The Power and Limits of Convex Relaxations:
Power Systems and Integer Programming

Matías Javier Villagra-Morales

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2026

Abstract

The Power and Limits of Convex Relaxations:

Power Systems and Integer Programming

Matías Javier Villagra-Morales

This dissertation studies the design and analysis of convex relaxations for optimization problems arising in power systems and integer programming. Convex relaxations play a central role in modern optimization, yet their effectiveness depends critically on structure, scalability, and modeling choices.

In the first chapter, we leverage structural properties of power flow equations to develop scalable linear cutting-plane methods that tightly approximate second-order cone relaxations of the Alternating-Current Optimal Power Flow (ACOPF) problem. The proposed approach yields relaxations that are accurate, numerically stable, and amenable to warm-starting, enabling the solution of large-scale multi-period instances with millions of variables and constraints. By preserving key physical features, such as nonnegative active power losses, the resulting relaxations are tight and produce high-quality bounds. These methods support applications such as real-time operations and the computation of nodal electricity prices at a fraction of the computational cost of existing convex nonlinear approaches.

In the second chapter, we take a complementary perspective by analyzing the performance of convex relaxations for structured integer programs. We investigate the role of symmetry in lift-and-project hierarchies and introduce a framework based on k-transitivity to quantify symmetry and characterize when these hierarchies succeed or fail to strengthen the standard linear programming

relaxation. Our results show that symmetry can fundamentally limit the strength of classical hierarchies. We identify structural properties that govern the effectiveness of convex relaxations in symmetric settings, with implications for integrality gaps and algorithmic performance.

Together, these contributions demonstrate that carefully designed linear relaxations can achieve both high accuracy and scalability, while revealing fundamental limitations of existing methods in structured settings. They provide new insights for developing structure-aware optimization techniques across both continuous and discrete domains.

Table of Contents

Acknowledgments x

Dedication xii

Chapter 1: Introduction 1

1.1 Convex Relaxations as a Unifying Framework 1

1.2 Research Overview 2

1.2.1 Mathematical Optimization for Power Systems 2

1.2.2 Symmetry and Relaxations in Integer Programming 4

1.3 Preliminaries and Notation 5

1.3.1 Linear Algebra 5

1.3.2 Computational Complexity 6

1.3.3 Optimization Problem Classes 8

Chapter 2: Mathematical Optimization for Power Systems 10

2.1 Introduction 10

2.2 Background on Power Systems 13

2.3 Formal Problem Statement 19

2.3.1 Multi-Period Formulation 21

2.4 Prior Work on ACOPF and Relaxations 22

2.5 SOCP Relaxations for ACOPF and Power Losses 27
2.5.1 Jabr Relaxation 27
2.5.2 A Stronger SOCP 28
2.5.3 Losses 29
2.5.4 Losses and Cutting-Planes 34
2.5.5 A Numerically Better-Behaved i2 Relaxation 36
2.6 Cutting-plane Framework: Cut Separation and Management 40
2.6.1 Two Simple Cut Procedures 41
2.6.2 Basic Cutting-Plane Algorithm 43
2.7 On the Accuracy of Lower Bounds 46
2.8 Computational Experiments 52
2.8.1 Single-Period Experiments 54
2.8.2 Warm-Started Single-Period Instances 57
2.8.3 Multi-Period Results 61
2.8.4 Stability of Dual Variables for Active-Power Balance Constraints 73
2.9 SDP-based Strengthening of the Relaxations 74
2.9.1 Revisiting ACOPF through QCQP Formulations 74
2.9.2 SDP Relaxations for ACOPF 77
2.9.3 On the Relationship between SDP-R and SDP-CR 81
2.9.4 Our Disciplined Approach to SDPs 86
2.9.5 SDP Cuts 87
2.9.6 Algorithm and Preliminary Computational Results 92
2.10 An Application to Electricity Markets 95

2.10.1 Algorithmic Price Formation 99
2.10.2 Cutting-Plane Pricing Algorithm 100
2.10.3 Computational Setup and Results 102
2.11 AC-Feasibility is Weakly $\mathbb{NP}$-hard on Trees: A Corrected Proof 109
2.12 Conclusions 114

Chapter 3: Symmetry and Relaxations in Integer Programming 116
3.1 Introduction 116
3.2 Background 120
3.2.1 Integer Programming Hierarchies 121
3.2.2 Symmetries 127
3.3 Known Results on Hierarchies and Integrality Gaps 130
3.4 Symmetry, Hierarchies, and Polyhedral Structure 133
3.4.1 Geometric Properties of LP Hierarchies 133
3.4.2 Geometric Properties of 2-transitive Polytopes 136
3.4.3 A Folklore Result: Sherali–Adams Dominates the Lift-and-Project Closure 138
3.4.4 Symmetric Operators 141
3.4.5 On k-Transitive Groups 146
3.5 On the Improvement of Integrality Gaps 147
3.5.1 k-Resistance 149
3.6 Examples and Applications 155
3.6.1 A 3-Dimensional Example 155
3.6.2 Steiner Triple Covering Problems 156

3.6.3 Parity and Knapsack-Cover 160

3.6.4 The Cropped Cube 162

3.7 Main Theorem: LP Hierarchies Collapse under Symmetry 163

3.7.1 Assumptions and Consequences of Theorem 3.1 163

3.7.2 A High-Level Description of the Proof of (A) $\Rightarrow$ (B) 168

3.7.3 Proof of (A) $\Rightarrow$ (B) 169

3.7.4 Remaining Proofs for Theorem 3.1 189

3.7.5 An Elementary Proof for 2-Transitive Polytopes 191

3.8 Extending the Main Theorem: Beyond Polyhedral Hierarchies 193

3.8.1 Lovász–Schrijver Operator LS_+ with PSD constraint 193

3.9 Extension: A Procedure for Analyzing the SA Rank 199

3.9.1 Proof of the Generalization by Revisiting Section 3.7 200

3.9.2 The Procedure 205

3.9.3 Maximum Stable Set 207

3.9.4 Minimum Knapsack 209

3.9.5 Knapsack-Cover Inequalities 210

3.10 Conclusions 214

Chapter 4: Conclusions and Future Work 215

References 218

Appendix A: Additional Proofs and Computational Results 232

A.1 Reformulating Maximum Angle Difference Constraints as Thermal Limits 232

A.2 i2 SOCP Stronger than Jabr SOCP 233

A.3 Definition of the i2 Variable 236

A.4 Superoptimality and SOCPs 238

A.5 Warm-Started Relaxation for $T = 12, 24$ 240

A.6 On the Inequivalence of the Matrices $L(X)$ and W 242

A.7 Proof of a Result on Valid Inequalities for Polytopes 243

A.8 Sherali-Adams Decomposition Lemma: A Complete Proof 244

A.8.1 Definitions 244

A.8.2 Simple Facts 246

A.8.3 The Proof 250

List of Figures

2.1 Π-model representation of a transmission line $\{k, m\}$ with a series admittance (red), a transformer at bus k (grey), and shunt elements (green). Power injected at bus k, S_{km}, flows with current I_{km} at voltage V_k before entering the transformer. The transformer scales the voltage to V_k/N, hence the current is correspondingly scaled to $N^* I_{km}$. The flow then passes through the series element $y_{\{k,m\}}$ and line charging devices $\frac{1}{2}y^{sh}$. The symbols "0" denote the ground reference. 17

2.2 ACOPF Polar Formulation 20

2.3 Circuit representation of case3 from MATPOWER. 33

2.4 The black circle represents a horizontal cross-section of a projected Jabr cone (2.12) onto the $(v_k^{(2)}, v_m^{(2)}, c_{km})$-space. The red and blue inequalities represent (2.27) for a simple transmission line, while the teal and cyan inequalities represent two Jabr outer-approximation cuts. 38

2.5 Linearly constrained base model M_0 43

2.6 Validity of bounds and superoptimality of ϵ-feasible solutions. 48

2.7 12hr load curve. 64

2.8 24hr load curve. 64

2.9 Cut management for $T = 4$ warm-started instances. 70

2.10 Lower bounds for ACTIVSg25k and number of cuts. 70

2.11 Distance between dual variables of active-power balance constraints of consecutive cutting-plane iterations. 74

2.12 ACOPF Complex-Voltage QCQP Formulation 76

2.13 ACOPF Rectangular QCQP Formulation 77

2.14 SDP Relaxation of (2.38) 78

2.15 SDP Relaxation of (2.39) 79

2.16 Power Flow Welfare Maximization 96

2.17 Average nodal price distance δ for the CP and DC relaxations relative to the Jabr SOCP relaxation. 104

3.1 Polytope $P \coloneqq \{x \in \mathbb{R}^2 : x_1 + x_2 \geq \frac{1}{4},\ 0 \leq x_1, x_2 \leq 1\}$ (shaded in light blue) together with $P_1(P)$ (left) and $\mathrm{LS}_0(P)$ (right), whose contours are shown in red dashed lines. 121

3.2 Polytope P_ϵ with $\epsilon = 1/6$. 140

3.3 A depiction of the optimal LP faces of (3.9) (left) and (3.10) (right). 149

3.4 A depiction of $P = \mathrm{conv}(\{\pi(0, 1/2, 1) : \pi \in \mathcal{S}_3\})$ (left), $\mathrm{SA}^1(P) = \mathrm{LS}^1(P)$ (right, violet), and $\mathrm{LS}_0^1(P)$ (right, gray). 156

3.5 A geometric representation of P_3^{CC} in $[0, 1]^3$. 162

3.6 On the left, the polytope $P \coloneqq \mathrm{conv}\left(\{\pi(1, 1/2, 0), \pi(1, 1, 1/10) : \pi \in C_3\}\right)$ from Section 3.7.1. On the right, a nonempty subset of $\mathrm{LS}_0(P)$. 166

3.7 Schematic representation of A_Δ^4, with the entries in bold corresponding to the surviving $\tilde{A}^2$ submatrix after deleting rows 1-2 and columns 1-2. 182

List of Tables

2.1 Average and minimum losses 100 repetitions of removing a randomly selected SOC constraint. 32

2.2 Average and minimum losses 100 repetitions of removing 100 randomly selected SOC constraints. 33

2.3 Active-power losses and optimality gaps. 34

2.4 Statistics of the coefficient α in (2.26) and of the coefficients in (2.27). 38

2.5 Performance of Cutting-Plane Algorithm (Not Warm-Started) 56

2.6 Solvers' Performance on Jabr SOCP 58

2.7 Solvers' Performance on i2 SOCP 59

2.8 Warm-Started Relaxations, Loads perturbed Gaussian $(\mu, \sigma) = (0.01 \cdot P_d, 0.01 \cdot P_d)$ 62

2.9 Warm-Started Relaxations, Transmission Line with largest flow turned off 63

2.10 Cutting-Plane Not Warm-Started, $T = 4$ and loads perturbed Unif$(0, 0.025) + 0.01 \cdot t$ for $t \in \{1, 2, 3\}$, ramping rates 50% 68

2.11 Warm-Started Relaxation versus Jabr SOCP, $T = 4$ and loads perturbed Unif $(0, 0.025) + 0.01 \cdot t$ for $t \in \{1, 2, 3\}$, ramping rates 50% 71

2.12 Warm-Started Relaxation versus i2 SOCP+, $T = 4$ and loads perturbed Unif $(0, 0.025) + 0.01 \cdot t$ for $t \in \{1, 2, 3\}$, ramping rates 50% 71

2.13 Warm-Started Relaxation versus DCOPF, $T = 12$ 72

2.14 Warm-Started Relaxation versus DCOPF, $T = 24$ 72

2.15 On the convergence of dual variables of active-power balance constraints $T = 4$ . . 73

2.16 New Lower Bounds with Linearly-Constrained Relaxations 94

2.17 Warm-Started Relaxations, Loads perturbed Gaussian $(\mu, \sigma) = (0.01 \cdot P_d, 0.01 \cdot P_d)$ 94

2.18 Price Statistics and Run Time for Jabr SOCP, DC and CP with IP 103

2.19 Economic Efficiency Metrics for Jabr SOCP, DC and CP Relaxations with IP after Redispatch 103

2.20 Price Statistics and Run Time for Jabr SOCP, DC and CP with CH Policy 105

2.21 Price Statistics and Run Time for Jabr SOCP, DC and CP with CH Policy under Contingencies 106

2.22 Reactive Power Prices for Jabr SOCP, DC and CP with CH 106

A.1 Warm-Started Relaxation versus i2 SOCP+, $T = 12$ hr active power load curve, ramping rates 50% 241

A.2 Warm-Started Relaxation versus i2 SOCP+, $T = 24$ hr active power load curve, ramping rates 50% 241

Acknowledgements

I am deeply grateful for the opportunities that life has presented to me and for the people who have helped shape them. In particular, I would like to thank the following wonderful people, and one very special companion.

My doctoral advisors, Dan Bienstock and Yuri Faenza, have both been generous, patient, and understanding through my mistakes and failures, while remaining endlessly supportive in celebrating my successes. True ride-or-dies. I appreciate every discussion we have shared – from proving lemmas to talking about Frito, life's curiosities, and insane workout routines. I also thank Santanu Dey, Bento Natura, and Andy Sun for serving on my dissertation committee and for their valuable comments and suggestions.

My family, to whom I would like to dedicate a few words in Spanish. *Quiero darles las gracias, en particular a mi mamá quien ha sido un pilar fundamental en mi carrera. Haber podido contar con tu compañía y con la de Sofía durante mi defensa y el tramo final de la tesis fue un verdadero regalo de la vida. A mis dos papás, Mario y Patricio, gracias por haberme enseñado, cada uno a su manera, el valor del esfuerzo y la perseverancia. También estoy muy agradecido por el apoyo de mis hermanas, en particular el de Sofía, cuya curiosidad por las cosas en las que trabajo siempre me ha conmovido. Y, por supuesto, a Frito, que estuvo en todas. Caminar contigo cada día me ayudó a recordar que había vida más allá de las demostraciones, sobre todo cuando estas se ponían esquivas.*

My old and new friends. La Jóse, a longtime friend who first inspired me to pursue a career in applied mathematics and whose encouragement, unwavering emotional support, and countless conversations have meant so much to me. Nicolás Figueroa, Jan Kiwi, and José Verschae, friends and mentors whose guidance during my master's theses in Economics and Mathematics helped lay the foundations of my academic career. My newer close friends, Mr. Cid, Cuxito, Darshan, Pipe Serrano, and Thomas, whose friendship and constant support I will always appreciate. Our conversations have challenged, grounded, inspired, and entertained me over the years. Chefi, Chengpi, Tianyu, and Viraat, whose friendship made Columbia 345 feel like home. Pipe Verástegui and Matías Romero, who believed in me when I proposed the pricing paper project and made my first PhD-led collaboration an especially memorable one.

The partners who accompanied me through these past years, in particular Carolina and Renata. Renata, thank you for taking care of Frito during my first few months at Columbia. You made my transition to this new country immeasurably smoother and gave me invaluable peace of mind. I will always be grateful for your companionship throughout the first year of my PhD. Carolina, thank you for your support and presence during Viaje2 in July 2025. I firmly believe that your love and care nourished me in ways that ultimately helped me prove a central technical result of Chapter 3, Lemma 3.41.

To all of you, thank you. I am the sum of my experiences, and this dissertation is no exception. In one way or another, each of you has helped shape this work.

Dedication

For my family and Frito, whose love and support made this dissertation possible.

Chapter 1: Introduction

1.1 Convex Relaxations as a Unifying Framework

This dissertation develops techniques for solving challenging continuous optimization problems and studies the capabilities and limitations of relaxation-based approaches for discrete optimization, combining practical algorithmic contributions with theoretical insights. The problems we consider differ significantly in their sources of difficulty. Continuous optimization problems arising in power systems are often large-scale, numerically challenging, and exhibit complex nonlinear and nonconvex behavior. In contrast, the difficulty of the discrete problems we study stems from more subtle structural properties, most notably symmetry.

Despite these differences, a unifying theme of this work is the use of convex relaxations as a tool for both computation and analysis. In the continuous setting, we develop new accurate and scalable relaxations for the Alternating-Current Optimal Power Flow problem. In the discrete setting, we study the behavior of classical relaxations for structured integer programs, and provide a new perspective on their limitations and capabilities. More broadly, our approach is guided by the principle that problem structure can be exploited in two complementary ways: (i) to design effective relaxations, and (ii) to understand the conditions under which relaxations are informative.

At a fundamental level, our work is driven by a central question that arises in solving hard optimization problems: given a candidate solution, how good is it? The "goodness" of a solution can refer to feasibility, whether it satisfies all constraints, or to optimality, how its cost compares to the (usually unknown) true optimal value. In practice, such assessments often need to be performed in real time, which calls for fast methods that provide certifiable guarantees.

The paradigm of a relaxation offers a powerful tool for addressing this challenge. A relaxation replaces a difficult optimization problem with a simpler version, typically by loosening some of its constraints, where, ideally, solutions can be obtained with greater speed and reliability. Crucially, the value of the relaxed problem provides a bound on the cost of the optimum: because constraints have been relaxed, the bound cannot be worse than the true optimum. This bound serves as a rigorous benchmark for assessing the quality of candidate solutions. Relaxations are therefore a central tool in modern optimization.

Overall, this dissertation highlights how relaxations provide both effective computational tools for bounding difficult optimization problems and a lens for understanding the structural features that govern their complexity.

1.2 Research Overview

1.2.1 Mathematical Optimization for Power Systems

Modernizing power systems has become a global priority, as reliable and efficient electricity networks are essential to the energy transition and to climate resilience. Around the world, electric grids are being reconfigured. In the U.S., for example, the Department of Energy (DOE) projected that by 2030 the electric power grid, one of the most complex engineered systems, would undergo unprecedented changes: active household participation in generation, greater reliance on renewables, the integration of storage technologies, and the emergence of more decentralized market structures [1]. These trends have continued to accelerate in recent years; see, e.g., [2]. At the same time, surges in demand from new industrial facilities, data centers, electric vehicles, and artificial intelligence are straining existing infrastructure. These developments highlight both the urgency and the global relevance of the problem. In the U.S. alone, it has been estimated that improvements in power system optimization could reduce total electricity costs by 5–10% annually, corresponding to savings of $6–19 billion USD per year [3].

At the heart of these challenges lies the Alternating-Current Optimal Power Flow (ACOPF) problem, which governs the secure and efficient dispatch of electricity across the network. ACOPF minimizes generation costs while ensuring that power flows satisfy both physical laws and operational constraints. The scale of modern power systems, combined with the highly intricate and nonconvex nature of AC power flows, makes this problem computationally challenging.

Despite this complexity, Independent System Operators (ISOs), the entities responsible for market clearing and system operation, must compute dispatch decisions, or approximations thereof, at high frequency, often every few minutes [4]. These approximations can be particularly inaccurate under stressed network conditions, where transmission constraints become binding, leading to significant economic inefficiencies.

Given these challenges, it is imperative to develop methods that produce high-quality feasible solutions while providing reliable assessments of their optimality in a timely manner. State-of-the-art interior-point methods compute high-quality solutions but provide no guarantees on solution quality, while strong lower bounds from nonlinear convex relaxations do not scale to large instances. This gap between scalability and certification is a central challenge in power systems optimization.

In addition to determining dispatch, the solution to ACOPF underlies the computation of electricity prices, which are used to signal scarcity and allocate resources in the market. However, the nonconvexities inherent in ACOPF complicate the derivation of prices that reflect true system marginal costs, since competitive equilibrium pricing relies on convexity and strong duality. In practice, pricing therefore relies on simplified models, most notably the Direct-Current (DC) approximation of ACOPF. While computationally tractable, these models fail to capture key aspects of network physics, leading to distorted price signals and a poor representation of congestion.

This work addresses these limitations by developing scalable and numerically robust linear relaxations of ACOPF that tightly approximate the underlying nonlinear constraints, enabling both accurate solution quality guarantees and the computation of price signals that more faithfully reflect network congestion and system marginal costs. The resulting methods support warm-starting under realistic system changes and extend naturally to multi-period settings, making them suitable for real-time and dynamic grid operations. We further provide theoretical justification for the strength of convex relaxations for ACOPF and demonstrate, through computational experiments, that our approach delivers tight and scalable bounds on large-scale instances where nonlinear convex relaxations are computationally prohibitive.

1.2.2 Symmetry and Relaxations in Integer Programming

Symmetries can be broadly understood as transformations that preserve the structural properties of an object. In optimization, such symmetries – whether inherent to the problem or induced by its formulation – typically act on the feasible region and the objective function. They are ubiquitous in optimization: for instance, at least 36% of the instances in MIPLIB 2010 exhibit nontrivial symmetries [5], [6]. It is well established that symmetry can adversely affect algorithmic performance in integer programming. The traditional and widely studied viewpoint is that symmetries degrade the efficiency of branch-and-bound–based algorithms by causing redundant exploration of equivalent solutions.

There is also substantial evidence that classical convex relaxations of well-known symmetric integer programs – such as the vertex formulation of Maximum Stable Set, Minimum Knapsack, and the edge formulation of Maximum Matching – are often weak, yielding fractional solutions far from the integer optimum. This has led to significant interest in understanding the power of *hierarchies* for such problems. These hierarchies are systematic operators that progressively strengthen linear programming relaxations, with the goal of closing the gap between fractional and integer solutions.

Motivated by these observations, we introduce a new and complementary perspective. Our work shows that symmetry can also have a fundamental impact on the strength of convex relaxations of binary integer programs. While the weakness of relaxations for symmetric problems has long been observed, there has been no unified framework to explain this phenomenon.

We develop such a framework by studying the behavior of lift-and-project operators under symmetry. By leveraging concepts from computational group theory, we identify a property that depends only on the symmetries of the linear programming relaxation and fully characterizes the strength of the relaxations produced by these operators, abstracting away from specific problem classes. This perspective further reveals that, although these operators can produce significantly different relaxations in general, under symmetry they collapse in strength. In particular, they become equivalent as cutting-plane proof systems: they yield the same certificates for improving linear programming bounds by eliminating the fractional (integer-empty) optimal LP face. This collapse exposes deep structural properties that can be further exploited algorithmically.

This work contributes to the structural understanding of symmetry in integer optimization and provides new insights into the behavior of classical convex relaxations in symmetric settings. Beyond its theoretical implications, it also informs the design of stronger formulations and is relevant for practitioners and solver developers.

1.3 Preliminaries and Notation

In this section we introduce some general notation and definitions that will be used in throughout this dissertation. We follow standard conventions in mathematical optimization and linear algebra.

1.3.1 Linear Algebra

We denote the set of natural numbers by $\mathbb{N}$, the set of rational numbers by $\mathbb{Q}$, the set of real numbers by $\mathbb{R}$, and the set of complex numbers by $\mathbb{C}$. We denote by $j := \sqrt{-1}$ the imaginary unit, and for $z = a + jb \in \mathbb{C}$ we write $\Re z := a$, $\Im z := b$ and $z^* := a - jb$. For $n \in \mathbb{N}$ we denote $\{1, \dots, n\}$ by $[n]$. We denote the empty set by $\emptyset$. For a vector $x \in \mathbb{K}^n$, where $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, we denote its i-th component by x_i for $i \in [n]$. We denote its Euclidean norm by $\|x\|_2$ and omit the subscript when it is clear from context. For a matrix $A \in \mathbb{K}^{m \times n}$, we denote its (i, j)-th entry by A_{ij} for $i \in [m]$ and $j \in [n]$. We denote the transpose of a matrix A by $A^\top$. We denote the all-ones and all-zeros vectors in $\mathbb{R}^n$ by $\mathbf{1}_n$ and $\mathbf{0}_n$. We denote the $n \times n$ identity matrix by I_n. For the previously defined objects, we drop the subscript when the dimension is clear from context.

Let $A \in \mathbb{R}^{n \times n}$ be a matrix. We say that A is symmetric if $A = A^\top$. Moreover, the matrix A is called positive semi-definite if it is symmetric and for all $x \in \mathbb{R}^n$ we have that $x^\top A x \geq 0$, while it is positive definite if $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $x^\top A x > 0$. Equivalently, a matrix is positive semi-definite if it is symmetric and all of its eigenvalues are non-negative, and positive definite if all its eigenvalues are positive. The set of symmetric, positive semi-definite and positive definite matrices in $\mathbb{R}^{n \times n}$ will be denoted by $\mathbb{S}^n$, $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$, respectively. We use $A \succeq 0$ to denote that A is positive semidefinite (PSD) over $\mathbb{R}$. For $A, B \in \mathbb{R}^{n \times n}$, we define the standard (Frobenius) inner-product

$\langle \cdot, \cdot \rangle$ as $\langle A, B \rangle := \text{trace}(A^\top B)$,

Given $X \subseteq \mathbb{R}^n$, we denote its convex hull as $\text{conv}(X)$. Recall that if X is finite than $\text{conv}(X)$ is a called a polytope. Moreover, recall that every affine set X in $\mathbb{R}^n$ can be expressed as

$$X = E + x_0 = \{s + x_0 \, : \, s \in E\},$$

where E is subspace of $\mathbb{R}^n$, and x_0 is any point in X. The dimension of the set X, denoted by $\dim(X)$, is defined as the dimension of the subspace $E = X - x_0$.

1.3.2 Computational Complexity

We follow the exposition in [7]. We adopt the standard asymptotic notation to describe the limiting behavior of functions. Let $f, g : \mathbb{N} \to \mathbb{R}_+$. We say that $f(n) = O(g(n))$ if there exist $C > 0$ and n_0 such that $f(n) \leq C\, g(n)$ for all $n \geq n_0$, and $f(n) = o(g(n))$ if for every $c > 0$ there exists n_1 such that $f(n) \leq c\, g(n)$ for all $n \geq n_1$. Analogously, $f(n) = \Omega(g(n))$ and $f(n) = \omega(g(n))$ are defined by reversing the inequality. Finally, we say that $f(n) = \Theta(g(n))$ if there exist constants $C_1, C_2 > 0$ and n_2 such that $C_1 g(n) \leq f(n) \leq C_2 g(n)$ for all $n \geq n_2$.

Size of Problems Consider the *alphabet* $\Sigma := \{0, 1\}$. Its elements are called *symbols*. An ordered finite sequence of symbols from Σ is called a *string*. Σ^* stands for the collection of all strings of symbols from Σ. The *size* of a string is the number of its components. The string of size 0 is the *empty string*, which we denote by $\emptyset$. Strings can have the form of finite sequences of rational numbers, matrices, graphs, systems of linear inequalities, and so on. Given a rational number $\alpha = p/q$, where p and q are relatively prime integers, using our definition of size, we can its size as the numbers of $\text{size}(\alpha) := 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil$, i.e., the number of $\{0, 1\}$ bits required to encode its *binary representation*. Naturally, the size of a rational $m \times n$ matrix $A = (\alpha_{i,j})_{i \in [m], j \in [n]}$ can be defined as $\text{size}(A) := mn + \sum_{i,j} \text{size}(\alpha_{i,j})$. Therefore, the size of a system $Ax \geq b$ is equal to $1 + \text{size}(A) + \text{size}(b)$. The size of a (directed or undirected) graph is equal to the size of its incidence matrix.

Problems, Algorithms, and Complexity Classes A *problem* is a subset Π of $\Sigma^* \times \Sigma^*$ defined as follows: given a string $z \in \Sigma^*$, find a string $y \in \Sigma^*$ such that $(z, y) \in \Pi$, or decide that no such string y exists. Here the string z is called an *instance* or the *input* of the problem, and y is called a *solution* or the *output*. A problem Π is a *decision problem* if for each (z, y) in Π, y is the empty string $\emptyset$. In that case, the problem is usually identified with the set $\mathcal{L}$ of strings in $z \in \Sigma^*$ for which $(z, \emptyset)$ belongs to Π. For a given input $z \in \Sigma^*$, *algorithm* R for problem $\Pi \subseteq \Sigma^* \times \Sigma^*$ determines an output y such that (z, y) is in Π, or stops without delivering an output if there exists no such y. We say that R *solves* problem Π, if for any instance z of Π, when giving the string (A, z) to a "universal Turing machine", the machine stops after a finite number of steps, while delivering a string y with (z, y) in Π, or delivering no string in the case where such a string does not exist. We define the *running time* of an algorithm R as the function $f : \mathbb{Z}_+ \to \mathbb{Z}_+$ defined by $f(\sigma) := \max_{z, \text{size}(z) \leq \sigma} t(A, z)$, there $t(A, z)$ is defined as the number of moves the "head" of a universal Turing machine makes before stopping, when it is given the input (A, z). An algorithm is called *polynomial-time* if its running time $f(\sigma) \in O(g(\sigma))$, where g is some polynomial function. A problem is said to be *solvable in polynomial time* if the problem can be solved by a polynomial-time algorithm.

The class of decision problems solvable in polynomial time is denoted by $\mathbb{P}$. Formally, a decision problem $\mathcal{L} \subseteq \Sigma^*$ belongs to the class $\mathbb{NP}$ if there exists a polynomially solvable decision problem $\mathcal{L}' \subseteq \Sigma^* \times \Sigma^*$ and a polynomial ϕ such that for each z in Σ^*: z in $\mathcal{L}$ if and only if there exists y in Σ^* such that (z, y) belongs to $\mathcal{L}'$ and the $\text{size}(y) \leq \phi(\text{size}(z))$. A decision problem $\mathcal{L} \subseteq \Sigma^*$ is called *Karp-reducible* to a decision problem $\mathcal{L}' \subseteq \Sigma^*$ if there is a polynomial-time algorithm R such that for every input string $z \in \Sigma^*$, R outputs a string x, such that z in $\mathcal{L}$ if and only if x in $\mathcal{L}'$. The class $\mathbb{NP}$-complete is defined as the subclass of problems in $\mathbb{NP}$ that are the hardest under Karp-reductions. In other words, a decision problem $\mathcal{L}$ is $\mathbb{NP}$-complete if $\mathcal{L}$ belongs to $\mathbb{NP}$ and each problem in $\mathbb{NP}$ is Karp-reducible to $\mathcal{L}$. Finally, a problem (not necessarily a decision problem) is $\mathbb{NP}$-hard if every problem in $\mathbb{NP}$ is Karp-reducible to it.

1.3.3 Optimization Problem Classes

We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is a *quadratic function* if there exists a triple $(Q, q, q_0) \in \mathbb{R}^{n\times n} \times \mathbb{R}^n \times \mathbb{R}$ such that $f(x) = x^\top Qx + q^\top x + q_0$ for all $x \in \mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is a *convex quadratic function* if it is a quadratic function and the matrix Q is positive semi-definite. We say that it is an *affine function*, or slightly abusing notation *linear function*, if it is a convex quadratic function with $Q = 0$.

Let $m \in \mathbb{N}$ and $f, g_i : \mathbb{R}^n \to \mathbb{R}$ for $i \in [m]$ be quadratic functions. Consider the following optimization problem:

$$\min \quad f(x) \tag{1.1}$$

$$\text{s.t.} \quad g_i(x) \leq 0 \quad i \in [m]. \tag{1.2}$$

If the functions f and g_i are arbitrary quadratic functions, the problem is called a *quadratically constrained quadratic program* (QCQP). If they are convex quadratic functions, the problem is called a *convex quadratically constrained quadratic program* (convex QCQP). This class includes *second-order cone programs* (SOCPs), as constraints of the form $\|x\|_2 \leq t$ can be expressed using convex quadratic functions. If f is quadratic while the functions g_i are affine we say that the problem is a *quadratic program* (QP). Finally, if all the functions are affine, we say that the problem is a *linear program* (LP).

Moreover, if f is linear the functions g_i are either affine or pairs of the form $x_i(1 - x_i) \leq 0$ and $-x_i(1 - x_i) \leq 0$, which are equivalent to the constraint $x_i(1 - x_i) = 0 \iff x_i \in \{0, 1\}$, the optimization problem (1.1) can model discrete decisions. In other words, the above problem has affine and integrality constraints. These problems are broadly called *mixed integer (binary) linear program* (MILP).

