPTER 5

This chapter developed a rigorous technique for designing a distributed anomaly detection framework for a large-scale uncertain nonlinear system. After decomposing such a system into a network of interconnected uncertain subsystems, our developments consist of three novel contributions. For each subsystem, we first provided a threshold design technique using the polynomial level-set approximation to determine a set that contains the minimum volume of healthy residuals in a probabilistic sense. Such a threshold set is then reshaped to be highly sensitive to the class of given anomalies. We then equipped each subsystem with a privatized setup based on the differential privacy concept to pre-process the overlapping variables before sending them to its neighboring subsystems. We finally developed a soft communication scheme and furnished each subsystem with a tuning knob associated with the reliability level of sending neighbors information to reduce the required communication burden.

6.2. RECOMMENDATIONS FOR FUTURE RESEARCH

In this section, we present open research problems that still have to be tackled along with some additional directions for future research in each chapter. In the following, we discuss two interesting open research problems related to our main contributions in this dissertation.

O1a **Scaling of the proposed distributed framework with the number of agents:** The scaling of the proposed distributed control approach with the number of agents remains an issue, and this can be considered as a direction for future work. In particular, in the proposed distributed scenario program, each agent $i \in \mathcal{N}$ can have a desired level of constraint violation ε_i and a desired confidence level $1-\beta_i$. To ensure a certain robustness level of the collection of solutions in a probabilistic sense, these choices have to follow a certain design rule, e.g. $\varepsilon = \sum_{i\in\mathcal{N}} \varepsilon_i \in (0,1)$ and $\beta = \sum_{i\in\mathcal{N}} \beta_i \in (0,1)$. This yields a fixed ε, β for the large-scale scenario program and the individual ε_i, β_i for each agent $i \in \mathcal{N}$. In order to maintain the violation level for the large-scale scenario program with many agents, i.e., $|\mathcal{N}| = N \uparrow$, the violation level of individual agents needs to decrease, i.e., $\varepsilon_i \downarrow$, which may lead to conservative results for each agent, since the number of required samples needs to increase, i.e., $N_{s_i} \uparrow$.

O1b **Analyzing the quality of the proposed distributed scenario-based solutions in term of optimum value (distributed performance bound):** Following the discussion in [25, Remark 1], even a small probability of constraint violation can significantly change the optimum value function of the robust counterpart problem. In [74], the authors derived an upper bound for the worst-case violation of the scenario program. In addition, they studied the relation between the robust/chance-constrained program and the worst-case violation of the scenario program in terms of their optimum value functions. To this end, it is an interesting research direction to investigate and analyze the statistical properties of optimum function value of the proposed distributed data-driven decision framework to provide a distributed performance bound by exploiting the results of [74].

We now present some interesting open problems for each chapter that ought to be considered as future research lines.

Chapter 2

O_{2a} **Refining ATES Model (from an application point of view):** Refining the proposed model of ATES system (2.1) in order to be able to predict more realistic situations, such as depletion of the stored warm/cold water in the wells, or new wells being installed. These situations may happen in reality, where water extraction is continued with the aquifer ambient temperature, leading to the thermal energy content of an ATES system becoming proportional to the difference between the temperatures of aquifer ambient and the return water from building. In [136], we developed such a model and currently the possibility of integration of such a development into STGs is under investigation.

O_{2b} **Developing Two-Layer Distributed Framework (from a theoretical point of view):** Developing a distributed setting to solve the tractable formulation (2.29). This could be achieved by a two-layer distributed framework. In the higher layer a distributed coordination problem of ATES systems is formulated, using for instance ADMM algorithm, to respect coupling constraints between neighboring agents. In the lower layer, agents (buildings) have their local climate comfort control problem independently. We proposed a distributed stochastic MPC setting in [146, 147] and plan to extend this framework to cope with such a large-scale mixed-integer stochastic program, e.g., (2.29).

Chapter 3

O_{3a} **Analyzing the recursive feasibility and stability of the closed-loop distributed systems:** It is important to point out that the discussion on scenario optimization throughout this dissertation was limited to constraint satisfaction probability. An important future work is to enhance the proposed framework with formal guarantees on the recursive feasibility and stability of the closed-loop. The main challenge is to design appropriate terminal cost-to-go functions for each agent, along with invariant sets that comply with the network structure. This could be achieved in a similar way as in [42] by extending to stochastic settings using the results in [52]. The results in [84] can be also considered for this research direction.

O_{3b} **Extension to stochastic hybrid systems:** The results in this chapter are based on a network of interconnected uncertain linear systems, which can be readily extended to a network of interconnected uncertain hybrid systems as a direct consequence of Theorem 1. This is due to the fact that the control problem of stochastic hybrid systems can be formulated using stochastic mixed-integer programs at each sampling time in a receding horizon fashion, and thus the proposed robust randomized approach for multiple chance constrained optimization problem in Chapter 1 can be adopted.

Chapter 4

O_{4a} **Extension to unit commitment (UC) problem:** The UC problem is one of the main tasks of the Transmission System Operator (TSO) with the objective to compute a binary vector that corresponds to the "on-off" status of the generating units, and the dispatch, which is the amount of power that each generator should produce to satisfy a given demand level. Solving the UC and RS problems in a unified framework is a challenging problem, even considering a DC model of power systems. This can be potentially addressed by adding the UC formulation into the proposed problem of this chapter, similarly to [132, Chapter 3] and the recent work in [47]. By exploiting the results in [47], where an approach to efficiently address OPF-UC problem based on sparsity-exploiting techniques (Lasserre relaxations [80]) is introduced, it could be possible to fit this formulation into the proposed distributed OPF-RS framework.

O_{4b} **Integrating the OPF management problem into the STG framework:** The main goal of STGs is to reduce global energy consumption, along with reduction in global greenhouse gas emission. Since buildings in STGs are physically connected to the power network, it is an interesting research objective to understand their coupling, and develop a distributed framework for STGs while maintaining stability of the power system. This idea would enable us to save significant amounts of energy by efficient usage of distributed energy resources of the buildings and achieve a more stable power network [60].

Chapter 5

O_{5a} **Extension to anomaly isolation and an active fault tolerant control framework:** The conducted research in this chapter focused on a set-based threshold design technique for anomaly detection in a distributed setting. This can be readily extended to an anomaly isolation setup by enhancing the proposed framework with multiple different anomaly classes. An interesting future research direction is to develop a framework to automatically detect anomalies and take corrective action. This is known as fault-tolerant control (FTC) systems. This can be achieved by modifying the control input to improve the detectability and isolability of potential anomalies. However, an important challenge is to coordinate controller modification policies, which involves discretely switching among controllers designed for different classes of potential anomalies. Such an extension could be built by combining the work in [124] with the proposed framework in this chapter.

O_{5b} **Developing an FDI framework using a nonparametric Gaussian process:** The method developed in this chapter relied on the availability of a mathematical model of the uncertain nonlinear system, in order to compare the output of the real system with the output of the faultless system model to generate the residual signal for fault detection purpose. Using such a parametric model of the system for accurate detection is difficult, due to the fact that modeling is hard in practice, and therefore, the prediction capability of these models will be limited. To overcome the aforementioned limitations, nonparametric Gaussian process (GP) regression models may be a promising technique [153] by integrating with efficient learning algorithms [43].

A

Proofs of Chapter 2

Proof of Corollary 1

The proof is straightforward by just substituting the corresponding relationships, we have

$$
\begin{aligned}
(\mathrm{V}^h_{\mathrm{a},k})_i + (\mathrm{V}^c_{\mathrm{a},k})_j &\leq \mathrm{V}_{ij} - \bar{\delta}_{ij,k}\,, \\
\frac{c_{aq}\pi\ell}{c_{pw}}\Big((\bar{r}^h_{\mathrm{a},k})^2_i + (\bar{r}^c_{\mathrm{a},k})^2_j\Big) &\leq \frac{c_{aq}\pi\ell}{c_{pw}}\Big((d_{ij})^2 - 2(\bar{r}^h_{\mathrm{a},k})_i\,(\bar{r}^c_{\mathrm{a},k})_j\Big), \\
(\bar{r}^h_{\mathrm{a},k})^2_i + (\bar{r}^c_{\mathrm{a},k})^2_j &\leq (d_{ij})^2 - 2(\bar{r}^h_{\mathrm{a},k})_i\,(\bar{r}^c_{\mathrm{a},k})_j\,, \\
\Big((\bar{r}^h_{\mathrm{a},k})_i + (\bar{r}^c_{\mathrm{a},k})_j\Big)^2 &\leq (d_{ij})^2, \\
(\bar{r}^h_{\mathrm{a},k})_i + (\bar{r}^c_{\mathrm{a},k})_j &\leq d_{ij}\,.
\end{aligned}
$$

The proof is completed by noting that the thermal radius is positive: $(\bar{r}^h_{\mathrm{a},k})_i \geq 0$, $(\bar{r}^c_{\mathrm{a},k})_i \geq 0$, $\forall i \in \{1, \cdots, N\}$. □

Proof of Theorem 1

Define $\mathrm{Vio}^p(\boldsymbol{y}^*_s)$, and $\mathrm{Vio}^c(\boldsymbol{y}^*_s)$ to be the violation probabilities of the private and common chance constraints as in (2.25b), and (2.25c), respectively, as follows:

$$
\begin{aligned}
\mathrm{Vio}^p(\boldsymbol{y}^*_s) &= \mathbb{P}_{\boldsymbol{w}}\left[\boldsymbol{w} \in \mathcal{W} : \boldsymbol{y}^*_s \notin \prod_{i=1}^{N} \mathcal{Y}_i(\boldsymbol{w}_i)\,,\ \boldsymbol{y}^*_s \in \mathcal{Y}\right], \\
\mathrm{Vio}^c(\boldsymbol{y}^*_s) &= \mathbb{P}_{\boldsymbol{\delta}}\left[\boldsymbol{\delta} \in \Delta : \boldsymbol{y}^*_s \notin \prod_{i=1}^{N} \bigcap_{j\in\mathcal{N}_i} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij})\,,\ \boldsymbol{y}^*_s \in \mathcal{Y}\right],
\end{aligned}
$$

A

where $\mathcal{Y}$ is the feasible region of the problem (2.29), and it can be characterized via

$$\mathcal{Y} := \left\{ \boldsymbol{y} \in \mathbb{R}^{n_y N} : \boldsymbol{y} \in \left\{ \prod_{i=1}^{N} \bigcap_{\boldsymbol{w}_i \in \{\mathcal{B}_i^* \bigcap \mathcal{W}_i\}} \mathcal{Y}_i(\boldsymbol{w}_i) \right\} \bigcap \left\{ \prod_{i=1}^{N} \bigcap_{j \in \mathcal{N}_i} \bigcap_{\boldsymbol{\delta}_{ij} \in \{\bar{\mathcal{B}}_{ij}^* \bigcap \Delta_{ij}\}} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij}) \right\} \right\}.$$

It is important to note that both definitions of $\mathrm{Vio}^p(\boldsymbol{y}_s^*)$ and $\mathrm{Vio}^c(\boldsymbol{y}_s^*)$ have to be conditional probabilities $\mathbb{P}_{\boldsymbol{w}}, [(\cdot)|\forall \boldsymbol{\delta} \in \Delta]$, and $\mathbb{P}_{\boldsymbol{\delta}}[(\cdot)|\forall \boldsymbol{w} \in \mathcal{W}]$, respectively. We however equivalently considered them in the above form, due to the independency of both random process following Assumption 6. Define $\mathcal{B}^* = \prod_{i=1}^{N} \mathcal{B}_i^*$, and consider now:

$$\boldsymbol{y}^* \in \prod_{i=1}^{N} \bigcap_{\boldsymbol{w}_i \in \{\mathcal{B}_i^* \bigcap \mathcal{W}_i\}} \mathcal{Y}_i(\boldsymbol{w}_i) \quad \Leftrightarrow \quad \boldsymbol{y}^* \in \bigcap_{\boldsymbol{w} \in \{\mathcal{B}^* \bigcap \mathcal{W}\}} \prod_{i=1}^{N} \mathcal{Y}_i(\boldsymbol{w}_i).$$

We also define $\bar{\mathcal{B}}^* = \prod_{i=1}^{N} \bar{\mathcal{B}}_i^*$ and $\bar{\mathcal{B}}_i^* = \prod_{j \in \mathcal{N}_i} \bar{\mathcal{B}}_{ij}^*$, and clearly, we can have:

$$\begin{aligned}
\boldsymbol{y}^* &\in \prod_{i=1}^{N} \bigcap_{j \in \mathcal{N}_i} \bigcap_{\boldsymbol{\delta}_{ij} \in \{\bar{\mathcal{B}}_{ij}^* \bigcap \Delta_{ij}\}} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij}) && \Leftrightarrow \\
\boldsymbol{y}^* &\in \prod_{i=1}^{N} \bigcap_{\boldsymbol{\delta}_i \in \{\bar{\mathcal{B}}_i^* \bigcap \Delta_i\}} \bigcap_{j \in \mathcal{N}_i} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij}) && \Leftrightarrow \\
\boldsymbol{y}^* &\in \bigcap_{\boldsymbol{\delta} \in \{\bar{\mathcal{B}}^* \bigcap \Delta\}} \prod_{i=1}^{N} \bigcap_{j \in \mathcal{N}_i} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij}).
\end{aligned}$$

Therefore,

$$\begin{cases}
\text{if} \quad \boldsymbol{w} \in \{\mathcal{B}^* \bigcap \mathcal{W}\} & \text{then} \quad \boldsymbol{y}^* \in \prod_{i=1}^{N} \mathcal{Y}_i(\boldsymbol{w}_i) \\
\text{if} \quad \boldsymbol{\delta} \in \{\bar{\mathcal{B}}^* \bigcap \Delta\} & \text{then} \quad \boldsymbol{y}^* \in \prod_{i=1}^{N} \bigcap_{j \in \mathcal{N}_i} \check{\mathcal{Y}}_{ij}(\boldsymbol{\delta}_{ij})
\end{cases}.$$