The class of QCQP is very general and expressive – it can model a myriad of problems and applications [8]. It is known that if the functions f, g_i are polynomials, then they can be represented as quadratic functions by introducing additional variables and constraints. QCQPs belong to the broader class of nonconvex optimization problems, which includes, for instance, problems with nondifferentiable functions.

Finally, we introduce *semidefinite programs* (SDPs), which are convex optimization problems of the form

$$\min \quad \langle C, X \rangle \tag{1.3}$$

$$\text{s.t.} \quad \langle A_i, X \rangle = b_i \quad i \in [m], \tag{1.4}$$

$$X \succeq 0. \tag{1.5}$$

where $C, A_i \in \mathbb{R}^{n \times n}$, and $b_i \in \mathbb{R}$ for $i \in [m]$. SDPs are highly expressive within convex optimization, as they can model linear programs, convex quadratic programs, and convex quadratically constrained quadratic programs, including second-order cone programs. Semidefinite programs play a central role in convex relaxation methods, as they often provide tight approximations for many nonconvex problems.

In general, the decision version of QCQP is $\mathbb{NP}$-hard, even in very restricted forms. For instance, QP is $\mathbb{NP}$-complete even when the objective has exactly one negative eigenvalue [9]. The decision version of MILP is $\mathbb{NP}$-complete [7].

In contrast, semidefinite programs admit polynomial-time algorithms that compute ε-optimal solutions, with complexity polynomial in the input size and $\log(1/\varepsilon)$; see, e.g., [10].

For background on convex optimization, we refer the reader to classical texts [10], [11]. For integer programming, see [7], [12], and for general QCQPs, see [8].

Chapter 2: Mathematical Optimization for Power Systems

2.1 Introduction

The Alternating-Current Optimal Power Flow problem is a well-known challenging computational task.[1] It is nonlinear, nonconvex and with feasible region that may be disconnected; see [17], [18] and Proposition 2.12. In [19], [20] it is shown that the feasibility version of the problem is strongly $\mathbb{NP}$-hard, and [21] proved that this decision problem is weakly $\mathbb{NP}$-hard on star-networks (see Section 2.11).

In the current state-of-the-art special-purpose interior point methods are empirically successful at computing very good solutions but cannot provide any bounds on solution quality. Strong lower bounds are available through second-order cone (SOC) relaxations [22], [23]; however all solvers struggle when handling such relaxations for large or even medium cases (see [24]; we will provide additional evidence for this point). Tighter semidefinite programming (SDP) relaxations can, in principle, yield improved bounds; however, they may also produce bounds that are invalid in practice due to numerical issues, as reported in [25], [26]. Other techniques, such as spatial-branch-and-bound methods applied to McCormick (linear) relaxations of quadratically-constrained formulations for ACOPF, tend to yield poor performance unless augmented by said SOC inequalities – interior point methods are still needed for upper bounds.

These challenges carry over directly to price formation in electricity markets. Rapid computation of prices is critical for real-time market operations. A common simplification is to use DC power flow models, which significantly reduce computational burden but may fail to accurately capture

[1]This chapter is based on joint work with Daniel Bienstock, Matías Romero, and Felipe Verástegui [13], [14], [15], [16], extended with additional exposition and results. I gratefully acknowledge support from the Advanced Research Projects Agency-Energy (ARPA-E) through a Grid Optimization Competition Award, as well as a travel award from the Mixed Integer Programming (MIP) Workshop, organized by the Mixed Integer Programming Society (MIPS), and a fellowship from the W. Edwards Deming Center at Columbia Business School.

congestion effects, leading to biased prices and nontrivial post-market adjustments by system operators. While convex nonlinear relaxations, such as those in [22], [24], can yield less biased price signals on small networks [27], their practical use in market applications is limited by the same scalability challenges described above.

In this chapter we present fast linear cutting-plane methods used to obtain tight relaxations for ACOPF, and its multi-period extension with ramping constraints, by appropriately handling convex nonlinear relaxations. Our approach scales well to the largest ACOPF instances currently available and multi-period extensions. Our research thrust is motivated by the following key observations:

(1) Linearly constrained convex quadratic optimization technology is, at this point, very mature – many solvers are able to handle massively large instances quickly and robustly; these attributes extend to the case where formulations are dynamically constructed and updated, as would be the case with a cutting-plane algorithm.

(2) Whereas the strength of the SOCP and SDP relaxations has been recognized for some time, an adequate theoretical understanding for this behavior was not available.

(3) In power engineering practice, it is often the case that a power flow problem (either in the AC or DC version) is solved on data that reflects a recent, and likely limited, update on a case that was previously handled. In power engineering language, a 'prior solution' was computed, and the problem is not solved 'from scratch.' In the context of our type of algorithm, this feature opens the door for the use of *warm-started formulations*, i.e., the application of a cutting plane procedure that leverages previously computed cuts to obtain sharp bounds more rapidly than from scratch.

(4) Finally, linearly constrained convex quadratic optimization enjoys favorable duality properties, which enable effective and scalable pricing schemes, i.e., extensions of the locational marginal pricing (LMP) framework currently used for real-time and day-ahead energy markets [28], [29], [30].

Contributions

- We describe very tight and numerically stable linearly constrained relaxations for ACOPF. The relaxations can be constructed and solved robustly and quickly via a cutting-plane algorithm that relies on proper cut management. On medium to very large instances our algorithm is competitive with or better than what was previously possible using nonlinear relaxations, both in terms of bound quality and solution speed. The robustness of our method is especially prominent in the multi-period setting, where our method is able to provide tight lower bounds with high accuracy, while the nonlinear relaxations are simply out of reach for nonlinear solvers (Sections 2.5.5 and 2.6).

- We provide a theoretical justification for the tightness of the SOCP relaxation for ACOPF as well as for the use of our linear relaxations. As we argue, the SOC constraints (or tight relaxations thereof, or equivalent formulations) are *necessary* in order to achieve a tight relaxation. More precisely, we point out that the (linear) active-power loss inequalities (introduced in [31]) are outer-cone envelope approximations for the SOC constraints. Moreover, the linear loss inequalities provide a fairly tight relaxation because they imply that every unit of demand and loss is matched by a corresponding unit of generation (a fact not otherwise guaranteed) via a flow-decomposition argument [32] (Section 2.5.4).

- We demonstrate, through extensive numerical testing, that the warm-start feature for our cutting-plane algorithm is indeed effective. It is worth noting that this capability stands in contrast to what is possible using nonlinear (convex) solvers, in particular interior point methods. We also show numerical convergence of the dual variables of active-power balance constraints; suggesting their stability for use in computing locational marginal prices (Section 2.8).

- The warm-start capability yields practicable and tight relaxations for *multi-period* formulations of ACOPF. The tightness of our formulations is certified through a heuristic for such problems (Section 2.8).

- We bring forth to the ACOPF literature a discussion on approximate solutions to convex relaxations and their accuracy as lower bounds. We argue that linearly constrained relaxations with convex quadratic objective possess robust theoretical bounding guarantees, in sharp contrast to what nonlinear relaxations such as SOCPs can offer (Section 2.7).

- We provide a rigorous understanding of SDP relaxations for ACOPF, offering new insights into two well-known SDP formulations and establishing theoretical justification for their tightness (Sections 2.9.1 and 2.9.3).

- We propose a numerically stable and scalable approach to SDP relaxations for ACOPF that exploits low-dimensional structure. In particular, we leverage SDP tightness via linear outer approximations, without solving any SDP, and present preliminary computational results (Sections 2.9.4–2.9.6).

- Building on our cutting-plane framework, we develop a scalable pricing algorithm for ACOPF, bridging a gap in the electricity pricing literature. Our approach achieves computational performance comparable to DCOPF, while producing active and reactive power prices that closely match those obtained from a tight SOCP relaxation (Section 2.10).

- We correct a gap in the weak $\mathbb{NP}$-hardness proof of [21] (Section 2.11).

2.2 Background on Power Systems

A *power system* is an undirected network $\mathcal{N} := (\mathcal{B}, \mathcal{E})$ of *buses* (nodes) $\mathcal{B}$ connected by *branches* (edges) $\mathcal{E}$, endowed with numerical parameters describing physical attributes. Each bus may house one or more *generators*, or have *loads* (demands) attached to it. In this interconnected system, *electric power* flows from generators to loads through the branches.

For modeling purposes, each branch $\{k, m\} \in \mathcal{E}$ is assigned a reference orientation, either (k, m) or (m, k), as specified by the data. We denote by $\mathcal{A}$ the resulting set of directed arcs, which contains exactly one arc per branch. This convention is standard, as network data formats inherently assume an orientation when defining branch parameters.

Electric power, measured in watts (W), corresponds to the rate at which energy is delivered. This quantity is computed as the product of *voltage*, which represents energy per unit of charge and is measured in volts (V), and *current*, which represents the flow of electric charge per unit of time and is measured in amperes (A). Depending on how voltage and current evolve over time, power can be delivered in the form of *direct current* (DC), where these quantities are constant, or *alternating current* (AC), where they vary sinusoidally. In practice, AC is used for most generation, transmission and consumption.

We begin with a brief introduction to the physics governing AC power flows, and then discuss the simpler DC case. This section closely follows the presentation of power engineering fundamentals in [32]. We refer the reader to [33] for further background on power system modeling.

AC Power Flows and Complex Numbers

In what follows, we show that voltage and current can be representated using complex numbers. Given a branch connecting two buses k and m, the voltage at bus k and current flowing from bus k to m can be represented, at time t, as

$$v_k(t) \coloneqq V_k^{\max} \cos(\omega t + \theta_k^V), \quad i_{km}(t) \coloneqq I_{km}^{\max} \cos(\omega t + \theta_{km}^I). \tag{2.1}$$

The constant ω is the angular frequency of the system (50 Hz or 60 Hz across different countries), $V_k^{\max}$ and $I_{km}^{\max}$ are constants representing the *amplitude* of voltage and current, respectively, and θ_k^V and θ_{km}^I are their respective *phase angles*. The instantaneous power flowing from bus k to m is given by the product of voltage and current, i.e., $p_{km}(t) \coloneqq v_k(t) i_{km}(t)$, which can be expressed as

$$p_{km}(t) = \frac{1}{2} V_k^{\max} I_{km}^{\max} \left[\cos(\theta_k^V - \theta_{km}^I) + \cos(2\omega t + \theta_k^V + \theta_{km}^I)\right]$$

using the identity $\cos(a)\cos(b) = \frac{1}{2}\left(\cos(a-b) + \cos(a+b)\right)$. Due to rapid system frequency, we consider the average value of $p_{km}(t)$ over a period of time $T = 2\pi/\omega$, i.e., the power flow in steady-state, which is given by

$$P_{km} := \frac{1}{T}\int_0^T p_{km}(t)dt = \frac{1}{2}V_k^{\max} I_{km}^{\max}\cos(\theta_k^V - \theta_{km}^I). \tag{2.2}$$

This quantity is called *active power* injected at bus k into line $\{k, m\}$. Next, we restate this quantity in terms of complex numbers. Indeed, we can rewrite (2.1) as

$$v_k(t) = V_k^{\max}\Re e^{j(\omega t+\theta_k^V)}, \quad i_{km}(t) = I_{km}^{\max}\Re e^{j(\omega t+\theta_{km}^I)}$$

where j is the imaginary unit. Then, by defining $V_k := \frac{V_k^{\max}}{\sqrt{2}}e^{j\theta_k^V}$ and $I_{km} := \frac{I_{km}^{\max}}{\sqrt{2}}e^{j\theta_{km}^I}$, we have

$$P_{km} = |V_k||I_{km}|\cos(\theta_k^V - \theta_{km}^I) = \Re(V_k I_{km}^*). \tag{2.3}$$

The term $S_{km} := V_k I_{km}^*$ is called *complex power* injected at bus k into line $\{k, m\}$. Its imaginary part $Q_{km} := \Im(V_k I_{km}^*)$ is called *reactive power*, and is of significant importance for the efficient and stable operation of the power grid.

The Π-Model of a Transmission Line

We describe a model that approximates the physical behavior of power flowing through a branch $\{k, m\}$; see Figure 2.1. Without loss of generality, we adopt the reference orientation (k, m), where the orientation determines the "from" bus, in this case bus k, and is used to define branch parameters. We use km as a shorthand for the arc (k, m) when indexing orientation-dependent quantities, while parameters independent of the branch orientation are indexed by $\{k, m\}$. As we will see, the AC power flow equations for a transmission line can be fully characterized by a 2×2 complex matrix, introduced below.

A transmission line is a *conductive* element characterized by a set of electrical parameters. The *series impedance* $z_{\{k,m\}} := r_{\{k,m\}} + jx_{\{k,m\}} \in \mathbb{C}$, where $r_{\{km\}}$ is called *resistance* and $x_{\{k,m\}}$ *reactance*, captures energy losses and voltage drops due to current flow. The *shunt admittance*

$y^{sh}_{\{k,m\}} \coloneqq g_{\{k,m\}} + jb_{\{k,m\}} \in \mathbb{C}$, where $g_{\{k,m\}}$ is the *shunt conductance* and $b_{\{k,m\}}$ *shunt susceptance*, models the conductive and capacitive effects between the line and the ground, capturing leakage currents and charging behavior along the line. In general, $x_{\{k,m\}}, r_{\{k,m\}} \geq 0$, and $x_{r\{k,m\}} \gg r_{\{k,m\}}$, i.e., $x_{r\{k,m\}}$ is an order of magnitude larger than $r_{\{k,m\}}$. Also, $b^{sh}_{\{k,m\}} \gg g^{sh}_{\{k,m\}}$ and $g^{sh}_{\{k,m\}}$ is often set to zero. For short transmission lines $b^{sh}_{\{k,m\}}$ is typically neglected, i.e., $b^{sh}_{\{k,m\}} = 0$. The reciprocal of $z_{\{k,m\}}$ is called the *series admittance* $y_{\{k,m\}} \coloneqq 1/z_{k\{k,m\}} = g_{\{k,m\}} + jb_{\{k,m\}}$, where $g_{\{k,m\}}$ is the *conductance* and $b_{\{k,m\}}$ the *susceptance* of the line. We remark that $g_{\{k,m\}} = r_{\{k,m\}}/(r^2_{\{k,m\}} + x^2_{\{k,m\}})$ and $b_{\{k,m\}} = -x_{\{k,m\}}/(r^2_{\{k,m\}} + x^2_{\{k,m\}})$.

Moreover, the branch may also include a *transformer* at bus k. A transformer is a device that can step up or step down the voltage level and introduce a phase shift in the power flowing through the line $\{k, m\}$, thereby facilitating efficient transmission over long distances. This device is modeled by a complex ratio $N_{km} \coloneqq \tau_{km} e^{j\sigma_{km}}$, where $\tau_{km} > 0$ is the *magnitude* of the transformer ratio and σ_{km} is its *phase shift*.

To see this, suppose that complex power $S_{km} = V_k I^*_{km}$ is injected at bus k, where current I_{km} flows at voltage V_k before entering the transformer; see Figure 2.1. The transformer scales the voltage to V_k/N_{km}, i.e., with ratio $N_{km} : 1$. Since complex power is preserved, the current is correspondingly scaled to $N^* I_{km}$, so that $(V_k/N_{km})(N^*_{km} I_{km})^* = V_k I^*_{km} = S_{km}$. Consistent with standard data formats, transformers are assumed to be located at the "from" bus, and thus their definition depends on the chosen branch. This assumption is without loss of generality, since a transformer at bus m can be equivalently represented by adjusting the ratio at bus k.[2]

All of this information – the series admittance, shunt elements, and transformer associated with the line $\{k, m\}$ – can be encoded in the following 2×2 complex *admittance matrix*:

$$Y_{km} \coloneqq \begin{pmatrix} \frac{y_{\{k,m\}}}{\tau^2_{km}} + \frac{y^{sh}_{\{k,m\}}}{2\tau^2_{km}} & -y_{\{k,m\}} \frac{1}{\tau_{km} e^{-j\sigma_{km}}} \\ -y_{\{k,m\}} \frac{1}{\tau_{km} e^{j\sigma_{km}}} & y_{\{k,m\}} + \frac{y^{sh}_{\{k,m\}}}{2} \end{pmatrix} = \begin{pmatrix} G_{kk} + jB_{kk} & G_{km} + jB_{km} \\ G_{mk} + jB_{mk} & G_{mm} + jB_{mm} \end{pmatrix}.$$

[2] Indeed, if transformers are present at both buses k and m, with ratios N_{km} and N_{mk}, respectively, they can be equivalently represented by a single transformer at bus k with ratio N_{mk}/N_{km}.

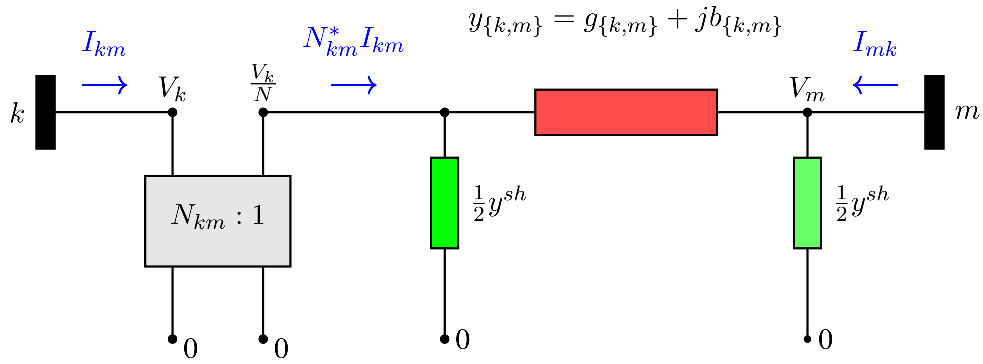

Figure 2.1: Π-model representation of a transmission line $\{k, m\}$ with a series admittance (red), a transformer at bus k (grey), and shunt elements (green). Power injected at bus k, S_{km}, flows with current I_{km} at voltage V_k before entering the transformer. The transformer scales the voltage to V_k/N, hence the current is correspondingly scaled to $N^* I_{km}$. The flow then passes through the series element $y_{\{k,m\}}$ and line charging devices $\frac{1}{2}y^{sh}$. The symbols "0" denote the ground reference.

Now we are ready to write the AC power flow equations for line $\{k, m\}$ in terms of these coefficients. We denote by $V_k := |V_k|e^{j\theta_k}$ and $V_m := |V_m|e^{j\theta_m}$ the complex voltages (also known as *voltage phasors*) at buses k and m, respectively, and by I_{km} and I_{mk} the complex currents flowing from bus k to m and from bus m to k, respectively. We make use of *Ohm's Law*, which states that the current flowing between the buses k and m along branch $\{k, m\}$ is linear in the bus voltages, and is given by:

$$\begin{pmatrix} I_{km} \\ I_{mk} \end{pmatrix} = Y_{km} \begin{pmatrix} V_k \\ V_m \end{pmatrix} = \begin{pmatrix} (G_{kk} + jB_{kk})V_k + (G_{km} + jB_{km})V_m \\ (G_{mk} + jB_{mk})V_k + (G_{mm} + jB_{mm})V_m \end{pmatrix}.$$

Since complex power injected at bus k into line $\{k, m\}$ is given by $S_{km} = V_k I^*_{km}$, we have

$$S_{km} = V_k V_k^* (G_{kk} - jB_{kk}) + V_k V_m^* (G_{km} - jB_{km}) \tag{2.4}$$

$$= G_{kk}|V_k|^2 - jB_{kk}|V_k|^2 + G_{km}|V_k||V_m|e^{j(\theta_k - \theta_m)} - jB_{km}|V_k||V_m|e^{j(\theta_k - \theta_m)} \tag{2.5}$$

Let $\theta_{km} := \theta_k - \theta_m$. Then, using Euler's formula $e^{j\theta_{km}} = \cos(\theta_{km}) + j\sin(\theta_{km})$, we can write

$$\begin{aligned}
S_{km} &= G_{kk}|V_k|^2 - jB_{kk}|V_k|^2 + (G_{km}|V_k||V_m| - jB_{km}|V_k||V_m|)(\cos(\theta_{km}) + j\sin(\theta_{km})) \\
&= G_{kk}|V_k|^2 + G_{km}|V_k||V_m|\cos(\theta_{km}) + B_{km}|V_k||V_m|\sin(\theta_{km}) \\
&+ j(-B_{kk}|V_k|^2 - B_{km}|V_k||V_m|\cos(\theta_{km}) + G_{km}|V_k||V_m|\sin(\theta_{km})) \\
&= P_{km} + jQ_{km}.
\end{aligned}$$

Analogously, we can compute $S_{mk} = V_m I_{mk}^*$ and its real and imaginary parts P_{mk} and Q_{mk}, respectively, yielding the full set of AC power flow equations in *polar coordinates*:

$$P_{km} = G_{kk}|V_k|^2 + G_{km}|V_k||V_m|\cos(\theta_{km}) + B_{km}|V_k||V_m|\sin(\theta_{km}) \tag{2.6a}$$

$$P_{mk} = G_{mm}|V_m|^2 + G_{mk}|V_k||V_m|\cos(\theta_{km}) - B_{mk}|V_k||V_m|\sin(\theta_{km}) \tag{2.6b}$$

$$Q_{km} = -B_{kk}|V_k|^2 - B_{km}|V_k||V_m|\cos(\theta_{km}) + G_{km}|V_k||V_m|\sin(\theta_{km}) \tag{2.6c}$$

$$Q_{mk} = -B_{mm}|V_m|^2 - B_{mk}|V_k||V_m|\cos(\theta_{km}) - G_{mk}|V_k||V_m|\sin(\theta_{km}). \tag{2.6d}$$

DC Power Flows

This power flow model is a linear approximation of AC power flows commonly used in practice [3], [34]. It is based on a number of simplifications that are approximately valid under normal system operations. The observations this model is based on are:

1. Under normal conditions voltage magnitudes are close to their nominal values, i.e., $|V_k| \approx 1$ for all $k \in \mathcal{B}$;

2. Under normal conditions voltage angle differences between connected buses are small, i.e., $\theta_{km} \approx 0$, hence $\cos(\theta_{km}) \approx 1$ and $\sin(\theta_{km}) \approx \theta_{km}$ for all $(k, m) \in \mathcal{A}$;

3. Since $r_{\{k,m\}} \ll x_{\{k,m\}}$ and $x_{\{k,m\}}$ is small, line resistances are assumed to be 0, hence $g_{\{k,m\}} = 0$ and $b_{\{k,m\}} = -1/x_{\{k,m\}}$ for all $\{k, m\} \in \mathcal{E}$;

4. Shunt elements are neglected, i.e., $g^{sh}_{\{k,m\}} = b^{sh}_{\{k,m\}} = 0$ for all $\{k, m\} \in \mathcal{E}$.

Under this model active power losses are zero and reactive power is completely ignored [33]. Indeed, by applying these assumptions to the AC power flow equations (2.6), we have[3]:

$$P_{km} = -P_{mk} = \frac{1}{\tau_{\{k,m\}} x_{\{k,m\}}} (\theta_k - \theta_m - \sigma_{km}) \tag{2.7a}$$

$$Q_{km} = Q_{mk} = 0 \tag{2.7b}$$

2.3 Formal Problem Statement

In standard (single-period) ACOPF we are given, as input, a power system given by an undirected network $\mathcal{N} := (\mathcal{B}, \mathcal{E})$ endowed with admittance matrices Y_{km} for every arc $(k, m) \in \mathcal{A}$, a set of generators, and a set of complex-valued loads. The objective of the problem is to set complex voltages at the buses, and generator outputs, so as to satisfy loads at minimum cost – cost is incurred by the generators – while transmitting power following laws of physics and equipment constraints. A mathematical description is given above in (2.8) using the so-called *polar representation*, see Section 2.2.

We denote by $\mathcal{G}$ the set of generators of the grid, each of which is located at some bus; for each bus $k \in \mathcal{B}$, we denote by $\mathcal{G}_k \subseteq \mathcal{G}$ the generators at bus k. Each bus k has a fixed load $P^d_k + jQ^d_k$, where $P^d_k \geq 0$ is termed active power load, and $-\infty < Q^d_k < +\infty$ is reactive power load; lower $V^{\min}_k \geq 0$ and upper $V^{\max}_k \geq 0$ voltage limits. For each branch $\{k, m\} \in \mathcal{E}$, we are given its thermal limit $0 \leq U_{\{k,m\}} \leq +\infty$ and maximum angle-difference $|\Delta_{\{k,m\}}| \leq \pi$, and we introduce active and reactive power flow variables P_{km}, P_{mk}, Q_{km}, and Q_{mk}.

Thus, the goal is to determine the voltage magnitudes $|V_k|$ and phase angles θ_k at each bus $k \in \mathcal{B}$, together with active and reactive generation outputs P^g_ℓ and Q^g_ℓ of generators $\ell \in \mathcal{G}$, so that the active and reactive power demands are satisfied at minimum cost. These quantities constitute the decision variables of the ACOPF formulation presented below; the remaining quantities are fixed network and operational parameters.

[3]This assumes that $\theta_k - \theta_m - \sigma_{km}$ is small as well.

Figure 2.2: ACOPF Polar Formulation

$$[\text{ACOPF}]: \qquad \min \quad \sum_{k \in \mathcal{G}} F_k(P_k^g) \tag{2.8a}$$

subject to:

$\forall$ bus $k \in \mathcal{B}$:

$$\sum_{\{k,m\} \in \delta(k)} P_{km} = \sum_{\ell \in \mathcal{G}_k} P_\ell^g - P_k^d \tag{2.8b}$$

$$\sum_{\{k,m\} \in \delta(k)} Q_{km} = \sum_{\ell \in \mathcal{G}_k} Q_\ell^g - Q_k^d \tag{2.8c}$$

$\forall$ branch $(k,m) \in \mathcal{A}$:

$$P_{km} = G_{kk}|V_k|^2 + |V_k||V_m|(G_{km}\cos(\theta_k - \theta_m) + B_{km}\sin(\theta_k - \theta_m)) \tag{2.8d}$$

$$P_{mk} = G_{mm}|V_m|^2 + |V_k||V_m|(G_{mk}\cos(\theta_k - \theta_m) - B_{mk}\sin(\theta_k - \theta_m)) \tag{2.8e}$$

$$Q_{km} = -B_{kk}|V_k|^2 - |V_k||V_m|(B_{km}\cos(\theta_k - \theta_m) - G_{km}\sin(\theta_k - \theta_m)) \tag{2.8f}$$

$$Q_{mk} = -B_{mm}|V_m|^2 - |V_k||V_m|(B_{mk}\cos(\theta_k - \theta_m) + G_{mk}\sin(\theta_k - \theta_m)) \tag{2.8g}$$

$\forall$ generator $k \in \mathcal{G}$:

$$P_k^{\min} \le P_k^g \le P_k^{\max} \tag{2.8h}$$

$$Q_k^{\min} \le Q_k^g \le Q_k^{\max} \tag{2.8i}$$

$\forall$ bus $k \in \mathcal{B}$:

$$(V_k^{\min})^2 \le |V_k|^2 \le (V_k^{\max})^2 \tag{2.8j}$$

$\forall$ branch $\{k,m\} \in \mathcal{E}$:

$$\max\left\{P_{km}^2 + Q_{km}^2, P_{mk}^2 + Q_{mk}^2\right\} \le U_{\{k,m\}} \tag{2.8k}$$

$$|\theta_k - \theta_m| \le \Delta_{\{k,m\}} \tag{2.8l}$$

In the above formulation (2.8), the physical parameters of each line $(k,m) \in \mathcal{A}$ are described by its complex admittance matrix Y_{km} (see Section 2.2). These parameters are used to model, in (2.8d)-(2.8g), active and reactive power flows. Moreover, inequalities (2.8k)-(2.8l) correspond to flow capacity constraints, and inequalities (2.8h)-(2.8j) impose operational limits on power generation and voltages. Constraints (2.8b)-(2.8c) impose active and reactive power balance; the left-hand side represents power injection at bus $k \in \mathcal{B}$, while the right-hand side represents net power generation (generation minus demand) at bus k. Finally, for each generator $k \in \mathcal{G}$, it is customary to assume the functions $F_k : \mathbb{R} \to \mathbb{R}$ in the objective (2.8a) are convex piecewise-linear or convex quadratic.

We remark that, often, constraint (2.8l) is not present, and, when explicitly given, concerns angle limits $\Delta_{\{k,m\}}$ that are *small* (smaller than $\pi/2$). Under such assumptions there are equivalent ways to restate (2.8l) (the same applies to convex relaxations); see Appendix A.1. Accordingly, we omit this constraint in what follows.[4]

Please refer to Section 2.9.1 and the surveys [35], [36] for alternative, but equivalent, ACOPF formulations.

2.3.1 Multi-Period Formulation

Our techniques apply to a multi-period setting for ACOPF; this extension is important because it corresponds to the way that energy markets are operated. We will rely on a simple multi-period setting for ACOPF, where in each period t there is a standard ACOPF formulation. The variables for this period are linked to those other periods $t' \neq t$ via generator ramping constraints[5] (defined below). Other variants of this problem include "uptime/downtime" costs for generators, energy storage reserve requirements, and network topology control, among others (see e.g. [37], [38], [39]).

The multi-period problem can be described as follows: given a time-horizon of $T \in \mathbb{N}$ periods, a time-invariant undirected network endowed with numerical parameters describing physical attributes, and complex network loads $P^d_{i,t} + jQ^d_{i,t}$ for $i \in \mathcal{B}$ and $t \in \{0, \ldots, T-1\}$, the goal is to set complex voltages $V_{i,t}$ at the buses, and generator outputs $P^g_{k,t} + jQ^g_{k,t}$ for $k \in \mathcal{G}$ and $t \in \{0, \ldots, T-1\}$, so as to satisfy loads at minimum cost while transmitting power following laws of physics and equipment costs, and satisfying generator ramping constraints. For every generator $k \in \mathcal{G}$, and any time-period

[4] In the above formulation, for simplicity, we omit *bus shunt* elements (see [33]), though they are included in our computational experiments. Under all well-known convex relaxations of ACOPF, such elements admit linear representations in terms of squared voltage magnitudes, which are treated as decision variables.

[5] As a simple case, $t' = t - 1$.

$t \in \{0, \ldots, T-1\}$, the associated ramping up and down constraints are given by[6]

$$(1 + r^u_{k,t}) P^g_{k,t} - P^g_{k+1,t} \geq 0 \tag{2.9a}$$

$$(1 - r^d_{k,t}) P^g_{k,t} - P^g_{k+1,t} \leq 0. \tag{2.9b}$$

where $r^u_{k,t}, r^d_{k,t} \in (0,1)$ are the ramping up and down rates, respectively. In summary, multi-period ACOPF for a time-invariant network $(\mathcal{B}, \mathcal{E})$ with rational data $(P^d_{i,t}, Q^d_{i,t})_{i \in \mathcal{B}, t \in \{0,\ldots,T-1\}}$ and $(Y_{km}, U_{\{k,m\}})_{\{k,m\} \in \mathcal{E}}$ can be defined by (2.8) and (2.9) on variables

$$(V_{i,t}, \theta_{i,t})_{i \in \mathcal{B}},\ (P_{km,t}, P_{mk,t}, Q_{km,t}, Q_{mk,t})_{\{k,m\} \in \mathcal{E}},\ (P^g_{k,t}, Q^d_{k,t})_{k \in \mathcal{G}}$$

for each $t \in \{0, \ldots, T-1\}$.

2.4 Prior Work on ACOPF and Relaxations

First, we briefly review the extensive literature on convex relaxations and approximations for ACOPF, covering both single- and multi-period formulations. We then discuss primal heuristics for computing AC-feasible solutions.