This yields the following relations:

$$\begin{aligned}
\mathrm{Vio}^p(\boldsymbol{y}_s^*) &\leq \mathbb{P}_{\boldsymbol{w}}\left[\boldsymbol{w} \in \mathcal{W} : \boldsymbol{w} \notin \mathcal{B}^*\right] = \mathrm{Vio}(\mathcal{B}^*), \\
\mathrm{Vio}^c(\boldsymbol{y}_s^*) &\leq \mathbb{P}_{\boldsymbol{\delta}}\left[\boldsymbol{\delta} \in \Delta : \boldsymbol{\delta} \notin \bar{\mathcal{B}}^*\right] = \mathrm{Vio}(\bar{\mathcal{B}}^*),
\end{aligned}$$

It is then sufficient to show that for $N_s = \max_{i=1,\cdots,N} N_{s_i}$, and $\bar{N}_s = \max_{i=1,\cdots,N} \max_{j \in N_j} \bar{N}_{s_{ij}}$:

$$\mathbb{P}_{\boldsymbol{w}}^{N_s}\left[\mathcal{S} \in \mathcal{W}^{N_s} : \mathrm{Vio}(\mathcal{B}^*) \geq \varepsilon\right] \leq \beta, \tag{A.1a}$$

$$\mathbb{P}_{\boldsymbol{\delta}}^{\bar{N}_s}\left[\bar{\mathcal{S}} \in \Delta^{\bar{N}_s} : \mathrm{Vio}(\bar{\mathcal{B}}^*) \geq \bar{\varepsilon}\right] \leq \bar{\beta}, \tag{A.1b}$$

where $\mathcal{S} = \prod_{i=1}^{N} \mathcal{S}_i$, and $\bar{\mathcal{S}} = \prod_{i=1}^{N} \prod_{j \in \mathcal{N}_i} \bar{\mathcal{S}}_{ij}$. To this end, we now break down the proof in the following steps to show (A.1a) and (A.1b):

a) Common chance constraint violation:

$$
\begin{aligned}
\mathrm{Vio}^c(\boldsymbol{y}_s^*) \le \mathrm{Vio}(\bar{\mathcal{B}}^*) &= \mathbb{P}_{\boldsymbol{\delta}}\left[\boldsymbol{\delta} \in \Delta : \boldsymbol{\delta} \notin \bar{\mathcal{B}}^*\right] \\
&= \mathbb{P}_{\boldsymbol{\delta}}\left[\boldsymbol{\delta} \in \Delta : \boldsymbol{\delta} \notin \prod_{i=1}^{N} \prod_{j \in \mathcal{N}_i} \bar{\mathcal{B}}_{ij}^*\right] \\
&= \mathbb{P}_{\boldsymbol{\delta}}\left[\boldsymbol{\delta} \in \Delta : \exists i \in \{1, \cdots, N\}, \boldsymbol{\delta}_i \notin \prod_{j \in \mathcal{N}_i} \bar{\mathcal{B}}_{ij}^*\right] \\
&= \mathbb{P}_{\boldsymbol{\delta}}\left[\bigcup_{i=1}^{N} \left\{\boldsymbol{\delta}_i \in \Delta_i : \boldsymbol{\delta}_i \notin \prod_{j \in \mathcal{N}_i} \bar{\mathcal{B}}_{ij}^*\right\}\right] \\
&\le \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{\delta}_i}\left[\boldsymbol{\delta}_i \in \Delta_i : \exists j \in \mathcal{N}_i, \boldsymbol{\delta}_{ij} \notin \bar{\mathcal{B}}_{ij}^*\right] \\
&= \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{\delta}_i}\left[\bigcup_{j \in \mathcal{N}_i} \left\{\boldsymbol{\delta}_{ij} \in \Delta_{ij} : \boldsymbol{\delta}_{ij} \notin \bar{\mathcal{B}}_{ij}^*\right\}\right] \\
&\le \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \mathbb{P}_{\boldsymbol{\delta}_{ij}}\left[\boldsymbol{\delta}_{ij} \in \Delta_{ij} : \boldsymbol{\delta}_{ij} \notin \bar{\mathcal{B}}_{ij}^*\right] \\
&= \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \,.
\end{aligned}
$$

This implies that $\mathrm{Vio}(\boldsymbol{y}_s^*) \le \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*)$, and thus, we have

$$
\begin{aligned}
\mathbb{P}_{\boldsymbol{\delta}}^{\tilde{N}_s}\left[\bar{\mathcal{S}} \in \Delta^{\tilde{N}_s} : \mathrm{Vio}(\boldsymbol{y}_s^*) \ge \bar{\varepsilon}\right] &\le \mathbb{P}_{\boldsymbol{\delta}}^{\tilde{N}_s}\left[\bar{\mathcal{S}} \in \Delta^{\tilde{N}_s} : \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \ge \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \bar{\varepsilon}_{ij}\right] \\
&= \mathbb{P}_{\boldsymbol{\delta}}^{\tilde{N}_s}\left[\bigcup_{i=1}^{N} \left\{\bar{\mathcal{S}}_i \in \Delta_i^{\tilde{N}_{s_i}} : \sum_{j \in \mathcal{N}_i} \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \ge \sum_{j \in \mathcal{N}_i} \bar{\varepsilon}_{ij}\right\}\right] \\
&\le \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{\delta}_i}^{\tilde{N}_{s_i}}\left[\bar{\mathcal{S}}_i \in \Delta_i^{\tilde{N}_{s_i}} : \sum_{j \in \mathcal{N}_i} \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \ge \sum_{j \in \mathcal{N}_i} \bar{\varepsilon}_{ij}\right] \\
&= \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{\delta}_i}^{\tilde{N}_{s_i}}\left[\bigcup_{j \in \mathcal{N}_i} \left\{\bar{\mathcal{S}}_{ij} \in \Delta_{ij}^{\tilde{N}_{s_{ij}}} : \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \ge \bar{\varepsilon}_{ij}\right\}\right] \\
&\le \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \mathbb{P}_{\boldsymbol{\delta}_{ij}}^{\tilde{N}_{s_{ij}}}\left[\bar{\mathcal{S}}_{ij} \in \Delta_{ij}^{\tilde{N}_{s_{ij}}} : \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \ge \bar{\varepsilon}_{ij}\right] \\
&\le \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \beta_{ij} = \beta \,.
\end{aligned}
$$

A

b) Private chance constraint violation:

$$\begin{aligned}
\mathrm{Vio}^p(\boldsymbol{y}_s^*) \le \mathrm{Vio}(\mathcal{B}^*) &= \mathbb{P}_{\boldsymbol{w}}\left[\boldsymbol{w} \in \mathcal{W} : \boldsymbol{w} \notin \mathcal{B}^*\right] \\
&= \mathbb{P}_{\boldsymbol{w}}\left[\boldsymbol{w} \in \mathcal{W} : \boldsymbol{w} \notin \prod_{i=1}^{N} \mathcal{B}_i^*\right] \\
&= \mathbb{P}_{\boldsymbol{w}}\left[\boldsymbol{w} \in \mathcal{W} : \exists i \in \{1, \cdots, N\}, \boldsymbol{w}_i \notin \mathcal{B}_i^*\right] \\
&= \mathbb{P}_{\boldsymbol{w}}\left[\bigcup_{i=1}^{N} \{\boldsymbol{w}_i \in \mathcal{W}_i : \boldsymbol{w}_i \notin \mathcal{B}_i^*\}\right] \\
&\le \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{w}_i}\left[\boldsymbol{w}_i \in \mathcal{W}_i : \boldsymbol{w}_i \notin \mathcal{B}_i^*\right] \\
&= \sum_{i=1}^{N} \mathrm{Vio}(\mathcal{B}_i^*) \,.
\end{aligned}$$

The last statement implies that $\mathrm{Vio}^p(\boldsymbol{y}_s^*) \le \sum_{i=1}^{N} \mathrm{Vio}(\mathcal{B}_i^*)$, and thus, we have

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{w}}^{N_s}\left[\mathcal{S} \in \mathcal{W}^{N_s} : \mathrm{Vio}^p(\boldsymbol{y}_s^*) \ge \varepsilon\right] &\le \mathbb{P}_{\boldsymbol{w}}^{N_s}\left[\mathcal{S} \in \mathcal{W}^{N_s} : \sum_{i=1}^{N} \mathrm{Vio}(\mathcal{B}_i^*) \ge \sum_{i=1}^{N} \varepsilon_i\right] \\
&= \mathbb{P}_{\boldsymbol{w}}^{N_s}\left[\bigcup_{i=1}^{N} \left\{\mathcal{S}_i \in \mathcal{W}_i^{N_{s_i}} : \mathrm{Vio}(\mathcal{B}_i^*) \ge \varepsilon_i\right\}\right] \\
&\le \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{w}_i}^{N_{s_i}}\left[\mathcal{S}_i \in \mathcal{W}_i^{N_{s_i}} : \mathrm{Vio}(\mathcal{B}_i^*) \ge \varepsilon_i\right] \\
&\le \sum_{i=1}^{N} \beta_i = \beta \,.
\end{aligned}$$

The obtained bounds in the above procedure are the desired assertions as it is stated in the theorem. It is important to mention that we use the existing results in [30] to determine N_{s_i} and $\bar{N}_{s_{ij}}$ and solve the tractable problems (2.28a) and (2.28b) for each agent $i = 1, \cdots, N$, $\forall j \in \mathcal{N}_i$, respectively. We thus have the following probabilistic guarantees:

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{w}_i}^{N_{s_i}}\left[\mathcal{S}_i \in \mathcal{W}_i^{N_{s_i}} : \mathrm{Vio}(\mathcal{B}_i^*) \ge \varepsilon_i\right] &\le \beta_i \,, \\
\mathbb{P}_{\boldsymbol{\delta}_{ij}}^{\bar{N}_{s_{ij}}}\left[\bar{\mathcal{S}}_{ij} \in \Delta_{ij}^{\bar{N}_{s_{ij}}} : \mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*) \ge \bar{\varepsilon}_{ij}\right] &\le \beta_{ij} \,.
\end{aligned}$$

The interpretation of the derivation of these bounds (A.1) is as follows. The probability of all violation probabilities $\mathrm{Vio}(\mathcal{B}_i^*)$ being simultaneously bounded by the corresponding ε_i is at least $1-\beta$, and $\mathrm{Vio}(\bar{\mathcal{B}}_{ij}^*)$ being simultaneously bounded by the corresponding $\bar{\varepsilon}_{ij}$ is at least $1-\bar{\beta}$. The proof is completed by noting that the feasible set $\mathcal{Y}$ of (2.29) has a non-empty interior:

$$\left\{\exists \rho \in \mathbb{R}_+ \,,\ \bar{\boldsymbol{y}} \in \mathcal{Y} \,:\, \|\boldsymbol{y} - \bar{\boldsymbol{y}}\| \le \rho, \forall \boldsymbol{y} \in \mathbb{R}^{n_y}\right\} \subset \mathcal{Y} \,,$$

and since the problem (2.29) has a non-empty interior feasible set, it admits at least one feasible solution $\boldsymbol{y}_s^*$. □

B

Proofs of Chapter 3

Proof of Proposition 1

Given Assumption 8 and following the proposed structure of decomposition Section 3.3, any optimizer of each subsystem $\boldsymbol{v}_i^*$ yields a feasible pair of the state and control input variables of its subsystem $\{\boldsymbol{x}_i^*, \boldsymbol{u}_i^*\} \in \mathbb{X}_i \times \mathbb{U}_i$ such that $\mathbb{X}_i = \mathcal{X}_i^T$, and $\mathbb{X}_i = \mathcal{U}_i^T$. Therefore, the collection of the optimizers $\boldsymbol{v}^* = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{v}_i^*)$ will yield the collection of feasible pairs of the state and control input variables of their subsystem:

$$\left\{\boldsymbol{x}^* = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{x}_i^*), \boldsymbol{u}^* = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{u}_i^*)\right\} \in \mathbb{X} \times \mathbb{U},$$

where $\mathbb{X} := \mathcal{X}^T = \prod_{i\in\mathcal{N}} \mathbb{X}_i$ and $\mathbb{U} := \mathcal{U}^T = \prod_{i\in\mathcal{N}} \mathbb{U}_i$, which eventually yields a feasible point for the optimization problem in (3.6). It is straightforward to use the above relation and to show that any optimizer of the optimization problem in (3.6) $\boldsymbol{v}^*$ also yields a feasible solution for the proposed optimization problem in (3.10). We then have to show that both optimization problems will have the same performance index in terms of their objective function values. Due to the proposed decomposition technique, it is easy to see that the objective function in (3.6) can be formulated as additive components such that each component represents the objective function of each subsystem $i \in \mathcal{N}$, and thus: $J(\boldsymbol{x}^*, \boldsymbol{u}^*) = \sum_{i\in\mathcal{N}} J_i(\boldsymbol{x}_i^*, \boldsymbol{u}_i^*)$. The proof is completed. □

Proof of Theorem 3

Define $\xi_{i,k} := (w_{i,k}, \delta_{i,k}, \{x_{j,k}\}_{j\in\mathcal{N}_i})$ to be a concatenated uncertain variable for each agent $i \in \mathcal{N}$ such that $\boldsymbol{\xi}_i := \{\xi_{i,k}\}_{k\in\mathcal{T}}$ is defined on probability space $(\Xi_i, \mathfrak{B}(\Xi_i), \mathbb{P}_{\boldsymbol{\xi}_i})$, where $\mathbb{P}_{\boldsymbol{\xi}_i}$ is a probability measure defined over $\Xi_i := \mathcal{W}_i \times \Delta_i \times \prod_{j\in\mathcal{N}_i} \mathbb{X}_j$ and $\mathfrak{B}(\cdot)$ denotes a Borel σ-algebra. Following this definition, it is straightforward to consider $\boldsymbol{\xi} = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{\xi}_i)$ and $\Xi = \prod_{i\in\mathcal{N}} \Xi_i$. Consider also the sample set $\mathcal{S}_{\boldsymbol{\xi}_i} = \mathcal{S}_{\boldsymbol{w}_i} \times \mathcal{S}_{\boldsymbol{\delta}_i} \times \prod_{j\in\mathcal{N}_i} \mathcal{S}_{\boldsymbol{x}_j}$ for each agent $i \in \mathcal{N}$ such that $\mathcal{S}_{\boldsymbol{\xi}} = \prod_{i\in\mathcal{N}} \mathcal{S}_{\boldsymbol{\xi}_i}$.