Linear Relaxations and Approximations

The simplest relaxations use, as a starting point, a rectangular formulation of the ACOPF problem (rather than the polar setup described above, see Section 2.9.1) yielding a QCQP and rely on the well known McCormick [40] reformulation to linearize bilinear expressions. This straightforward relaxation has long been known to provide very weak bounds.

We next turn to stronger linear relaxations for ACOPF. [31], [41] (also see [35]) introduce the so-called active-power loss linear inequalities which impose the fact that on any branch the active power loss is nonnegative. The resulting relaxation, which we term the linear-loss-relaxation, is

[6] Since P^g_k can potentially be negative for some generator $g \in \mathcal{G}$, then the latter constraints need to be formulated in terms of the absolute value of P^g_k.

shown to yield good lower bounds. In a similar vein, [42] propose the network flow and the copper plate relaxations. The network flow relaxation amounts to the linear loss-relaxation with additional sparse linear inequalities that lower bound net reactive power losses in appropriate cases[7]. Moreover, the “copper plate" relaxation is obtained from the network flow relaxation by neglecting the power flow equations entirely via aggregation of all active and reactive power injections in the network. Along these lines [43] provides a relaxation which enforces a (valid) linear relationship between active and reactive power losses by relaxing linear combinations of (2.8d)-(2.8g).

The technique in [31], inspired by Glover [44], ϵ-approximates the products of continuous variables (arising from the rectangular formulation of ACOPF [35]), to arbitrary precision, using binary expansions and McCormick inequalities. This process yields a mixed integer linear ϵ-approximation for ACOPF. Another linear ϵ-approximation, which is based on the Jabr relaxation, is used in [45]. Their main contribution is using a polyhedral approximation of SOCs developed in [46] which requires $O(k_\ell \log(1/\epsilon))$ linear constraints and additional variables to ϵ-approximate a conic constraint of row size $k_\ell \in \mathbb{N}$. One of the algorithms in [47] is a successive linear programming (SLP) method – originally introduced by Kelley in [48] to address convex problems – which is applied to the Jabr relaxation. This process yields a linear relaxation for ACOPF, which was [47] on instances with up to 3375 buses.

The well-known, and widely used in practice, Direct Current Optimal Power Flow (DCOPF) model is a linear approximation for ACOPF [3], [34]. It is based on a number of simplifications that are approximately valid under normal system operations; under this formulation active power losses are zero and reactive power is completely ignored [33]. A main drawback of this formulation is that AC-feasible solutions might not even be feasible for DCOPF [49] – losses are the main cause of infeasibility – in contrast to the previously mentioned linear ϵ-approximations. In any case, this linear model is the canonical approximation of ACOPF for many extensions and applications such as welfare maximization and pricing in energy markets, day-ahead security-constrained unit commitment (SCUC), real-time security-constrained dispatch (SCD) and transmission switching

[7]Note that it is physically possible to have reactive power gains, i.e., negative reactive losses, see [33]).

(TS) among many others.

SOCP Relaxations

The SOCP relaxation introduced in [22], which is widely known as the *Jabr relaxation* (see Section 2.5.1), has had significant impact due to its effectiveness as a lower bounding technique. Briefly, it linearizes the power flow definitions (2.8d)-(2.8g) using $|\mathcal{B}| + 2|\mathcal{E}|$ additional variables and adds $|\mathcal{E}|$ valid SOC inequalities (see Section 2.5.1 for a complete derivation). Moreover, in [22] it is shown that for radial transmission networks, i.e., networks that have a tree structure, the load flow problem can be solved via an SOCP whose conic constraints correspond to the latter SOC inequalities. While on the one hand the SOCP relaxation is strong, it also yields formulations that, in the case of large ACOPF instances, are very challenging even for the best solvers.

A wide variety of techniques have been proposed to strengthen the Jabr relaxation. In [23] *arctangent constraints* are associated with cycles, with the goal of capturing the relationship between the additional variables in the Jabr relaxation and phase angles – these are formulated as bilinear constraints, and then linearized via McCormick inequalities. The two other strengthened SOCP formulations proposed in [23] add polyhedral envelopes for *arctangent functions*, and dynamically generate semi-definite cuts for cycles in the network. [25] proposes a *minor-based* formulation for ACOPF (which is a reformulation of the rank-one constraints in the QCQP formulation for ACOPF [35]), which is relaxed to generate cutting-planes improving on the tightness of the Jabr relaxation.

A stronger SOCP relaxation than the Jabr relaxation arises from a rotated-cone inequality derived from a convex relaxation of the relationship between complex power, voltage, and current. This inequality was introduced in [50] and, to the best of our knowledge, this was its first use in an optimization context. We refer to this inequality as the *i2 inequality*, and to the SOCP relaxation obtained by incorporating it as the *i2 relaxation* (see Section 2.5.2). The Quadratic Convex (QC) relaxation of [24] incorporates the i2 inequality and strengthens it through polyhedral envelopes for the sine, cosine, and bilinear terms appearing in the power flow definitions (2.8d)-(2.8g). More recently, [51] proposed a novel machine learning-based approach for obtaining fast dual bounds for the Jabr relaxation on instances with up to 2869 buses.

SDP Relaxations

There is an extensive literature on semidefinite programming relaxations for ACOPF. In [52], a real SDP relaxation was introduced, based on the *Shor relaxation* [53], and later further developed in [54]. This formulation is at least as tight as the Jabr relaxation at the expense of higher computational cost [25]. In [54] it is proven that the SDP relaxation achieves zero duality gap under appropriate network assumptions. Moreover, a complex SDP relaxation was proposed in [55] and subsequently used in [56] to establish the equivalence between the branch-flow and bus-injection models of ACOPF (see [35]). We refer to Section 2.9.2 for a presentation of SDP formulations.

Experiments with SDP relaxations have been constrained by current SDP technology capabilities. However, promising efforts have been made to improve scalability through clique decomposition techniques that leverage the sparsity of the network [57], [58], [59], [60]. Finally, moment relaxations have been studied in [61], [62], [63]. In all of these works, a small minimum resistance is enforced on all transmission lines, which appears to be necessary for achieving numerical convergence of the corresponding SDP implementations.

We also note that in [25], cuts are introduced to strengthen an SOCP relaxation of ACOPF, although their computation requires solving an SDP. Other types of cuts have also been proposed, such as the polynomial determinant cuts in [64], which, however, are nonlinear polynomial inequalities.

We refer the reader to the comprehensive survey in [35] for further background on convex relaxations for single-period ACOPF; see also [36], [65].

Multi-Period Relaxations and Approximations

Single-period nonlinear convex relaxations can naturally be extended to multi-period ACOPF. To the best of our knowledge, all of the prior computational experiments address small to medium-sized cases – which is not surprising given that the single-period (nonlinear) convex relaxations are already challenging enough (see e.g. Section 2.8.1). It is worth mentioning that most research in the multi-period setting has focused on obtaining AC primal bounds for models incorporating a variety of operational constraints and extensions, rather than on computing strong lower bounds. Multi-period Jabr SOCPs are solved in: [66] to find approximate solutions to a robust multi-period ACOPF; [67] and [68] as subproblems of Benders decomposition algorithms for AC-constrained UC; [69] for SOC-constrained UC; and a multi-period mixed integer Jabr SOCP is solved [70] for AC-constrained UC.

Moreover, [71] solves a sequence of Jabr SOCP relaxations using penalty terms to enforce AC feasibility. On the other hand, notable linear approximations for multi-period ACOPF are the copper-plate linear, DCOPF, and strengthened versions of DCOPF [72] which account for reactive power and voltage magnitudes. These approximations are key ingredients for solving scalable UC or Security Constrained ACOPF (SC-ACOPF). In [73] a linearized multi-period Security Constrained Stochastic ACOPF based on the strengthened DC approximation [72] is presented.

Primal Heuristics

Next, we succinctly review literature on heuristics – mostly based on convex relaxations or linear approximations – used to find AC-feasible solutions. In the single-period setting, the linear programming Hot-Start and Warm-Start approximation models are proposed in [74] for finding locally optimal single-period AC solutions. The Hot-Start model assumes there is a solved AC base-point solution available, and it uses to fix voltage magnitudes in the power AC power flow definitions. Moreover, they approximate $\sin(\theta)$ by θ and use a linear convex approximation of $\cos(\theta)$ (under the assumption $\theta \in (-\pi/2, \pi/2)$). Since fixing the voltages might be too restrictive in some cases, their Warm-Start model chooses takes a subset of the buses as target buses, and linearizes deviations from the targeted voltage magnitudes. In [37] and [47] SLP algorithms are proposed for finding locally optimal single-period AC solutions on small-to medium-sized instances. These SLP-based methods were shown to produce solutions with runtimes competitive with contemporary NLP solvers, although convergence was assessed using relatively loose feasibility tolerances (typically on the order of 10^{-3}), and feasibility statistics were reported primarily in terms of average residuals.

With respect to the multi-period setting, [75] presents an outer-approximation iterative algorithm for the UC problem with AC network constraints (i.e, an multi-period ACOPF is being solved) – this method solves a finite sequence of MILP master problems and nonlinear subproblems (ACOPF with fixed binary variables). The nonlinear subproblems are tackled using the SLP presented in [76]. Moreover, in [67] an SLP is used as a penalty method to recover a UC problem with AC-constraints feasible solution from the multi-period Jabr relaxation. This is different from the technique used in [71] where a multiobjective SOCP relaxation is used to obtain an AC-feasible solution. In [77] a Benders decomposition algorithm is used for AC-constrained UC – the subproblem is a multi-period ACOPF on a small network. Note that all of the previously mentioned references on convex relaxations of some multi-period ACOPF variant are used for AC feasibility.

Two general purpose nonlinear solvers which can be highly effective toward the single and multi-period ACOPF problems are Knitro [78] and IPOPT [79]; see Section 2.8.

2.5 SOCP Relaxations for ACOPF and Power Losses

2.5.1 Jabr Relaxation

A well-known convex relaxation of ACOPF is the Jabr relaxation [22]. It linearizes the power flow definitions (2.8d)-(2.8g) using $|\mathcal{B}| + 2|\mathcal{E}|$ additional variables and adds $|\mathcal{E}|$ rotated-cone inequalities. A simple derivation is as follows: For any line $(k, m) \in \mathcal{A}$, we define:

$$v_k^{(2)} := |V_k|^2, \quad v_m^{(2)} := |V_m|^2, \quad c_{km} := |V_k||V_m|\cos(\theta_{km}), \quad s_{km} := |V_k||V_m|\sin(\theta_{km}). \quad (2.10)$$

Clearly we have the following valid nonconvex quadratic relation

$$c_{km}^2 + s_{km}^2 = v_k^{(2)} v_m^{(2)}, \quad (2.11)$$

which in Jabr [22] is relaxed into the (convex) inequality

$$c_{km}^2 + s_{km}^2 \leq v_k^{(2)} v_m^{(2)}. \tag{2.12}$$

This is a rotated-cone inequality hence it can be represented as a second-order cone constraint. Note that (2.10) can be used to represent the power flow equations in (2.8d)-(2.8g) as, $\forall (k,m) \in \mathcal{A}$,

$$P_{km} = G_{kk} v_k^{(2)} + G_{km} c_{km} + B_{km} s_{km} \tag{2.13a}$$

$$P_{mk} = G_{mm} v_m^{(2)} + G_{mk} c_{km} - B_{mk} s_{km} \tag{2.13b}$$

$$Q_{km} = -B_{kk} v_k^{(2)} - B_{km} c_{km} + G_{km} s_{km} \tag{2.13c}$$

$$Q_{mk} = -B_{mm} v_m^{(2)} - B_{mk} c_{km} - G_{mk} s_{km}. \tag{2.13d}$$

In summary, the Jabr relaxation can be obtained from the formulation (2.8) by: (i) adding the $c_{km}, s_{km}, v_k^{(2)}$ variables; (ii) replacing (2.8d)-(2.8e) with (2.13); and (iii) adding the constraints (2.12).[8]

2.5.2 A Stronger SOCP

Recall that complex power injected into branch $(k,m) \in \mathcal{A}$ at bus $k \in \mathcal{B}$ is defined by $S_{km} := V_k I_{km}^*$, hence

$$|S_{km}|^2 = |V_k|^2 |I_{km}|^2 \tag{2.14}$$

holds. Moreover, since complex power can be decomposed into active and reactive power as $S_{km} = P_{km} + jQ_{km}$, and recalling that $v_k^{(2)} := |V_k|^2$ while denoting $i_{km}^{(2)} := |I_{km}|^2$, we have

$$P_{km}^2 + Q_{km}^2 = v_k^{(2)} i_{km}^{(2)}. \tag{2.15}$$

[8] We stress that (2.10) is *not* added.

By relaxing the equality (2.15) we obtain the *i2 rotated-cone inequality* [50]

$$P_{km}^2 + Q_{km}^2 \leq v_k^{(2)} i_{km}^{(2)}. \tag{2.16}$$

Since the variable $i_{km}^{(2)}$ can be defined by

$$i_{km}^{(2)} = \alpha_{km} v_k^{(2)} + \beta_{km} v_m^{(2)} + \gamma_{km} c_{km} + \zeta_{km} s_{km} \tag{2.17}$$

for appropriate constants $\alpha_{km}, \beta_{km}, \gamma_{km}$ and ζ_{km} (Appendix A.3), we obtain an alternative SOCP relaxation. This formulation, which we call the *i2 relaxation*, is comprised by (2.8a)-(2.8c), (2.8h)-(2.8k), the definition of $i_{km}^{(2)}$ (2.17), and the rotated-cone inequalities (2.16).

It is known [24], [80], [81] that for each branch $\{k, m\}$ the system defined by the linearized power flows (2.13) plus the Jabr inequality (2.12), and the system defined by the linearized power flows (2.13), the rotated-cone inequality (2.16) and the linear definition of $i_{km}^{(2)}$ (A.12), are equivalent. It is to be noted that by this we mean that for each feasible solution to one system there is a feasible solution to the other one. However, equivalence may fail to hold if $i^{(2)}$ is upper bounded.

Proposition 2.1. *The Jabr and the $i2$ relaxations are equivalent if the $i^{(2)}$ variables are not upper bounded, and otherwise the $i2$ relaxation can be strictly stronger.*

Proof. Sufficiency is established in [24]. For an example where the i2 relaxation is strictly stronger than the Jabr relaxation see Appendix A.2. □

Our computational experiments corroborate this fact; we have found that linear cuts for the rotated-cone inequalities (2.12) and (2.16) have significantly different impact in lower bounding ACOPF (Section 2.8).

2.5.3 Losses

Transmission line losses are a structural feature of power systems [33]. It is a physical phenomenon in which the (complex) power sent from bus k to bus m along branch km will not necessarily be equal to the power received at bus m. This feature of power grids is captured (in steady state) by the power flow equations (2.8d)-(2.8g). Indeed, active-power loss on line $\{k,m\}$ can be defined as the quantity

$$\ell_{\{k,m\}} := P_{km} + P_{mk}. \tag{2.18}$$

First, we will focus on active-power losses. It is well-known (see, e.g., [31], [50]) that for a simple transmission line, i.e., a non-transformer branch without shunts, active-power loss equals

$$\ell_{\{k,m\}} = g_{\{k,m\}}|V_k - V_m|^2 \tag{2.19}$$

where $g_{\{k,m\}}$ denotes line conductance (for simple lines we have $G_{kk} = G_{mm} =: g_{\{k,m\}}$). Thus, if $g_{\{k,m\}} \geq 0$ (i.e., the standard case [33], [82]), the linear inequality $P_{km} + P_{mk} \geq 0$ is valid for ACOPF. This inequality still holds under (traditional) branch shunts and transformers provided that $g_{\{k,m\}} \geq 0$. In an arbitrary relaxation to ACOPF we might have $\ell_{\{k,m\}} < 0$, and, as we demonstrate below; this feature will make the relaxation weak. Moreover, critically, Proposition 2.4 below shows that the Jabr inequalities imply nonnegative active power losses as outer-envelope inequalities.

Additionally, the following theorem in [32] sheds light on the relationship between active-power losses and (global) active power balance. Consider an ACOPF instance on an undirected network $\mathcal{N} = (\mathcal{B}, \mathcal{E})$. Let us subdivide each branch $\{k,m\}$ by introducing a new *node*, denoted n_{km}. To properly place this result, we consider a solution to any relaxation for ACOPF. *Any* relaxation is of interest (even an exact relaxation) so long as variables P_{km} and P_{mk} are present, and active power flow balance constraints (2.8b) are given (or implied). Now, given a feasible solution to the relaxation we assign real flow values to the edges of the subdivided network, and orient those edges, as follows. Consider an arbitrary branch $\{k,m\}$.

(1) If $P_{km} \geq 0$ the edge between k and n_{km} is oriented from k to n_{km} and assigned flow P_{km}. If $P_{km} < 0$ the edge between k and n_{km} is oriented from n_{km} to k and assigned flow $-P_{km}$.

(2) Similarly, if $P_{mk} \geq 0$ the edge between m and n_{km} is oriented from m to n_{km} and assigned

flow P_{mk}. If $P_{mk} < 0$ the edge between m and n_{km} is oriented from n_{km} to m and assigned flow $-P_{mk}$.

Thus, each edge in the subdivided network is oriented in the direction of the flow value that was assigned. Using the well-known network flow-decomposition theorem [83] yields that the flows in the subdivided network can be decomposed into a sum of *flow-carrying paths*, i.e., directed paths where

- Each path starts from a *source* (a bus with positive net generation or a node on a branch with negative loss) and ends at a *sink* (a bus with negative net generation or a node on a branch with positive loss),
- Each path carries a fixed positive flow amount.

In summary, we have the following theorem.

Theorem 2.2 (Theorem 1.2.11 [32])**.** *Consider any relaxation to an ACOPF instance that includes, or implies, active power flow balance constraints* (2.8b)*. Then*

(a) The sum of active-power generation minus negative active power losses amounts equals the sum of active-power loads plus positive active-power losses. Furthermore, there exists a family of flow-carrying paths that accounts for all the flows in the subdivided network.

(b) There is a similar statement regarding reactive power flows (paraphrased here for brevity). The sum of reactive-power generation amounts plus the sum of reactive-power generated by line shunt elements, equals the sum of reactive-power loads and losses. Furthermore, there exist a family of paths that account for all reactive-power generation, loads, and losses. Each path has as origin a generator (with positive reactive-power generation) or a shunt element, and as destination a bus with a positive reactive-power load or a line loss. If all shunt susceptances are zero, then each path has a bus with positive reactive-power generation as origin.

The relevance of (a) is that when negative losses are present, then total generation may be smaller than total loads – effectively, the negative losses amount to a source of free generation and directly contribute to a lower objective value for ACOPF than AC-feasible. While this point is also made in [24] regarding a particular relaxation, the above theorem provides a clear explanation for this fact that applies to *any* relaxation.

As an experiment, suppose we remove, from the Jabr relaxation, just one of the SOC constraints, say for branch $\{\hat{k}, \hat{m}\}$, thus obtaining a weaker relaxation; and suppose that, furthermore, the branch limits U_{km} are large (the usual case). Then, as we substantiate next, it is quite likely that the loss on $\{\hat{k}, \hat{m}\}$ will be negative in the weakened relaxation and the value of the relaxation will be (much) weaker – precisely because that phenomenon would allow a solution with less generation, and thus, lower cost, than otherwise possible.

Table 2.1 provides empirical verification; it reports on 100 experiments in which exactly one Jabr constraint is removed from the formulation. "Avg Loss" represents the average loss across all branches; "Avg (br)" and "Min (br)" are the average and minimum (resp.) loss at the branch whose SOC constraint was bypassed; and "Jabr" denotes the value of the Jabr relaxation, while "weak Jabr" is the average value of the 100 weakened Jabr SOCs. For reference, "AC-L" represents the total loss of an AC locally optimal solution, while "SOC-L" is the total loss of an optimal solution to the Jabr SOC.

Table 2.1: Average and minimum losses 100 repetitions of removing a randomly selected SOC constraint.

Case	Avg Loss	Avg (br)	Min (br)	Jabr	weak Jabr	AC-L	SOC-L
case14	-0.3808	-0.4906	-1.7443	8075.12	6292.78	0.0929	0.0918
case118	0.1084	-0.7046	-5.1803	129340.00	126982.72	0.7740	0.7125
case300	1.8485	-1.1652	-6.1421	718654.00	714858.26	3.0274	2.8064

In Table 2.2 we provide additional evidence in the same direction. We report on 100 experiments, for larger cases, where we remove 100 Jabr constraints randomly selected using the uniform distribution. To highlight the relative orders of magnitude of the losses, we exhibit the total load for each case. The column "Sample" is an average over all the samples of the 100 repetitions.

Table 2.2: Average and minimum losses 100 repetitions of removing 100 randomly selected SOC constraints.

		Weak Jabr				Jabr		AC
		Avg Losses		Min Losses				
Case	Load	Total	Sample	Sample	Objective	Objective	Losses	Losses
1354pegase	730.60	-150.38	-379.60	-764.52	58021.55	74008.58	9.49	10.09
2869pegase	1324.37	-197.83	-402.59	-1119.85	112663.19	133872.69	14.24	15.51
3375wp	483.63	-65.50	-117.11	-343.72	6467352.82	7385372.70	6.90	8.30

The following example sheds light on the importance of imposing the non-negative active-power loss requirement.

Remark 2.3. Solving the McCormick relaxation of the rectangular QCQP formulation [35] of case3 from MATPOWER, we obtain an optimal solution with zero active-power generation. In particular[9], the active-power losses are

$$\ell_{12} = 0.0789, \quad \ell_{13} = -1.1941, \quad \ell_{23} = 0.0331,$$

i.e., total active-power losses equals -1.0821.

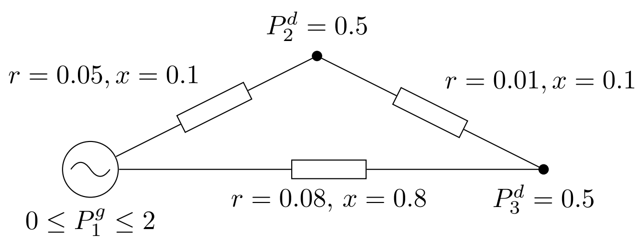

Figure 2.3: Circuit representation of case3 from MATPOWER.

Given that we obtain active power loss as a linear combination of the power flow equations it is natural to wonder if the SOCP relaxations for ACOPF (or outer approximations thereof) accurately approximate active-power losses. In Table 2.3 we provide empirical evidence that the quality of the

[9]Since total active-power demand is 1.00 and there is a bus shunt at bus 2 with shunt conductance 0.1, and the magnitude of voltage at bus 2 squared is 0.8211 in our solution, we have that generation equals losses plus the sum of the loads.

approximations to total losses provided by the SOCP relaxations is still very weak – even though the objective lower bounds proved by the SOCP relaxations are quite close to ACOPF primal bounds.

Table 2.3: Active-power losses and optimality gaps.

	AC		Jabr SOCP			
Case	Objective	Losses	Objective	Losses	Opt Gap	Loss Gap
1354pegase	74069.35	10.09	74008.58	9.48	0.08%	6.04%
ACTIVSg10k	2485898.75	23.01	2466666.10	13.88	0.77%	39.68%
ACTIVSg70k	16439499.83	156.83	16217263.66	121.31	1.35%	22.65%

It must be noted, however, that from a theoretical perspective the existence of active-power losses is not what makes ACOPF intractable; the $\mathbb{NP}$-hardness proof in [19], [20] relies on a lossless system.

2.5.4 Losses and Cutting-Planes

Next we provide theoretical justification for using an outer-approximation cutting-plane framework on the Jabr and $i2$ relaxations. We show that for transmission lines with $G_{kk} > 0 > G_{km} = G_{mk} \geq -G_{kk}$ and $B_{km} = B_{mk}$, in particular lines with no transformer or shunt elements, active-power loss inequalities are implied by the Jabr inequalities, and also by the definition of the $i^{(2)}$ variable. We begin with two simple technical observations.

First, consider a (generic) rotated cone inequality

$$x^2 + y^2 \leq wz \tag{2.20}$$

where $x, y \in \mathbb{R}$ and $w, z \in \mathbb{R}_+$. Then

$$x^2 + y^2 \leq wz \iff (2x)^2 + (2y)^2 \leq (w+z)^2 - (w-z)^2 \tag{2.21a}$$

$$\iff \|(2x, 2y, w-z)^\top\|_2 \leq w + z. \tag{2.21b}$$

Next, let $\lambda \in \mathbb{R}^3$ satisfy $\|\lambda\|_2 = 1$. Then, by (2.21b),

$$2\lambda_1 x + 2\lambda_2 y + \lambda_3(w - z) \leq \|\lambda\|_2 \, \|(2x, 2y, w - z)^\top\|_2 \leq w + z. \tag{2.22}$$

Inequality (2.22) provides a generic recipe to obtain outer-envelope inequalities for the rotated cone (2.20). As a result of these developments, we have:

Proposition 2.4. *For a transmission line $(k,m) \in \mathcal{A}$ with $G_{kk} > 0 > G_{km} = G_{mk} \geq -G_{kk}$ and $B_{km} = B_{mk}$, the Jabr inequality $c_{km}^2 + s_{km}^2 \leq v_k^{(2)} v_m^{(2)}$ implies, as an outer envelope approximation inequality, that $\ell_{\{k,m\}} \geq 0$.*

Proof. By (2.21b) we have that the Jabr inequality $c_{km}^2 + s_{km}^2 \leq v_k^{(2)} v_m^{(2)}$ can be written as $\|(2c_{km}, 2s_{km}, v_k^{(2)} - v_m^{(2)})\|_2 \leq v_k^{(2)} + v_m^{(2)}$. Hence, taking $\lambda = (1,0,0)^\top$, and using (2.22) we have

$$v_k^{(2)} + v_m^{(2)} - 2c_{km} \geq 0. \tag{2.23}$$

On the other hand, by summing up equations (2.13b) and (2.13a) we have

$$\begin{aligned} \ell_{\{k,m\}} &= G_{kk} v_k^{(2)} + G_{mm} v_m^{(2)} - 2G_{km} c_{km} \\ &\geq \min\{G_{kk}, G_{mm}\}(v_k^{(2)} + v_m^{(2)} - 2c_{km}) \end{aligned}$$

given our assumptions on $G_{kk}, G_{mm}, G_{km}, G_{mk}$. Therefore, (2.23) implies $\ell_{\{k,m\}} \geq 0$. $\square$

Moreover, for simple transmission lines the definition of the variable $i^{(2)}$ implies the active-power loss inequalities [80], [84]. Actually, if a line (k,m) has no shunts, then it is folklore (and follows from basic physics) that $i_{km}^{(2)} \geq 0$ implies $\ell_{\{k,m\}} \geq 0$.

Proposition 2.5. *For any transmission line $(k,m) \in \mathcal{A}$ with $y_{\{k,m\}}^{sh} = 0$, the definition of the variable $i_{km}^{(2)}$ implies the active-power loss inequality associated with that line.*

Proof. First, we note that if $y^{sh}_{km} = 0$, a simple calculation yields

$$\ell_{km} = g_{\{km\}} \left| \frac{1}{\tau_{km}} V_k - e^{j\sigma_{km}} V_m \right|^2 ,$$
$$I_{km} = \frac{y_{\{km\}}}{\tau_{km}} \left(\frac{1}{\tau_{km}} V_k - e^{j\sigma_{km}} V_m \right)$$

see A.3. Therefore $i^{(2)}_{km} \geq 0$ clearly implies $\ell_{\{k,m\}} \geq 0$. $\square$

2.5.5 A Numerically Better-Behaved i2 Relaxation

The i2 relaxation is at least as strong as the Jabr relaxation, and in congested networks it can be strictly stronger (Proposition 2.1). Thus, the i2 relaxation is a natural candidate for outer-approximation via a cutting-plane framework (Section 2.6); we numerically verify the strength of a linear outer-approximation of the i2 relaxation in Section 2.8.1. This additional strength comes at a price: numerical instability. To be precise, the potentially large coefficients in the definition of the $i^{(2)}$ variables can result in an ill-conditioned SOCP (see, e.g., [85]). We illustrate this issue with a simple case and show how the Jabr outer-envelope inequalities complement the strength of the $i^{(2)}$ definitions. We conclude this subsection with a new, numerically stable modified i2 relaxation, which we term i2+(ρ), where $\rho > 0$.

Let (k, m) be a transmission line with a limit $U_{\{k,m\}} < +\infty$, and without any transformers or shunt elements. By (A.12) we have

$$i^{(2)}_{km} = \left(\frac{1}{r^2_{\{k,m\}} + x^2_{\{k,m\}}} \right) \left(v^{(2)}_k + v^{(2)}_m - 2c_{km} \right) \tag{2.24}$$

where $r_{\{k,m\}}$ and $x_{\{k,m\}}$ denote the line's resistance and reactance. Moreover, equation (2.14) implies that $i^{(2)}_{km}$ can be upper-bounded by the constant $H_{\{k,m\}} := U_{\{k,m\}}/(V^{\min}_k)^2$, where $V^{\min}_k$ corresponds to the lower limit for voltage magnitude at bus k (both voltage limits are close to 1 under standard data conditions).

Suppose that this simple line (k, m) has small resistance, e.g., on the order of 10^{-5}. Then

$1/(r^2_{\{k,m\}} + x^2_{\{k,m\}})$ can be on the order of 10^8. In addition, assume that the limit U_{km} is *small*. Thus, $(r^2_{\{k,m\}} + x^2_{\{k,m\}})H_{\{k,m\}}$ can be fairly small, which yields

$$v_k^{(2)} + v_m^{(2)} - 2c_{km} \leq (r^2_{\{k,m\}} + x^2_{\{k,m\}})H_{\{k,m\}} \approx 0. \tag{2.25}$$

Since $v_k^{(2)} + v_m^{(2)} - 2c_{km} \geq 0$ is a Jabr outer-envelope cut (proof of Proposition 2.4), inequality (2.25) is forcing our solutions to lie near the surface of the rotated-cone $c_{km}^2 + s_{km}^2 \leq v_k^{(2)} v_m^{(2)}$, i.e., it is cutting-off solutions which are in the interior of the Jabr cone for line $\{k, m\}$. Notice that the more capacitated the transmission line is, the stronger the effect of the $i^{(2)}$ variables (Figure 2.4).