Consider now $\boldsymbol{v}^*$ to be the optimizer of the centralized scenario MPC problem (3.6) and define $\mathrm{Vio}(\boldsymbol{v}^*)$ to be the violation probability of the chance constraint (3.5c) as fol-

lows:

$$\mathrm{Vio}(\boldsymbol{v}^*) := \mathbb{P}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi} \in \Xi : g(\boldsymbol{v}^*, \boldsymbol{\xi}) \notin \mathbb{X}\right], \tag{B.1a}$$

where $g(\boldsymbol{v}^*, \boldsymbol{\xi})$ represents the predicted state trajectory of large-scale system dynamics (3.1) in a more compact form. In particular, the violation probability can be precisely written as $\mathrm{Vio}(\boldsymbol{v}^*) := \mathbb{P}[\,\boldsymbol{w} \in \mathcal{W}, \boldsymbol{\delta} \in \Delta : x_{k+\ell} = A_{cl}(\delta_k)x_k + B(\delta_k)v_k^* + C(\delta_k)w_k \notin \mathcal{X}, \ell \in \mathbb{N}_+ \,|\, x_k = x_0\,]$, where $A_{cl}(\delta_k) = A(\delta_k) + B(\delta_k)K$. Following Theorem 2, we have

$$\mathbb{P}_{\boldsymbol{\xi}}^{N_s}\left[\mathcal{S}_{\boldsymbol{\xi}} \in \Xi^{N_s} : \mathrm{Vio}(\boldsymbol{v}^*) \le \varepsilon\right] \ge 1 - \beta. \tag{B.1b}$$

Given Proposition 1, the problem (3.10) is an exact decomposition of the problem (3.6). This yields the following equivalence relations:

$$\boldsymbol{v}^* := \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{v}_i^*)\,,\ \mathbb{X} := \prod_{i\in\mathcal{N}} \mathbb{X}_i\,, \tag{B.2}$$

where $\boldsymbol{v}^*$ is the optimizer of the problem (3.6) with $\boldsymbol{v}_i^*$ as the optimizer of each agent $i \in \mathcal{N}$ obtained via the problem (3.10). To this end, it is necessary to prove that the above statements (B.1) still hold under the aforementioned relations (B.2). We now break down the proof into two steps. We first show the results for each agent $i \in \mathcal{N}$, and then extend into the large-scale scenario MPC problem (3.6).

1) Define $\mathrm{Vio}(\boldsymbol{v}_i^*)$ to be the violation probability of each agent $i \in \mathcal{N}$ fo the chance constraint (3.11) as follows:

$$\mathrm{Vio}(\boldsymbol{v}_i^*) := \mathbb{P}_{\boldsymbol{\xi}_i}\left[\boldsymbol{\xi}_i \in \Xi_i : g_i(\boldsymbol{v}_i^*, \boldsymbol{\xi}_i) \notin \mathbb{X}_i\right], \tag{B.3}$$

where $g_i(\boldsymbol{v}_i^*, \boldsymbol{\xi}_i)$ corresponds to the predicted state trajectory of subsystem dynamics (3.9) for each agent $i \in \mathcal{N}$. Applying the existing results in Theorem 2 for each agent $i \in \mathcal{N}$, we have

$$\mathbb{P}_{\boldsymbol{\xi}_i}^{N_{s_i}}\left[\mathcal{S}_{\boldsymbol{\xi}_i} \in \Xi_i^{N_{s_i}} : \mathrm{Vio}(\boldsymbol{v}_i^*) \le \varepsilon_i\right] \ge 1 - \beta_i\,. \tag{B.4}$$

2) Following the relations (B.2), it is easy to rewrite $\mathrm{Vio}(\boldsymbol{v}^*)$ in the following form:

$$\mathrm{Vio}(\boldsymbol{v}^*) = \mathbb{P}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi} \in \Xi : g(\boldsymbol{v}^*, \boldsymbol{\xi}) \notin \prod_{i\in\mathcal{N}} \mathbb{X}_i\right].$$

It is then sufficient to show that for $N_s = \max_{i\in\mathcal{M}} N_{s_i}$:

$$\mathbb{P}_{\boldsymbol{\xi}}^{N_s}\left[\mathcal{S}_{\boldsymbol{\xi}} \in \Xi^{N_s} : \mathrm{Vio}(\boldsymbol{v}^*) \ge \varepsilon\right] \le \beta. \tag{B.5}$$

where $\varepsilon = \sum_{i\in\mathcal{N}} \varepsilon_i \in (0,1)$ and $\beta = \sum_{i\in\mathcal{N}} \beta_i \in (0,1)$. Hence

$$\begin{aligned}
\mathrm{Vio}(\boldsymbol{v}^*) &= \mathbb{P}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi} \in \Xi : g(\boldsymbol{v}^*, \boldsymbol{\xi}) \notin \prod_{i\in\mathcal{N}} \mathbb{X}_i\right] \\
&= \mathbb{P}_{\boldsymbol{\xi}}\left[\boldsymbol{\xi} \in \Xi : \exists\, i \in \mathcal{N},\, g(\boldsymbol{v}_i^*, \boldsymbol{\xi}_i) \notin \mathbb{X}_i\right] \\
&= \mathbb{P}_{\boldsymbol{\xi}}\left[\bigcup_{i\in\mathcal{N}} \left\{\boldsymbol{\xi}_i \in \Xi_i : g(\boldsymbol{v}_i^*, \boldsymbol{\xi}_i) \notin \mathbb{X}_i\right\}\right] \\
&\le \sum_{i\in\mathcal{N}} \mathbb{P}_{\boldsymbol{\xi}_i}\left[\boldsymbol{\xi}_i \in \Xi_i : g(\boldsymbol{v}_i^*, \boldsymbol{\xi}_i) \notin \mathbb{X}_i\right] \\
&= \sum_{i\in\mathcal{N}} \mathrm{Vio}(\boldsymbol{v}_i^*)\,.
\end{aligned}$$

The last statement implies that $\mathrm{Vio}(\boldsymbol{v}^*) \leq \sum_{i\in\mathcal{N}} \mathrm{Vio}(\boldsymbol{v}_i^*)$, and thus, we have

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{\xi}}^{N_s}\left[\mathcal{S}_{\boldsymbol{\xi}} \in \Xi^{N_s} : \mathrm{Vio}(\boldsymbol{v}^*) \geq \varepsilon\right] &\leq \mathbb{P}_{\boldsymbol{\xi}}^{N_s}\left[\mathcal{S}_{\boldsymbol{\xi}} \in \Xi^{N_s} : \sum_{i\in\mathcal{N}} \mathrm{Vio}(\boldsymbol{v}_i^*) \geq \sum_{i\in\mathcal{N}} \varepsilon_i\right] \\
&= \mathbb{P}_{\boldsymbol{\xi}}^{N_s}\left[\bigcup_{i\in\mathcal{N}} \left\{\mathcal{S}_{\boldsymbol{\xi}_i} \in \Xi_i^{N_{s_i}} : \mathrm{Vio}(\boldsymbol{v}_i^*) \geq \varepsilon_i\right\}\right] \\
&\leq \sum_{i\in\mathcal{N}} \mathbb{P}_{\boldsymbol{\xi}_i}^{N_{s_i}}\left[\mathcal{S}_{\boldsymbol{\xi}_i} \in \Xi_i^{N_{s_i}} : \mathrm{Vio}(\boldsymbol{v}_i^*) \geq \varepsilon_i\right] \\
&\leq \sum_{i\in\mathcal{N}} \beta_i = \beta\,.
\end{aligned}$$

The obtained bounds in the above procedure are the desired assertions as it is stated in the theorem. The proof is completed by noting that the feasible set $\mathbb{X} = \prod_{i\in\mathcal{N}} \mathbb{X}_i$ of the problem (3.6) and (3.10) has a non-empty interior, and it thus admits at least one feasible solution $\boldsymbol{v}^* = \mathrm{col}_{i\in\mathcal{N}}(\boldsymbol{v}_i^*)$. □

Proof of Theorem 4

Equation (3.15) is a direct result of the scenario approach theory in [27], if $\tilde{\beta}_j$ is chosen such that

$$\binom{\tilde{N}_{s_i}}{n_j} \tilde{\alpha}_j^{\tilde{N}_{s_i}-n_j} \leq \tilde{\beta}_j\,. \tag{B.6}$$

Considering the worst-case equality in the above relation and some algebraic manipulations, one can obtain the above assertion. The proof is completed. □

Proof of Theorem 5

The proof consists of two main steps. We first provide an analytical expression for the robustness of the solution in agent i by taking into account the effect of just one neighboring agent $j \in \mathcal{N}_j$, and then extend the obtained results for the case when the agent i interacts with more neighboring agents, e.g. for all $j \in \mathcal{N}_j$.

Following Remark 5, we have the following updated situation:

$$\alpha_i \leq \mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\,,\ \boldsymbol{x}_j \in \tilde{\mathcal{B}}_j\right]\,,$$

which is a joint probability of $\boldsymbol{x}_i \in \mathbb{X}_i$ and $\boldsymbol{x}_j \in \tilde{\mathcal{B}}_j$. Such a joint probability can be equivalently written as a joint cumulative distribution function (CDF):

$$\begin{aligned}
\alpha_i &\leq \mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\,,\ \boldsymbol{x}_j \in \tilde{\mathcal{B}}_j\right] \\
&= \int_{\mathbb{X}_i}\int_{\tilde{\mathcal{B}}_j} p(\boldsymbol{x}_i\,,\boldsymbol{x}_j)\,\mathrm{d}\boldsymbol{x}_i\,\mathrm{d}\boldsymbol{x}_j \\
&= F_{\boldsymbol{x}_i,\boldsymbol{x}_j}(\mathbb{X}_i\,,\tilde{\mathcal{B}}_j)\,,
\end{aligned} \tag{B.7}$$

where $p(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is the joint probability density function (PDF) of $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. Our goal is to calculate:

$$\mathbb{P}[\boldsymbol{x}_i \in \mathbb{X}_i] = \int_{\mathbb{X}_i} p(\boldsymbol{x}_i)\,\mathrm{d}\boldsymbol{x}_i = F_{\boldsymbol{x}_i}(\mathbb{X}_i)\,,$$

where $p(\boldsymbol{x}_i)$ is the PDF of $\boldsymbol{x}_i$. To transform the joint CDF into the marginal CDF of $\boldsymbol{x}_i$, one can take the limit of the joint CDF as $\tilde{\mathcal{B}}_j$ approaches $\mathbb{R}^{n_j}$:

$$\begin{aligned}
\mathbb{P}[\boldsymbol{x}_i \in \mathbb{X}_i] &= F_{\boldsymbol{x}_i}(\mathbb{X}_i) \\
&= \lim_{\tilde{\mathcal{B}}_j \to \mathbb{R}^{n_j}} F_{\boldsymbol{x}_i, \boldsymbol{x}_j}(\mathbb{X}_i\,,\, \tilde{\mathcal{B}}_j) \\
&= \lim_{\tilde{\mathcal{B}}_j \to \mathbb{R}^{n_j}} F_{\boldsymbol{x}_i \mid \boldsymbol{x}_j}(\mathbb{X}_i \mid \tilde{\mathcal{B}}_j) F_{\boldsymbol{x}_j}(\tilde{\mathcal{B}}_j) \\
&= F_{\boldsymbol{x}_i}(\mathbb{X}_i) \lim_{\tilde{\mathcal{B}}_j \to \mathbb{R}^{n_j}} F_{\boldsymbol{x}_j}(\tilde{\mathcal{B}}_j)\,,
\end{aligned} \tag{B.8}$$

where the last equality is due to the fact that $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are conditionally independent.

To determine $\lim_{\tilde{\mathcal{B}}_j \to \mathbb{R}^{n_j}} F_{\boldsymbol{x}_j}(\tilde{\mathcal{B}}_j)$, one can calculate:

$$\begin{aligned}
\lim_{\tilde{\mathcal{B}}_j \to \mathbb{R}^{n_j}} F_{\boldsymbol{x}_j}(\tilde{\mathcal{B}}_j) &= \int_{\mathbb{R}^{n_j}} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j \\
&= \int_{\mathbb{R}^{n_j} \setminus \tilde{\mathcal{B}}_j} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j + \int_{\tilde{\mathcal{B}}_j} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j \\
&= \mathbb{P}\left[\boldsymbol{x}_j \notin \tilde{\mathcal{B}}_j\right] + \mathbb{P}\left[\boldsymbol{x}_j \in \tilde{\mathcal{B}}_j\right] \\
&= (1 - \tilde{\alpha}_j) + \tilde{\alpha}_j = 1\,,
\end{aligned} \tag{B.9}$$

where $p(\boldsymbol{x}_j)$ is the PDF of $\boldsymbol{x}_j$, and the last equality is a direct result of Theorem 4. We now put all the steps together as follows:

$$\begin{aligned}
\alpha_i &\leq \mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\,,\, \boldsymbol{x}_j \in \tilde{\mathcal{B}}_j\right] \\
&= F_{\boldsymbol{x}_i, \boldsymbol{x}_j}(\mathbb{X}_i\,,\, \tilde{\mathcal{B}}_j) \\
&\leq F_{\boldsymbol{x}_i}(\mathbb{X}_i) \lim_{\tilde{\mathcal{B}}_j \to \mathbb{R}^{n_j}} F_{\boldsymbol{x}_j}(\tilde{\mathcal{B}}_j) \\
&= \mathbb{P}[\boldsymbol{x}_i \in \mathbb{X}_i] \left(\int_{\mathbb{R}^{n_j} \setminus \tilde{\mathcal{B}}_j} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j + \int_{\tilde{\mathcal{B}}_j} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j \right) \\
&\leq \int_{\mathbb{R}^{n_j} \setminus \tilde{\mathcal{B}}_j} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j + \mathbb{P}[\boldsymbol{x}_i \in \mathbb{X}_i] \int_{\tilde{\mathcal{B}}_j} p(\boldsymbol{x}_j)\mathrm{d}\boldsymbol{x}_j \\
&= (1 - \tilde{\alpha}_j) + \tilde{\alpha}_j\, \mathbb{P}[\boldsymbol{x}_i \in \mathbb{X}_i]\,,
\end{aligned}$$

where the first inequality and equality is due to (B.7), the second inequality is due to (B.8), the second and last equality is due to (B.9), and the last inequality is considering the worst-case situation, e.g. $\mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i \,\middle|\, \boldsymbol{x}_j \notin \tilde{\mathcal{B}}_j\right] = 1$.