This key observation motivates a numerically more stable relaxation. Consider the definition of $i_{km}^{(2)}$ in Appendix A.3 for appropriate constants $\alpha_{km}, \beta_{km}, \gamma_{km}$ and ζ_{km}:

$$i_{km}^{(2)} = \alpha_{km} v_k^{(2)} + \beta_{km} v_m^{(2)} + \gamma_{km} c_{km} + \zeta_{km} s_{km}. \tag{2.26}$$

Let $\rho > 0$ be some fixed parameter, and consider the following heuristic: for every transmission line $(k, m) \in \mathcal{A}$,

(i) if $\alpha_{km} > \rho$, we say that the branch induces a *bad-i2*, and we define the following inequalities (which are valid by the previous discussion)

$$v_k^{(2)} + \frac{\beta_{km}}{\alpha_{km}} v_m^{(2)} + \frac{\gamma_{km}}{\alpha_{km}} c_{km} + \frac{\zeta_{km}}{\alpha_{km}} s_{km} \geq 0, \tag{2.27a}$$

$$v_k^{(2)} + \frac{\beta_{km}}{\alpha_{km}} v_m^{(2)} + \frac{\gamma_{km}}{\alpha_{km}} c_{km} + \frac{\zeta_{km}}{\alpha_{km}} s_{km} \leq \frac{H_{\{k,m\}}}{\alpha_{km}}; \tag{2.27b}$$

(ii) otherwise, we say that the branch induces a *good-i2*, and we define

$$i_{km}^{(2)} = \alpha_{km} v_k^{(2)} + \beta_{km} v_m^{(2)} + \gamma_{km} c_{km} + \zeta_{km} s_{km}.$$

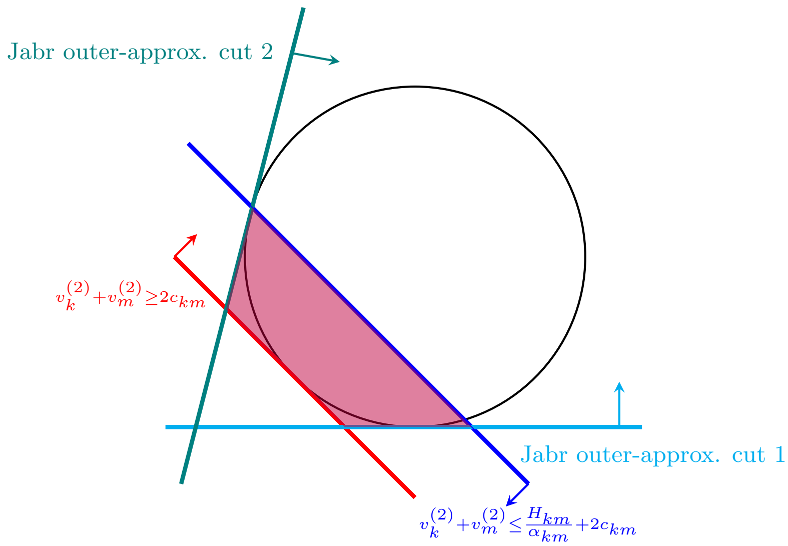

Figure 2.4: The black circle represents a horizontal cross-section of a projected Jabr cone (2.12) onto the $(v_k^{(2)}, v_m^{(2)}, c_{km})$-space. The red and blue inequalities represent (2.27) for a simple transmission line, while the teal and cyan inequalities represent two Jabr outer-approximation cuts.

Empirically, we see in Table 2.4 that the coefficients of the inequalities (2.27) are relatively small; clearly closer to the range of coefficients of the linear constraints of the Jabr SOCP, and to the coefficients[10] of the Jabr outer-envelope cuts which will be deployed in our cutting-plane algorithm (refer to the linearly constrained relaxation M in Section 2.6.2). We observe that for $\rho = 100$ roughly 90% of the branches correspond to *bad-i2*s.

Table 2.4: Statistics of the coefficient α in (2.26) and of the coefficients in (2.27).

		α		$\frac{\beta}{\alpha}$		$\frac{\gamma}{\alpha}$		$\frac{\zeta}{\alpha}$		$\frac{H}{\alpha}$		
Case	%bad-i2s	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Min	Avg
ACTIVSg10k	91.33	1.00e10	1.09e06	1.47	1.01	-1.80	-2.01	0.88	0.00	18.52	2.57e-09	3.06
10000goc-sad	88.50	2.23e06	2.41e05	1.21	1.00	-1.90	-2.00	0.01	0.00	1.89	4.64e-06	0.48
30000goc-api	96.13	2.41e07	3.66e05	1.31	1.00	-1.94	-2.00	0.01	5.90e-06	1.89	2.16e-06	0.71

By making reasonable assumptions about transmission line parameters we can derive the following upper bounds for the coefficients in inequalities (2.27).

[10]The size of coefficients of these cuts are linear in the values of variables c, s and v in any solution to a linear relaxation of the Jabr SOCP; see Proposition 2.7.

Proposition 2.6. *Let* $(k, m) \in \mathcal{A}$ *be a transmission line that satisfies* $|b_{\{k,m\}}| > b^{sh}_{\{k,m\}}$, $g^{sh}_{\{k,m\}} = 0$ *and* $|b_{\{k,m\}}| > g_{\{k,m\}}$*. Then, the coefficients* $\frac{\beta_{km}}{\alpha_{km}}$, $\frac{\gamma_{km}}{\alpha_{km}}$, *and* $\frac{\zeta_{km}}{\alpha_{km}}$ *in* (2.27) *are bounded by* $\max\{\tau^2_{km}, 3\tau_{km}\}$, *where* τ_{km} *is the transformer tap ratio located at bus* k.[11]

Proof. Let $\{k, m\} \in \mathcal{E}$ be an arbitrary transmission line, and denote by $A := g^2_{\{k,m\}} + b^2_{\{k,m\}}$, $B := g_{\{k,m\}} g^{sh}_{\{k,m\}} + b_{\{k,m\}} b^{sh}_{\{k,m\}}$ and $C := (g^{sh2} + b^{sh2})/4$, see (A.12). We drop the subscripts for clarity of exposition. Then, the coefficients in (2.27) can be written as

$$\begin{aligned}
\frac{\beta}{\alpha} &= \tau^2 \frac{A}{A+B+C} \\
\frac{\gamma}{\alpha} &= \tau \frac{-\cos(\sigma)(2A+B) + \sin(\sigma)(bg^{sh} - gb^{sh})}{A+B+C} \\
\frac{\zeta}{\alpha} &= -\tau \frac{\sin(\sigma)(2A+B) + \cos(\sigma)(bg^{sh} - gb^{sh})}{A+B+C}
\end{aligned}$$

Since we assumed $|b| > b^{sh}$, $g^{sh} = 0$, and $|b| > g$ we have that $A + B > 0$, hence the ratios $\frac{\gamma}{\alpha}$ and $\frac{\zeta}{\alpha}$ can be upper-bounded by

$$\tau \left(\frac{2A + |bb^{sh}|}{A+B+C} + \frac{|gb^{sh}|}{A+B+C} \right).$$

In consequence, we have the following upper-bounds

$$\frac{\beta}{\alpha} \le \tau^2, \quad \frac{\gamma}{\alpha} \le 3\tau, \quad \frac{\zeta}{\alpha} \le 3\tau.$$

$\square$

In consequence, given $\rho > 0$, we define our numerically better-behaved i2+(ρ) relaxation by (2.8a)-(2.8c), (2.8h)-(2.8k), the definition of $i^{(2)}_{km}$ (A.12) and the rotated-cone inequalities (2.16) for every good-i2 branch (k, m), and inequalities (2.27) for every bad-i2 branch (k, m). We use the i2+(ρ) relaxation for our multi-period experiments (see discussion and results in Section 2.8.3).

[11] If the line does not have a transformer, then $\tau_{km} = 1$; generally, tap ratios do not take extreme values.

2.6 Cutting-plane Framework: Cut Separation and Management

In this paper we use a dynamically generated linearly constrained relaxation as a lower bounding procedure for ACOPF; relying on a perspective sometimes associated with integer programming.

Given a set X in $\mathbb{R}^n$, we say that a convex inequality $g(x) \leq d$ is *valid* for X if $g(x) \leq d$ holds for all $x \in X$. Given X and $\overline{x} \notin X$, then we say that $c^\top x \leq d$ is a (linear) *cut* for X if the inequality is valid for X but it is not satisfied by $\overline{x}$. A (linear) cutting-plane algorithm [48], [86] for a set X is an iterative procedure in which, starting from an initial (linear) relaxation, in every round (linear) cuts are added to approximate X. Typically, these cuts are computed in iterative fashion; at each round an optimal solution $\overline{x} \notin X$ to the current relaxation is separated using a computed cut. We note that in addition to Kelley's cutting-plane algorithm [48] other related algorithms have been proposed in the literature. See [87] for background.

In the case of single-period experiments our target set X is the i2 relaxation of ACOPF, while we use the i2+(ρ) relaxation for the multi-period setting. The reason is simple: the multi-period i2 relaxation for medium to large networks is not numerically tractable.

In support of this statement, we note that direct solution of the Jabr and i2 relaxations of ACOPF, for large instances, is computationally prohibitive and often results in non-convergence (see tables 2.6, 2.7, 2.8, 2.9, 2.12, A.1 and A.2). Our empirical evidence further shows that outer-approximation of the rotated-cone inequalities (in either case) requires a large number of cuts in order to achieve a tight relaxation value. Moreover, employing such large families of cuts yields a relaxation that, while linearly constrained, still proves challenging – both from the perspective of running time and numerical tractability. Nonetheless, a characteristic feature of our iterative procedure is its robustness to potentially suboptimal termination of the oracle used to solve the LPs or convex QPs; *independent* of the quality of the primal solution obtained, our linear cuts will always be valid.

However, as we show, adequate cut management proves successful, yielding a procedure that is (a) rapid, (b) numerically stable, and (c) provides a very tight relaxation (Tables 2.8 and 2.9).

The critical ingredients in this procedure are: (1) fast cut separation; (2) appropriate violated cut selection; and (3) effective dynamic cut management, including rejection of *nearly-parallel* cuts and removal of *expired* cuts, i.e., previously added cuts that have become slack (see Section 2.6.2).

Our procedure possesses efficient warm-starting capabilities – this is a central goal of our work. Previously computed cuts, for some given instance, can be reutilized and loaded into new runs of *related* instances, hence leveraging previous computational effort. It is worth noting that this reoptimization feature stands in sharp contrast to what is possible using nonlinear (convex) solvers. In Section 2.8.2 we justify the validity of this feature and Tables 2.8 and 2.9 summarize extensive numerical evidence on its performance relative to solving the SOCPs 'from scratch'. We remark that adequate cut management is what enables this capability in the case of large single and multi-period ACOPF instances.

2.6.1 Two Simple Cut Procedures

The following (routine) results provide a fast procedure for separating over the rotated-cone inequalities

$$c_{km}^2 + s_{km}^2 \leq v_k^{(2)} v_m^{(2)}, \qquad P_{km}^2 + Q_{km}^2 \leq v_k^{(2)} i_{km}^{(2)}. \tag{2.28}$$

Proposition 2.7. *Consider the second-order cone* $C := \{(x, s) \in \mathbb{R}^n \times \mathbb{R}_+ \,:\, \|x\|_2 \leq s\}$*. Suppose* $(x', s') \notin C$ *with* $s' \geq 0$*. Then the projection of* (x', s') *onto* C *is given by*

$$x_0 := s_0 \frac{x'}{\|x'\|}, \quad s_0 := \frac{\|x'\| + s'}{2}.$$

Moreover, the hyperplane which achieves the maximum distance from (x', s') *to any hyperplane separating* C *and* (x', s') *is given by*

$$(x')^\top x \leq s\|x'\|.$$

Proof. First, we show that for every $(x, s) \in C$ the inequality $(x')^\top x \leq \|x'\|s$ holds. Indeed, by

Cauchy-Schwartz inequality we have that $(x')^\top x \leq \|x'\|\|x\|$, and since $\|x\| \leq s$ the inequality follows. Next, we show that the point (x_0, s_0) defined by $s_0 := (\|x'\| + s')/2$ and $x_0 := (s_0/\|x'\|)x'$ lies on $(x')^\top x = \|x'\|s$. Certainly, $(x')^\top x_0 = (s_0/\|x'\|)\|x'\|^2 = s_0\|x'\|$. Moreover, (x_0, s_0) lies on C since $\|x_0\| = \|(s_0/\|x'\|)x'\| = s_0$. Finally, we show that the vector $(x_0 - x', s_0 - s')$ is orthogonal to the plane (it suffices to take the vector (x_0, s_0) since $(0, 0)$ lies on C). It can be readily checked that $\|(x_0, s_0)\|^2 - (x')^\top x_0 - s's_0 = 0$. $\square$

Corollary 2.8. *Let* $C := \{(x, y, w, z) \in \mathbb{R}^2 \times \mathbb{R}^2_+ \;:\; x^2 + y^2 \leq wz\} \subseteq \mathbb{R}^4$ *and suppose that* $(x', y', w', z') \notin C$ *where* $w' + z' > 0$. *Then the hyperplane separating* C *from* (x', y', w', z') *and at maximum distance from this point is given by*

$$(4x')x + (4y')y + ((w' - z') - n_0)w + (-(w' - z') - n_0)z \leq 0, \tag{2.29}$$

where $n_0 := \|(2x', 2y', w' - z')^\top\|_2$.

Proof. Rewriting the rotated-cone inequality as (2.21b) and applying Proposition 2.7 provides the desired separating hyperplane. $\square$

The following proposition gives us a simple procedure for computing linear cuts for violated limit inequalities

$$P_{km}^2 + Q_{km}^2 \leq U_{\{k,m\}}. \tag{2.30}$$

Proposition 2.9. *Consider the Euclidean ball in* $\mathbb{R}^2$ *of radius* r, $S_r := \{(x, y) \in \mathbb{R}^2 \;:\; x^2 + y^2 \leq r^2\}$, *and let* $(x', y') \notin S_r$. *Then the separating hyperplane that is at maximum distance from* (x', y') *is given by*

$$(x')x + (y')y \leq r\|(x', y')\|_2. \tag{2.31}$$

Proof. Since $(x')^2 + (y')^2 > r^2$, there exists some $0 < t_0 < 1$ such that $(t_0x')^2 + (t_0y')^2 = r^2$. It can be readily checked that $t_0(x', y')$ is the projection of (x', y') onto S_r. Therefore, the normal to the separating hyperplane is $(1 - t_0)(x', y')$ and the RHS is $(1 - t_0)t_0\|(x', y')\|_2^2$, in other words, $(x')x + (y')y \leq r\|(x', y')\|_2$ is the desired valid inequality since $0 < t_0 < 1$. $\square$

2.6.2 Basic Cutting-Plane Algorithm

Single-period

We describe our cutting-plane algorithm for the single-period setting. First, we define a linearly constrained base model M_0 as follows:

Figure 2.5: Linearly constrained base model M_0

$$[M_0]: \qquad \min \quad \sum_{k \in \mathcal{G}} F_k(P^g_k) \tag{2.32a}$$

subject to:

$\forall$ bus $k \in \mathcal{B}$:

$$\sum_{\{k,m\} \in \delta(k)} P_{km} = \sum_{\ell \in \mathcal{G}_k} P^g_\ell - P^d_k, \qquad \sum_{\{k,m\} \in \delta(k)} Q_{km} = \sum_{\ell \in \mathcal{G}_k} Q^g_\ell - Q^d_k \tag{2.32b}$$

$\forall$ branch $(k,m) \in \mathcal{A}$:

$$P_{km} = G_{kk} v^{(2)}_k + G_{km} c_{km} + B_{km} s_{km} \tag{2.32c}$$

$$P_{mk} = G_{mm} v^{(2)}_m + G_{mk} c_{km} - B_{mk} s_{km} \tag{2.32d}$$

$$Q_{km} = -B_{kk} v^{(2)}_k - B_{km} c_{km} + G_{km} s_{km} \tag{2.32e}$$

$$Q_{mk} = -B_{mm} v^{(2)}_m - B_{mk} c_{km} - G_{mk} s_{km} \tag{2.32f}$$

$$i^{(2)}_{km} = \alpha_{km} v^{(2)}_k + \beta_{km} v^{(2)}_m + \gamma_{km} c_{km} + \zeta_{km} s_{km} \tag{2.32g}$$

$\forall$ generator $k \in \mathcal{G}$:

$$P^{\min}_k \leq P^g_k \leq P^{\max}_k, \quad Q^{\min}_k \leq Q^g_k \leq Q^{\max}_k \tag{2.32h}$$

$\forall$ bus $k \in \mathcal{B}$:

$$(V^{\min}_k)^2 \leq v^{(2)}_k \leq (V^{\max}_k)^2 \tag{2.32i}$$

$\forall$ branch $(k,m) \in \mathcal{A}$:

$$-\sqrt{U_{\{k,m\}}} \leq P_{km}, P_{mk} \leq \sqrt{U_{\{k,m\}}}, \quad -\sqrt{U_{\{k,m\}}} \leq Q_{km}, Q_{mk} \leq \sqrt{U_{\{k,m\}}} \tag{2.32j}$$

$$0 \leq i^{(2)}_{km} \leq \frac{U_{\{k,m\}}}{(V^{\min}_k)^2} \tag{2.32k}$$

In other words, we consider the linearized power flow equations of the Jabr SOCP, together with all linear constraints in (2.8), the definition of the variable $i^{(2)}$ in (2.26), and simple box constraints on the power flow and $i^{(2)}$ variables. Recall that for the single-period ACOPF problem, our target set X is the i2 relaxation.

In every round of our iterative procedure, linear constraints will be added to and removed from the current relaxation, which is initialized as M_0. We will denote by M the relaxation at a generic iteration of our cutting-plane algorithm.

Given a feasible solution $\bar{x}$ to M, and letting $f_{km}(x) \leq 0$ be some valid convex inequality (2.28) or (2.30), our measure of *cut-quality* is $\max\{f_{km}(\overline{x}), 0\}$, i.e., the amount by which $\overline{x}$ violates the inequality. Let $\epsilon > 0$. For each type $\tau \in \{\text{Jabr}, \text{i2}, \text{limit}\}$ of inequality we select the top p_τ percent –a fixed parameter– most highly violated inequalities of type τ whose violation is greater than ϵ – these are the τ-candidates (Line 7 of the pseudocode).

For each list of τ-candidates, we compute cuts for the corresponding branches using the cut procedures described in Section 2.6.1. Candidate cuts will be rejected if they are *nearly parallel* to incumbent cuts in M – near-linear dependency is a common source of ill-conditioning in optimization models [85], [88]. To be precise, given $\epsilon_{par} > 0$, we say that two linear inequalities $c^\top x \leq 0$ and $d^\top x \leq 0$ are ϵ_{par}*-parallel* if the cosine of the angle between their normal vectors c and d is strictly more that $1 - \epsilon_{par}$.

Finally, we describe a heuristic for *cleaning-up* our formulation. For each added cut, we keep track of its *cut-age*, i.e., the count of rounds since it was added. Then, in every iteration, if a cut $c^\top x \leq d$ has age greater or equal than a fixed parameter $T_{age} \in \mathbb{N}$, and it is ϵ*-slack*, i.e., $d - c^\top \overline{x} > \epsilon$, then it is removed from M.

In addition to M_0 and the parameters $p_\tau, \epsilon, \epsilon_{par}, T_{age}$, other inputs for our procedure are: a time limit $T > 0$; the number of admissible iterations without sufficient objective improvement $T_{ftol} \in \mathbb{N}$; and a threshold for objective relative improvement $\epsilon_{ftol} > 0$.

Algorithm 1 Cutting-Plane Algorithm

procedure CUTPLANE
    Initialize $r \leftarrow 0$, $M \leftarrow M_0$, $z_0 \leftarrow +\infty$
    while $t < T$ and $r < T_{ftol}$ **do**
        $z \leftarrow \min M$ and $\bar{x} \leftarrow$ argmin M
        Check for violated inequalities by solution $\overline{x}$.
        Sort inequalities by violation.
        Compute cuts for the most violated inequalities.
        Add computed cuts if they are not ϵ-parallel to existing cuts in M.
        Drop cuts of age $\geq T_{age}$ whose slack is $\geq \epsilon_j$.
        if $z - z_0 < z_0 \cdot \epsilon_{ftol}$ **then**
            $r \leftarrow r + 1$
        else
            $r \leftarrow 0$
        end if
        $z_0 \leftarrow z$
    end while
end procedure

Multi-Period

We outer-approximate the i2+ relaxation in the multi-period setting. Thus, the linearly constrained base model M_0^T is given by

$$[M_0^T]: \qquad \min \qquad \sum_{t=0}^{T-1} \sum_{k \in \mathcal{G}} F_k(P^g_{k,t}) \tag{2.33a}$$

subject to:

$$\text{constraints (2.8b), (2.8c), (2.13), (2.8h)} - \text{(2.8i), (2.8j)} \tag{2.33b}$$

$$\text{for every good-i2 branch } (k, m), \text{ (A.12), (2.16)} \tag{2.33c}$$

$$\text{for every bad-i2 branch } (k, m), \text{ (2.27).} \tag{2.33d}$$

In other words, we consider all the linear constraints of our i2+ relaxation.

For each time-period t, we sort violated inequalities $\tau \in \{\text{Jabr}, \text{i2}, \text{limit}\}$ by violation, as in the single-period setting, and pick as (τ, t)-candidates branches the top p_τ percentage of the most violated branches per period t. The rest follows as in Algorithm 1.

Our experiments show that our cut selection is robust, and that Gurobi version 10.0.1 [89] is able to handle surprisingly well large volume of cuts, see tables 2.10 2.11, 2.12.

2.7 On the Accuracy of Lower Bounds

In this section we address an issue of fundamental importance in numerical optimization, and which, in our opinion, has not received adequate attention in the ACOPF literature.

Most optimization solvers, commercial and academic, work with finite precision, floating-point arithmetic (some notable exceptions are the LP solvers [90] and [91]) and are subject to roundoff errors, usually small. To be precise, whenever an optimization instance is “solved" and declared “optimal" by a solver, it is highly likely that the provided solution will have small infeasibilities, i.e., the solution will exhibit constraint violations up to some tolerance. Usually, this *infeasibility tolerance* can be controlled by the user.

A natural question then is: given an ϵ-feasible $\tilde{x}$ solution to a convex optimization problem $[P]$ with optimal solution $\overline{x}$ (likely unknown by the user), how *superoptimal* can the approximate solution be? In other words, are there any general guarantees which permit us bound the superoptimality of approximate solutions? For example, if f denotes the objective function for $[P]$, we seek a guarantee of the form

$$f(\tilde{x}) \geq f(\overline{x}) - h_P(\epsilon)$$

for a certain function h_P which only depends on the instance's data. Ideally, we would like for h_P to be monotonic on ϵ and polynomial on the size (bit-length) of the problem $[P]$ data. If $[P]$ is a linear program then h_P is linear function whose slope coefficient has bit-length polynomial on the problem's bit encoding [86]. Moreover, if $[P]$ is a convex QP we also have a good guarantee as in the LP case. Our proof leverages an argument used to show that QP is in NP [92].

Proposition 2.10. *Let $[P]$ be a convex QP with objective given by $f(x) := x^\top Hx + c^\top x$, where $H \in \mathbb{Q}^{n\times n}$ is a positive definite matrix and $c \in \mathbb{Q}^n$, and feasible set $\mathcal{X} := \{x \in \mathbb{R}^n \,:\, Ax \geq b\}$, where $A \in \mathbb{Q}^{m\times n}$ and $b \in \mathbb{Q}^m$. Let $\overline{x} \in \mathbb{Q}^n$ be an optimal solution to this convex QP. Suppose there* *is a point $\tilde{x} \in \mathbb{Q}^n$ that is ϵ-feasible for $[P]$. Then,*

$$f(\tilde{x}) \geq f(\overline{x}) - \|\overline{\lambda}\|_1 \epsilon. \tag{2.34}$$

where $\overline{\lambda}$ is an optimal solution to the dual of $[P]$.

Moreover, $\|\overline{\lambda}\|_1$ can be bounded by a constant $g(A, b, H, c)$ whose bit-length is polynomial on the bit-complexity of the input (A, b, H, c).

Proof. Let $\tilde{x} \in \mathbb{R}^n$ such that $A\tilde{x} \geq b - \epsilon e$, where e denotes the all-ones vector. Let $[D]$ denote the Lagrangean Dual problem of $[P]$ on variables $\lambda \in \mathbb{R}^m$. Consider a primal-dual optimal pair $(\overline{x}, \overline{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$ such that $\overline{\lambda}$ is a rational vector that satisfies $\|\overline{\lambda}\|_1 \leq g(A, b, H, c)$, where g is a polynomial in the length of the input A, b, H, c. Such a solution $\overline{\lambda}$ exists by strong duality for convex QPs, since we assumed $[P]$ is feasible and bounded; see [92, Section 2]. Strong duality also implies that $(\overline{x}, \overline{\lambda})$ satisfies KKT conditions (i) Lagrangean optimality $\nabla f(\overline{x}) = A^\top \overline{\lambda}$ and (ii) complementary slackness $\overline{\lambda}^\top (b - A\overline{x}) = 0$. In addition, given that f is convex and differentiable we have that $f(\overline{x}) + \nabla f(\overline{x})^\top (x - \overline{x}) \leq f(x)$ for every $x \in \mathbb{R}^n$. In consequence,

$$\begin{aligned}
f(\tilde{x}) &\geq f(\overline{x}) + \nabla f(\overline{x})^\top (\tilde{x} - \overline{x}) \\
&= f(\overline{x}) + (A^\top \overline{\lambda})^\top (\tilde{x} - \overline{x}) \\
&= f(\overline{x}) + \overline{\lambda}^\top (A\tilde{x}) - \overline{\lambda}^\top (A\overline{x}) \\
&\geq f(\overline{x}) + \overline{\lambda}^\top (b - A\overline{x}) - \epsilon \|\overline{\lambda}\|_1 \\
&= f(\overline{x}) - \epsilon \|\overline{\lambda}\|_1 \\
&\geq f(\overline{x}) - \epsilon g(A, b, H, c).
\end{aligned}$$

□

Remark 2.11. This proposition can be readily applied to the dual of the convex QP which is again a convex QP (the quadratic term in the objective is negative definite). Indeed, if we denote by q the dual objective and by $\tilde{\lambda}$ an ϵ-feasible dual solution, we can apply our result to the convex QP with objective $q' := -q$ which yields

$$q'(\tilde{\lambda}) \geq q'(\overline{\lambda}) - \|\overline{x}\|_1 \epsilon \iff q(\tilde{\lambda}) \leq q(\overline{\lambda}) + \|\overline{x}\|_1 \epsilon.$$

Next, we elaborate on the relevance of Proposition 2.10. Suppose we are given:

- A (not necessarily convex) optimization problem $[Z]$ whose optimal objective value is denoted by $\overline{z}$;
- A convex relaxation $[P]$ of $[Z]$, with objective function f, and its dual $[D]$, with objective function q;
- An ϵ-feasible primal-dual pair $(\tilde{x}, \tilde{\lambda})$ and an optimal primal-dual pair $(\overline{x}, \overline{\lambda})$ for $[P]$-$[D]$.

Assume that strong duality holds[12] for the pair of convex problems $[P]$-$[D]$. Figure 2.6 depicts two inherent risks when dealing with ϵ-feasible solutions: incurring in (1) invalid lower bounds, and (2) poor primal bounds for $[Z]$. Indeed, (1) an *infeasible* (even if ϵ-feasible) dual solution could provide an invalid lower bound for $[Z]$, i.e., $\overline{z} < q(\tilde{\lambda})$. On the other hand, (2) if a primal ϵ-solution $\tilde{x}$ to the relaxation $[P]$ is used as an approximate or even an exact solution[13] to $[Z]$ and we do not have guarantees such as those provided by Proposition 2.10, then the value $f(\tilde{x})$ may be significantly off (far from $\overline{z}$).

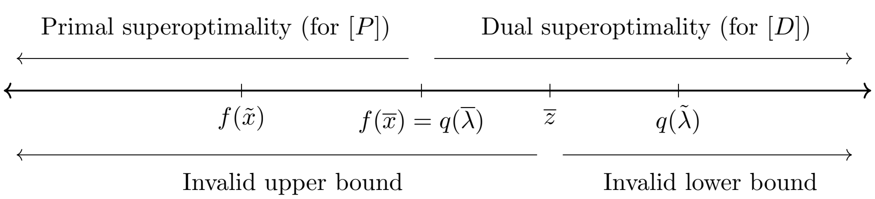

Figure 2.6: Validity of bounds and superoptimality of ϵ-feasible solutions.

Results such as inequality (2.34) are appealing since the bound has a linear dependence on ϵ. Unfortunately, small dual violations are not the only ingredient for accurate bounds; problem

[12] An analogous analysis can be done is there exists a duality gap; we assumed strong duality for clarity of explanation.
[13] For instance, for radial networks the Jabr SOCP relaxation is exact.

structure is key. In Appendix A.4 we show that if $[P]$ and $[D]$ are a primal-dual pair of SOCPs for which strong duality holds, then any ϵ-feasible solution $\tilde{x}$ satisfies $c^\top \tilde{x} \geq c^\top \overline{x} - (\|\overline{\lambda}\|_1 + \|\overline{u}\|_1)\epsilon$ where c denotes the primal objective and $(\overline{x}, (\overline{\lambda}, \overline{u}))$ a primal-dual optimal pair. Thus, a natural question is: can we always provide a 'reasonable' bound on $\|\overline{\lambda}\|_1 + \|\overline{u}\|_1$? From a fundamental perspective, such a bound should be polynomial on the bit-length of the problem data – as in the case for LPs or convex QPs[14]. In what follows we provide a partially negative answer to the latter question even for the case of the Jabr SOCP for the simplest possible network – two buses and one transmission line.

Proposition 2.12. *There exists an ACOPF instance on a two-bus network with standard line parameters and rational data such that:*

(a) the only AC-feasible solutions are irrational;

(b) the Jabr SOCP relaxation has a unique minimizer, which is irrational; and

(c) the dual of the Jabr SOCP relaxation has a unique maximizer, which is irrational.

Proof. We first prove (a) and (b). Consider a network consisting of two buses and a line joining them. There is one generator located at bus 1 and a load located at bus 2. The transmission line is simple (i.e., no shunts or transformers) and has series admittance $y = g + jb$ with $g > 0$, $b < 0$ and $-b > g$. Assume that the voltages at both buses are fixed to 1, that reactive power generation is unconstrained and that the line limit is very large. Finally, suppose that the objective consists in minimizing active generation. Additionally, there is no explicit reactive power load. We can make the following observations:

1. Active power balance constraints: $P_{1,2} = P_1^g$ and $P_{2,1} = -P_2^d$;

2. Power flows: $P_{1,2} = g - gc_{1,2} - bs_{1,2}$ and $P_{2,1} = g - gc_{1,2} + bs_{1,2}$;

3. Jabr inequality: $c_{1,2}^2 + s_{1,2}^2 \leq 1$.

[14]The existence of SOCPs whose unique solution is irrational are folklore results, see [93] for examples.

Since reactive power generation is unconstrained and the line limit is very large, reactive power flows can be ignored. Therefore, we can write the Jabr SOCP as

$$\begin{aligned} \text{OPT} := \min \quad & g - gc_{1,2} - bs_{1,2} \\ \text{subject to:} \quad & \\ & g - gc_{1,2} + bs_{1,2} = -P^d_2 && \text{(2.35a)} \\ & c^2_{1,2} + s^2_{1,2} \leq 1 && \text{(2.35b)} \end{aligned}$$

Since the network is radial, we know that imposing (2.35b) as an *equation* yields an exact ACOPF formulation [22], [94]. Given that (2.35a) is linear in $c_{1,2}$, $s_{1,2}$ and constraint (2.35b) represents the unit circle, for a particular choice of parameters g, b, and P^d_2, the line represented by (2.35a) will intersect the boundary of the unit circle at exactly two points. These points will be the unique two AC-feasible solutions for this instance, while their convex hull represents the feasible region of the SOCP. Finally unless the objective is orthogonal to the hyperplane defined by constraint (2.35a) the SOCP optimum will be exactly one of the two AC-feasible points. We derive analytically these points. Indeed, by substituting the expression for $s_{1,2}$ in (2.35a) into (2.35b) we obtain

$$\left(\frac{g^2+b^2}{b^2}\right) c^2_{1,2} - 2\left(\frac{g}{b}\right)\gamma c + (\gamma^2 - 1) = 0, \quad \gamma := \frac{P^d + g}{b}$$
$$\implies c_{1,2} = \frac{(g/b)\gamma \pm \sqrt{\alpha}}{(g^2+b^2)/b^2}, \quad \alpha := 1 + \left(\frac{g}{b}\right)^2 - \gamma^2.$$

Moreover, solving for P^d in terms of α we have $P^d = -g \pm \sqrt{g^2 - b^2(\alpha - 1)}$. Letting $g = 3$, $b = -8$, $\alpha = 3/4$, gives us $P^d = 2$ and $\gamma = -5/8$. This choice of parameters implies that the $s_{1,2}$-intercept in the $c_{1,2} - s_{1,2}$ plane of the line (2.35a) is strictly between $(-1, 1)$, hence there are exactly two AC-feasible solutions and moreover they are irrational

$$\tilde{c}_{1,2} = \frac{(15/64) + (\sqrt{3}/2)}{(9/64) + 1},\ \tilde{s}_{1,2} = -(3/8)\left(\frac{(15/64) + (\sqrt{3}/2)}{(9/64) + 1}\right) + (5/8)$$
$$\hat{c}_{1,2} = \frac{(15/64) - (\sqrt{3}/2)}{(9/64) + 1},\ \hat{s}_{1,2} = -(3/8)\left(\frac{(15/64) - (\sqrt{3}/2)}{(9/64) + 1}\right) + (5/8)$$

Given that the objective is not orthogonal to the linear constraint, and that it lies in the second quadrant, we have that the optimal solution to the Jabr SOCP is $(\tilde{c}_{1,2}, \tilde{s}_{1,2}) \approx (0.9647, 0.2632)$. This proves (a) and (b).

Next, we establish (c) by analyzing the dual of (2.35). Let $\lambda \in \mathbb{R}$ and $\delta \in \mathbb{R}_+$ be the dual variables associated to the constraints (2.35a) and (2.35b), respectively, and consider the following Lagrangean dual:

$$\begin{aligned} L(c_{1,2}, s_{1,2}; \lambda, \delta) :&= g - gc_{1,2} - bs_{1,2} + \lambda(g - gc_{1,2} + bs_{1,2} + P_2^d) \\ &\quad + \delta(c_{1,2}^2 + s_{1,2}^2 - 1) \end{aligned}$$

The function $L(c_{1,2}, s_{1,2}; \lambda, \delta)$ is convex quadratic in the variables $c_{1,2}, s_{1,2}$. Given that Slater's condition holds for (2.35), we have that an optimal primal-dual pair must satisfy the KKT condition (Lagrangean optimality)

$$\text{OPT} = \min_{c_{1,2}, s_{1,2}} \quad L(c_{1,2}, s_{1,2}; \lambda^*, \delta^*) = L(c_{1,2}^*, s_{1,2}^*; \lambda^*, \delta^*)$$

where $(c_{1,2}^*, s_{1,2}^*, \lambda^*, \delta^*)$ is an optimal primal-dual pair. It can be readily checked that the first-order stationarity condition implies that

$$2\delta^* c_{1,2}^* - g(1 + \lambda^*) = 0 \tag{2.36}$$

$$2\delta^* s_{1,2}^* + b(\lambda^* - 1) = 0. \tag{2.37}$$

Since we previously showed that the unique minimizer of (2.35) is the irrational solution $(\tilde{c}_{1,2}, \tilde{s}_{1,2})$, we have that $c^*_{1,2} = \tilde{c}_{1,2}$ and $s^*_{1,2} = \tilde{s}_{1,2}$. Therefore, equations (2.36) and (2.37) imply that (λ^*, δ^*) must also be irrational. $\square$

Remark 2.13. This is an example of an ACOPF instance (in the simplest possible non-trivial setting) whose feasible region is disconnected while its convex hull is described by the feasible region of the Jabr SOCP (since the network is radial [22]). Also see the classical example in [17]. Moreover, this example shows that to answer whether AC-Feasibility [20], [21] is in NP or not one cannot rely on a rational AC-feasible solution as a polynomial-size certificate.

In summary, LP and convex QP relaxations possess robust theoretical guarantees for bounding challenging optimization problems, in sharp contrast to what nonlinear relaxations such as SOCPs can offer. Given that this work focuses on lower bounds for ACOPF, we are particularly interested in dual infeasibility of primal-dual solutions provided by the solvers (the primal being a convex relaxation of ACOPF). In Section 2.8.3 we report on dual infeasibilities attained by solvers run on our linearly constrained models.

2.8 Computational Experiments

We ran all of our experiments on an Intel(R) Xeon(R) 64-bit Linux machine, with 2 E5-2687W v3 3.10GHz CPUs, 20 physical cores, 40 logical processors, and 256 GB RAM. We used three commercial solvers: Gurobi version 10.0.1 [89], Artelys Knitro versions 13.2.0 and 14.1.0 beta [78], and Mosek 10.0.43 [95]. As a common interface for all SOCP and ACOPF models we used AMPL [96] via Python. We note that unlike Gurobi and Knitro, Mosek does not accept a constraint of the form $x^2 + y^2 \leq z^2$ or $x^2 + y^2 \leq wz$ as a conic constraint; thus we reformulated the SOCPs using a format that Mosek-AMPL was able to read.

We report extensive numerical experiments on instances from the following data sets: medium and large-sized standard IEEE instances available from MATPOWER [82], the Pan European Grid Advanced Simulation and State Estimation (PEGASE) project [97] [98], ACTIVSg synthetic cases developed as part of the US ARPA-E GRID DATA research project [99], [100], and the largest instances from the Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms [101] (PGLIB). Our main focus are cases with 9000 buses or more.

Our cutting-plane algorithm is implemented in Python 3 and uses Gurobi 10.0.1 to solve LPs and convex QPs. All of our reported single-period experiments were obtained with the following parameter setup (Section 2.6.2): barrier convergence tolerance and absolute feasibility and optimality tolerances equal to 10^{-6}, $\epsilon = 10^{-5}$, $p_{Jabr} = 0.55$, $p_{i2} = 0.15$, $p_{limit} = 1$, $T_{age} = 5$, $\epsilon_{par} = 10^{-5}/2$, $\epsilon_{ftol} = 10^{-5}$, and $T_{ftol} = 5$. For our multi-period experiments we used the numerically better-behaved i2+(ρ) relaxation with $\rho = 10^2$, and, as a result, we were able to increase solution accuracy: we set absolute feasibility and optimality tolerances equal to 10^{-8}, and we relaxed T_{ftol} from 5 to 3 for $T = 12$ and to 2 for $T = 24$. All of our code, AMPL model files, and solution files can be downloaded from `www.github.com/matias-vm`.

Next we describe the symbols used in our tables. The character " $-$ " denotes that the solver did not converge, while "TLim" means that the solver did not converge within our time limit. By convergence we mean that the solver *declares* to have obtained an *optimal* solution, within the previously defined tolerances. Further, "INF" means that the instance was declared infeasible by the solver, while "LOC INF", used by KNITRO, means that the solver converged to an infeasible point. Moreover, if Gurobi declares *numerical trouble* while solving our LPs or convex QPs at some iteration of our cutting-plane algorithm, we report the objective value of the previous iteration followed by the character " $*$ ". The objective values and running times are reported with 2 decimal places.

We remark that, to the best of our knowledge, this is the first computational study which compares the performance of three leading commercial solvers on the Jabr SOCP using a common framework (AMPL). We evaluate the *solvers* on the Jabr SOCP, and compare their performance with that of our warm-started i2+ formulations. When running the solvers, we used the Jabr rather than the i2 SOCP because the former is numerically better behaved from the solvers' perspective (Section 2.5.5). We report on these numerical issues in Section 2.8.1; also compare tables 2.11 and 2.12.

2.8.1 Single-Period Experiments

SOCPs: Solver Parameters We set a time limit of $1,000$ seconds for all of our SOCP experiments[15]. Next we describe additional parameter specifications for each solver.

- We use Gurobi's default homogeneous self-dual embedding interior-point algorithm (barrier method without *Crossover*, and *Bar Homogeneous* set to 1), and we set the parameter *Numeric Focus* equal to 1. Barrier convergence tolerance and absolute feasibility and optimality tolerances were set to 10^{-6}. Gurobi was allowed 20 threads.

- We use Knitro's default Interior-Point/Barrier Direct Algorithm, with absolute feasibility and optimality tolerances equal to 10^{-6}. We used the HSL MA57 sparse symmetric indefinite linear solver, and the Intel Math Kernel Library (MKL) functions for Basic Linear Algebra Subroutines (BLAS), i.e., for basic vector and matrix computations. Moreover, 20 threads were used with Knitro. SOCPs were explicitly input to Knitro as convex problems. When computing primal bounds, we employed the linear solver HSL MA97 whenever Knitro under MA57 was not converging.

- We use Mosek's default homogeneous and self-dual interior-point algorithm for conic optimization. We set the relative termination tolerance, as well as primal and dual feasibility tolerances to 10^{-6}. We used 20 threads with Mosek.

Remark 2.14. We did not set a time limit for computing single-period ACOPF primal bounds.

Cut Computations

Table 2.5 summarizes outcomes on cut computations for a substantial number of instances from the libraries described above. Under "Cutting-Plane," "Objective" reports the objective of the last iteration of our algorithm, "Time" reports the total running time (in seconds) of our method; "Computed" reports the number of cuts computed throughout the entire procedure; "Added" exhibits

[15] One iteration of our algorithm requires approximately $60 - 100$ seconds.

the total number of cuts in our linearly constrained relaxation at the last round; and “Rnd" is the number of rounds or our algorithm. Finally, “Primal bound" reports the objective value of a feasible solution to ACOPF and the amount of time (in seconds) required by Knitro for this purpose.

Overall, our cut management heuristics yield very tight linearly constrained relaxations using a relatively small number of cuts - note that we could potentially add $3|\mathcal{E}|$ cuts per round (for each branch $\{k, m\}$ there are three inequalities (2.28) and (2.30) that might be violated). For instance, case ACTIVSg70k has 88207 branches and after 10 rounds of cuts we only keep 123431 out of the 350572 linear cuts computed throughout the course of our algorithm. Thus, fewer than 1.5 linear cuts per branch give us a relaxation with *optimality gap*[16] at most 0.69%[17].

We remark that for some instances the objective value of our procedure can be higher than the objective value of the Jabr SOCP since our algorithm is outer-approximating the feasible region of the i2 SOCP (Proposition 2.1).

We also note that 30000goc-sad was the only instance for which Gurobi experienced numerical difficulties while solving the linearly constrained relaxation (indicated by the character “ $*$ ” next to the objective value). In this case the reported objective value corresponds to the previous iteration – setting a more aggressive cut management heuristic, for instance decreasing T_{age} from 5 to 4, gave us numerically more stable cuts and a better bound.

Finally, we obtained our primal bounds by running Knitro with a *flat-start*, i.e., we provided as initial point voltage magnitudes set to 1 and all $\theta_{km} = 0$.

Performance on SOCPs

In Table 2.6, we observe that for the cases in which at least two solvers converge, the reported bounds for the Jabr SOCP agree on the first 2 to 3 most significant digits. These differences in bounds reflect how numerically challenging the given instances are. We remind the readers of the parameter choices that we made in order for the solvers to achieve termination – the solvers

[16]Given a primal bound of a minimization problem, we define the optimality gap of a relaxation of the given problem as $\frac{z_p - z_r}{z_p}$, where z_p denotes the objective value of the primal bound and z_r denotes the objective value of the relaxation.
[17]We assume, thoughout, that the upper bounds provided by Knitro are valid.

Table 2.5: Performance of Cutting-Plane Algorithm (Not Warm-Started)

	Cutting-Plane Algorithm					Primal bound	
Case	Objective	Time	Computed	Added	Rounds	Objective	Time
9241pegase	309221.81	378.82	135599	29875	23	315911.56	96.74
9241pegase-api	6924650.57	277.32	128316	30230	21	7068721.98	73.85
9241pegase-sad	6141202.28	386.51	113686	27273	21	6318468.57	33.92
9591goc-api	1346373.10	187.26	87812	22469	22	1570263.74	42.85
9591goc-sad	1055493.25	246.87	90153	20514	27	1167400.79	28.15
ACTIVSg10k	2476851.62	132.16	60803	18183	19	2485898.75	76.71
10000goc-api	2502026.03	147.12	73084	19666	24	2678659.51	23.46
10000goc-sad	1387303.02	114.97	58984	18528	17	1490209.66	103.06
10192epigrids-api	1849488.30	152.87	97921	24882	22	1977687.11	117.15
10192epigrids-sad	1672819.53	185.02	95740	23726	23	1720194.13	23.74
10480goc-api	2708819.18	200.48	114967	29805	21	2863484.4	38.71
10480goc-sad	2287314.69	270.38	118122	28004	24	2314712.14	27.93
13659pegase	379084.55	841.83	176962	37297	22	386108.81	1184.15
13659pegase-api	9270988.77	326.57	147479	34390	19	9385711.45	44.43
13659pegase-sad	8868216.24	301.87	130682	32662	19	9042198.49	42.08
19402goc-api	2448812.41	440.67	213564	52388	22	2583627.35	87.33
19402goc-sad	1954047.79	488.33	218291	49749	25	1983807.59	64.01
20758epigrids-api	3042956.88	464.17	189436	46124	25	3126508.30	61.39
20758epigrids-sad	2612551.03	379.36	180790	44624	24	2638200.23	58.11
24464goc-api	2560407.12	471.14	226595	57162	22	2683961.9	533.03
24464goc-sad	2605128.51	506.39	222908	55242	23	2653957.66	73.87
ACTIVSg25k	5993266.85	592.39	156285	43851	28	6017830.61	56.69
30000goc-api	1531110.84	464.16	142385	41840	24	1777930.63	134.71
30000goc-sad	1130733.51*	147.74	76546	76546	6	1317280.55	565.05
ACTIVSg70k	16326225.66	1065.76	350572	123431	13	16439499.83	240.55
78484epigrids-api	15877674.54	1007.99	556893	240576	10	16140427.68	1079.03
78484epigrids-sad	15175077.19	1062.55	501202	313587	8	15315885.86	343.45

otherwise would often fail.

As we mentioned at the beginning of this section, the i2 SOCP is numerically even more challenging for the solvers than the Jabr SOCP. Indeed, in Table 2.7 we can see that the solvers do struggle (see also Table 2.12). We studied in detail some cases where Gurobi AMPL declared optimality –for example case ACTIVSg70k– while reporting variable bound max scaled violation equal to 8.43 as well as large primal and dual residuals (0.0128 and 3.25, respectively). Moreover, we noticed inconsistent termination status for cases 10192epigrids-sad, 10480goc-api, 20758epigrids-sad, and 30000goc-sad on Gurobi and Gurobi through AMPL (Gurobi-AMPL) using the same model; Gurobi AMPL declares optimal termination for these instances while Gurobi does not. Cases for which we were able to identify *low quality* solutions or inconsistencies have been denoted with the character “†" next to their reported objective value in Table 2.7.

2.8.2 Warm-Started Single-Period Instances

As described in the introduction, in power engineering practice it is of interest to address problem instances where a limited change in data occurs after solving a previous problem. In the context of this paper, we can therefore assume that we have a *warm-started formulation*, i.e., one where we leverage previously computed cuts. Here we present this warm-starting feature of our algorithm; we justify its validity and show via numerical experiments its appealing lower bounding capabilities.

We note that the convex inequalities (2.28), which we use to develop cuts, do not depend on input data such as loads or operational limits. Any such cut remains valid and can be used if the associated branch remains operational. This will be our strategy, below.

We created two kinds of perturbed instances: **a)** Instances were the load of each bus was perturbed by a Gaussian $(\mu, \sigma) = (0.01 \cdot P_d, 0.01 \cdot P_d)$, where P_d denotes the original load, subject to the newly perturbed load being non negative; and **b)** instances were the transmission line which carries the largest amount of active power in an ACOPF solution is turned off. With regards to cases of type a), we focus on relatively small individual load changes to model re-optimization after

Table 2.6: Solvers' Performance on Jabr SOCP

	Objective			Time		
Case	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek
9241pegase	-	309234.16	-	82.11	34.68	31.11
9241pegase-api	-	6840612.84	-	116.32	23.39	72.29
9241pegase-sad	-	6083747.85	-	111.05	26.01	75.99
9591goc-api	1346480.71	1348107.89	1345869.72	38.25	23.74	36.60
9591goc-sad	1055698.54	1058606.56	1054379.58	49.29	32.83	37.61
ACTIVSg10k	-	2468172.93	2466666.10	40.18	21.48	26.08
10000goc-api	-	2507034.94	2498948.00	48.63	35.19	30.13
10000goc-sad	1387288.49	1388679.63	1386041.07	23.58	26.27	23.68
10192epigrids-api	-	1849684.14	1848873.47	75.82	42.69	29.09
10192epigrids-sad	-	1672989.96	1672534.72	83.85	28.33	28.63
10480goc-api	-	2708973.58	2707828.26	75.94	27.21	56.82
10480goc-sad	-	2286454.3	2285547.23	149.93	38.17	59.48
13659pegase	379135.73	379144.11	-	33.61	43.26	34.92
13659pegase-api	-	9198542.14	-	162.21	30.64	105.11
13659pegase-sad	8826902.31	8826958.23	8787429.86	83.75	31.84	108.74
19402goc-api	-	2449020.25	2447799.72	158.12	152.89	103.04
19402goc-sad	-	1954331.70	1952550.06	203.56	155.89	104.88
20758epigrids-api	-	-	3040421.02	143.99	TLim	93.46
20758epigrids-sad	-	-	2610196.94	98.30	TLim	75.88
24464goc-api	2548335.96	-	2558631.63	603.95	TLim	129.90
24464goc-sad	-	-	2603525.46	333.50	TLim	128.50
ACTIVSg25k	5956787.54	5964417.54	5955368.56	169.66	87.14	87.18
30000goc-api	-	1531256.65	1529197.81	207.60	118.80	123.38
30000goc-sad	-	-	1130868.71	191.22	TLim	84.90
ACTIVSg70k	-	16221577.73	16217263.66	553.26	320.98	232.47
78484epigrids-api	-	-	-	756.00	TLim	637.48
78484epigrids-sad	15180775.21	-	15169401.54	463.17	TLim	601.04

Table 2.7: Solvers' Performance on i2 SOCP

	Objective			Time		
Case	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek
10192epigrids-api	1849683.44	1849684.2	-	37.44	19.4	30.74
10192epigrids-sad	1672998.72†	1672998.73	-	23.36	21.48	24.53
10480goc-api	2709110.52†	2709110.71	-	41.44	27.78	35.28
10480goc-sad	2287736.73	2287715.33	-	41.11	28.46	28.48
13659pegase	379142.67	-	-	52.12	TLim	36.49
13659pegase-api	9287242.7	9287244.72	-	66.39	236.65	30.01
13659pegase-sad	8878803.69	-	-	63.11	TLim	30.48
19402goc-api	2449100.15†	2449102.05	-	79.38	55.6	54.18
19402goc-sad	1954367.11†	1954367.2	-	180.97	59.46	82.55
20758epigrids-api	-	3043275.95	-	79.03	64.17	56.02
20758epigrids-sad	2612841.71†	2612841.8	-	48.32	84.07	58.87
24464goc-api	2560829.65†	-	-	132.34	TLim	88.79
24464goc-sad	2605532.65†	-	-	74.43	916.1	65.41
ACTIVSg25k	5994727.45	-	-	70.61	TLim	52.38
30000goc-api	1531320.78	1531322.2	-	96.91	593.73	71.38
30000goc-sad	1132242.88†	1132256.94	-	78.0	325.11	74.61
ACTIVSg70k	16333807.38†	-	-	300.98	TLim	209.3
78484epigrids-api	15882668.49	15882668.46	15882654.42	216.15	315.31	203.81
78484epigrids-sad	15180792.15	15180792.0	15180763.6	250.43	376.82	222.17

a relatively small time change. In the multi-period setting, below, we consider much larger load changes, to capture hourly models. Cases of type b) do change the structure of the network and in their warm-start we removed any cuts associated with the switched-off branch.

Tables 2.8 and 2.9 summarize our warm-started experiments on perturbed instances from our data set in Table 2.5 and compare our algorithm and solvers' performance on the Jabr SOCP. “First Round" reports the objective value and running time of the relaxation M_0 loaded with the cuts computed in Table 2.5, i.e., our warm-started relaxation. Moreover, under “Cutting-plane", we report on the objective value (at termination) and the total running time of our cutting-plane procedure (on the warm-started relaxation). “Jabr SOCP" and “Primal bound" report, respectively, on the objective value and running time of the Jabr SOCP for the three solvers, and ACOPF primal solutions.

We stress that after just *one* round, the bound provided by our algorithm is already quite good – and also point out the comparison between the running time for our first round, and the solvers' running time.

Loads Perturbed – Gaussian Deltas. For most of these cases, our procedure proves very tight lower bounds in less than 25 seconds (“First Round" column). Judging by the time required by Knitro and by the number of cases in which the solvers converge running the SOCPs, it appears that these instances are, overall, more challenging than their unperturbed counterparts.

Our procedure also stands out in quickly lower bounding the largest cases. For instance, a very sharp bound for case ACTIVSg70k is obtained in 102.25 seconds, taking less than half of the time it takes the fastest SOCP solver to converge. Similar performance is achieved on the largest epigrids cases where our method is 3 to 5 times faster.

An interesting empirical fact is that our cuts are robust with respect to load perturbations. Indeed, our evidence shows that there is not a considerable improvement from the “First Round" to the “Last Round" objectives. Thus, the previously pre-computed cuts loaded to M_0 in the first iteration are, already, accurately outer-approximating the SOCP relaxations.

Our linearly constrained relaxations are able to prove infeasibility for the modified case 9241pegase-api in 23.10 seconds while none of the three solvers were able to provide a certificate of infeasibility for the Jabr SOCP. Knitro required 1845.42 seconds to declare convergence to a locally infeasible solution. Similar results are obtained for case 24464ogc-api.

The only case where our method fails to provide a valid lower bound is case 30000goc-sad – our minimization oracle reports numerical trouble and fails to provide a solution to our warm-started relaxation. This is not surprising since difficult numerical behavior was noticed when computing cuts for this case.

Transmission Line with Largest Flow Switched Off. Overall, our method achieves a similar performance on this set of perturbed instances as in a); sharp lower bounds are obtained in about 25 seconds for most of the cases.

For this data set, our method and all of the SOCP solvers are able to prove infeasibility relatively quickly. On the other hand, our method proves a lower bound for ACTIVSg70k relatively quickly in the first round, but fails to converge in the next round due to numerical trouble caused by the newly added cuts.

As when perturbing loads, our warm-started formulation achieves a good performance on the largest epigrid cases – bounds are sharp with respect to the SOCP relaxations and it is at least 3x faster.

2.8.3 Multi-Period Results

Data At present, there is a dearth of publicly-available, realistic multi-period AC load and generation data. One partial source is the set of cases developed for the GO3 Competition [39][18]. Another partial source is the data set from the optimization package “UnitCommitment.jl" for the SCUC problem [102]; however, the largest instance contains less than 14,000 nodes, which is not among the largest publicly available instances. We are also aware that creating *good* ACOPF data

[18]The GO3 Competition does not provide explicit load data, since a welfare maximization problem needs to be solved to determine the active and reactive power consumption. We purposely avoided this dependence on the solution of another challenging optimization problem.

Table 2.8: Warm-Started Relaxations, Loads perturbed Gaussian $(\mu, \sigma) = (0.01 \cdot P_d, 0.01 \cdot P_d)$

	Cutting-Plane				Jabr SOCP						Primal bound	
	First Round				Objective			Time				
Case	Objective	Time	Objective	Time	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek	Objective	Time
9241pegase	309288.32	13.78	309299.97	160.28	-	309302.67	-	73.12	32.21	36.04	315979.53	101.48
9241pegase-api	INF	23.10	INF	23.10	-	-	-	134.53	TLim	72.96	LOC INF	1845.92
9241pegase-sad	6153913.91	16.18	6154117.59	136.78	-	6096743.03	-	97.51	26.07	83.43	6333763.92	43.71
9591goc-api	1343642.47	11.06	1343670.62	56.36	1343767.43	1345384.57	1343190.29	39.36	25.36	35.30	1571582.59	54.16
9591goc-sad	1058124.48	12.62	1058157.44	65.37	1058337.76	1061275.83	1057323.31	51.85	34.04	37.52	1178895.53	29.53
ACTIVSg10k	2475041.43	9.52	2475078.69	50.51	-	2466383.20	-	42.31	21.75	29.33	2484093.15	57.24
10000goc-api	2502049.28	8.51	2502098.01	36.48	2501946.30	2507074.78	2499373.75	31.91	43.44	32.33	LOC INF	1677.21
10000goc-sad	1388833.86	8.70	1388859.09	44.50	1388824.91	1390230.41	1387588.17	25.96	29.31	23.67	1493481.44	93.72
10192epigrids-api	1848085.36	10.27	1848133.48	45.84	-	1848285.26	1847120.93	65.38	41.17	25.99	LOC INF	1458.35
10192epigrids-sad	1672358.89	10.33	1672398.61	53.37	-	1672533.02	1671364.67	73.64	28.61	35.66	1717429.36	23.89
10480goc-api	2704157.29	12.43	2704252.95	58.45	-	2704373.73	2703432.85	197.17	27.57	55.92	2868495.28	36.89
10480goc-sad	2294908.37	12.81	2294990.69	70.93	-	2294080.35	2292830.56	185.22	35.90	58.31	2322198.81	27.34
13659pegase	379742.62	60.74	379794.51	426.88	379799.37	379804.43	-	34.21	43.17	32.75	386765.25	370.23
13659pegase-api	9253539.07	21.25	9253773.43	109.20	9181205.93	9181269.20	-	97.11	30.41	118.31	9368277.57	62.20
13659pegase-sad	8865733.59	21.28	8865892.49	113.04	8824442.20	8824486.03	-	86.49	33.19	102.59	9039904.52	40.02
19402goc-api	2452185.69	23.55	2452270.83	120.10	-	2452448.33	2451708.50	146.87	120.39	103.32	LOC INF	4440.99
19402goc-sad	1956255.19	23.28	1956313.91	113.89	-	1956570.60	1955018.07	231.90	172.82	102.19	1986936.95	66.02
20758epigrids-api	3043006.76	22.34	3043076.56	104.06	-	-	3032919.24	134.60	TLim	78.32	LOC INF	12425.89
20758epigrids-sad	2610197.53	20.46	2610261.88	93.09	-	-	2608090.26	143.69	TLim	72.19	2635892.81	49.25
24464goc-api	2561680.14	26.28	INF	50.38	-	LOC INF	-	223.07	573.37	118.6	-	19444.54
24464goc-sad	2606391.76	26.78	2606473.78	133.54	-	-	2604708.86	423.12	TLim	128.84	2655942.01	72.48
ACTIVSg25k	5988886.18	28.24	5989016.75	198.58	5952404.50	5960068.30	5949381.04	138.01	73.75	109.39	6013477.05	57.87
30000goc-api	1527412.96	25.35	1527487.45	151.75	-	1528338.73	1525625.64	243.61	369.83	119.92	LOC INF	3407.47
30000goc-sad	-	46.33	-	46.33	-	-	1132715.53	257.94	TLim	75.20	1318389.55	620.27
ACTIVSg70k	16316572.42	102.25	16317886.35	536.51	-	16210682.53	16206290.43	498.80	309.56	229.07	16428367.50	243.84
78484epigrids-api	15862318.24	115.76	15865624.98	883.93	-	-	15859950.52	757.64	TLim	642.24	-	8113.53
78484epigrids-sad	15176866.00	151.77	15180592.27	1118.02	15182602.75	-	15174716.43	420.56	TLim	589.46	15316872.94	353.13

Table 2.9: Warm-Started Relaxations, Transmission Line with largest flow turned off

	Cutting-Plane				Jabr SOCP							
	First Round				Objective			Time			Primal bound	
Case	Objective	Time	Objective	Time	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek	Objective	Time
9241pegase	INF	10.86	INF	10.86	INF	INF	INF	8.36	7.55	10.22	INF	8.37
9241pegase-api	INF	7.55	INF	7.55	INF	INF	INF	7.92	8.02	10.22	INF	8.23
9241pegase-sad	INF	7.33	INF	7.33	INF	INF	INF	8.14	8.10	10.31	INF	8.46
9591goc-api	1346470.95	10.42	1346859.06	60.76	1346969.75	1348591.44	1346437.99	39.30	17.98	36.89	1395829.51	28.08
9591goc-sad	1055823.53	11.51	1056267.57	101.64	1056447.48	1059382.10	1055501.31	45.09	35.41	37.18	1199276.44	29.90
ACTIVSg10k	2477043.05	9.94	2477537.79	75.85	-	2468821.96	2466981.35	44.81	21.60	17.52	LOC INF	7092.45
10000goc-api	2506671.15	8.06	2509971.69	46.10	2509846.00	2514991.16	2506236.75	31.03	33.95	32.15	2692320.35	23.28
10000goc-sad	1387382.65	8.76	1387515.89	66.14	1387480.33	1388870.68	1386283.75	26.53	34.24	24.14	1506187.88	108.19
10192epigrids-api	1849901.82	9.47	1850621.81	68.73	-	1850788.76	1849821.44	69.81	38.01	25.30	2021493.05	117.18
10192epigrids-sad	1673575.50	11.08	1674274.99	74.49	-	1674417.21	1673564.57	69.91	43.91	28.54	1734014.50	24.11
10480goc-api	2710040.46	11.33	2711100.23	73.85	-	2711224.27	2710520.15	95.40	27.99	56.58	2862699.50	225.06
10480goc-sad	2288069.64	13.47	2288969.47	98.88	-	2288069.23	2286864.08	106.54	37.65	59.43	2318279.76	26.13
13659pegase	379102.13	53.29	379163.58	199.96	379177.99	379182.14	-	33.18	345.83	31.90	386126.93	394.93
13659pegase-api	INF	10.81	INF	10.81	INF	INF	INF	10.94	8.90	13.79	INF	11.55
13659pegase-sad	INF	10.63	INF	10.63	INF	INF	INF	11.02	10.80	13.76	INF	11.73
19402goc-api	2450110.09	23.93	2451621.60	171.31	-	2451793.39	2450488.01	154.42	132.03	104.58	2587915.50	403.20
19402goc-sad	1954365.39	23.70	1954881.06	191.52	-	1955116.35	1953676.05	258.93	156.10	102.63	1985954.83	63.38
20758epigrids-api	3043482.21	20.44	3044690.46	133.92	-	-	3041974.74	112.78	TLim	96.19	3132571.31	52.82
20758epigrids-sad	2612646.70	20.62	2612786.37	115.89	-	-	2610315.41	169.67	TLim	73.02	2638560.64	47.87
24464goc-api	2560669.11	25.94	2561110.03	161.80	2550118.22	-	2559240.55	440.81	TLim	118.05	2684708.93	1663.63
24464goc-sad	2605179.75	26.98	2605369.03	166.81	-	2605474.23	2603609.34	564.27	74.69	124.19	2654344.45	76.39
ACTIVSg25k	6045885.88	27.52	6048122.86	238.67	6009656.52	6018875.03	6009500.57	144.28	65.42	81.41	LOC INF	1634.38
30000goc-api	1531110.55	25.02	1531159.95	130.30	-	1532013.41	1529195.12	195.74	135.71	119.77	LOC INF	3203.26
30000goc-sad	-	45.60	-	45.60	-	-	1130917.78	218.12	TLim	74.27	1324622.71	186.43
ACTIVSg70k	16426522.74	98.51	16426522.74*	98.51	-	-	-	150.77	TLim	129.60	LOC INF	3160.81
78484epigrids-api	15888353.48	104.32	15892229.11	<916.88	15894055.55	-	15880422.78	322.22	TLim	625.74	16169740.92	2328.79
78484epigrids-sad	15179882.22	149.65	15185980.69	1151.33	15188085.99	-	15182701.65	437.59	TLim	594.84	15330674.69	272.41

is a non-trivial task [103], [99], but given that there were no publicly available multi-period load time-series to match the standard largest AC cases we developed our own data sets[19].

We created and worked with three data sets: (1) $T = 4$ with increasing[20] (on expectation) active power loads, to be precise, at each time-period $t \in \{1, 2, 3\}$, $P^d_{k,t}$ is set to $(1 + 0.01 \cdot t + u_{k,t}) \cdot P^d_k$, where $u_{k,t}$ is a uniformly distributed random variable with range $[0, 0.025]$; $T = 12$ where a half-a-day active power load curved is simulated, see Figure 2.7; and (3) $T = 24$ where a 24 hr active power load curve is simulated, see Figure 2.8. Moreover, we assumed 50% generator ramp up r^u and down r^d rates for consecutive time-periods (see Section 2.3.1) in all of the three settings.