By rearranging the last equation in above result:

$$\frac{\alpha_i - (1-\tilde{\alpha}_j)}{\tilde{\alpha}_j} = 1 - \frac{1-\alpha_i}{\tilde{\alpha}_j} = \tilde{\alpha}_i \ \leq \mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\right] \ . \tag{B.10}$$

This completes the proof of first part. We now need to show the effect of having more than one neighboring agent. To this end, the most straightforward step, in order to extend the current results, is to use the fact that all neighboring agents are independent from each other. We therefore can apply the previous results for a new situation where the agent i with the probabilistic level of feasibility $\tilde{\alpha}_i$ have another neighboring agent $v \in \mathcal{N}_i$ with $\tilde{\alpha}_v$ the level of reliability of $\tilde{\mathcal{B}}_v$. By using Equation (B.10), we have the following relations for $j, v \in \mathcal{N}_i$

$$1 - \frac{1-\tilde{\alpha}_i}{\tilde{\alpha}_v} = 1 - \frac{1-\left(1-\frac{1-\alpha_i}{\tilde{\alpha}_j}\right)}{\tilde{\alpha}_v} = 1 - \frac{1-\alpha_i}{\tilde{\alpha}_j\,\tilde{\alpha}_v} \ \leq \mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\right] \ .$$

By continuing the similar arguments for all neighboring agents, one can obtain $\tilde{\alpha}_i = 1 - \frac{1-\alpha_i}{\tilde{\alpha}_i} \leq \mathbb{P}\left[\boldsymbol{x}_i \in \mathbb{X}_i\right]$ such that $\tilde{\alpha}_i = \tilde{\alpha}_1 \cdots \tilde{\alpha}_j\,\tilde{\alpha}_v \cdots \tilde{\alpha}_{|\mathcal{N}_i|} = \prod_{j \in \mathcal{N}_i}(\tilde{\alpha}_j)$. The proof is completed. □

C

Proofs of Chapter 4

Proof of Theorem 6

P-OPF-RS problem is an inner approximation version of C-OPF-RS problem [67, Definition 12.2.13], which is exact if we show that the gap between their objective function values is zero. This can be shown using the fact that the difference between optimizers Ξ^* of C-OPF-RS and the projection of Ξ^* onto the feasible set of P-OPF-RS is zero, since C-OPF-RS and P-OPF-RS have clearly similar objective functions.

In order to show the difference between the optimizers Ξ^* and the optimizer $\hat{\Xi}^*$ (the projection of Ξ^* onto the feasible set of P-OPF-RS), we need to establish equivalence between constraints of both problems using the proposed policy in (4.13). It is important to notice that the proposed affine policy (4.13) is obtained by algebraic manipulation of the reserve power definition in (4.10).

By comparing C-OPF-RS and P-OPF-RS, it is clear that the first and second constraints in P-OPF-RS are the same as (4.7b) and (4.7c), respectively, where W_t is replaced by $\hat{W}_t(\mathbf{p}_t^m)$, following the proposed affine policy in (4.13). As for the constraint (4.7d), we use the following equivalent constraints using a conic combination concept[1] of matrix variables, called coefficient matrices:

$$\hat{W}_t(\mathbf{p}_t^m) \succeq 0 \quad \longleftrightarrow \quad W_t^f \succeq 0\,,\, W_t^{\text{us}} \succeq 0\,,\, W_t^{\text{ds}} \succeq 0\,.$$

Imposing PSD constraints on the coefficient matrices is equivalent to imposing a PSD constraint on the proposed policy for the network state in (4.13), since

$$\max(\mathbf{p}_t^m, 0) \succeq 0\,,\ \max(-\mathbf{p}_t^m, 0) \succeq 0\,, \forall \mathbf{p}_t^m \in \mathcal{P}\,,$$

together with the fact that any conic combination of PSD matrices is a PSD matrix [23, Section 2.2.5], thus the approximated network state $\hat{W}_t(\mathbf{p}_t^m)$ is guaranteed to be PSD.

[1] A *conic combination* is a linear combination with only non-negative coefficients, see also [23, Section 2.1.5].

We now examine the definition of reserve power in (4.8) expressed with the new parametrization $\hat{W}_t(\mathbf{p}_t^m)$:

$$\begin{aligned} r_{k,t} &= p_{k,t}^{G} - p_{k,t}^{G,f} \\ &= \mathrm{Tr}\Big(Y_k\big(\hat{W}(\mathbf{p}_t^m) - W_t^f\big)\Big) \\ &= \mathrm{Tr}\Big(Y_k\big(\max(-\mathbf{p}_t^m,0)W_t^{\mathrm{us}} + \max(\mathbf{p}_t^m,0)W_t^{\mathrm{ds}}\big)\Big) \\ &= -\mathrm{Tr}\big(Y_k W_t^{\mathrm{us}}\big)\min(\mathbf{p}_t^m,0) + \mathrm{Tr}\Big(Y_k W_t^{\mathrm{ds}}\Big)\max(\mathbf{p}_t^m,0)\,, \end{aligned}$$

where the last equation is similar to the above assertion by using the linearity of the trace operator and the fact that $\forall \alpha \in \mathbb{R}, \max(-\alpha,0) = -\min(\alpha,0)$.

The last two constraints are similar to the constraints (4.11) to ensure that the reserve powers will always be the exact opposite of the mismatch power. By summing the previous result over all generators $k \in \mathcal{G}$:

$$\begin{aligned} \sum\nolimits_{k\in\mathcal{G}} r_{k,t} &= -\overbrace{\sum\nolimits_{k\in\mathcal{G}} \mathrm{Tr}\big(Y_k W_t^{\mathrm{us}}\big)}^{=1}\min(\mathbf{p}_t^m,0) + \overbrace{\sum\nolimits_{k\in\mathcal{G}} \mathrm{Tr}\Big(Y_k W_t^{\mathrm{ds}}\Big)}^{=-1}\max(\mathbf{p}_t^m,0), \\ &= -\min(\mathbf{p}_t^m,0) - \max(\mathbf{p}_t^m,0) \\ &= -\mathbf{p}_t^m\,. \end{aligned}$$

The proof is completed by noting that $\tilde{W}^{\mathrm{us}}$ and $\tilde{W}^{\mathrm{ds}}$ are related to changes in the network state by the distribution of up- and downspinning reserve power, respectively. Moreover, the proposed equality constraints are feasible due to fact that Y_k is indefinite $\forall k \in \mathcal{G}$. □

PROOF OF THEOREM 9

Consider $\hat{\Xi}^*$ to be the optimizer of the SP-OPF-RS problem and define $\mathrm{Vio}(\hat{\Xi}^*)$ to be the violation probability of the P-OPF-RS constraints as follows:

$$\mathrm{Vio}(\hat{\Xi}^*) := \mathbb{P}\left[\,\tilde{\mathbf{p}}^m \in \mathcal{P} : \hat{\Xi}^* \notin \tilde{\mathcal{X}}(\tilde{\mathbf{p}}^m)\,\right], \tag{C.1a}$$

where $\tilde{\mathcal{X}}(\tilde{\mathbf{p}}^m)$ is the uncertain feasible region of the P-OPF-RS problem, and it can be characterized via its constraints. Following Theorem 7, we have

$$\mathbb{P}^{N_s}\left[\,\mathcal{S} \in \mathcal{P}^{N_s} : \mathrm{Vio}(\hat{\Xi}^*) \le \varepsilon\,\right] \ge 1-\beta\,. \tag{C.1b}$$

Given Assumption 10 and Assumption 11 together with Proposition 4, the MASP-OPF-RS problem is an exact decomposition of the SP-OPF-RS problem. This yields the following equivalence relations:

$$\begin{cases} \hat{\Xi}^* := \Xi^*_{\mathrm{ma}} = \{\Xi^*_a\}_{\forall a\in\mathcal{A}} \\ \tilde{\mathcal{X}}(\tilde{\mathbf{p}}^m) := \prod_{a\in\mathcal{A}} \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m) \end{cases}, \tag{C.2}$$

where $\Xi^*_{\mathrm{ma}} = \{\Xi^*_a\}_{\forall a\in\mathcal{A}}$ is the set of optimizers of the MASP-OPF-RS problem with Ξ^*_a as the optimizer of each control area $a \in \mathcal{A}$. Moreover, $\tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)$ represents the uncertain

feasible set of each control area $a \in \mathcal{A}$ and it can be characterized using constraints of the MASP-OPF-RS problem for each control area $a \in \mathcal{A}$. To this end, it is necessary to prove that the above statements (C.1) are still hold under the aforementioned relations (C.2). We now break down the proof into two steps. We first show the results for each control area $a \in \mathcal{A}$, and then extend into a multi-area power system problem.

1) Define $\mathrm{Vio}(\Xi_a^*)$ to be the violation probability of each control area $a \in \mathcal{A}$ for the P-OPF-RS problem as follows:

$$\mathrm{Vio}(\Xi_a^*) := \mathbb{P}\left[\tilde{\mathbf{p}}^m \in \mathcal{P} : \Xi_a^* \notin \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)\right]. \tag{C.3}$$

Applying the existing results in Theorem 7 for each control area $a \in \mathcal{A}$, we have

$$\mathbb{P}^{N_{s_a}}\left[\mathcal{S}_a \in \mathcal{P}^{N_{s_a}} : \mathrm{Vio}(\Xi_a^*) \le \varepsilon_a\right] \ge 1 - \beta_a. \tag{C.4}$$

2) Following the relations (C.2), it is easy to rewrite $\mathrm{Vio}(\hat{\Xi}^*)$ in the following form:

$$\mathrm{Vio}(\hat{\Xi}^*) = \mathrm{Vio}(\Xi_{\mathrm{ma}}^*) = \mathbb{P}\left[\tilde{\mathbf{p}}^m \in \mathcal{P} : \Xi_{\mathrm{ma}}^* \notin \prod_{a\in\mathcal{A}} \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)\right],$$

It is then sufficient to show that for $N_s = \max_{a\in\mathcal{A}} N_{s_a}$:

$$\mathbb{P}^{N_s}\left[\mathcal{S} \in \mathcal{P}^{N_s} : \mathrm{Vio}(\Xi_{\mathrm{ma}}^*) \ge \varepsilon\right] \le \beta, \tag{C.5}$$

where $\varepsilon = \sum_{a\in\mathcal{A}} \varepsilon_a \in (0,1)$ and $\beta = \sum_{a\in\mathcal{A}} \beta_a \in (0,1)$. Hence

$$\begin{aligned}
\mathrm{Vio}(\hat{\Xi}^*) = \mathrm{Vio}(\Xi_{\mathrm{ma}}^*) &= \mathbb{P}\left[\tilde{\mathbf{p}}^m \in \mathcal{P} : \Xi_{\mathrm{ma}}^* \notin \prod_{a\in\mathcal{A}} \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)\right] \\
&= \mathbb{P}\left[\tilde{\mathbf{p}}^m \in \mathcal{P} : \exists a \in \mathcal{A}, \Xi_{\mathrm{a}}^* \notin \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)\right] \\
&= \mathbb{P}\left[\bigcup_{a\in\mathcal{A}} \{\tilde{\mathbf{p}}^m \in \mathcal{P} : \Xi_{\mathrm{a}}^* \notin \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)\}\right] \\
&\le \sum_{a\in\mathcal{A}} \mathbb{P}\left[\tilde{\mathbf{p}}^m \in \mathcal{P} : \Xi_{\mathrm{a}}^* \notin \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)\right] \\
&= \sum_{a\in\mathcal{A}} \mathrm{Vio}(\Xi_a^*).
\end{aligned}$$

The last statement implies that $\mathrm{Vio}(\Xi_{\mathrm{ma}}^*) \le \sum_{a\in\mathcal{A}} \mathrm{Vio}(\Xi_a^*)$, and thus, we have

$$\begin{aligned}
\mathbb{P}^{N_s}\left[\mathcal{S} \in \mathcal{P}^{N_s} : \mathrm{Vio}(\Xi_{\mathrm{ma}}^*) \ge \varepsilon\right] &\le \mathbb{P}^{N_s}\left[\mathcal{S} \in \mathcal{P}^{N_s} : \sum_{a\in\mathcal{A}} \mathrm{Vio}(\Xi_a^*) \ge \sum_{a\in\mathcal{A}} \varepsilon_a\right] \\
&= \mathbb{P}^{N_s}\left[\bigcup_{a\in\mathcal{A}} \{\mathcal{S}_a \in \mathcal{P}^{N_{s_a}} : \mathrm{Vio}(\Xi_a^*) \ge \varepsilon_a\}\right] \\
&\le \sum_{a\in\mathcal{A}} \mathbb{P}^{N_{s_a}}\left[\mathcal{S}_a \in \mathcal{P}^{N_{s_a}} : \mathrm{Vio}(\Xi_a^*) \ge \varepsilon_a\right] \\
&\le \sum_{a\in\mathcal{A}} \beta_a = \beta.
\end{aligned}$$

C

The obtained bounds in the above procedure are the desired assertions as it is stated in the theorem. The proof is completed by noting that the feasible set $\tilde{\mathcal{X}}(\tilde{\mathbf{p}}^m) = \prod_{a\in\mathcal{A}} \tilde{\mathcal{X}}_a(\tilde{\mathbf{p}}^m)$ of the MASP-OPF-RS has a non-empty interior, and it thus admits at least one feasible solution $\Xi^*_{\text{ma}} = \{\Xi^*_a\}_{\forall a\in\mathcal{A}}$. □