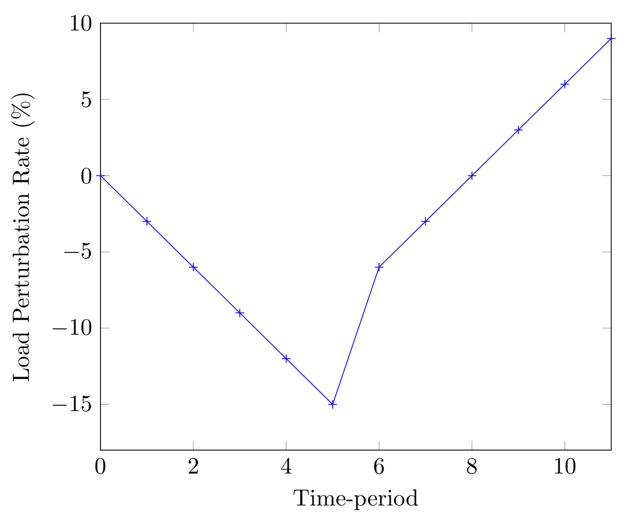

Figure 2.7: 12hr load curve.

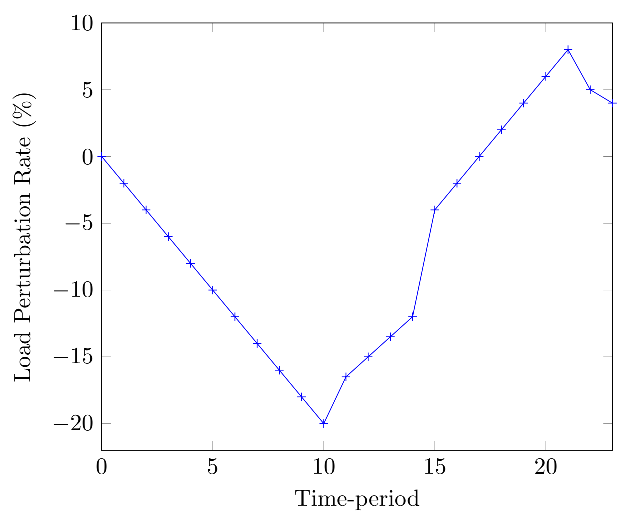

Figure 2.8: 24hr load curve.

Upper Bound Heuristics In order to evaluate how tight a given lower bound is, an upper bound is needed. We relied on two simple heuristics to compute upper bounds. The first heuristic, which we term H1, consists in providing to Knitro the complete multi-period formulation. The parameter setup for H1 is: absolute feasibility tolerance equal to 10^{-6}, while we set relative optimality tolerance to 10^{-3}; simple variables bounds are enforced throughout the optimization; and $T_{ftol} = 3$; and time limit of 2400 seconds for $T = 4$, and $7,200$ seconds for $T = 12, 24$. If H1 fails, then we

[19]These data sets can be downloaded from www.github.com/matias-vm.

[20]Given that the PGLIB library [101] contains congested instances, we applied decreasing (on expectation) active power loads perturbations to the PGLIB instances.

proceed with heuristic H2 which solves one single-period ACOPF instance at a time, and upon convergence, fixes the upper and lower limits of active power generation of the next instance as a function of generator outputs in the current period. To be precise, for every generator $k \in \mathcal{G}$ and any time-period $t \in \{0, \ldots, T-2\}$, and provided that the nonlinear solver converges for the single-period instance at time t, then the output of generator k for the time-period $t+1$ is bounded by $P^{min}_{k,t+1} \leq P^g_{k,t+1} \leq P^{max}_{k,t+1}$ where $P^{min}_{k,t+1} := (1 - r^d_{k,t+1})P^g_{k,t}$ and $P^{max}_{k,t+1} := (1 + r^u_{k,t+1})P^g_{k,t}$. If the solver fails to find a solution for some time-period, then we declare failure of H2. Overall, H2 succeeded in most of the instances for which we were able to obtain a primal bound. The parameter setup for H2 was the same as for H1 except that we set the time limit for single-period ACOPFs to $1,200$ seconds.

SOCPs: Solver Parameters for Multi-period Instances With respect to our SOCP experiments, we set a time limit of $2,400$ seconds for $T = 4$ instances, $4,800$ seconds for $T = 12$, while we set $8,000$ seconds for $T = 24$ instances. In what follows we describe the parameter specifications for each solver for the SOCP runs.

- We use Gurobi's default homogeneous self-dual embedding interior-point algorithm (barrier method without *Crossover*, and *Bar Homogeneous* set to 1), and we set the parameter *Numeric Focus* equal to 1. Barrier convergence tolerance and absolute feasibility and optimality tolerances were set to 10^{-6}. Gurobi was allowed 20 threads.

- We use Knitro's default Interior-Point/Barrier Direct Algorithm, with absolute feasibility and optimality tolerances equal to 10^{-6}. We used the Intel MKL PARDISO sparse symmetric indefinite solver, and the Intel Math Kernel Library (MKL) functions for Basic Linear Algebra Subroutines (BLAS), i.e., for basic vector and matrix computations. Moreover, we gave Knitro 20 threads to use for parallel computing features (10 threads were given to the linear solver). When solving the SOCPs, we explicitly instructed Knitro to handle the problem as convex, and to enforce simple variable bounds throughout the optimization.

- We use Mosek's default homogeneous and self-dual interior-point algorithm for conic optimization. We set the relative termination tolerance, as well as primal and dual absolute feasibility tolerances to 10^{-6}. We assigned 20 threads to Mosek.

We recall the reader that we use the i2+(ρ) relaxation, with $\rho = 100$, as target set for our cutting-plane multi-period experiments. For our cutting-plane algorithm, when $T = 4$ a $1,200$ seconds time limit was enforced prior to *starting* a new round of cutting; the time limit was $3,600$ seconds for $T = 12$, and $6,000$ seconds for $T = 24$.

Not Warm-Started Cutting-Plane Algorithm: T = 4

Table 2.10 summarizes our not warm-started cutting-plane experiments on multi-period instances with $T = 4$. Columns “# Vars", “# Cons", and “FTime" denote the number of variables, number of constraints and formulation time of the base model M_0^T. Under “First Iteration" we report on the objective value, running time, and maximum dual infeasibility error (see Section 2.7) of the relaxation M_0^T, i.e., without any cuts. On the other hand, “Last Iteration" presents the objective value, running time, total number of cuts held by the model M_{last}^T on the last iteration of the algorithm, and maximum dual infeasibility error. “Total" presents the total running time of our algorithm and the total number of rounds of cuts, and “ACOPF" describes the number of variables, number of constraints, and the objective value and running time of each ACOPF primal solution.

Overall, we can see that lower bounds are fairly tight; even our base model M_0^T without any cuts, whose objective value is shown under “First Iteration", attains quick and sharp lower bounds for these multi-period instances. For instance, an optimality gap of 6.91% is achieved by M_0^T for case ACTIVSg70k, and, after three rounds of cuts, nearly 800,000 cuts drive the gap down to 1.75%.

Our cut management heuristics prove to be empirically successful: dual infeasibility reported by Gurobi is either maintained or slightly worsened (compare “DInfs" of “First Iteration" versus that of “Last Iteration") after adding a significant number of cuts. This means that our cuts are keeping our linearly constrained relaxation numerically stable and that our reported lower bounds are reasonably accurate (Section 2.7). Moreover, we observe that instances which require a significant number of rounds of cuts, say 10 or more, under the given time limit, notoriously benefit from our cut clean-up heuristic. We see that a small number of cuts (roughly 10% of the total number of constraints) is good enough to outer-approximate the rotated-cones and attain sharp ACOPF lower bounds (as in the single-period case; Section 2.8.1).

Warm-Started Cutting-Plane Algorithm: T = 4, 12, 24

We use a simple methodology to warm-start the algorithm described in Section 2.6.2: we propagate cuts for the original instances to every time-period. To be precise, we create a pool that includes, for each branch $\{k, m\} \in \mathcal{E}$, all the associated single-period cuts – which are assumed to have been previously computed (Section 2.8.1). The cuts in the pool are propagated to every time-period $t \in \{0, \ldots, T-1\}$ and are added to the base model M_0^T.

In Tables 2.11 and 2.12 we present the results of our warm-started experiments with $T = 4$. In Table 2.11 we compare our lower bounds (obtained by warm-starting our base model with precomputed cuts) against the lower bounds obtained by solving the Jabr SOCPs with each of the three commercial solvers. In Tables 2.12, A.1, and A.2 we compare our bounds against the performance of the solvers on the i2+(ρ) relaxation plus Jabr inequalities for every branch whose i2 inequality is bad; this relaxation is numerically better behaved than the i2 SOCP (see discussion in Section 2.5.5).

In both tables the columns "# Vars", "# Cons", and "FTime" denote the number of variables, number of constraints and formulation time of the base model M_0^T (before adding the previously computed cuts). 'First Iteration" reports on the objective value, running time, the number of cuts with which the instance was warm-started, and the maximum dual infeasibility error (Section 2.7) of the warm-started relaxation M_T^0. Multi-column "Last Iteration" presents the objective value, running time, total number of cuts held by the model M_{last}^T on the last iteration of the algorithm, and maximum dual infeasibility error of model M_{last}^T. "Total" presents the total running time of our algorithm and the total number of rounds of cuts. Under the header "Jabr SOCP", columns "#Vars" and "#Cons" denote the number of variables, number of constraints of the Jabr SOCP (which are the same for the three solvers given that we used a common AMPL model file), and "Obj" and "Time"

Table 2.10: Cutting-Plane Not Warm-Started, $T = 4$ and loads perturbed $\text{Unif}(0, 0.025) + 0.01 \cdot t$ for $t \in \{1, 2, 3\}$, ramping rates 50%

	Cutting-Plane												ACOPF			
				First Iteration			Last Iteration				Total				PBound	
Case	#Vars	#Cons	FTime	Obj	Time	DInfs	Obj	Time	#Cuts	DInfs	Time	Rnds	#Vars	#Cons	Obj	Time
1354pegase	66897	72564	1.35	299460.87	2.33	5.80e-10	303439.90	2.31	11043	3.32e-09	37.32	12	53772	70220	303727.82	7.60
2869pegase	150431	161568	3.31	543415.33	6.16	1.23e-07	549381.91	7.83	26938	6.00e-10	154.55	14	120712	158388	549988.85	92.99
3375wp	148217	159692	3.18	29810390.36	6.79	5.13e-08	30318794.35	10.92	16457	1.67e-07	171.70	16	117364	151844	30410712.08	63.05
6468rte	319517	323396	6.57	349845.41	13.77	1.29e-04	355995.14	11.83	46902	2.72e-06	354.29	15	247296	321888	356969.89	2596.76
9241pegase	524028	538308	10.93	1250007.83	27.06	1.92e-07	1269609.76	258.55	318602	5.63e-05	1403.50	7	412248	543530	1297205.96	172.18
ACTIVSg10k	456684	490360	9.98	10035162.94	24.82	7.07e-09	10222977.88	27.19	52082	2.10e-08	399.43	13	363940	470558	10265810.55	228.28
10000goc-api	465676	495636	10.04	8916264.44	24.76	1.50e-09	9495292.75	20.41	61212	5.28e-07	492.16	19	368928	478650	10019708.68	1049.69
10000goc-sad	465676	495636	10.03	5373568.66	28.48	8.39e-07	5456260.77	29.84	61388	1.12e-06	403.37	11	368928	478650	5866821.40	241.57
13659pegase	723245	735740	15.46	1528552.09	31.02	1.29e-08	1551553.61	422.84	469991	9.30e-07	1380.08	6	567716	739636	1584844.81	3083.66
13659pegase-api	723245	735740	15.71	34787242.54	39.17	2.84e-05	35703095.45	207.32	316521	2.24e-08	1207.94	9	567716	739636	36163923.41	1288.60
13659pegase-sad	723245	735740	15.7	33913060.97	40.32	1.46e-08	34547415.27	97.98	243868	9.98e-08	1218.18	10	567716	739636	35199898.78	545.24
20758epigrids-api	1106271	1127168	23.52	11540550.41	175.25	3.38e-08	11834277.30	373.71	453981	2.33e-06	1294.19	5	859924	1131168	12088587.29	1018.17
20758epigrids-sad	1106271	1127168	24.30	10060842.01	86.20	4.97e-08	10219196.77	306.29	410645	1.22e-07	1294.96	7	859924	1131168	10317105.84	709.85
ACTIVSg25k	1145783	1212392	24.89	24097056.99	80.20	1.93e-08	24710933.54	120.53	239603	9.12e-08	1228.25	8	902588	1170088	24835784.33	451.45
30000goc-api	1253699	1366264	27.91	4916231.92	73.67	5.00e-07	5705500.71	82.73	112230	1.31e-06	1238.19	12	990172	1280368	6563873.25	1669.13
30000goc-sad	1253699	1366264	27.62	4151693.26	87.16	7.53e-06	4340911.74	84.31	104180	7.58e-07	1213.47	11	990172	1280368	-	3666.85
ACTIVSg70k	3115935	3316972	68.93	65302843.92	269.38	1.85e-08	68917427.01	389.02	799462	2.03e-08	1268.46	3	2448820	3175256	70148300.24	2034.37
78484epigrids-api	4179320	4225032	89.55	60235793.42	347.62	1.49e-04	61087514.34	677.62	1496304	1.51e-04	1816.21	3	3230648	4252514	62011780.96	4831.70
78484epigrids-sad	4179320	4225032	89.43	57816998.73	387.48	1.23e-04	58689459.71	571.38	1081287	1.18e-04	1361.42	2	3230648	4252514	-	-

show the objective value and running time of the Jabr SOCP for each solver. Finally, under column “ACOPF" we report objective value of an ACOPF primal solution.

As in the single-period experiments (Tables 2.8 and 2.9), “First Iteration" shows that our warm-started relaxation, i.e., model M_0^T loaded with the propagated precomputed cuts, attains sharp bounds fairly quickly. A notable example, in which at least 2 solvers converge in Table 2.11, is case ACTIVSg25k for which Gurobi solves the convex QP in 94 seconds resulting in 2.1% optimality gap. Moreover, after 9 rounds of cuts our algorithm is able to drive the down the optimality gap to 0.5% at very high accuracy (Figure 2.10).

Dual infeasibility in Table 2.11, as in the not warm-started experiments, is not significantly worsened by our cuts, i.e., they do not seem to ill-condition our instances. On the contrary, we see that for some instances, our cut management heuristics are able to improve on this metric. In Figure 2.9 we illustrates the evolution of the number of outer-envelope cuts with respect to iterations of our cutting-plane algorithm for two $T = 4$ warm-started cases. For instance, ACTIVSg10k was warm-started with 28,184 cuts, and peaked 111,323 cuts at iteration 6 which corresponds to the chosen threshold T_{age} (Section 2.6.2). From the 7th iteration onward, there is a steady cut-cleansing ending up with $51,188$ cuts. As expected, both instances end up with more cuts than were present when warm-started.

With respect to the solvers' performance on the SOCPs, we see that all of the solvers struggle. In Table 2.11 we see differences up to the second digit in the objective values of instances on which at least 2 solvers converge, which corroborates that these are numerically challenging instances. Table 2.12 exhibits even more striking results; here we ran the solvers on the i2+(ρ) relaxation (the relaxation that we are outer-approximating with our cutting-plane algorithm).

In Appendix A.5, we present tables A.1 and A.2 which show computational results for our $T = 12$ and $T = 24$ instances. We see that Gurobi is able to robustly handle our very large LPs and convex QPs, such as case ACTIVSg70k, with almost 19 million variables and 20 million constraints. For this instance, we load 1.5 million precomputed cuts, and after two rounds of cuts we end up with 3.35 million cuts, and the solver outputs a lower bound with very small dual infeasibilities.

For these large instances, we were only able to compute an AC primal bound for case 1354pegase $T = 12$; our heuristics failed in all the other instances with $T = 12$ and $T = 24$.

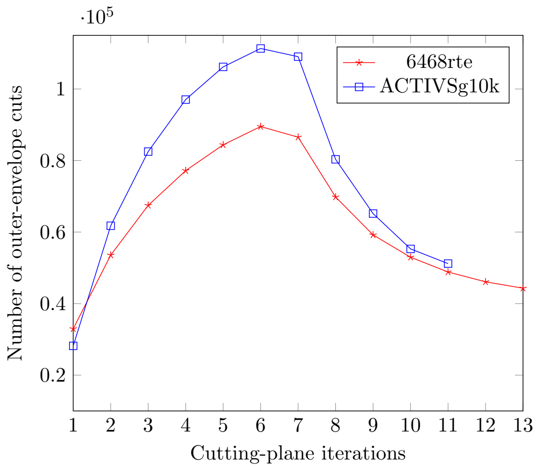

Figure 2.9: Cut management for $T = 4$ warm-started instances.

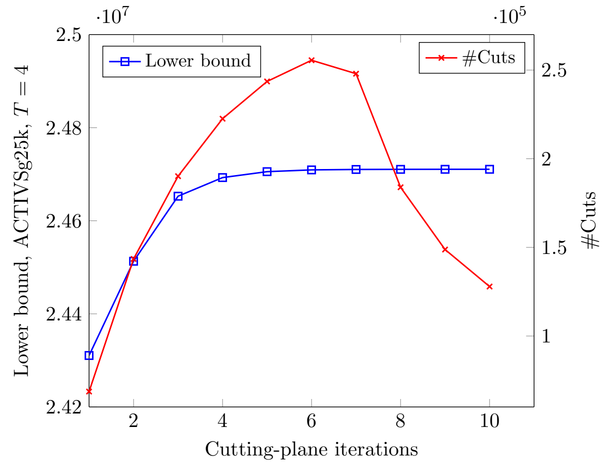

Figure 2.10: Lower bounds for ACTIVSg25k and number of cuts.

On tables 2.13 and 2.14 we compare our linearly constrained warm-started relaxation against multi-period DCOPF[21]. We report on the number of variables and constraints for each instance, as well as the objective, total running time, and dual infeasibility[22]. We remind the reader that DCOPF is a linearly constrained approximation of ACOPF, not a relaxation [49]; hence, DCOPF infeasibility does not necessarily imply that the associated ACOPF is infeasible, in sharp contrast to what any relaxation for ACOPF guarantees.

Our ACOPF relaxations quadruple the size (in terms of number of variables and constraints) of the corresponding multi-period DCOPF instances. Gurobi clearly runs faster on multi-period DCOPF. Interestingly, there are some DCOPF instances for which Gurobi does not converge[23] – either because of suboptimal termination or because it runs into numerical trouble – while it does converge for our ACOPF relaxations attaining small dual infeasibilities.

[21] Multi-period DCOPF is solved with Gurobi using the same parameter setup as our cutting-plane algorithm.

[22] For our relaxation, we report on the last iteration of our cutting-plane algorithm. Moreover, the number of constraints includes the total number of cuts kept during the last iteration.

[23] We remind the reader that there are alternative formulations of DCOPF. Rather than directly solving an LP or convex QP, such alternative approaches could significantly enhance computational efficiency by exploiting problem structure (deploying a delayed constrained generation algorithm where the decision variables are nodal injections).

Table 2.11: Warm-Started Relaxation versus Jabr SOCP, $T = 4$ and loads perturbed Unif $(0, 0.025) + 0.01 \cdot t$ for $t \in \{1, 2, 3\}$, ramping rates 50%

	Cutting-Plane													Jabr SOCP								ACOPF
				First Iteration				Last Iteration				Total				Obj			Time			
Case	#Vars	#Cons	FTime	Obj	Time	#Cuts	DInfs	Obj	Time	#Cuts	DInfs	Time	Rnds	#Vars	#Cons	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek	PBound
1354pegase	66897	72564	1.41	301493.85	3.42	6000	2.55e-08	303438.51	1.93	10345	2.95e-08	36.04	12	55280	66149	-	303420.78	303415.82	15.34	11.65	14.79	303727.82
3375wp	148217	159692	3.17	30132477.23	8.76	10448	1.42e-06	30318337.40	8.59	15786	4.44e-07	124.03	13	118128	142915	-	-	30233269.55	60.00	TLim	36.2	30410712.08
6468rte	319517	323396	6.81	352645.45	18.24	32960	2.68e-06	355986.22	11.76	42631	2.72e-06	205.53	13	252240	302520	356099.86	-	355981.27	66.75	TLim	57.88	356969.89
9241pegase	524028	538308	11.36	1258422.42	48.99	68472	3.61e-06	1270138.75	100.47	189019	5.18e-08	1268.90	7	433700	515921	1262979.52	-	-	182.32	TLim	89.64	1297205.96
ACTIVSg10k	456684	490360	10.27	10117471.23	30.65	28184	5.72e-09	10222978.04	23.94	51188	5.01e-08	306.72	10	364824	437972	-	9894331.31	9876639.32	243.69	357.69	98.62	10265810.55
10000goc-api	465676	495636	9.88	9129427.52	28.01	35856	6.70e-09	9495973.72	20.46	59211	3.92e-07	301.67	11	383978	467735	-	-	9479414.94	304.47	TLim	132.55	10019708.68
10000goc-sad	465676	495636	10.21	5414194.06	30.94	35224	1.29e-06	5456194.0	26.47	53520	1.31e-06	336.21	10	383978	467735	-	-	5448060.51	316.29	TLim	127.01	5866821.40
13659pegase	723245	735740	15.67	-	365.52	97036	-	-	365.52	97036	-	384.05	0	578580	686433	1488477.49	-	-	308.57	TLim	100.23	1584844.81
13659pegase-api	723245	735740	15.28	35162915.42	74.46	85584	2.42e-06	35703716.41	178.18	218790	4.06e-07	1348.20	8	594950	719171	35070776.73	-	-	717.60	TLim	479.87	36163923.41
13659pegase-sad	723245	735740	15.60	34093210.18	52.16	78612	2.89e-08	34547590.80	98.42	136983	5.16e-07	1273.65	12	594950	719171	34196027.45	-	33517709.2	523.11	TLim	466.32	35199898.78
20758epigrids-api	1106271	1127168	23.34	-	195.65	80904	5.54	11833435.09	658.47	260055	9.29e-05	1575.77	3	918078	1105639	-	-	11830133.74	1093.22	TLim	322.49	12088587.29
20758epigrids-sad	1106271	1127168	23.64	10138699.01	92.13	77228	2.76e-08	10219195.88	181.38	185780	8.84e-08	1374.00	8	918078	1105639	-	-	10211969.97	398.31	TLim	497.46	10317105.84
ACTIVSg25k	1145783	1212392	25.46	24310520.83	93.95	68676	1.01e-07	24710662.67	92.3	127912	1.61e-08	1206.14	9	912168	1099187	23293596.46	-	23640456.52	768.78	TLim	395.86	24835784.33
30000goc-api	1253699	1366264	30.24	5185673.79	97.36	59732	3.54e-07	5706550.06	80.15	105806	9.79e-07	1146.25	12	1013234	1246465	-	-	5637560.70	TLim	TLim	644.65	6563873.25
30000goc-sad	1253699	1366264	28.03	4316717.55	138.08	179704	7.84e-07	4341119.26	77.37	91071	1.06e-06	994.70	9	1013234	1246465	-	-	4325675.85	985.49	TLim	569.21	-
ACTIVSg70k	3115935	3316972	68.7	67721489.41	350.58	258228	3.04e-08	69422696.57	369.05	597271	1.11e-07	1301.25	3	2480088	3003929	-	-	65293153.18	TLim	TLim	925.00	70148300.24
78484epigrids-api	4179320	4225032	96.42	60947456.94	494.72	625780	1.53e-04	61271585.98	563.69	1047809	1.52e-04	1523.35	2	3420774	4126501	-	-	-	TLim	TLim	1109.78	62011780.96
78484epigrids-sad	4179320	4225032	97.61	58930587.37	743.6	888916	1.20e-04	59051305.24	845.85	1133203	1.20e-04	1569.46	1	3420774	4126501	59124882.16	-	-	1586.33	TLim	1103.22	-

Table 2.12: Warm-Started Relaxation versus i2 SOCP+, $T = 4$ and loads perturbed Unif $(0, 0.025) + 0.01 \cdot t$ for $t \in \{1, 2, 3\}$, ramping rates 50%

	Cutting-Plane													i2 SOCP+								ACOPF
				First Iteration				Last Iteration				Total				Obj			Time			
Case	#Vars	#Cons	FTime	Obj	Time	#Cuts	DInfs	Obj	Time	#Cuts	DInfs	Time	Rnds	#Vars	#Cons	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek	PBound
1354pegase	66897	72564	1.41	301493.85	3.42	6000	2.55e-08	303438.51	1.93	10345	2.95e-08	36.04	12	56326	84155	303424.53	-	-	41.94	TLim	7.86	303727.82
3375wp	148217	159692	3.17	30132477.23	8.76	10448	1.42e-06	30318337.40	8.59	15786	4.44e-07	124.03	13	121822	179665	-	-	-	60.00	TLim	15.28	30410712.08
6468rte	319517	323396	6.81	352645.45	18.24	32960	2.68e-06	355986.22	11.76	42631	2.72e-06	205.53	13	269070	373246	355828.72	-	-	468.73	TLim	41.89	356969.89
9241pegase	524028	538308	11.36	1258422.42	48.99	68472	3.61e-06	1270138.75	100.47	189019	5.18e-08	1268.90	7	451546	643811	-	-	-	411.55	TLim	92.33	1297205.96
ACTIVSg10k	456684	490360	10.27	10117471.23	30.65	28184	5.72e-09	10222978.04	23.94	51188	5.01e-08	306.72	10	364824	437972	-	-	-	243.69	TLim	79.58	10265810.55
10000goc-api	465676	495636	9.88	9129427.52	28.01	35856	6.70e-09	9495973.72	20.46	59211	3.92e-07	301.67	11	390042	567215	-	-	-	559.76	TLim	148.77	10019708.68
10000goc-sad	465676	495636	10.21	5414194.06	30.94	35224	1.29e-06	5456194.0	26.47	53520	1.31e-06	336.21	10	390042	567215	-	-	-	419.79	TLim	104.86	5866821.40
13659pegase	723245	735740	15.67	-	365.52	97036	-	-	365.52	97036	-	384.05	0	618066	859791	-	-	-	457.45	TLim	117.18	1584844.81
13659pegase-api	723245	735740	15.28	35162915.42	74.46	85584	2.42e-06	35703716.41	178.18	218790	4.06e-07	1348.20	8	618066	859791	-	-	-	1093.04	TLim	155.84	36163923.41
13659pegase-sad	723245	735740	15.60	34093210.18	52.16	78612	2.89e-08	34547590.80	98.42	136983	5.16e-07	1273.65	12	618066	859791	-	-	-	891.76	TLim	150.4	35199898.78
20758epigrids-api	1106271	1127168	23.34	-	195.65	80904	5.54	11833435.09	658.47	260055	9.29e-05	1575.77	3	950270	1340191	-	-	-	1529.93	TLim	356.82	12088587.29
20758epigrids-sad	1106271	1127168	23.64	10138699.01	92.13	77228	2.76e-08	10219195.88	181.38	185780	8.84e-08	1374.00	8	950270	1340191	-	-	-	1397.93	TLim	247.43	10317105.84
ACTIVSg25k	1145783	1212392	25.46	24310520.83	93.95	68676	1.01e-07	24710662.67	92.3	127912	1.61e-08	1206.14	9	961854	1387817	-	-	-	589.52	TLim	200.22	24835784.33
30000goc-api	1253699	1366264	30.24	5185673.79	97.36	59732	3.54e-07	5706550.06	80.15	105806	9.79e-07	1146.25	12	1018714	1524129	-	-	-	1761.14	TLim	286.2	6563873.25
30000goc-sad	1253699	1366264	28.03	4316717.55	138.08	179704	7.84e-07	4341119.26	77.37	91071	1.06e-06	994.70	9	1018714	1524129	-	-	-	727.32	TLim	247.6	-
ACTIVSg70k	3115935	3316972	68.7	67721489.41	350.58	258228	3.04e-08	69422696.57	369.05	597271	1.11e-07	1301.25	3	2584042	3765747	-	-	-	1312.16	TLim	405.88	70148300.24
78484epigrids-api	4179320	4225032	96.42	60947456.94	494.72	625780	1.53e-04	61271585.98	563.69	1047809	1.52e-04	1523.35	2	3558322	4997073	-	-	-	TLim	TLim	1117.23	62011780.96
78484epigrids-sad	4179320	4225032	97.61	58930587.37	743.6	888916	1.20e-04	59051305.24	845.85	1133203	1.20e-04	1569.46	1	3558322	4997073	-	-	-	TLim	TLim	1117.35	-

Table 2.13: Warm-Started Relaxation versus DCOPF, $T = 12$

	Warm-Started Relaxation					DCOPF				
Case	#Vars	#Cons	Obj	Time	DInfs	#Vars	#Cons	Obj	Time	Dinfs
1354pegase	201209	250008	860664.27	92.63	1.06e-08	62369	67828	850023.48	2.44	1.72e-10
9241pegase	1574972	1980693	3565252.49	3811.42	4.32e-04	447608	477952	3627501.88	18.06	3.76e-10
ACTIVSg10k	1375020	1639853	28690135.28	1308.80	3.38e-07	449628	501812	28187885.77	19.50	2.34e-07
10000goc-api	1401204	1686928	28459989.11	1066.39	2.40e-08	446364	490232	INF	20.15	-
10000goc-sad	1401204	1647684	16474320.51	1450.12	3.36e-06	446364	490232	16114986.72	19.96	2.21e-07
ACTIVSg25k	3447015	4472632	69534077.06	4073.97	9.75e-07	1097943	1199456	67992943.20	48.82	7.88e-07
30000goc-api	3768147	4484763	17079448.81	3667.40	1.06e-05	1225815	1299860	-	75.87	-
30000goc-sad	3768147	4373646	13190466.44	3901.76	2.45e-06	1225815	1299860	12846216.92	53.33	7.65e-08
ACTIVSg70k	9368583	11913750	188736994.5	4702.30	3.66e-08	2977455	3195644	-	175.35	-
78484epigrids-api	12551704	14607420	INF	447.39	-	3555448	3699780	INF	4.19	-
78484epigrids-sad	12551704	16185978	177335668.05	5465.67	1.36e-04	3555448	3699780	-	175.91	-

Table 2.14: Warm-Started Relaxation versus DCOPF, $T = 24$

	Warm-Started Relaxation					DCOPF				
Case	#Vars	#Cons	Obj	Time	DInfs	#Vars	#Cons	Obj	Time	Dinfs
1354pegase	402677	503633	1666609.97	180.61	2.97e-08	124997	136696	1646661.83	4.61	2.15e-10
9241pegase	3151388	4613025	6925029.02	8419.66	1.25e-03	896660	961684	7013508.62	36.02	5.25e-10
ACTIVSg10k	2752524	3290672	55184650.09	2869.25	7.18e-08	901740	1013564	54241448.25	37.77	5.44e-08
10000goc-api	2804496	3403822	53800966.46	2273.58	3.18e-07	894816	988820	INF	40.25	-
10000goc-sad	2804496	3262316	32678842.78	3099.33	7.97e-06	894816	988820	32120062.03	41.02	1.97e-06
ACTIVSg25k	6898863	8756843	134142620.00	6991.16	2.08e-06	2200719	2418248	131325272.36	95.18	5.69e-09
30000goc-api	7539819	9332691	31616476.13	6799.92	2.06e-05	2455155	2613824	-	171.42	-
30000goc-sad	7539819	9661317	25811075.96	6249.53	1.23e-05	2455155	2613824	25279287.29	108.86	2.22e-08
ACTIVSg70k	18747555	23469443	357083446.24	8496.71	2.98e-08	5965299	6432848	-	829.66	-
78484epigrids-api	25110280	24855336	-	2138.78	-	7117768	7427052	INF	380.67	-
78484epigrids-sad	25110280	30188832	344805634.00	5399.62	1.34e-04	7117768	7427052	-	513.08	-

Table 2.15: On the convergence of dual variables of active-power balance constraints $T = 4$

Case	$\|\lambda^k - \lambda^{k-1}\|_2$								
	$k = 2$	$k = 4$	$k = 6$	$k = 8$	$k = 10$	$k = 12$	$k = 14$	$k = 16$	$k = 18$
1354pegase	42.5161	9.2024	2.9110	2.0099	1.0506	0.4326	0.189	0.1038	0.0494
ACTIVSg2000	2590.985	481.1567	133.0116	67.344	38.3696	13.2228	7.9999	2.8572	2.5617
2869pegase	84.6365	14.806	5.7585	3.4687	2.4777	0.9821	0.8868	0.4770	1.8608
6468rte	232.8488	42.6396	14.7121	7.3318	4.1112	1.5908	1.2435	0.3554	0.6279
ACTIVSg10k	6394.2841	1712.8186	475.1414	104.5385	54.3611	17.6938	8.4489	7.7985	4.2688

2.8.4 Stability of Dual Variables for Active-Power Balance Constraints

In current energy pricing rules [28], [29], dual variables associated with active-power balance constraints (2.8b) of a welfare maximization problem constitute a key input for computing Locational Marginal Prices (LMPs). In general, linear models such as DCOPF are used for pricing because of their computational tractability. The trade-offs of using such approximations are well-known. As stated in [30], “If DCOPF does not adequately reflect the physics of the power grid, the prices retrieved from the DCOPF solution will not be accurate. Thus, tighter relaxations of ACOPF could lead to prices that better reflect scarcity in the physical network." Among their findings, the authors empirically show that nonlinear tighter relaxations such as the Jabr SOCP or the QC relaxation mitigate biasedness in price signals, in particular in congested networks, and lead to higher welfare.

Given the theoretical and empirical evidence presented in this paper, we believe that our linearly constrained relaxation is well-placed as a potential substitute for DCOPF in electricity market pricing. A first step in evaluating its potential is to understand whether the dual variables associated to active-power balance constraints of our dynamically built linearly constrained relaxations numerically converge to some vector of duals. In Table 2.15 and Figure 2.11 below we provide preliminary evidence of numerical convergence for medium-sized $T = 4$ instances. At each iteration $k > 1$ of our cutting-plane algorithm, we obtain the corresponding vector of duals λ^k and compute its Euclidean distance to its predecessor λ^{k-1}.

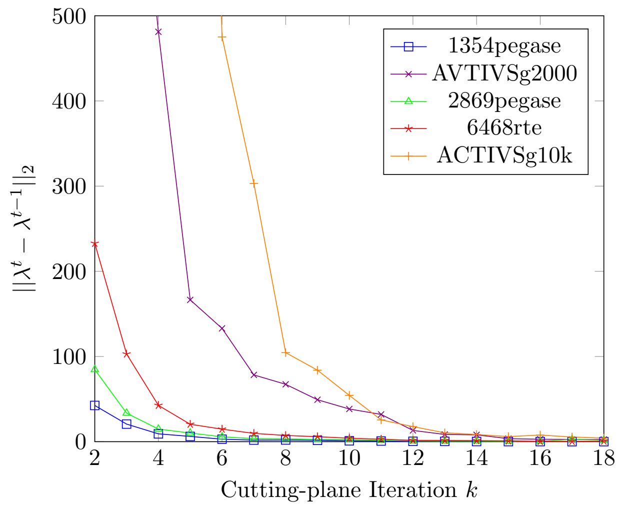

Figure 2.11: Distance between dual variables of active-power balance constraints of consecutive cutting-plane iterations.

2.9 SDP-based Strengthening of the Relaxations

Building on the cutting-plane framework introduced in Section 2.6, we propose a new approach to leverage semidefinite programming (SDP) relaxations for the Alternating-Current Optimal Power Flow (ACOPF) problem. Our goal is to combine the tightness of SDP relaxations with the scalability and numerical robustness of linear programming by constructing linear outer-approximations that avoid solving SDPs directly.

We note that in existing approaches to SDP relaxations for ACOPF, a strictly positive lower bound on transmission line resistances is typically enforced to ensure numerical convergence of the corresponding SDP implementations. In contrast, our algorithms make no such assumptions on the data.

2.9.1 Revisiting ACOPF through QCQP Formulations

In Section 2.2, we derived (2.4) which shows that the complex power injected at bus k into line (k, m) can be written as

$$S_{km} = V_k V_k^*(G_{kk} - jB_{kk}) + V_k V_m^*(G_{km} - jB_{km}).$$

Since $P_{km} = \Re S_{km}$ and $Q_{km} = \Im S_{km}$ we obtain

$$\begin{aligned} P_{km} &= G_{kk} V_k V_k^* + G_{km} \Re V_k V_m^* - B_{km} \Re j V_k V_m^* \\ &= G_{kk} |V_k|^2 + G_{km} \Re V_k V_m^* + B_{km} \Im V_k V_m^* \\ Q_{km} &= -B_{kk} \Re j V_k V_k^* + G_{km} \Im V_k V_k^* - B_{km} \Im j V_k V_m^* \\ &= -B_{kk} \Re j |V_k|^2 - B_{km} \Re V_k V_m^* + G_{km} \Im V_k V_m^* \end{aligned}$$

by the linearity of $\Re, \Im$ and the fact that $\Re jz = -\Im z$ and $\Im jz = \Re z$ for $z \in \mathbb{C}$. It is straightforward to verify that P_{mk} and Q_{mk} also depend on $\Re V_k V_m^*$ and $\Im V_k V_m^*$ since $\Re V_m V_k^* = \Re V_k V_m^*$ and $\Im V_m V_k^* = -\Im V_k V_m^*$. Following the approach in in Section 2.5.1, we introduce the auxiliary variables $v_k^{(2)} := |V_k|^2$, $v_m^{(2)} := |V_m|^2$, $c_{km} := \Re V_k V_m^*$ and $s_{km} := \Im V_k V_m^*$. Moreover, recalling that

$$\Re V_k V_m^* = \frac{V_k V_m^* + V_m V_k^*}{2}, \quad \Im V_k V_m^* = \frac{V_k V_m^* - V_m V_k^*}{2j}$$

we obtain the so-called *complex-voltage representation* [55]. This leads to the following quadratically constrained optimization problem in complex variables[24]:

[24]We remind the reader that angle-difference constraints have been omitted; see Section 2.3.

Figure 2.12: ACOPF Complex-Voltage QCQP Formulation

$$\text{[ACOPF-C]}: \qquad \min \quad \sum_{k \in \mathcal{G}} F_k(P_k^g) \tag{2.38a}$$

subject to:

$\forall$ bus $k \in \mathcal{B}$:

$$\sum_{\{k,m\} \in \delta(k)} P_{km} = \sum_{\ell \in \mathcal{G}_k} P_\ell^g - P_k^d \tag{2.38b}$$

$$\sum_{\{k,m\} \in \delta(k)} Q_{km} = \sum_{\ell \in \mathcal{G}_k} Q_\ell^g - Q_k^d \tag{2.38c}$$

$\forall$ branch $(k, m) \in \mathcal{A}$:

$$P_{km} = G_{kk} v_k^{(2)} + G_{km} c_{km} + B_{km} s_{km} \tag{2.38d}$$

$$P_{mk} = G_{mm} v_m^{(2)} + G_{mk} c_{km} - B_{mk} s_{km} \tag{2.38e}$$

$$Q_{km} = -B_{kk} v_k^{(2)} - B_{km} c_{km} + G_{km} s_{km} \tag{2.38f}$$

$$Q_{mk} = -B_{mm} v_m^{(2)} - B_{mk} c_{km} - G_{mk} s_{km} \tag{2.38g}$$

$$c_{km} = \frac{V_k V_m^* + V_k^* V_m}{2}, \quad s_{km} = \frac{V_k V_m^* - V_k^* V_m}{2j} \tag{2.38h}$$

$\forall$ generator $k \in \mathcal{G}$:

$$P_k^{\min} \leq P_k^g \leq P_k^{\max} \tag{2.38i}$$

$$Q_k^{\min} \leq Q_k^g \leq Q_k^{\max} \tag{2.38j}$$

$\forall$ bus $k \in \mathcal{B}$:

$$(V_k^{\min})^2 \leq v_k^{(2)} \leq (V_k^{\max})^2 \tag{2.38k}$$

$$v_k^{(2)} = V_k V_k^* \tag{2.38l}$$

$\forall$ branch $\{k, m\} \in \mathcal{E}$:

$$\max\left\{P_{km}^2 + Q_{km}^2, P_{mk}^2 + Q_{mk}^2\right\} \leq U_{\{k,m\}} \tag{2.38m}$$

On the other hand, by writing the complex voltage phasors using *rectangular coordinates*, i.e., $V_k = e_k + j f_k$ for $k \in \mathcal{B}$, we have the following equivalent QCQP formulation in real variables for

ACOPF [52]:

Figure 2.13: ACOPF Rectangular QCQP Formulation

$$[\text{ACOPF-R}] : \qquad \min \quad \sum_{k \in \mathcal{G}} F_k(P_k^g) \tag{2.39a}$$

subject to: (2.38b) – (2.38g), (2.38i) – (2.38m)

$$\forall\, k \in \mathcal{B} :$$

$$v_k^{(2)} = e_k^2 + f_k^2, \tag{2.39b}$$

$$\forall (k, m) \in \mathcal{A} : \tag{2.39c}$$

$$c_{km} = e_k e_m + f_k f_m, \quad s_{km} = e_m f_k - e_k f_m, \tag{2.39d}$$

2.9.2 SDP Relaxations for ACOPF

We consider three SDP relaxations of the ACOPF problem. The first, which we denote by SDP-C [55], is a complex-valued SDP relaxation of (2.38). The second formulation, which we denote by SDP-R [53], arises from applying the Shor relaxation [53] to the QCQP (2.39) in rectangular coordinates, where only second-order terms are considered. Finally, SDP-CR is obtained by converting SDP-C into an equivalent real-valued formulation. We refer the reader to Section 2.4 for a brief overview on SDP relaxations of ACOPF.

Before we introduce the relaxations we fix some notation. For the sake of simplicity, we assume that there $n \in \mathbb{N}$ buses labeled as $\mathcal{B} = \{1, \ldots, n\}$. We denote by $(\mathbb{R}^{2n \times 2n}, \langle \cdot, \cdot \rangle)$ and $(\mathbb{C}^{n \times n}, \langle \cdot, \cdot \rangle)$ the Euclidean spaces of real matrices and complex matrices, respectively. We denote by $(\cdot)^*$ the Hermitian transpose. We recall that the standard (Frobenius) inner-product $\langle \cdot, \cdot \rangle$ is defined as $\langle A, B \rangle \coloneqq \text{trace}(A^* B)$ for matrices $A, B \in \mathbb{H}^n$, where $\mathbb{H}^n \subseteq \mathbb{C}^{n \times n}$ denotes the subspace of $(n \times n)$-dimensional Hermitian matrices. This inner-product induces the Euclidean norm $\|A\| \coloneqq \sqrt{\langle A, A \rangle}$. We use $A \succeq 0$ to denote that A is PSD, interpreting this over $\mathbb{R}$ or $\mathbb{C}$ according to the nature of A.

Remark 2.15. We note that the following semidefinite programming formulations are not presented in standard form. By means of an epigraph reformulation and the Schur complement, see e.g. [36], [54], the objective can be expressed as a linear function in the matrix variable. Moreover, by applying the Schur complement again, (2.38m) can be reformulated as a PSD constraint.

SDP-C

Following ideas from [55], the constraints (2.38l) and (2.38h) can be written as

$$v_k^{(2)} = \langle R_k^v, VV^* \rangle \quad c_{km} = \langle R_{km}^c, VV^* \rangle \quad s_{km} = \langle R_{km}^s, VV^* \rangle$$

where the matrices $R_k^v := \phi_k \phi_k^\top, R_{km}^c := \frac{1}{2}(\phi_k \phi_m^\top + \phi_m \phi_k^\top), R_{km}^s := \frac{1}{2j}(\phi_k \phi_m^\top - \phi_m \phi_k^\top) \in \mathbb{C}^{n \times n}$ are Hermitian, and $\phi_i \in \mathbb{R}^n$ denotes the i-th canonical basis vector in $\mathbb{R}^n$. Therefore, we can write $v_k^{(2)} = \langle R_k^v, VV^* \rangle$, $c_{km} = \langle R_{km}, X \rangle$ and $s_{km} = \langle T_{mk}, X \rangle$ with $X = VV^*$, i.e., X has rank one. By relaxing the rank constraint and approximating the outer-product VV^* by a PSD matrix $X \in \mathbb{C}^{n \times n}$, we obtain the following SDP relaxation in complex variables:

Figure 2.14: SDP Relaxation of (2.38)

$$[\text{SDP-C}]: \qquad \min \qquad \sum_{k \in \mathcal{G}} F_k(P_k^g) \tag{2.40a}$$

subject to: (2.38b) – (2.38g), (2.38i) – (2.38m)

$\forall$ bus $k \in \mathcal{B}$:

$$v_k^{(2)} = X_{kk}, \tag{2.40b}$$

$$\forall (k,m) \in \mathcal{A}: \tag{2.40c}$$

$$c_{km} = \frac{X_{km} + X_{mk}}{2}, \quad s_{km} = \frac{X_{km} - X_{mk}}{2j}, \tag{2.40d}$$

$$X \succeq 0, \quad X \in \mathbb{C}^{n \times n} \tag{2.40e}$$

SDP-R

On the other hand, by approximating the outer-product $(e, f)^\top(e, f)$ by a PSD matrix $W \in \mathbb{R}^{2n\times 2n}$, i.e., by relaxing the rank constraint in (2.39), we obtain the following SDP relaxation:

Figure 2.15: SDP Relaxation of (2.39)

$$[\text{SDP-R}] : \qquad \min \qquad \sum_{k\in\mathcal{G}} F_k(P_k^g) \tag{2.41a}$$

subject to: (2.38b) – (2.38g), (2.38i) – (2.38m)

$\forall\, k \in \mathcal{B}$:

$$v_k^{(2)} = W_{kk} + W_{k'k'}, \tag{2.41b}$$

$$\forall (k, m) \in \mathcal{A} : \tag{2.41c}$$

$$c_{km} = W_{km} + W_{k'm'},\ s_{km} = W_{mk'} - W_{km'}, \tag{2.41d}$$

$$W \succeq 0, \quad W \in \mathbb{R}^{2n\times 2n} \tag{2.41e}$$

where $k' := n + k$ for $k \in \mathcal{B}$.

SDP-CR

Next, we present SDP-CR. We first elaborate on a procedure, known as *realification* of an Hermitian matrix, which reduces complex semidefinite programs to real ones. Let $\mathbb{B}^{2n} \subseteq \mathbb{S}^{2n}$ be the subspace of $(2n \times 2n)$-dimensional real symmetric matrices of the form

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

where $A, B \in \mathbb{R}^{n\times n}$ with $A = A^\top$ and $B^\top = -B$. Let $L : \mathbb{H}^n \to \mathbb{B}^{2n}$ be the map defined by:

$$L(X) \coloneqq \begin{pmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{pmatrix} \quad \text{for } X \in \mathbb{H}^n.$$

It is straightforward to verify that this map is a linear bijection. The following facts are standard: a proof of Lemma 2.16 can be found in any linear algebra text, while a proof of Lemma 2.17 can be found in [104].

Lemma 2.16. *Let $X \in \mathbb{C}^n$. Then the quadratic form $\mathbb{C}^n \to \mathbb{C}$ defined as $u \mapsto u^* X u$ is real-valued.*

Lemma 2.17. *Let $X \in \mathbb{H}^{n\times n}$. Then $X \succeq 0$ if and only if $L(X) \succeq 0$.*

Finally, we observe that the only constraints affected by the realification of X in SDP-C are (2.40b) and (2.40d). Recall that these constraints can be written as $v_k^{(2)} = \langle R_k^v, X\rangle$, $c_{km} = \langle R_{km}^c, X\rangle$, and $s_{km} = \langle R_{km}^s, X\rangle$ where $R_k^v, R_{km}^c, R_{km}^s \in \mathbb{H}^{n\times n}$. We make use of the following fact [62] concerning the realification of Hermitian matrices: for any $Z_1, Z_2 \in \mathbb{H}^{n\times n}$, $\langle Z_1, Z_2\rangle = \frac{1}{2}\langle L(Z_1), L(Z_2)\rangle$. By this observation we have that

$$\begin{aligned}
v_k^{(2)} &= \frac{1}{2}\langle L(R_k^v), L(X)\rangle \\
&= \frac{1}{2}\langle \delta_k\delta_k^\top + \delta_{k'}\delta_{k'}^\top, L(X)\rangle = \frac{1}{2}(L(X)_{kk} + L(X)_{k'k'}) && \text{(2.42a)} \\
c_{km} &= \frac{1}{2}\langle L(R_{km}^c), L(X)\rangle \\
&= \frac{1}{4}\langle \delta_k\delta_m^\top + \delta_m\delta_k^\top + \delta_{k'}\delta_{m'}^\top + \delta_{m'}\delta_{k'}^\top, L(X)\rangle = \frac{1}{2}(L(X)_{km} + L(X)_{k'm'}) && \text{(2.42b)} \\
s_{km} &= \frac{1}{2}\langle L(R_{km}^s), L(X)\rangle \\
&= \frac{1}{4j}(\delta_m\delta_{k'}^\top - \delta_k\delta_{m'}^\top + \delta_m\delta_{k'}^\top - \delta_k\delta_{m'}^\top), L(X)\rangle = \frac{1}{2}(L(X)_{mk'} - L(X)_{km'}) && \text{(2.42c)}
\end{aligned}$$

where δ_i denotes the i-th canonical basis vector in $\mathbb{R}^{2n}$ and $i' \coloneqq n + i$. Therefore, the formulation SDP-CR with variable $\tilde{X} \coloneqq \frac{1}{2}L(X)$ is obtained by applying the realification operator to X in SDP-C. As shown in [104] these two SDP relaxations are equivalent, in the sense that an optimal solution to SDP-C can be obtained from an optimal solution to SDP-CR, and vice versa.

2.9.3 On the Relationship between SDP-R and SDP-CR

A widespread belief is that SDP-C and SDP-R are equivalent. To the best of our knowledge, nearly all papers on SDP relaxations for the ACOPF problem [24], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64] study one of these formulations, typically SDP-C or its equivalent form SDP-CR, without comparing them, implicitly treating them as reformulations of one another. Some works even refer to either formulation as *the* SDP relaxation for ACOPF.

In [105], an effort was made to unify various SDP relaxations within a common framework. In particular, [105] claims that the matrix variable W in SDP-R is equivalent, up to an orthogonal similarity transformation, to the matrix variable $L(X)$ in SDP-CR, i.e., that there exists some orthogonal $Q \in \mathbb{R}^{2n \times 2n}$ such that $W = Q^\top L(X) Q$. This statement has subsequently been cited in works such as [23], [25]. Unfortunately, this claim is not correct in general (see Appendix A.3 for a proof). It holds only when W already has the structure of $L(X)$, in which case the orthogonal transformation is trivial, i.e., $Q = I$; see (2.42). This commonly held belief, though not formally established in the literature, has played a central role in work on SDP relaxations and, more broadly, on convex relaxations of the ACOPF problem.

This issue is significant because relaxations are typically evaluated relative to one another. For example, the perceived superiority of *the* SDP relaxation over the Jabr SOCP relaxation [22] rests on the fact that the positive semidefiniteness of the matrix variable in SDP-C implies the Jabr inequality. However, to the best of our knowledge at the time of the initial version of this work, no proof had been established showing that SDP-R also implies this inequality. This distinction affects all works comparing convex SDP formulations for ACOPF, particularly those that rely on the presumed strength of SDP-R under the assumption that it enforces fundamental inequalities such as the Jabr inequality.

In this section, we analyze the properties of the SDP-R relaxation and show that SDP-R and SDP-C attain the same optimal objective value. We first establish that SDP-R is at least as strong as SDP-C. After the completion of the initial version of this work, Daniel Molzahn brought to our attention an unpublished manuscript [106] establishing an analogous objective value equivalence for the Shor relaxation of general complex-valued QCQPs. The reverse inequality, namely that SDP-C is at least as strong as SDP-R, follows from his argument, which we include here for completeness; see Theorem 2.20.

We now establish the equivalence between SDP-R and SDP-C. To this end, we first introduce a definition and prove an auxiliary result.

Definition 2.18. Let $W \in \mathbb{R}^{2n\times 2n}$ be an arbitrary symmetric matrix. We define a Hermitian matrix $X^W \in \mathbb{C}^{n\times n}$ associated to W as follows:

$$\begin{aligned} X^W_{kk} &:= W_{kk} + W_{k'k'} \\ \Re X^W_{km} &:= W_{km} + W_{k'm'} \\ \Im X^W_{km} &:= W_{km'} - W_{mk'}. \end{aligned}$$

Theorem 2.19. *Let $W \in \mathbb{R}^{2n\times 2n}$ be an arbitrary symmetric matrix. Then $W \succeq 0$ implies $X^W \succeq 0$.*[25]

Proof. Since W is PSD over the reals, the following bilinear map is a semi-inner product in $\mathbb{C}^n$:

$$\langle u, v\rangle_W \coloneqq u^* W v, \qquad \forall u, v \in \mathbb{C}^n.$$

This follows because $\langle \cdot, \cdot\rangle_W$ satisfies: (i) *linearity* in the second argument; (ii) *conjugate symmetry*, i.e., $\langle u, v\rangle_W = \langle v, u\rangle_W^*$; and (iii) *semipositivity*, i.e., it holds that $\langle u, u\rangle_W \geq 0$ for every $u \in \mathbb{C}^n$, by Lemma 2.17 since $L(W)$ is a block-diagonal PSD matrix.

Consider the following points in $\mathbb{C}^n$: $z_k \coloneqq \delta_k + j\delta_{k'}$ for $k \in \mathcal{B}$. Recall that δ_k denotes the kth

[25] A generative AI tool assisted in producing this proof. The final proof was verified and refined by the author.

canonical basis vector in $\mathbb{R}^{2n}$. We note that

$$
\begin{aligned}
\langle z_k, z_k \rangle_W &= \delta_k^* W \delta_k + \delta_{k'}^* W \delta_{k'} = W_{kk} + W_{k'k'} \\
\langle z_k, z_m \rangle_W &= \delta_k^* W \delta_m + j\delta_k^* W \delta_{m'} - j\delta_{k'}^* W \delta_m + \delta_{k'}^* W \delta_{m'} \\
&= (W_{km} + W_{k'm'}) + j(W_{km'} - W_{mk'}) \\
\langle z_m, z_k \rangle_W &= \delta_m^* W \delta_k + j\delta_m^* W \delta_{k'} - j\delta_{m'}^* W \delta_k + \delta_{m'}^* W \delta_{k'} \\
&= (W_{km} + W_{k'm'}) + j(W_{mk'} - W_{km'})
\end{aligned}
$$

where we used the fact that W is symmetric. Next, we construct the following "semi" Gram matrix in $\mathbb{C}^{n\times n}$: for $k, m \in \mathcal{B}$,

$$
G_{km} := \begin{cases} \langle z_k, z_k \rangle_W & \text{if } k = m, \\ \langle z_k, z_m \rangle_W & \text{if } k \neq m. \end{cases}
$$

Since $\langle z_m, z_k \rangle_W = \langle z_k, z_m \rangle_W^*$, we deduce that G is Hermitian. We claim that G is PSD over $\mathbb{C}^n$. Indeed, let $y \in \mathbb{C}^n$ be arbitrary, then

$$
\begin{aligned}
y^* G y &= \sum_{k=1}^{n} \sum_{m=1}^{n} y_k^* y_m \langle z_k, z_m \rangle_W \\
&= \sum_{k=1}^{n} \sum_{m=1}^{n} \langle y_k z_k, y_m z_m \rangle_W \\
&= \sum_{k=1}^{n} \langle y_k z_k, \sum_{m=1}^{n} y_m z_m \rangle_W \\
&= \langle \sum_{k=1}^{n} y_k z_k, \sum_{k=1}^{n} y_k z_k \rangle_W \\
&\geq 0
\end{aligned}
$$

where the second equality follows since $\langle \cdot, \cdot \rangle_W$ satisfies conjugate symmetry and it is linear in the second argument, and nonnegativity follows by the semipositivity of $\langle \cdot, \cdot \rangle_W$. Thus, given that $G = X^W$, we conclude that X^W is PSD. $\square$

Theorem 2.20. *Let z_C and z_R denote the optimal values of SDP-C and SDP-R, respectively. Then $z_C = z_R$.*

Proof. We first prove $z_C \le z_R$. If SDP-R is infeasible, then the claim is vacuously true since $z_R = +\infty$. Next, suppose that SDP-R is feasible and let $y_R := (\hat{P}^g, \hat{Q}^g, \hat{W}, \hat{P}, \hat{Q}, \hat{c}, \hat{s}, \hat{v}^{(2)})$ be a feasible solution. We show how to obtain a feasible solution to SDP-C from it. Indeed, using $\hat{W}$ we can construct by Definition 2.18 the Hermitian matrix $\hat{X}^{\hat{W}} \in \mathbb{C}^{n\times n}$. We claim that $y_C := (\hat{P}^g, \hat{Q}^g, \hat{X}^{\hat{W}}, \hat{P}, \hat{Q}, \hat{c}, \hat{s}, \hat{v}^{(2)})$ satisfies all the constraints of SDP-C. Indeed, clearly $\hat{P}^g, \hat{Q}^g, \hat{P}, \hat{Q}, \hat{c}, \hat{s}, \hat{v}^{(2)}$ satisfy constraints (2.38b)-(2.38g), (2.38i)-(2.38m). Moreover, since $\hat{X}^{\hat{W}}_{kk} = \hat{W}_{kk} + \hat{W}_{k'k'} = v^{(2)}_k$, then constraint (2.40b) holds. Next, we note that constraints (2.40d) are equivalent to $c_{km} = \Re X_{km}$, $s_{km} = \Im X_{km}$. Hence, since $\hat{c}_{km} = \hat{W}_{km} + \hat{W}_{k'm'}$ and $s_{km} = \hat{W}_{km'} - \hat{W}_{mk'}$, we have that $\hat{c}_{km} = \Re\hat{X}^{\hat{W}}_{km}$ and $\hat{s}_{km} = \Im\hat{X}^{\hat{W}}_{km}$, i.e., (2.40d) is satisfied. Finally, given that $\hat{W} \succeq 0$, Theorem 2.19 implies that $\hat{X}^{\hat{W}} \succeq 0$ and (2.40e) holds. Since y_C and y_R have the same objective value, we conclude that $z_C \le z_R$.

Next, we show that $z_C \ge z_R$ using the argument in [106]; this relies on the fact that SDP-R is a relaxation of SDP-CR. Indeed, any feasible $y_{CR} := (\hat{P}^g, \hat{Q}^g, \hat{\hat{X}}, \hat{P}, \hat{Q}, \hat{c}, \hat{s}, \hat{v}^{(2)})$ satisfies constraints (2.38b)-(2.38g), (2.38i)-(2.38m) as well as (2.42a)-(2.42c). However, the real $2n \times 2n$ PSD matrix $\hat{\hat{X}}$ possesses additional structure, i.e., $\hat{\hat{X}} \in \mathbb{B}^{2n}$, that is not explicitly enforced on W in (2.41). Since SDP-C and SDP-CR are equivalent, the inequality follows. $\square$

Moreover, from our results above we can deduce that the Jabr inequality is implied by the set of feasible solutions to SDP-W. In particular, it is implied by the constraint $W \succeq 0$. Interestingly, we uncover what we term *permuted Jabr inequalities*, by exploiting the permutation symmetries of (2.44). This provides a theoretical justification for the tightness of the SDP-R relaxation since the *loss inequalities* are implied by the Jabr inequality; see Section 2.5.4 for a discussion.

Let $A \in \mathbb{K}^{r\times r}$ where $K \in \{\mathbb{R}, \mathbb{C}\}$ and $r \in \mathbb{N}$. We denote by A_R the principal submatrix of A given by a set $R \subseteq [r]$ of row and column indices.

Corollary 2.21. *Consider any 2×2 principal submatrix of X^W*

$$X^W_{\{k,m\}} := \begin{pmatrix} X^W_{kk} & \Re X^W_{km} + j\Im X^W_{km} \\ \Re X^W_{mk} + j\Im X^W_{km} & X^W_{mm} \end{pmatrix}.$$

If $W \succeq 0$, then $X^W_{\{k,m\}} \succeq 0$, i.e., $X^W_{kk}, X^W_{mm} \geq 0$ and

$$(\Re X^W_{km})^2 + (\Im X^W_{km})^2 \leq X^W_{kk} X^W_{mm}. \tag{2.43}$$

In particular, the following inequality holds

$$(W_{km} + W_{k'm'})^2 + (W_{mk'} - W_{km'})^2 \leq (W_{kk} + W_{k'k'})(W_{mm} + W_{m'm'}). \tag{2.44}$$

Moreover, for any row-column permutation σ of the submatrix $W_{\{k,m,k',m'\}}$, the following inequality also holds:

$$c_\sigma^2 + s_\sigma^2 \leq r_\sigma t_\sigma. \tag{2.45}$$

where $c_\sigma := W_{\sigma(k)\sigma(m)} + W_{\sigma(k')\sigma(m')}$, $s_\sigma := W_{\sigma(k)\sigma(m')} - W_{\sigma(m)\sigma(k')}$, $r_\sigma := W_{\sigma(k)\sigma(k)} + W_{\sigma(k')\sigma(k')}$, $t_\sigma := W_{\sigma(m)\sigma(m)} + W_{\sigma(m')\sigma(m')}$.

Proof. The first claim follows by a direct application of Theorem 2.19 to the principal submatrix $W_{\{k,m,k',m'\}}$ of W. The inequality 2.44 follows immediately by (2.43) and the definition of X^W. Finally, inequality (2.45) holds for any permutation σ on $\{k, m, k', m'\}$ since any row-column permutation of $W_{\{k,m,k',m'\}}$ can be represented by a similarity transformation $Q^T W_{\{k,m,k',m'\}} Q$ where Q is a permutation matrix (in particular orthogonal). Since similarity transformations are PSD-preserving the claim follows. □

We remark that by taking σ as the identity permutation gives in (2.45) the *standard* Jabr inequality. Consider the relaxations of SDP-C and SDP-R, denoted SOC-C and SOC-R, where we relax $X \succeq 0$ and $W \succeq 0$ to $X_{\{k,m\}} \succeq 0$ and $W_{\{k,m\}} \succeq 0$ for every $k \neq m$, respectively. Since the matrices $W_{\{k,m\}}$ are 4×4, the constraints $W_{\{k,m\}} \succeq 0$ in SOC-R are not rotated-cone constraints.

Nevertheless, by the arguments presented in the proof of Theorem 2.20, these relaxations attain the same objective value. In particular, permuted Jabrs, which are implied by constraints $W_{\{k,m\}} \succeq 0$, do not improve the objective beyond SOC-C (i.e., the standard Jabr relaxation). We corroborate these observations empirically. We emphasize that this equivalence is only in terms of objective values: the standard Jabr inequality does not imply the permuted Jabr inequalities, which may be violated by points feasible for the former.