D

Proofs of Chapter 5

Proof of Theorem 11

Due to the non-convexity introduced by the Chebyshev distance, we have to recast the second stage problem (5.14b) into ξ sub-programs. By denoting with Ψ_j the feasible solution set of the j–th subproblem it is clear that the optimizer of (5.14b) can be found in $\bigcup_{j=1}^{\xi} \Psi_j$ [107]. For clarity the proof will be broken down into three steps: a) application of the scenario approach of [30] to each individual sub-program; b) extension to the ξ sub-programs; c) theoretical conditions for the optimizer $v^* := [\theta_b^*, \gamma^*]^\top$ to be a feasible solution of (5.13). Let us now denote with $\mathcal{T}(\theta_b^*)$ the threshold set $\mathcal{T}_k$ obtained when $\mathbf{1}_{\mathcal{T}_k}$ is parameterized by a given θ_b^*, and recall that $\mathcal{V}(\mathcal{T}(\theta_b^*))$ is the violation probability as in Definition 11.

a) Applying the existing results in [30] to each sub-program, we have $\forall j \in \{1, \cdots, \xi\}$:

$$\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_{b_j}^*)) \leq 1-\alpha\right] \leq \sum_{i=0}^{\ell-1} \binom{N_s}{i} (1-\alpha)^i \alpha^{N_s-i}.$$

b) Considering that $\mathcal{V}(\mathcal{T}(\theta_b^*)) \subseteq \bigcup_{j=1}^{\xi} \mathcal{V}(\mathcal{T}(\theta_{b_j}^*))$, we can readily extend the aforesaid results to ξ sub-programs as follows:

$$\begin{aligned}
\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha\right] &\leq \mathbb{P}_s^N\left[\bigcup_{j=1}^{\xi} \mathcal{V}(\mathcal{T}(\theta_{b_j}^*)) \leq 1-\alpha\right] \\
&\leq \sum_{j=1}^{\xi} \mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_{b_j}^*)) \leq 1-\alpha\right] \\
&< \xi \sum_{i=0}^{\ell-1} \binom{N_s}{i} (1-\alpha)^i \alpha^{N_s-i} \leq \beta.
\end{aligned}$$

Notice that the obtained bound is the desired assertion as it is stated in the theorem. However, the most important part of the proof is to extend this result to the cascade setup of the present optimization problem in (5.14).

c) In order to proceed let us first define another indicator function $\mathbf{1}_{\{\cdot\}} : [0,1] \mapsto \{0,1\}$ that indicates whether the inequality in its argument, which is a function of a random variable, holds or not. We now have to provide a new bound for the following N_s-fold product conditional probability $\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha \middle| \gamma^*\right]$ which is a random variable with respect to γ^* due to the fact that γ^* is an optimal solution of the first step optimization problem and it depends on specific random samples. To this end consider the following N_s-fold product conditional expectation problem:

$$\begin{aligned}\mathbb{E}^{N_s}\left[\mathbf{1}_{\{\mathcal{V}(\mathcal{T}(\theta_b^*))\leq 1-\alpha\}}\middle|\gamma^*\right] &= 1\cdot\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha\middle|\gamma^*\right] \\ &\quad + 0\cdot\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha\middle|\gamma^*\right] \\ &= \mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha\middle|\gamma^*\right].\end{aligned}$$

The best approximation of $\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha|\gamma^*\right]$ is given by

$$\mathbb{E}^{N_s}\left[\mathbf{1}_{\{\mathcal{V}(\mathcal{T}(\theta_b^*))\leq 1-\alpha\}}\middle|\gamma^*\right],$$

which is a function of random variable γ^*. The best here means that one cannot do any better than this due to the fact that $\mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha|\gamma^*\right]$ is itself a function of random variable γ^*. Finally, we calculate the above quantity by the law of the unconscious [73] as follows:

$$\begin{aligned}\mathbb{E}^{N_s}\left[\mathbb{E}^{N_s}\left[\mathbf{1}_{\{\mathcal{V}(\mathcal{T}(\theta_b^*))\leq 1-\alpha\}}\middle|\gamma^*\right]\right] &= \sum\nolimits_{\nu}\mathbb{E}^{N_s}\left[\mathbf{1}_{\{\mathcal{V}(\mathcal{T}(\theta_b^*))\leq 1-\alpha\}}\middle|\gamma^*=\nu\right]\mathbb{P}^{N_s}\left[\gamma^*=\nu\right] \\ &= \mathbb{E}^{N_s}\left[\mathbf{1}_{\{\mathcal{V}(\mathcal{T}(\theta_b^*))\leq 1-\alpha\}}\right] = \mathbb{P}^{N_s}\left[\mathcal{V}(\mathcal{T}(\theta_b^*)) \leq 1-\alpha\right],\end{aligned}$$

where the last equation is due to the partition theorem.

The proof is completed by noting that the final expression is already bounded in part (b) of the proof. □

PROOF OF THEOREM 12

Following Proposition 8 together with Proposition 9, let $p_{u_{\mathcal{N}}}$ and $p_{u'_{\mathcal{N}}}$ denote the probability density function of $M_u(u_{\mathcal{N}})$ and $M_u(u'_{\mathcal{N}})$, respectively, and let $p_{y_{\mathcal{N}}}$ and $p_{y'_{\mathcal{N}}}$ denote the probability density function of $M_y(y_{\mathcal{N}})$ and $M_y(y'_{\mathcal{N}})$, respectively. We now compare $p_{u_{\mathcal{N}}}$ and $p_{u'_{\mathcal{N}}}$ at some arbitrary point $z \in \mathbb{R}^{n_{\mathcal{N}}}$ in order to show the first inequality in the

above assertion as follows:

$$\begin{aligned}\frac{p_{u_{\mathcal{N}}}(z)}{p_{u'_{\mathcal{N}}}(z)} &= \frac{\exp\left(\frac{-\epsilon_u \|g_{\mathcal{N}}(\psi_{\mathcal{N}},u_{\mathcal{N}})-z\|}{2\zeta L}\right)}{\exp\left(\frac{-\epsilon_u \|g_{\mathcal{N}}(\psi_{\mathcal{N}},u'_{\mathcal{N}})-z\|}{2\zeta L}\right)} \\ &= \frac{\exp\left(\frac{-\epsilon_u \|y_{\mathcal{N}}-z\|}{2\zeta L}\right)}{\exp\left(\frac{-\epsilon_u \|y'_{\mathcal{N}}-z\|}{2\zeta L}\right)} \\ &= \frac{\exp\left(\frac{-\epsilon_y \|y_{\mathcal{N}}-z\|}{2\xi}\right)}{\exp\left(\frac{-\epsilon_y \|y'_{\mathcal{N}}-z\|}{2\xi}\right)} = \frac{p_{y_{\mathcal{N}}}(z)}{p_{y'_{\mathcal{N}}}(z)}\end{aligned}$$

where the second equality follows from choosing $\epsilon_y = \frac{\xi\epsilon_u}{\zeta L}$. Observe that $\epsilon_y \le \epsilon_u$ holds for $\xi \le \zeta L$. The rest of the proof follows the same steps as in [49, Theorem 3.6]:

$$\begin{aligned}\frac{p_{y_{\mathcal{N}}}(z)}{p_{y'_{\mathcal{N}}}(z)} &= \frac{\exp\left(\frac{-\epsilon_y \|y_{\mathcal{N}}-z\|}{2\xi}\right)}{\exp\left(\frac{-\epsilon_y \|y'_{\mathcal{N}}-z\|}{2\xi}\right)} \\ &= \exp\left(\frac{-\epsilon_y\left(\|y_{\mathcal{N}}-z\| - \|y'_{\mathcal{N}}-z\|\right)}{2\xi}\right) \\ &\le \exp\left(\frac{\epsilon_y\left(\|y'_{\mathcal{N}}-y_{\mathcal{N}}\|\right)}{2\xi}\right) \\ &\le \exp\left(\epsilon_y\right) \\ &\le \exp\left(\epsilon_u\right)\ ,\end{aligned}$$

where the first inequality follows from the inverse triangle inequality, the second follows from the definition of sensitivity and the last is due to $\xi \le \zeta L$.

Proof of Theorem 14

Following Definition 11, we have the following updated situation:

$$\alpha \le \mathbb{P}\left[\, r_{k+1} \in \mathcal{T}_{k+1}\ ,\ \tilde{d} \in \tilde{\mathcal{B}}_{\mathcal{N}} \,\right]\ ,$$

which is a joint probability of $r_{k+1} \in \mathcal{T}_{k+1}$ and $\tilde{d} \in \tilde{\mathcal{B}}_{\mathcal{N}}$. Such a joint probability can be equivalently written as a joint cumulative distribution function (CDF):

$$\begin{aligned}\alpha &\le \mathbb{P}\left[\, r_{k+1} \in \mathcal{T}_{k+1}\ ,\ \tilde{d} \in \tilde{\mathcal{B}}_{\mathcal{N}} \,\right] \\ &= \int_{\mathcal{T}_{k+1}} \int_{\tilde{\mathcal{B}}_{\mathcal{N}}} p(r_{k+1}, \tilde{d})\, \mathrm{d}r_{k+1}\, \mathrm{d}\tilde{d} \\ &= F_{r_{k+1},\tilde{d}}(\mathcal{T}_{k+1}\ ,\ \tilde{\mathcal{B}}_{\mathcal{N}})\ ,\end{aligned} \tag{D.1}$$

D

where $F_{r_{k+1},\tilde{d}}(\mathcal{T}_{k+1}\,,\,\tilde{\mathcal{B}}_{\mathcal{N}})$ and $p(r_{k+1}\,,\,\tilde{d})$ are a joint CDF and a joint probability density function (PDF) of r_{k+1} and $\tilde{d}$, respectively. Our goal is to calculate:

$$\begin{aligned}\mathbb{P}\left[\,r_{k+1}\in\mathcal{T}_{k+1}\,\right] &= \int_{\mathcal{T}_{k+1}} p(r_{k+1})\,\mathrm{d}r_{k+1}\\ &= F_{r_{k+1}}(\mathcal{T}_{k+1})\,,\end{aligned}$$

where $p(r_{k+1})$ is the PDF of r_{k+1}. In order to transform the joint CDF into the marginal CDF of r_{k+1}, one can take the limit of the joint CDF as $\tilde{\mathcal{B}}_{\mathcal{N}}$ approaches $\mathbb{R}^{n_{\mathcal{N}}}$:

$$\begin{aligned}\mathbb{P}[\,r_{k+1}\in\mathcal{T}_{k+1}\,] &= F_{r_{k+1}}(\mathcal{T}_{k+1})\\ &= \lim_{\tilde{\mathcal{B}}_{\mathcal{N}}\to\mathbb{R}^{n_{\mathcal{N}}}} F_{r_{k+1},\tilde{d}}(\mathcal{T}_{k+1}\,,\,\tilde{\mathcal{B}}_{\mathcal{N}})\\ &= \lim_{\tilde{\mathcal{B}}_{\mathcal{N}}\to\mathbb{R}^{n_{\mathcal{N}}}} F_{r_{k+1}|\tilde{d}}(\mathcal{T}_{k+1}\mid\tilde{\mathcal{B}}_{\mathcal{N}})F_{\tilde{d}}(\tilde{\mathcal{B}}_{\mathcal{N}})\\ &= F_{r_{k+1}}(\mathcal{T}_{k+1})\lim_{\tilde{\mathcal{B}}_{\mathcal{N}}\to\mathbb{R}^{n_{\mathcal{N}}}} F_{\tilde{d}}(\tilde{\mathcal{B}}_{\mathcal{N}})\,,\end{aligned} \tag{D.2}$$

where the last equality is due to the independency of r_{k+1} and $\tilde{d}$. In order to determine $\lim_{\tilde{\mathcal{B}}_{\mathcal{N}}\to\mathbb{R}^{n_{\mathcal{N}}}} F_{\tilde{d}}(\tilde{\mathcal{B}}_{\mathcal{N}})$, one can calculate:

$$\begin{aligned}\lim_{\tilde{\mathcal{B}}_{\mathcal{N}}\to\mathbb{R}^{n_{\mathcal{N}}}} F_{\tilde{d}}(\tilde{\mathcal{B}}_{\mathcal{N}}) &= \int_{\mathbb{R}^{n_{\mathcal{N}}}} p(\tilde{d})\mathrm{d}\tilde{d}\\ &= \int_{\mathbb{R}^{n_{\mathcal{N}}}\setminus\tilde{\mathcal{B}}_{\mathcal{N}}} p(\tilde{d})\mathrm{d}\tilde{d} + \int_{\tilde{\mathcal{B}}_{\mathcal{N}}} p(\tilde{d})\mathrm{d}\tilde{d}\\ &= \mathbb{P}\left[\,\tilde{d}\notin\tilde{\mathcal{B}}_{\mathcal{N}}\,\right] + \mathbb{P}\left[\,\tilde{d}\in\tilde{\mathcal{B}}_{\mathcal{N}}\,\right]\\ &= (1-\tilde{\alpha}_{\mathcal{N}}) + \tilde{\alpha}_{\mathcal{N}}\\ &= 1\,,\end{aligned} \tag{D.3}$$

where $p(\tilde{d})$ is the PDF of $\tilde{d}$, and the last equality is a direct result of Th. 13. We now put all the steps together as follows:

$$\begin{aligned}\alpha &\le \mathbb{P}\left[\,r_{k+1}\in\mathcal{T}_{k+1}\,,\,\tilde{d}\in\tilde{\mathcal{B}}_{\mathcal{N}}\,\right]\\ &= F_{r_{k+1},\tilde{d}}(\mathcal{T}_{k+1}\,,\,\tilde{\mathcal{B}}_{\mathcal{N}})\\ &\le F_{r_{k+1}}(\mathcal{T}_{k+1})\lim_{\tilde{\mathcal{B}}_{\mathcal{N}}\to\mathbb{R}^{n_{\mathcal{N}}}} F_{\tilde{d}}(\tilde{\mathcal{B}}_{\mathcal{N}})\\ &= \mathbb{P}[\,r_{k+1}\in\mathcal{T}_{k+1}]\left(\int_{\mathbb{R}^{n_{\mathcal{N}}}\setminus\tilde{\mathcal{B}}_{\mathcal{N}}} p(\tilde{d})\mathrm{d}\tilde{d} + \int_{\tilde{\mathcal{B}}_{\mathcal{N}}} p(\tilde{d})\mathrm{d}\tilde{d}\right)\\ &\le \int_{\mathbb{R}^{n_{\mathcal{N}}}\setminus\tilde{\mathcal{B}}_{\mathcal{N}}} p(\tilde{d})\mathrm{d}\tilde{d} + \mathbb{P}[\,r_{k+1}\in\mathcal{T}_{k+1}]\int_{\tilde{\mathcal{B}}_{\mathcal{N}}} p(\tilde{d})\mathrm{d}\tilde{d}\\ &= (1-\tilde{\alpha}_{\mathcal{N}}) + \tilde{\alpha}_{\mathcal{N}}\,\mathbb{P}[\,r_{k+1}\in\mathcal{T}_{k+1}]\,,\end{aligned}$$

where the first inequality and equality is due to (D.1), the second inequality is due to (D.2), the second and last equality is due to (D.3), and the last inequality is due to the fact that $\mathbb{P}[\, r_{k+1} \in \mathcal{T}_{k+1} \,] \leq 1$. Rearranging the last equation results in:

$$\frac{\alpha - (1 - \tilde{\alpha}_{\mathcal{N}})}{\tilde{\alpha}_{\mathcal{N}}} = 1 - \frac{1 - \alpha}{\tilde{\alpha}_{\mathcal{N}}} = \bar{\alpha} \leq \mathbb{P}[\, r_{k+1} \in \mathcal{T}_{k+1} \,] \ .$$

The proof is completed by noting that the final equation is our desired assertion. □

BIBLIOGRAPHY

[1] A. Abur and M. K. Celik. Least absolute value state estimation with equality and inequality constraints. *IEEE Transactions on Power Systems*, 8(2):680–686, 1993.

[2] A. Afram and F. Janabi-Sharifi. Theory and applications of HVAC control systems–a review of model predictive control (MPC). *Building and Environment*, 72:343–355, 2014.

[3] T. Alamo, R. Tempo, A. Luque, and D. R. Ramirez. Randomized methods for design of uncertain systems: Sample complexity and sequential algorithms. *Automatica*, 52:160–172, 2015.

[4] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47–97, 2002.

[5] W. Ananduta. Distributed energy management in smart thermal grids with uncertain demands. Master's thesis, DCSC Department of 3ME Faculty, Delft University of Technology, The Netherlands, 2016.

[6] E. Andersen and K. Andersen. The MOSEK optimization toolbox for Matlab manual, Mosek ApS, 2015.

[7] G. Andersson. *Modelling and analysis of electric power systems*. ETH Zürich University, 2008.

[8] P. Attaviriyanupap, H. Kita, E. Tanaka, and J. Hasegawa. New bidding strategy formulation for day-ahead energy and reserve markets based on evolutionary programming. *International Journal of Electrical Power & Energy Systems*, 27(3):157–167, 2005.

[9] A. Bemporad and M. Morari. Robust model predictive control: A survey. In *Robustness in Identification and Control*, pages 207–226. Springer, 1999.

[10] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[11] D. P. Bertsekas. *Convex optimization theory*. Athena Scientific Belmont, 2009.

[12] D. Bertsimas and M. Sim. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107(1-2):5–36, 2006.

[13] J. Blesa, V. Puig, and J. Saludes. Robust fault detection using polytope-based set-membership consistency test. *IET Control Theory & Applications*, 6(12):1767–1777, 2012.

[14] J. Bloemendal. *The hidden side of cities: Methods for governance, planning and design for optimal use of subsurface space with ATES*. PhD thesis, Department of Civil Engineering, Delft University of technology, 2018.

[15] M. Bloemendal, M. Jaxa-Rozen, and T. Olsthoorn. Methods for planning of ATES systems. *Applied Energy*, 216:534–557, 2018.

[16] M. Bloemendal, M. Jaxa-Rozen, and V. Rostampour. Use it or lose it, adaptive ATES planning. In *International Energy Agency (IEA) Heat Pump Conference*. International Energy Agency, 2017.

[17] F. Boem, R. M. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou. A distributed networked approach for fault detection of large-scale systems. *IEEE Transactions on Automatic Control*, 62(1):18–33, 2017.

[18] F. Boem, R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou. Optimal topology for distributed fault detection of large-scale systems. *IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*, 48(21):60–65, 2015.

[19] S. Boersma, V. Rostampour, B. Doekemeijer, W. van Geest, and J.-W. van Wingerden. A centralized model predictive wind farm controller in palm providing power reference tracking: an LES study. In *to appear in Journal of Physics Conference Series (Torque)*. Elsevier, 2018.

[20] S. Boersma, V. Rostampour, B. Doekemeijer, J.-W. van Wingerden, and T. Keviczky. Wind farm active power tracking using nonlinear model predictive control. *Submitted to IFAC Conference on Nonlinear Model Predictive Control (NMPC)*, 2018.

[21] F. Borrelli, A. Bemporad, and M. Morari. *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[22] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[23] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[24] S. Bubeck et al. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.

[25] G. Calafiore and M. C. Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.

[26] G. C. Calafiore. Scenario optimization methods in portfolio analysis and design. In *Optimal Financial Decision Making under Uncertainty*, pages 55–87. Springer, 2017.

[27] G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[28] G. C. Calafiore and L. Fagiano. Stochastic model predictive control of LPV systems via scenario optimization. *Automatica*, 49(6):1861–1866, 2013.

[29] G. C. Calafiore and D. Lyons. Random convex programs for distributed multi-agent consensus. In *European Control Conference (ECC)*, pages 250–255. IEEE, 2013.

[30] M. C. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.

[31] M. C. Campi and S. Garatti. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *Journal of Optimization Theory and Applications*, 148(2):257–280, 2011.

[32] M. C. Campi, S. Garatti, and M. Prandini. The scenario approach for systems and control design. *Annual Reviews in Control*, 33(2):149–157, 2009.

[33] M. Cannon, B. Kouvaritakis, S. V. Rakovic, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1):194–200, 2011.

[34] J. Carpentier. Contribution a l'etude du dispatching economique. *Bulletin de la Societe Francaise des Electriciens*, 3(1):431–447, 1962.

[35] M. Chamanbaz, F. Dabbene, and C. Lagoa. AC optimal power flow in the presence of renewable sources and uncertain loads. *arXiv*, 2017.

[36] S. Chawla, C. Dwork, F. McSherry, A. Smith, and H. Wee. Toward privacy in public databases. In *Theory of Cryptography Conference*, pages 363–385. Springer, 2005.

[37] R. Christie. Power systems test case archive, University of Washington, 2000.

[38] C. Conte, M. N. Zeilinger, M. Morari, and C. N. Jones. Robust distributed model predictive control of linear systems. In *European Control Conferance (ECC)*, pages 2764–2769. IEEE, 2013.

[39] F. Dabbene and D. Henrion. Set approximation via minimum-volume polynomial sublevel sets. In *European Control Conference (ECC)*, pages 1114–1119. IEEE, 2013.

[40] F. Dabbene, D. Henrion, C. Lagoa, and P. Shcherbakov. Randomized approximations of the image set of nonlinear mappings with applications to filtering. *IFAC Symposium on Robust Control Design*, 48(14):37–42, 2015.

[41] L. Dai, Y. Xia, Y. Gao, and M. Cannon. Distributed stochastic MPC of linear systems with additive uncertainty and coupled probabilistic constraints. *IEEE Transactions on Automatic Control*, 62(99):3474–3481, 2016.

[42] G. Darivianakis, S. Fattahi, J. Lygeros, and J. Lavaei. High-performance cooperative distributed model predictive control for linear systems. In *to appear in American Control Conference (ACC)*, 2018.

[43] M. P. Deisenroth. *Efficient reinforcement learning using Gaussian processes*, volume 9. KIT Scientific Publishing, 2010.

[44] S. X. Ding. *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science & Business Media, 2008.

[45] B. Dong. *Integrated bulding heating, cooling and ventilation control*. PhD thesis, Department of Architecture, Carnegie Mellon University, 2010.

[46] J. Dopazo, O. Klitin, and A. Sasson. Stochastic load flows. *IEEE Transactions on Power Apparatus and Systems*, 94(2):299–309, 1975.

[47] E. Dumon, M. Ruiz, H. Godard, and J. Maeght. SDP resolution techniques for the optimal power flow with unit commitment. In *Power Tech Conference*, pages 1–6. IEEE, 2017.

[48] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*, pages 265–284. Springer, 2006.

[49] C. Dwork, A. Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[50] I. Fagarasan, S. Ploix, and S. Gentil. Causal fault detection and isolation based on a set-membership approach. *Automatica*, 40(12):2099–2110, 2004.

[51] S. S. Farahani, Z. Lukszo, T. Keviczky, B. De Schutter, and R. M. Murray. Robust model predictive control for an uncertain smart thermal grid. In *European Control Conference (ECC)*, pages 1195–1200. IEEE, 2016.

[52] S. Fattahi, G. Fazelnia, J. Lavaei, and M. Arcak. Transformation of optimal centralized controllers into near-globally optimal static distributed controllers. *to appear in IEEE Transactions on Automatic Control*, 2018.

[53] R. M. Ferrari, T. Parisini, and M. M. Polycarpou. A robust fault detection and isolation scheme for a class of uncertain input-output discrete-time nonlinear systems. In *American Control Conference (ACC)*, pages 2804–2809. IEEE, 2008.

[54] R. M. Ferrari, T. Parisini, and M. M. Polycarpou. Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach. *IEEE Transactions on Automatic Control*, 57(2):275–290, 2012.

[55] P. M. Frank and X. Ding. Survey of robust residual generation and evaluation methods in observer-based fault detection systems. *Journal of Process Control*, 7(6):403–424, 1997.

[56] W. Fu and J. D. McCalley. Risk based optimal power flow. In *Power Tech Conference*, volume 3, pages 1–6. IEEE, 2001.

[57] X. Ge and Q. Han. Distributed fault detection over sensor networks with Markovian switching topologies. *International Journal of General Systems*, 43(3-4):305–318, 2014.

[58] R. Grone, C. R. Johnson, E. M. Sá, and H. Wolkowicz. Positive definite completions of partial Hermitian matrices. *Linear Algebra and Its Applications*, 58:109–124, 1984.

[59] L. J. Guibas, A. Nguyen, and L. Zhang. Zonotopes as bounding volumes. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 803–812. Society for Industrial and Applied Mathematics, 2003.

[60] J. Hagerman. Buildings-to-grid technical opportunities: Introduction and vision. Technical report, EERE Publication and Product Library, 2014.

[61] S. Han, U. Topcu, and G. J. Pappas. Differentially private convex optimization with piecewise affine objectives. In *Conference on Decision and Control (CDC)*, pages 2160–2166. IEEE, 2014.

[62] S. Han, U. Topcu, and G. J. Pappas. Differentially private distributed protocol for electric vehicle charging. In *Conference on Communication, Control, and Computing*, pages 242–249. IEEE, 2014.

[63] S. Han, U. Topcu, and G. J. Pappas. Differentially private distributed constrained optimization. *IEEE Transactions on Automatic Control*, 62(1):50–64, 2017.

[64] A. W. Harbaugh. *MODFLOW, the US Geological Survey modular ground-water model: the ground-water flow process*. US Department of the Interior, US Geological Survey Reston, VA, USA, 2005.

[65] Z. A. H. Hassan and C. Udriste. Equivalent reliability polynomials modeling eas and their geometries. *Annals of West University of Timisoara-Mathematics and Computer Science*, 53(1):177–195, 2015.

[66] B. He and X. Yuan. On non-ergodic convergence rate of Douglas–Rachford alternating direction method of multipliers. *Numerische Mathematik*, 130(3):567–577, 2015.

[67] O. Hernández-Lerma and J. B. Lasserre. *Further topics on discrete-time Markov control processes*, volume 42. Springer Science & Business Media, 2012.

[68] P. Hokayem, E. Cinquemani, D. Chatterjee, F. Ramponi, and J. Lygeros. Stochastic receding horizon control with output feedback and bounded controls. *Automatica*, 48(1):77–88, 2012.

[69] A. Ingimundarson, J. M. Bravo, V. Puig, T. Alamo, and P. Guerra. Robust fault detection using zonotope-based set-membership consistency test. *International Journal of Adaptive Control and Signal Processing*, 23(4):311–330, 2009.

[70] R. A. Jabr. Adjustable robust OPF with renewable energy sources. *IEEE Transactions on Power Systems*, 28(4):4742–4751, 2013.

[71] M. Jaxa-Rozen, M. Bloemendal, V. Rostampour, and J. Kwakkel. Assessing the sustainable application of aquifer thermal energy storage. In *European Geothermal Congress (EGC)*, 2016.

[72] A. Kalbat and J. Lavaei. A fast distributed algorithm for decomposable semidefinite programs. In *Conference on Decision and Control (CDC)*, pages 1742–1749. IEEE, 2015.

[73] O. Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 2006.

[74] T. Kanamori and A. Takeda. Worst-case violation of sampled convex programs for optimization with uncertainty. *Journal of Optimization Theory and Applications*, 152(1):171–197, 2012.

[75] B. Kouvaritakis, M. Cannon, and V. Tsachouridis. Recent developments in stochastic MPC and sustainable development. *Annual Reviews in Control*, 28(1):23–35, 2004.

[76] E. Kyriakides and M. Polycarpou. *Intelligent Monitoring, Control, and Security of Critical Infrastructure Systems*, volume 565. Springer, 2014.

[77] A. Y. Lam, B. Zhang, and N. T. David. Distributed algorithms for optimal power flow problem. In *Conference on Decision and Control (CDC)*, pages 430–437. IEEE, 2012.

[78] G. K. Larsen, N. D. van Foreest, and J. M. Scherpen. Distributed control of the power supply-demand balance. *IEEE Transactions on Smart Grid*, 4(2):828–836, 2013.

[79] G. K. Larsen, N. D. van Foreest, and J. M. Scherpen. Distributed MPC applied to a network of households with micro-CHP and heat storage. *IEEE Transactions on Smart Grid*, 5(4):2106–2114, 2014.

[80] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[81] J. Lavaei and A. G. Aghdam. Control of continuous-time LTI systems by means of structurally constrained controllers. *Automatica*, 44(1):141–148, 2008.

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

[83] K. Lehmann, A. Grastien, and P. Van Hentenryck. AC-feasibility on tree networks is NP-hard. *IEEE Transactions on Power Systems*, 31(1):798–801, 2016.

[84] D. Limon, J. Calliess, and J. Maciejowski. Learning-based nonlinear model predictive control. *IFAC World Congress*, 50(1):7769–7776, 2017.

[85] S. Liu, A. Sadowska, H. Hellendoorn, and B. De Schutter. Scenario-based distributed model predictive control for freeway networks. In *International Conference on Intelligent Transportation Systems (ITSC)*, pages 1779–1784. IEEE, 2016.

[86] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *International Symposium on Computer Aided Control Systems Design*, pages 284–289. IEEE, 2004.

[87] J. Löfberg. Oops! I cannot do it again: Testing for recursive feasibility in MPC. *Automatica*, 48(3):550–555, 2012.

[88] Y. Long, S. Liu, L. Xie, and K. H. Johansson. A scenario-based distributed stochastic MPC for building temperature regulation. In *Conference on Automation Science and Engineering (CASE)*, pages 1091–1096. IEEE, 2014.

[89] M. Lorenzen, F. Allgöwer, F. Dabbene, and R. Tempo. Scenario-based stochastic MPC with guaranteed recursive feasibility. In *Conference on Decision and Control (CDC)*, pages 4958–4963. IEEE, 2015.

[90] H. Lund, S. Werner, R. Wiltshire, S. Svendsen, J. E. Thorsen, F. Hvelplund, and B. V. Mathiesen. 4th generation district heating (4GDH): Integrating smart thermal grids into future sustainable energy systems. *Energy*, 68:1–11, 2014.

[91] Z. Luo, W. Ma, A. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.

[92] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, and P. Haves. Model predictive control for the operation of building cooling systems. *IEEE Transactions on Control Systems Technology*, 20(3):796–803, 2012.

[93] Y. Ma, S. Richter, and F. Borrelli. DMPC for building temperature regulation. In L. T. Biegler, S. L. Campbell, and V. Mehrmann, editors, *Control and Optimization with Differential-Algebraic Constraints*, chapter 14, pages 293–314. Springer International Publishing, 2012.

[94] M. Maasoumy Haghighi. Modeling and optimal control algorithm design for HVAC systems in energy efficient buildings. Master's thesis, EECS Department, University of California, Berkeley, 2011.

[95] R. Madani, M. Ashraphijuo, and J. Lavaei. Promises of conic relaxation for contingency-constrained optimal power flow problem. *IEEE Transactions on Power Systems*, 31(2):1297–1307, 2016.

[96] R. Madani, A. Kalbat, and J. Lavaei. ADMM for sparse semidefinite programming with applications to optimal power flow problem. In *Conference on Decision and Control (CDC)*, pages 5932–5939. IEEE, 2015.

[97] R. Madani, S. Sojoudi, and J. Lavaei. Convex relaxation for optimal power flow problem: Mesh networks. *IEEE Transactions on Power Systems*, 30(1):199–211, 2015.

[98] K. Margellos, P. Goulart, and J. Lygeros. On the road between robust optimization and the scenario approach for chance constrained optimization problems. *IEEE Transactions on Automatic Control*, 59(8):2258–2263, 2014.

[99] K. Margellos, M. Prandini, and J. Lygeros. On the connection between compression learning and scenario based single-stage and cascading optimization problems. *IEEE Transactions on Automatic Control*, 60(10):2716–2721, 2015.

[100] K. Margellos, V. Rostampour, M. Vrakopoulou, M. Prandini, G. Andersson, and J. Lygeros. Stochastic unit commitment and reserve scheduling: A tractable formulation with probabilistic certificates. In *European Control Conference (ECC)*, pages 2513–2518. IEEE, 2013.

[101] R. T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6):369–395, 2004.

[102] G. R. Marseglia, J. Scott, L. Magni, R. D. Braatz, and D. M. Raimondo. A hybrid stochastic-deterministic approach for active fault diagnosis using scenario optimization. *IFAC World Congress*, 47(3):1102–1107, 2014.

[103] A. Mesbah. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems*, 36(6):30–44, 2016.

[104] I. T. Michailidis, S. Baldi, M. F. Pichler, E. B. Kosmatopoulos, and J. R. Santiago. Proactive control for solar energy exploitation: A German high-inertia building case study. *Applied Energy*, 155:409–420, 2015.

[105] M. Milanese, J. Norton, H. Piet-Lahanier, and É. Walter. *Bounding approaches to system identification*. Springer Science & Business Media, 2013.

[106] A. Mirakhorli and B. Dong. Occupancy behavior based model predictive control for building indoor climate — A critical review. *Energy and Buildings*, 129:499–513, 2016.

[107] P. Mohajerin Esfahani and J. Lygeros. A tractable fault detection and isolation approach for nonlinear systems with probabilistic performance. *IEEE Transactions on Automatic Control*, 61(3):633–647, 2016.

[108] P. Mohajerin Esfahani, T. Sutter, D. Kuhn, and J. Lygeros. From infinite to finite programs: Explicit error bounds with applications to approximate dynamic programming. *arXiv*, 2017.

[109] P. Mohajerin Esfahani, T. Sutter, and J. Lygeros. Performance bounds for the scenario approach and an extension to a class of non-convex programs. *IEEE Transactions on Automatic Control*, 60(1):46–58, 2015.

[110] P. Mohajerin Esfahani, M. Vrakopoulou, G. Andersson, and J. Lygeros. A tractable nonlinear fault detection and isolation technique with application to the cyber-physical security of power systems. In *Conference on Decision and Control (CDC)*, pages 3433–3438. IEEE, 2012.

[111] 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 Transactions on Power Systems*, 28(4):3987–3998, 2013.

[112] R. D. Monteiro. First-and second-order methods for semidefinite programming. *Mathematical Programming*, 97(1-2):209–244, 2003.

[113] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz. Economic valuation of reserves in power systems with high penetration of wind power. *IEEE Transactions on Power Systems*, 24(2):900–910, 2009.

[114] P.-D. Moroşan, R. Bourdais, D. Dumur, and J. Buisson. A distributed MPC strategy based on Benders' decomposition applied to multi-source multi-zone temperature regulation. *Journal of Process Control*, 21(5):729–737, 2011.

[115] E. Noursadeghi and I. A. Raptis. Reduced-order distributed fault diagnosis for large-scale nonlinear stochastic systems. *Journal of Dynamic Systems, Measurement, and Control*, 140(5):051009, 2018.

[116] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45:15–27, 2012.

[117] G. Papaefthymiou and B. Klockl. MCMC for wind power simulation. *IEEE Transactions on Energy Conversion*, 23(1):234–240, 2008.

[118] A. Papavasiliou, S. Oren, and R. O'Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems*, 26(4):2197–2206, 2011.

[119] N. R. Patel, M. J. Risbeck, J. B. Rawlings, M. J. Wenzel, and R. D. Turney. Distributed economic model predictive control for large-scale building temperature regulation. In *American Control Conference (ACC)*, pages 895–900. IEEE, 2016.

[120] K. M. Powell and T. F. Edgar. An adaptive-grid model for dynamic simulation of thermocline thermal energy storage systems. *Energy Conversion and Management*, 76:865–873, 2013.

[121] M. Prandini, S. Garatti, and J. Lygeros. A randomized approach to stochastic model predictive control. In *Conference on Decision and Control (CDC)*, pages 7315–7320. IEEE, 2012.

[122] J. A. Primbs. A soft constraint approach to stochastic receding horizon control. In *Conference on Decision and Control (CDC)*, pages 4797–4802. IEEE, 2007.

[123] J. Qi, Y. Kim, C. Chen, X. Lu, and J. Wang. Demand response and smart buildings: A survey of control, communication, and cyber-physical security. *ACM Transactions Cyber-Physical Systems*, 1(3):1–28, 2016.

[124] D. M. Raimondo, G. R. Marseglia, R. D. Braatz, and J. K. Scott. Fault-tolerant model predictive control with active fault isolation. In *Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 444–449. IEEE, 2013.

[125] M. Razmara, G. R. Bharati, M. Shahbakhti, S. Paudyal, and R. D. Robinett III. Bilevel optimization framework for smart Building-to-Grid systems. *IEEE Transactions on Smart Grid*, 9(99):582–593, 2016.

[126] A. Richards and J. P. How. Robust distributed model predictive control. *International Journal of Control*, 80(9):1517–1531, 2007.

[127] M. Rivarolo, A. Greco, and A. Massardo. Thermo-economic optimization of the impact of renewable generators on poly-generation smart-grids including hot thermal storage. *Energy Conversion and Management*, 65:75–83, 2013.

[128] S. Riverso, F. Boem, G. Ferrari-Trecate, and T. Parisini. Plug-and-play fault detection and control-reconfiguration for a class of nonlinear large-scale constrained systems. *IEEE Transactions on Automatic Control*, 61(12):3963–3978, 2016.

[129] S. Riverso, M. Farina, and G. Ferrari-Trecate. Plug-and-play decentralized model predictive control for linear systems. *IEEE Transactions on Automatic Control*, 58(10):2608–2614, 2013.

[130] L. Roald and G. Andersson. Chance-constrained AC optimal power flow: Reformulations and efficient algorithms. *arXiv*, 2017.

[131] L. Roald, F. Oldewurtel, B. Van Parys, and G. Andersson. Security constrained optimal power flow with distributionally robust chance constraints. *arXiv*, 2015.

[132] V. Rostampour. Tractable reserve scheduling formulations for power systems with uncertain generation. Master's thesis, DEIB Department, Politecnico di Milano, Italy, 2013.

[133] V. Rostampour, D. Adzkiya, S. E. Z. Soudjani, B. De Schutter, and T. Keviczky. Chance-constrained model predictive controller synthesis for stochastic max-plus linear systems. In *Conference on Systems, Man, and Cybernetics (SMC)*, pages 3581–3588. IEEE, 2016.

[134] V. Rostampour, W. Ananduta, and T. Keviczky. Distributed stochastic thermal energy management in smart thermal grids. In P. Palensky, M. Cvetkovic, and T. Keviczky, editors, *Submitted to PowerWeb: Intelligent Energy Systems*, chapter 6. Springer International Publishing, 2018.

[135] V. Rostampour, M. Bloemendal, M. Jaxa-Rozen, and T. Keviczky. A control-oriented model for combined building climate comfort and aquifer thermal energy storage system. In *European Geothermal Congress (EGC)*, 2016.

[136] V. Rostampour, M. Bloemendal, and T. Keviczky. Aquifer thermal energy storage for seasonal thermal energy balance. In *European Geosciences Union (EGU) General Assembly Conference*, volume 19, page 12077, 2017.

[137] V. Rostampour, M. Bloemendal, and T. Keviczky. A model predictive framework of ground source heat pump coupled with aquifer thermal energy storage system in heating and cooling equipment of a building. In *International Energy Agency (IEA) Heat Pump Conference*, pages 1–12, 2017.

[138] V. Rostampour, P. M. Esfahani, and T. Keviczky. Stochastic nonlinear model predictive control of an uncertain batch polymerization reactor. *IFAC Conference on Nonlinear Model Predictive Control (NMPC)*, 48(23):540–545, 2015.

[139] V. Rostampour, R. Ferrari, and T. Keviczky. A set based probabilistic approach to threshold design for optimal fault detection. In *American Control Conference (ACC)*, pages 5422–5429. IEEE, 2017.

[140] V. Rostampour, R. Ferrari, A. Teixeira, and T. Keviczky. Differentially private distributed fault diagnosis for large-scale nonlinear uncertain systems. *to appear in IFAC Conference on Fault Detection, Supervision and Safety (SAFEPROCESS)*, 2018.

[141] V. Rostampour, R. Ferrari, A. Teixeira, and T. Keviczky. Privatized distributed anomaly detection for large-scale nonlinear uncertain systems. *Submitted to IEEE Transactions on Automatic Control*, 2018.

[142] V. Rostampour, M. Jaxa-Rozen, M. Bloemendal, and T. Keviczky. Building climate energy management in smart thermal grids via aquifer thermal energy storage systems. *Energy Procedia*, 97:59–66, 2016.

[143] V. Rostampour, M. Jaxa-Rozen, M. Bloemendal, J. Kwakkel, and T. Keviczky. Aquifer thermal energy storage (ates) smart grids: Large-scale seasonal energy storage as a distributed energy management solution. *Submitted to Applied Energy*, 2018.

[144] V. Rostampour and T. Keviczky. Robust randomized model predictive control for energy balance in smart thermal grids. In *European Control Conference (ECC)*, pages 1201–1208. IEEE, 2016.

[145] V. Rostampour and T. Keviczky. Energy management for building climate comfort in uncertain smart thermal grids with aquifer thermal energy storage. *IFAC World Congress*, 50:13156–13163, 2017.

[146] V. Rostampour and T. Keviczky. Distributed stochastic model predictive control synthesis for large-scale uncertain linear systems. In *to appear in American Control Conference (ACC)*. IEEE, 2018.

[147] V. Rostampour and T. Keviczky. Distributed stochastic MPC for large-scale systems with private and common uncertainty sources. *Submitted to IEEE Transactions on Control of Networked Systems*, 2018.

[148] V. Rostampour and T. Keviczky. Probabilistic energy management for building climate comfort in smart thermal grids with seasonal storage systems. *IEEE Transactions on Smart Grid*, 2018.

[149] V. Rostampour, K. Margellos, M. Vrakopoulou, M. Prandini, G. Andersson, and J. Lygeros. Reserve requirements in AC power systems with uncertain generation. In *Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, pages 1–5. IEEE, 2013.

[150] V. Rostampour, O. ter Haar, and T. Keviczky. Distributed stochastic reserve scheduling in AC power systems with uncertain generation. *Submitted to IEEE Transactions on Power Systems*, 2017.

[151] V. Rostampour, O. ter Haar, and T. Keviczky. Tractable reserve scheduling in AC power systems with uncertain wind power generation. In *Conference on Decision and Control (CDC)*, pages 2647–2654. IEEE, 2017.

[152] V. Rostampour, O. ter Haar, and T. Keviczky. Computationally tractable reserve scheduling for AC power systems with wind power generation. In P. Palensky, M. Cvetkovic, and T. Keviczky, editors, *Submitted to PowerWeb: Intelligent Energy Systems*, chapter 6. Springer International Publishing, 2018.

[153] B. Safarinejadian and E. Kowsari. Fault detection in non-linear systems based on GP-EKF and GP-UKF algorithms. *Journal of Systems Science & Control Engineering*, 2(1):610–620, 2014.

[154] M. Sameti and F. Haghighat. Optimization approaches in district heating and cooling thermal network. *Energy and Buildings*, 140:121–130, 2017.

[155] L. Sankar, S. Kar, R. Tandon, and H. V. Poor. Competitive privacy in the smart grid: An information-theoretic approach. In *Conference on Smart Grid Communications (SmartGridComm)*, pages 220–225. IEEE, 2011.

[156] H. Satyavada and S. Baldi. An integrated control-oriented modelling for HVAC performance benchmarking. *Journal of Building Engineering*, 6:262–273, 2016.

[157] G. Schildbach, L. Fagiano, C. Frei, and M. Morari. The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations. *Automatica*, 50(12):3009–3018, 2014.

[158] G. Schildbach, L. Fagiano, and M. Morari. Randomized solutions to convex programs with multiple chance constraints. *SIAM Journal on Optimization*, 23(4):2479–2501, 2013.

[159] G. Schildbach and M. Morari. Scenario MPC for linear time-varying systems with individual chance constraints. In *American Control Conference (ACC)*, pages 415–421. IEEE, 2015.

[160] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin. On the linear convergence of the ADMM in decentralized consensus optimization. *IEEE Transaction on Signal Processing*, 62(7):1750–1761, 2014.

[161] M. Slater. Lagrange multipliers revisited. Technical report, Cowles Foundation for Research in Economics, Yale University, 1959.

[162] G. Stamatescu, I. Stamatescu, N. Arghira, V. Calofir, and I. Fagarasan. Building cyber-physical energy systems. *arXiv*, 2016.

[163] L. Standardi. *Economic model predictive control for large-scale and distributed energy systems*. PhD thesis, Department of Applied Mathematics and Computer Science , Technical University of Denmark, 2015.

[164] B. Stott, J. Jardim, and O. Alsaç. DC power flow revisited. *IEEE Transactions on Power Systems*, 24(3):1290–1300, 2009.

[165] A. F. Taha, N. Gatsis, B. Dong, A. Pipri, and Z. Li. Building-to-Grid integration framework. *IEEE Transactions on Smart Grid*, 2017.

[166] C. Ummenhofer, J. Olsen, J. Page, and T. Roediger. How to improve peak time coverage through a smart-controlled MCHP unit combined with thermal and electric storage systems. *Energy and Buildings*, 139:78–90, 2017.

[167] E. Van Vliet. Flexibility in heat demand at the TU Delft campus smart thermal grid with phase change materials. Master's thesis, Department P&E, 3ME Faculty, Delft University of Technology, The Netherlands, 2013.

[168] L. Vandenberghe and M. S. Andersen. Chordal graphs and semidefinite optimization. *Foundations and Trends® in Optimization*, 1(4):241–433, 2015.

[169] A. N. Venkat, I. A. Hiskens, J. B. Rawlings, and S. J. Wright. Distributed MPC strategies with application to power system automatic generation control. *IEEE Transactions on Control Systems Technology*, 16(6):1192–1206, 2008.

[170] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis. Convex relaxations of chance constrained AC optimal power flow. *arXiv*, 2017.

[171] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. A probabilistic framework for reserve scheduling and N-1 security assessment of systems with high wind power penetration. *IEEE Transactions on Power Systems*, 28(4):3885–3896, 2013.

[172] T. Wada, R. Morita, T. Asai, I. Masubuchi, and Y. Fujisaki. A randomized algorithm for chance constrained optimal power flow with renewables. *Journal of Control Measurement and System Integration*, 10(4):303–309, 2017.

[173] S. Wang and X. Xu. Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm. *Energy conversion and management*, 47(13):1927–1941, 2006.

[174] J. Warrington, P. J. Goulart, S. Mariéthoz, and M. Morari. Robust reserve operation in power systems using affine policies. In *Conference on Decision and Control (CDC)*, pages 1111–1117. IEEE, 2012.

[175] A. Yahiaoui, J. Hensen, L. Soethout, and D. Van Paassen. Model based optimal control for integrated building systems. In *International Postgraduate Research Conference in the Built and Human Environment*, pages 322–332, 2006.

[176] B. E. Ydstie. New vistas for process control: Integrating physics and communication networks. *AIChE Journal*, 48(3):422–426, 2002.

[177] H. Ye. Surrogate affine approximation based co-optimization of transactive flexibility, uncertainty, and energy. *IEEE Transactions on Power Systems*, 2018.

[178] L. A. Zadeh. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy sets and systems*, 90(2):111–127, 1997.

[179] D. Zhang, Q. Wang, L. Yu, and H. Song. Fuzzy-model-based fault detection for a class of nonlinear systems with networked measurements. *IEEE Transactions on Instrumentation and Measurement*, 62(12):3148–3159, 2013.

[180] Q. Zhang and X. Zhang. Distributed sensor fault diagnosis in a class of interconnected nonlinear uncertain systems. *Annual Reviews in Control*, 37(1):170–179, 2013.

[181] X. Zhang, K. Margellos, P. Goulart, and J. Lygeros. Stochastic model predictive control using a combination of randomized and robust optimization. In *Conference on Decision and Control (CDC)*, pages 7740–7745. IEEE, 2013.

[182] H. Zhu and G. B. Giannakis. Multi-area state estimation using distributed SDP for nonlinear power systems. In *Conference on Smart Grid Communication (SmartGridComm)*, pages 623–628. IEEE, 2012.

[183] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, 2011.

[184] R. Zippel. *Effective polynomial computation*, volume 241. Springer Science & Business Media, 2012.

AUTHOR BIOGRAPHY

Vahab Rostampour received the M.Sc. degree in Automation and Control Engineering from Politecnico di Milano, Italy in 2013. He has been employed as a Ph.D. candidate at the Delft Center for Systems and Control, Delft University of Technology from September 2014, under the supervision of Dr. Tamás Keviczky and Prof. Nathan van de Wouw. He also obtained the certificate of the Dutch Institute of Systems and Control.

His research interests include optimal control synthesis, analysis, stochastic model predictive control, distributed data-driven and learning optimization of dynamical systems, and decision making in uncertain dynamical environments, with applications to large-scale systems such as thermal grids and power networks. His main research focuses on various centralized and decentralized problems with and without communication between subsystems, and finding tractable ways of solving these problems.

List of Publications

International Journal Articles

J_6 **V. Rostampour**, M. Jaxa-Rozen, M. Bloemendal, J. Kwakkel, T. Keviczky, *Aquifer Thermal Energy Storage (ATES) Smart Grids: Large-Scale Seasonal Energy Storage as A Distributed Energy Management Solution*, Submitted to Applied Energy, August 2018.

J_5 **V. Rostampour**, R. Ferrari, A. H. Teixeira, T. Keviczky, *Privatized Distributed Anomaly Detection for Large-Scale Nonlinear Uncertain Systems*, Submitted to IEEE Transactions on Automatic Control, July 2018.

J_4 **V. Rostampour**, T. Keviczky, *Distributed Stochastic MPC for Large-Scale Systems with Private and Common Uncertainty Sources*, Submitted to IEEE Transactions on Control of Networked Systems, February 2018.

J_3 **V. Rostampour**, O. ter Haar, T. Keviczky, *Distributed Stochastic Reserve Scheduling in AC Power Systems With Uncertain Generation*, Submitted to IEEE Transactions on Power Systems, December 2017.

J_2 **V. Rostampour**, T. Keviczky, *Probabilistic Energy Management for Building Climate Comfort in STGs with Seasonal Storage Systems*, To appear in IEEE Transactions on Smart Grid, 2018.

J_1 **V. Rostampour**, M. Jaxa-Rozen, M. Bloemendal, T. Keviczky, *Building Climate Energy Management in Smart Thermal Grids via ATES Systems*, Energy Procedia, Vol. 97, pp. 59-66, 2016.

Book Chapters

B_2 **V. Rostampour**, O. ter Haar, T. Keviczky, *Computationally Tractable Reserve Scheduling for AC Power Systems with Wind Power Generation*, Submitted to PowerWeb: Intelligent Energy Systems, Chapter 6, Springer, March 2018.

B_1 **V. Rostampour**, W. Ananduta, T. Keviczky, *Distributed Stochastic Thermal Energy Management in Smart Thermal Grids*, Submitted to PowerWeb: Intelligent Energy Systems, Chapter 9, Springer, March 2018.

Peer-Reviewed Conference Papers

C_{16} S.Boersma, **V. Rostampour**, B.Doekemeijer, J.W. van Wingerden, T. Keviczky, *Wind Farm Active Power Tracking Using Nonlinear Model Predictive Control*, Submitted to IFAC Conference on Nonlinear Model Predictive Control (NMPC), Madison, Wisconsin, US, Aug 2018.

C_{15} **V. Rostampour**, R. Ferrari, A. H. Teixeira, T. Keviczky, *Differentially Private Distributed Fault Diagnosis for Large-Scale Nonlinear Uncertain Systems*, To appear in IFAC Conference on Fault Detection, Supervision and Safety (SAFEPROCESS), Warsaw, Poland, Aug 2018.

C_{14} **V. Rostampour**, T. Keviczky, *Distributed Stochastic Model Predictive Control Synthesis for Large-Scale Uncertain Linear Systems*, to appear in Proceedings of American Control Conference (ACC), Milwaukee, Wisconsin, US, Jun 2018.

C_{13} S.Boersma, **V. Rostampour**, B.Doekemeijer, J.W. van Wingerden, *A Centralized Model Predictive Wind Farm Controller in PALM Providing Power Reference Tracking*, to appear in Proceedings of Journal of Physics — Conference Series (Torque), Milan, Italy, Jun 2018.

C_{12} **V. Rostampour**, O. ter Haar, T. Keviczky, *Tractable Reserve Scheduling in AC Power Systems With Uncertain Wind Power Generation*, in Proceedings of Conference on Decision and Control (CDC), pp. 2647-2654, Melbourne, Australia, Dec 2017.

C_{11} **V. Rostampour**, T. Keviczky, *Energy Management for Building Climate Comfort in Uncertain Smart Thermal Grids with ATES*, in International Federation of Automatic Control World Congress (IFAC), pp. 13698-13705, Toulouse, France, Jul 2017.

C_{10} M. Bloemendal, M. Jaxa-Rozen, **V. Rostampour**, *Use it or Lose it: Adaptive Governance of Aquifers with ATES*, in Proceedings of International Energy Agency (IEA) Conference on Heat Pump, Rotterdam, The Netherlands, May 2017.

C_9 **V. Rostampour**, R. Ferrari, T. Keviczky, *Set Based Probabilistic Approach to Threshold Design for Optimal Fault Detection*, in Proceedings of American Control Conference (ACC), pp. 5422-5429, Seattle, WA, US, May 2017.

C_8 **V. Rostampour**, T. Keviczky, *A MPC Framework of GSHP coupled with ATES System in Heating and Cooling Networks of a Building*, in Proceedings of International Energy Agency (IEA) Conference on Heat Pump, Rotterdam, The Netherlands, May 2017.

C_7 **V. Rostampour**, Dieky Adzkiya, Sadegh Soudjani, Bart De Schutter, T. Keviczky, *Chance Constrained MPC Synthesis for Stochastic Max-Plus Linear Systems*, in Proceedings of Systems, Man, and Cybernetics Conference (SMC), pp. 3581-3588, Budapest, Hungary, Oct 2016.

C_6 **V. Rostampour**, M. Bloemendal, T. Keviczky, *A Control-Oriented Model For Combined Building Climate Comfort and ATES System*, in Proceedings of European Geothermal Congress (EGC), Strasbourg, France, Sep 2016.

C_5 M. Jaxa-Rozen, M. Bloemendal, **V. Rostampour**, J. Kwakkel, *Assessing the Sustainable and Optimal Application of Aquifer Thermal Energy Storage*, in Proceedings of European Geothermal Congress (EGC), Strasbourg, France, Sep 2016.

C_4 **V. Rostampour**, T. Keviczky, *Robust Randomized Model Predictive Control for Energy Balance in Smart Thermal Grids*, in Proceedings of European Control Conference (ECC), pp. 1201-1208, Aalborg, Denmark, Jun 2016.

C_3 **V. Rostampour**, P. Mohajerin Esfahani, T. Keviczky, *Stochastic Nonlinear Model Predictive Control of an Uncertain Batch Polymerization Reactor*, in IFAC Conference on Nonlinear Model Predictive Control (NMPC), pp. 540-545, Seville, Spain, Sep 2015.

C_2 **V. Rostampour**, K. Margellos, M. Vrakopoulou, M. Prandini, G. Andersson and J. Lygeros, *Stochastic Nonlinear Model Predictive Control of an Uncertain Batch Polymerization Reactor*, in Proceedings of Innovative Smart Grid Technologies Europe (ISGT), pp. 1-5, Copenhagen, Denmark, Oct 2013.

C_1 K. Margellos, **V. Rostampour**, M. Vrakopoulou, M. Prandini, G. Andersson, J. Lygeros, *Stochastic Unit Commitment and Reserve Scheduling: A Tractable Formulation with Probabilistic Certificates*, in Proceedings of European Control Conference (ECC), pp. 2513-2518, Züriuch, Switzerland, June 2013.