A particularly insteresting empirical observation, which merits further investigation, is the the fact that the permuted Jabr corresponding to $\sigma = (km)$ has exactly the same impact on the objective as the loss inequalities $P_{km} + P_{mk} \geq 0$ when the constraint $W \succeq 0$ is removed. In other words, the relaxation of SPD-R without $W \succeq 0$, together with

$$(W_{km} + W_{k'm'})^2 + (W_{kk'} - W_{mm'})^2 \leq (W_{mm} + W_{k'k'})(W_{kk} + W_{m'm'}).$$

attains the same objective as the linear-loss-relaxation, highlighting a close connection between permuted Jabr inequalities, which are nonlinear, and nonnegative losses-based constraints, which are linear.

2.9.4 Our Disciplined Approach to SDPs

An SDP relaxation with matrix variable $X \succeq 0$ can be reformulated using a clique decomposition technique applied to the underlying power system network [107]. This approach results in an equivalent SDP problem but with PSD constraints $X_i \succeq 0$ on much smaller matrices than X and with linear constraints. The size of these matrices is relatively small (e.g., for case ACTIVSg70k with $70,000$ nodes, a clique decomposition may find a clique of size 114). In [57], [58], [60] different clique merging techniques and reformulations have been studied.

Our approach can be outlined as follows. Given some decomposition $\{X_i\}_{i \in \mathcal{D}}$, for each matrix

X_i, we derive sparse, numerically valid, and stable linear inequalities which outer-approximate the PSD cone $X_i \succeq 0$. To maintain speed and numerical robustness we apply modern cut-management techniques, as in [14], to several cut families, including eigenvector-based cuts [108]. Our goal is to outer-approximate the feasible region of SDP-C. As in prior works in sparse SDPs, the philosophy is to keep the formulation as *light* as possible, without adding too many *dense* cuts. Currently, we have experiments with two decomposition heuristics: (i) identifying all the 3-cliques in the network so that we do not have to add extra variables; and (ii) using a chordal decomposition [107] of the network $(\mathcal{B}, \mathcal{E})$.

In what follows, we also study the underlying geometry and related algebraic issues, with the aim of delivering *fully valid* linear inequalities.

2.9.5 SDP Cuts

We denote by $\mathbb{H}_+^n$ the PSD cone in $\mathbb{C}^{n \times n}$. Let $X_0 \in \mathbb{C}^n$ be Hermitian. Suppose that X_0 is not positive semidefinite, and consider an eigen-decomposition (normalized) of X_0, i.e., $X_0 = Q \Lambda Q^*$ where $\{q_1, \ldots, q_n\} \subseteq \mathbb{C}^n$ are eigenvectors of X_0 which correspond to the columns of Q, and Λ is a diagonal matrix whose elements along the diagonal are the eigenvalues of X_0 sorted in non-increasing order $\lambda_1 \geq \cdots \geq \lambda_n$.

Suppose that we want to separate X_0 from the PSD cone $\mathbb{H}_+^n$. We discuss two types of cuts. For further discussion of (sparse) cutting planes for QCQPs, see [108], [109].

Eigenvector-based Cuts

Consider the eigenvector $q_n \in \mathbb{C}^n$ associated with the eigenvalue λ_n, i.e. $X_0 q = \lambda_n q$ and $\|q\|_2 = 1$. Since we assumed that X_0 is indefinite, we have $\lambda_n < 0$. Recall that $X \in \mathbb{H}_+^n$ if and only if $x^* X x \geq 0$ for all $x \in \mathbb{C}^n$. Let $A \coloneqq qq^*$. Hence, the following inequality defines a valid linear cut, which we refer to as an *eigen-cut* for $\mathbb{H}_+^n$,

$$\langle A, X \rangle = q^* X q \geq 0, \quad \forall X \in \mathbb{H}^n \tag{2.46}$$

Clearly, $\langle A, X_0\rangle = q^\top X_0 q = \lambda_n < 0$. In other words, X_0 violates this inequality by $-\lambda_n$ and its distance to (2.46) is $|\lambda_n|$.

Next, we show that eigen-cuts derived from 2×2 principal submatrices of X_0 correspond to the standard Jabr outer-approximation cuts (2.29).

At an arbitrary iteration, let $\overline{X}_{\{k,m\}}$ be the principal 2×2 Hermitian matrix associated to the branch $\{k,m\}$. Suppose that

$$\overline{X}_{\{k,m\}} = \begin{pmatrix} \overline{v}_k^{(2)} & \overline{c}_{km} + j\overline{s}_{km} \\ \overline{c}_{km} - j\overline{s}_{km} & \overline{v}_m^{(2)} \end{pmatrix}$$

is not PSD, i.e., $\overline{v}_k^{(2)}\overline{v}_m^{(2)} < \overline{c}_{km}^2 + \overline{s}_{km}^2$ since $\overline{v}_k^{(2)}, \overline{v}_m^{(2)} > 0$.

We first compute the smallest eigenvalue of $\overline{X}_{\{k,m\}}$, which is real since $\overline{X}_{\{k,m\}}$ is Hermitian. Indeed,

$$\begin{aligned} \det(\overline{X}_{\{k,m\}} - \lambda I) = 0 \iff \lambda^2 - \lambda(\overline{v}_k^{(2)} + \overline{v}_m^{(2)}) + (\overline{v}_k^{(2)}\overline{v}_m^{(2)} - \overline{c}_{km}^2 - \overline{s}_{km}^2) = 0 \\ \implies \lambda = \frac{1}{2}\left[(\overline{v}_k^{(2)} + \overline{v}_m^{(2)}) \pm \sqrt{(\overline{v}_k^{(2)} + \overline{v}_m^{(2)})^2 - 4(\overline{v}_k^{(2)}\overline{v}_m^{(2)} - \overline{c}_{km}^2 - \overline{s}_{km}^2)}\right] \\ = \frac{1}{2}\left[(\overline{v}_k^{(2)} + \overline{v}_m^{(2)}) \pm \sqrt{(\overline{v}_k^{(2)} - \overline{v}_m^{(2)})^2 + 4(\overline{c}_{km}^2 + \overline{s}_{km}^2)}\right]. \end{aligned}$$

Hence, $\lambda_{\min}$ equals $\frac{1}{2}\left[(\overline{v}_k^{(2)} + \overline{v}_m^{(2)}) - \sqrt{(\overline{v}_k^{(2)} - \overline{v}_m^{(2)})^2 + 4(\overline{c}_{km}^2 + \overline{s}_{km}^2)}\right] < 0$, i.e., one half of the negative of the violation of the Jabr inequality (2.12) at the current iterate $(\overline{v}^{(2)}, \overline{c}, \overline{s})$.

Let $\alpha := \sqrt{(\overline{v}_k^{(2)} - \overline{v}_m^{(2)})^2 + 4(\overline{c}_{km}^2 + \overline{s}_{km}^2)}$ which is strictly positive. Next, we compute an eigenvector associated to $\lambda_{\min}$, i.e., we solve for $q \in \mathbb{C}^2$ in

$$\begin{pmatrix} \frac{1}{2}(\overline{v}_k^{(2)} - \overline{v}_m^{(2)}) + \frac{1}{2}\alpha & \overline{c}_{km} + j\overline{s}_{km} \\ \overline{c}_{km} - j\overline{s}_{km} & -\frac{1}{2}(\overline{v}_k^{(2)} - \overline{v}_m^{(2)}) + \frac{1}{2}\alpha \end{pmatrix} q = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We note that this eigenspace must have dimension 1 and that the diagonal of the latter matrix is strictly positive. Therefore, an eigenvector is given by

$$q = \begin{pmatrix} \frac{-2(\overline{c}_{km}+j\overline{s}_{km})}{(\overline{v}_k^{(2)}-\overline{v}_m^{(2)})+\alpha} q_2 \\ q_2 \end{pmatrix} \in \mathbb{C}^2,$$

and a normalized eigenvector is given by $\overline{q} \coloneqq q/\|q\|$, i.e.,

$$\overline{q} = \frac{1}{|q_2|\sqrt{4+(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)^2}} \begin{pmatrix} -2(\overline{c}_{km}+j\overline{s}_{km})q_2 \\ (\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)q_2 \end{pmatrix}$$

An eigen-cut induced by an eigenvector $\overline{q}$ for a 2×2 matrix can be written as

$$|\overline{q}_1|^2 v_k^{(2)} + |\overline{q}_2|^2 v_m^{(2)} + 2\Re\overline{q}_1\overline{q}_2^* c_{km} + 2\Im\overline{q}_1\overline{q}_2^* s_{km} \geq 0.$$

Since

$$\begin{aligned} \Re\overline{q}_1\overline{q}_2^* &= -\frac{2(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)}{4+(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)^2}\overline{c}_{km} \\ \Im\overline{q}_1\overline{q}_2^* &= -\frac{2(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)}{4+(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)^2}\overline{s}_{km} \end{aligned}$$

we obtain, after multiplying both sides by the positive common denominator $4+(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)^2$,

$$\begin{aligned} & 4(\overline{c}_{km}^2+\overline{s}_{km}^2)v_k^{(2)} + (\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)^2 v_m^{(2)} - 4(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)(\overline{c}_{km}c_{km}+\overline{s}_{km}s_{km}) \geq 0 \\ \iff & \frac{4(\overline{c}_{km}^2+\overline{s}_{km}^2)}{(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)} v_k^{(2)} + (\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha)v_m^{(2)} - (4\overline{c}_{km})c_{km} - (4\overline{s}_{km})s_{km} \geq 0. \end{aligned} \tag{2.47}$$

since $(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}+\alpha) > 0$. Next, we compare this cut with the standard Jabr outer-approximation cut:

$$(\alpha-(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}))v_k^{(2)} + (\alpha+(\overline{v}_k^{(2)}-\overline{v}_m^{(2)}))v_m^{(2)} - (4\overline{c}_{km})c_{km} - (4\overline{s}_{km})s_{km} \geq 0. \tag{2.48}$$

It straightforward to verify that the Jabr cut (2.48) is equivalent to the 2×2 eigen-cut (2.47), since

$$\frac{4(\overline{c}_{km}^2 + \overline{s}_{km}^2)}{(\overline{v}_k^{(2)} - \overline{v}_m^{(2)}) + \alpha} = (\alpha - (\overline{v}_k^{(2)} - \overline{v}_m^{(2)})) \iff 4(\overline{c}_{km}^2 + \overline{s}_{km}^2) = \alpha^2 - (v_k^{(2)} - v_m^{(2)})^2$$

$$= 4(\overline{c}_{km}^2 + \overline{s}_{km}^2).$$

Projection-based Cuts

Next, we derive a cut using the Euclidean projection of X_0 onto $\mathbb{H}_+^n$. This is also known as the *maximum-distance cut*. Consider the problem of minimizing $\|X - X_0\|_F^2$ subject to $X \in \mathbb{H}_+^n$. It is known that the unique minimizer to this SDP is given by $X^* := \sum_{\lambda_i > 0} \lambda_i q_i q_i^*$, see e.g. [110]. Using a well-known characterization of the projection operator, we can derive the following valid linear inequality for $\mathbb{H}_+^n$,

$$\langle X_0 - X^*, X - X^* \rangle \leq 0, \qquad X \in \mathbb{H}_+^n. \tag{2.49}$$

The left-hand side of the inequality is real, as both $X_0 - X^*$ and $X - X^*$ are Hermitian. Thus, the inequality is well defined. Let $A := X^* - X_0 = \sum_{\lambda_i < 0} (-\lambda_i) q_i q_i^*$. Then, (2.49) can be rewritten as $\langle X^* - X_0, X \rangle \geq \langle X^* - X_0, X^* \rangle$, i.e.,

$$\langle A, X \rangle \geq 0 \tag{2.50}$$

since $\langle X^* - X_0, X^* \rangle = 0$. The matrix X_0 violates this inequality by $\sum_{\lambda_i < 0} \lambda_i^2$ and its distance to (2.50) is $\sqrt{\sum_{\lambda_i < 0} \lambda_i^2}$.

Cut Computations

Let $A, X \in \mathbb{H}^n$. Recall that $\langle A, X\rangle = \sum_{k,m} A_{km}^* X_{km}$ is real and consider the summands $A_{km}^* X_{km}$ and $A_{mk}^* X_{mk}$. We note that

$$\begin{aligned}
A_{km}^* X_{km} &= (\Re A_{km} - j\Im A_{km})(\Re X_{km} + j\Im X_{km}) \\
&= \Re A_{km}\Re X_{km} + \Im A_{km}\Im X_{km} \\
&+ j(\Re A_{km}\Im X_{km} - \Re X_{km}\Im A_{km}) \\
A_{mk}^* X_{mk} &= (\Re A_{mk} - j\Im A_{mk})(\Re X_{mk} + j\Im X_{mk}) \\
&= \Re A_{mk}\Re X_{mk} + \Im A_{mk}\Im X_{mk} \\
&+ j(\Re A_{mk}\Im X_{mk} - \Re X_{mk}\Im A_{mk})
\end{aligned}$$

Since $\Re X_{km} = \Re X_{mk}$, $\Re A_{km} = \Re A_{mk}$, $\Im X_{km} = -\Im X_{mk}$, $\Im A_{km} = -\Im A_{mk}$ we have that $A_{km}^* X_{km} + A_{mk}^* X_{mk}$ equals $2(\Re A_{km}\Re X_{km} + \Im A_{km}\Im X_{km})$. Therefore, $\langle A, X\rangle \geq 0$ can be written as

$$\textstyle\sum_{k=1}^n A_{kk} v_k^{(2)} + \sum\limits_{1\leq k<m\leq n} 2(\Re A_{km} c_{km} + \Im A_{km} s_{km}) \geq 0. \tag{2.51}$$

using definitions (2.40b) and (2.40d). We conclude this section with a result on the numerical robustness of our methods.

Proposition 2.22. *Let $X_0 \in \mathbb{C}^{3\times 3}$ be a complex Hermitian matrix that approximates the outer product of three complex voltage phasors VV^*, where $V \in \mathbb{C}^3$ (hence the entries of X_0 are well-behaved). Let $\epsilon > 0$, and suppose that the smallest eigenvalue of X_0, which we denote as λ_3, satisfies $\lambda_3 > -\epsilon$. Then the distance of X_0 to $\mathbb{H}_+^3$ is at most $\sqrt{2}\epsilon$.*

Proof. Let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ be the eigenvalues of X_0. Since the trace of X_0 is always strictly positive, we have that X_0 has at most two negative eigenvalues. Suppose that $\lambda_2, \lambda_3 < 0$. Then by assumption $-\lambda_2, -\lambda_3 < \epsilon$. Therefore, by our previous discussion we have that the distance of X_0 to the $\mathbb{H}_+^3$ is at most $\sqrt{\epsilon^2 + \epsilon^2} = \sqrt{2}\epsilon$. □

2.9.6 Algorithm and Preliminary Computational Results

First, we define the linearly constrained model M_0 as model (2.40) with only linear constraints and no explicit dependence on X, i.e., without (2.40b), (2.40d), and (2.40e). We denote by M our dynamic relaxation at some iteration of our algorithm, see Algorithm 2. We set $\mathcal{D}$ to be the set of all the 3-cycles of the network $(\mathcal{B}, \mathcal{E})$. Hence, X_i is a 3×3 Hermitian matrix for each $i \in \mathcal{D}$.

As in Algorithm 1, in every round of our procedure, linear constraints will be added to and removed from M. Given a solution $\overline{y}$ to M, line (6) of Algorithm 2 refers to the application of cut management techniques to separate $\overline{y}$ from the Jabr, i2 and Limit inequalities. Next, we check whether the matrices $\{X_i(\overline{y})\}_{i \in D}$ are PSD by calling NumPy's SVD implementation. If $X_i(\overline{y})$ is not PSD, we separate it from the PSD cone $X_i \succeq 0$ using the linear inequalities (2.51). We use projection-based cuts with at most two negative eigenvalues. If the round r surpasses the threshold r^*, we augment $\mathcal{D}$ by computing a chordal decomposition [107] of the network. For further details on cut management please refer to Section 2.6.2.

Finally, other input parameters for our procedure are: a time limit $T > 0$; a number of admissible iterations without sufficient objective improvement $T_{ftol} \in \mathbb{N}$; and a threshold for relative objective improvement $\epsilon_{ftol} > 0$.

Algorithm 2 Cutting-Plane Algorithm

1: **procedure** CUTPLANE
2: Initialize $r \leftarrow 0$, $M \leftarrow M_0$, $z_0 \leftarrow +\infty$
3: $\mathcal{D} \leftarrow$ all the 3-cycles of the network $(\mathcal{B}, \mathcal{E})$
4: **while** $t < T$ and $r < T_{ftol}$ **do**
5: $z \leftarrow \min M$ and $\bar{y} \leftarrow$ argmin M
6: Apply cut management techniques from Section 2.6
7: Check whether matrices $\{X_i(\overline{y})\}_{i \in D}$ are PSD
8: Compute cuts (2.51) for $X_i(\overline{y})$ nonPSD
9: Drop cuts of age $\geq T_{age}$ whose slack is $\geq \epsilon_j$
10: **If** $r \geq r^*$ **do** CYCLEHIERARCHY
11: **If** $z - z_0 < z_0 \cdot \epsilon_{ftol}$ **do** $r \leftarrow r + 1$
12: **Else do** $r \leftarrow 0$
13: $z_0 \leftarrow z$
14: **end while**
15: **end procedure**

All our experiments were conducted on an Apple M2 Pro machine with 16 GB RAM. We used Gurobi 11.0.0 with its default homogeneous self-dual embedding interior-point algorithm, without *Crossover*. We set the parameters *Bar Homogeneous* and *Numeric Focus* equal to 1. Barrier convergence tolerance and absolute feasibility and optimality tolerances were set to 10^{-6}. The Jabr SOCP bounds in Table 2.17, and the AC primal bounds in Tables 2.16-2.17 are reported Tables 2.5 and 2.8.

Tighter Lower Bounds In Table 2.16 we report on bounds obtained by Algorithm 2 with a time limit of 1200 seconds as in Section 2.8.1. After $r^* = 5$ rounds, the set $\mathcal{D}$ is augmented by cliques of length at most 5. The column #Cliques reports the number of 3-, 4-, and 5-cliques, respectively, handled in formulation M at the final iteration. The column 'EigRatio' shows the smallest ratio, among the matrices $X_i(\overline{y})$, between the largest and the second-largest eigenvalues – a commonly used proxy for assessing rank-one behavior. Column 'DInfs' reports on the maximum dual infeasibility error for the last solve; see Section 2.7.

Loads Perturbed The load of each bus was perturbed by a Gaussian $(\mu, \sigma) = (0.01 \cdot P_d, 0.01 \cdot P_d)$, where P_d denotes the original load, subject to the newly perturbed load being non negative [14]. We warm-start Algorithm 2 with precomputed cuts (Jabr, i2, and limit) in [14] with a time limit of 600 seconds. Our SDP cuts are only for 3-cycles in the network. Column 'DInfs' reports on the maximum dual infeasibility error for the last solve, while column 'Added' reports on the number of cuts in the formulation at the last solve.

Table 2.16: New Lower Bounds with Linearly-Constrained Relaxations

	Advanced Cutting-Plane				Cutting-Plane Section 2.6	
Case	Objective	#Cliques	DInf	EigRatio	Objective	AC
14	8080.96	(5,0,0)	2.36e-11	682972.8	8074.70	8081.18
1354pegase	74013.67	(87,111,44)	3.31e-09	1.11	74011.00	74069.35
9241pegase	310226.80	(8673,1029,484)	4.30e-07	1.46	309221.81	315911.56
13659pegase	380055.79	(8673,0,0)	3.70e-05	1.83	379084.55	386108.81

Table 2.17: Warm-Started Relaxations, Loads perturbed Gaussian $(\mu, \sigma) = (0.01 \cdot P_d, 0.01 \cdot P_d)$

	Advanced Cutting-Plane				Cutting-Plane Section 2.6					
Case	Objective	Time (s)	DInfs	Added	Objective	Time (s)	DInfs	Added	Jabr SOCP	AC
9241pegase	310225.93	307.67	8.01e-10	73940	309288.30	15.01	9.52e-07	29875	-	315979.53
ACTIVSg10k	2475135.37	7.43	5.06e-07	22605	2475040.71	4.94	2.55e-08	18183	-	2484093.15
13659pegase	380518.18	429.05	5.46e-08	73889	379742.50	31.09	3.22e-06	37297	-	386765.25
ACTIVSg25k	5989222.32	32.36	7.21e-08	60690	5988885.07	20.03	1.34e-08	43851	5949381.04	6013477.05
78484epigrids-api	15866507.73	118.11	1.27e-05	315589	15862318.35	82.28	4.04e-07	240576	15859950.52	-
78484epigrids-sad	15175077.19	124.12	5.33e-06	397189	15176865.38	96.74	3.86e-06	313587	15174716.43	15316872.94

2.10 An Application to Electricity Markets

In this section, we leverage our techniques to develop a scalable pricing algorithm for real-time market computations and conduct experiments under well-known pricing rules. On small grids ($< 1,000$ nodes), our algorithm yields allocations with very low redispatch costs compared to DC and produces prices that closely match those of the Jabr SOC. On large grids ($\geq 15,000$ nodes), it scales similarly to DC when cuts are warm-started, while maintaining high accuracy relative to the Jabr SOC.

We consider a power system where its underlying undirected network is given $\mathcal{N} := (\mathcal{B}, \mathcal{E})$. Price formation starts with a welfare problem, where *buyers* or loads, denoted by $\mathcal{L}$, (*sellers* or generators, denoted by $\mathcal{G}$), are assumed to truthfully submit piecewise linear bidding functions $\pi^\tau : \mathbb{R}_+ \mapsto \mathbb{R}$, $\tau \in \{d, g\}$ mapping consumption (production, unit commitments) to profits (costs). For each bus $k \in \mathcal{B}$, $\mathcal{G}_k \subseteq \mathcal{G}$ and $\mathcal{L}_k \subseteq \mathcal{L}$ denotes the generators and loads at bus k. The welfare problem [30], [111] can be written as:

Figure 2.16: Power Flow Welfare Maximization

$$[\text{WF}]: \qquad \max \qquad \sum_{\ell \in \mathcal{L}} \pi^d_\ell(P^d_\ell) + \sum_{\ell \in \mathcal{G}} \pi^g_\ell(P^g_\ell, x^{on}_\ell, x^{su}_\ell, x^{sd}_\ell) \tag{2.52a}$$

subject to:

$\forall$ bus $k \in \mathcal{B}$:

$$\sum_{\{k,m\} \in \delta(k)} P_{km} = \sum_{\ell \in \mathcal{G}_k} P^g_\ell - \sum_{\ell \in \mathcal{L}_k} P^d_\ell \quad \forall k \in \mathcal{B} \tag{2.52b}$$

$$\sum_{\{k,m\} \in \delta(k)} Q_{km} = \sum_{\ell \in \mathcal{G}_k} Q^g_\ell - \sum_{\ell \in \mathcal{L}_k} Q^d_\ell \quad \forall k \in \mathcal{B} \tag{2.52c}$$

$$(x^{on}, x^{su}, x^{sd}) \in \{0,1\}^{3 \times |\mathcal{G}|} \tag{2.52d}$$

$$(P^g_\ell, Q^g_\ell)_{\ell \in \mathcal{G}} \in \mathcal{P}_\mathcal{G}(x^{on}, x^{su}, x^{sd}) \tag{2.52e}$$

$$(P^d_\ell, Q^d_\ell)_{\ell \in \mathcal{L}} \in \mathcal{P}_\mathcal{L} \tag{2.52f}$$

$$(y, (P_{km}, P_{mk}, Q_{km}, Q_{mk})_{\forall (k,m) \in \mathcal{A}}) \in \mathcal{Q}^{\text{PF}} \tag{2.52g}$$

where a social planner maximizes surplus subject to: active and reactive power balance constraints (2.52b)-(2.52c); production and commitment constraints for sellers (2.52d) and (2.52e); consumption constraints for buyers (2.52f); and a representation of steady-state power flows and operational constraints (2.52g).

We next describe each component of the formulation in detail, following the welfare model of the ARPA-E Grid Optimization Competition 2 (GOC2) [111] documentation.

Production and Unit Commitment Constraints for Sellers (2.52d) **and** (2.52e) The set $\mathcal{P}_\mathcal{G}$ is a mixed-integer linear set, in (2.52e) we make explicit its dependence on (x^{on}, x^{su}, x^{sd}). When these variables are fixed, the set $\mathcal{P}_\mathcal{G}$ is polyhedral. For each generator $\ell \in \mathcal{G}$, the start-up and shut-down indicator variables $x^{su}_\ell, x^{sd}_\ell \in \{0,1\}$ are defined in terms of changes in the generator commitment status $x_\ell^{on} \in \{0,1\}$ relative to the prior base commitment status:

$$x_\ell^{on} - x_\ell^{on,0} = x_g^{su} - x_g^{sd}$$
$$x_\ell^{su} + x_\ell^{sd} \le 1$$

Here, the prior status $x_\ell^{on,0}$ is part of the data. In addition, $\mathcal{P}_\mathcal{G}$ includes generator bound constraints

$$P_k^{\min} x_\ell^{on} \le P_\ell^g \le P_\ell^{\max} x_\ell^{on}, \quad Q_\ell^{\min} x_\ell^{on} \le Q_\ell^g \le Q_\ell^{\max} x_\ell^{on},$$

as well as the ramping constraints

$$(P_\ell^{g,0} - P_\ell^{g,rd})(x_\ell^{on} - x_\ell^{su}) \le P_\ell^g \le (P_\ell^{g,0} + P_\ell^{g,ru})(x_\ell^{on} - x_\ell^{su}) + (P_\ell^{\min} + P_\ell^{g,ru}) x_g^{su}$$

where $P_\ell^{g,ru}$ and $P_\ell^{g,rd}$ are generator ramp-up and ramp-down limits, and $P_\ell^{g,0}$ is the prior base power output. The total real power output P_ℓ^g is decomposed into quantities $P_{\ell h}^g$ from the blocks $h \in H_\ell$:

$$P_\ell^g := \sum_{h \in H_\ell} P_{\ell h}^g$$

with bounds $0 \le P_{\ell h}^g \le \overline{P}_{\ell h}^g$.

Consumption Constraints for Buyers (2.52f) The set $\mathcal{P}_\mathcal{L}$ is polyhedral and characterized, for load $\ell \in \mathcal{L}$, by the constraints

$$P_\ell^{d,0} - P_\ell^{d,rd} \le P_\ell^d \le P_\ell^{d,0} + P_\ell^{d,ru}$$

where $P_\ell^{d,ru}$ and $P_\ell^{d,rd}$ are load ramp-up and ramp-down limits, and $P_\ell^{d,0}$ is prior base power consumption. The real power consumption P_ℓ^d is decomposed into blocks $P_{\ell h}^d$ as

$$P_\ell^d := \sum_{h \in H_\ell} P_{\ell h}^d.$$

with bounds $0 \leq P^d_{\ell h} \leq \overline{P}^d_{\ell h}$ for $h \in H_\ell$.

The Objective (2.52a) The benefit (cost) π^d (π^g) are modeled as piecewise linear functions with constant marginal benefit (cost) blocks, where $c^d_{\ell h}$ and $c^g_{\ell h}$ denote the corresponding coefficients. In other words,

$$\begin{aligned}
\pi^d(P^d_\ell) &:= \sum_{h \in H_\ell} c^d_{\ell h} P^d_{\ell h} \\
\pi^g(P^g_\ell, x_\ell) &:= \left(-\sum_{h \in H_\ell} c^g_{\ell h} P^g_{\ell h} - c^{on}_g x^{on}_g \right) - c^{su}_\ell x^{su}_\ell - c^{sd}_\ell x^{sd}_\ell.
\end{aligned}$$

where $x_\ell \coloneqq (x^{on}_\ell, x^{su}_\ell, x^{sd}_\ell)$.

The Power Flow Models (2.52g) The set $\mathcal{Q}^{\text{PF}}$ is bounded and denotes the feasible region associated with a given power model. In what follows, we consider several such models. We denote by $\mathcal{Q}^{\text{AC}}$ the AC power flow model in polar coordinates which can be described by (2.8d)-(2.8g), together with voltage magnitude constraints (2.8j) and thermal limit constraints (as used in [111]) given by

$$P^2_{km} + Q^2_{km} \leq \bar{i}^2_{\{k,m\}} |V_k|^2, \quad P^2_{mk} + Q^2_{mk} \leq \bar{i}^2_{\{k,m\}} |V_m|^2 \tag{2.53}$$

where $\bar{i}^2_{\{k,m\}}$ is a bound on current-squared. We denote by $\mathcal{Q}^{\text{Jabr}}$ the power flows in the variables $v^{(2)}_k, c_{km}, s_{km}$, defined by the Jabr inequalities (2.12), linearized power flows constraints (2.13), voltage magnitude constraints (2.8j), and thermal limit constraints (2.53). Next, we denote by $\mathcal{Q}^{\text{CP}}$ the region obtained by outer-approximating the nonlinear convex constraints in $\mathcal{Q}^{\text{Jabr}}$ using the linear cuts in Section 2.6. Finally, we denote by $\mathcal{Q}^{\text{DC}}$ the DC power flow model described by the constraints (2.7). The vector of variables y collects the additional variables, beyond the power flows P and Q, required to define the corresponding feasible region.

The use of $\mathcal{Q}^{\text{AC}}$ in (2.52) poses a challenge for competitive equilibrium pricing since $\mathcal{Q}^{\text{AC}}$ is nonconvex [27], [112]. In a recent paper [27], it is empirically shown that the relaxation $\mathcal{Q}^{\text{Jabr}}$

yields less biased price signals than $\mathcal{Q}^{\text{DC}}$ under two pricing rules widely used in practice. These findings are based on small cases in the GOC2 dataset. However, nonlinear relaxations are typically out of reach for nonlinear solvers on medium-to-large instances [24], [27], see also Section (2.8). Motivated by this, we explore a linear relaxation $\mathcal{Q}^{\text{CP}}$, based on outer-approximating Q^{Jabr} via linear inequalities and compare it against the power flow models $\mathcal{Q}^{\text{Jabr}}$ and $\mathcal{Q}^{\text{DC}}$. This approach aims to retain the computational advantages of linear programming while capturing key features of nonlinear power flow models.

2.10.1 Algorithmic Price Formation

A full description of an electricity price formation algorithm consists of three components [113]: (1) an operational model; (2) a parametrization, used to compute forward-looking estimates of demand and generator availability; and (3) a pricing rule. We do not address parametrization here. Instead, we focus on the trade-offs arising from the use of different power flow models, which play a critical for in the third component, namely the pricing rule.

Convex Hull and Integer-Programming Pricing

After computing an optimal allocation to the welfare problem (2.52), the market operator must determine electricity prices at each bus that ideally support the resulting equilibrium. At such prices, competitive suppliers and loads have no incentive to deviate from their bids and are incentivized to follow the prescribed dispatch. When this occurs, the price vector is said to clear the market, meaning that there is no excess demand or supply and that both buyers and sellers receive nonnegative payoffs.

In nonconvex markets, however, achieving this is challenging. Prices arise as dual variables, whose existence and economic interpretation rely on properties such as convexity and strong duality, which may fail in this setting. As a result, various pricing heuristics have been proposed to address these issues through side payments, known as uplift payments; see [114] for a detailed discussion of alternative pricing rules.

Two well-known rules designed to address nonconvexities imposed by binary commitment decisions are *convex hull* (CH) and *integer programming* (IP) pricing rules. CH pricing [29], [115] replaces nonconvex cost functions and constraints by their convex hull. Under piecewise linear bids, CH prices can be efficiently computed by relaxing the binary constraints (2.52d) and obtaining the dual variables of the associated convex problem [116]. On the other hand, IP pricing [28] requires solving a mixed-integer convex program, fixing the binary variables to their welfare-maximizing value $(x^{on*}, x^{su*}, x^{sd*})$, and then computing the dual variables associated with the power balance constraints (2.52b)-(2.52c) of the resulting convex program; see [27], [114].

A key requirement of both CH and IP pricing is that the underlying operational model, namely the Cartesian product $\mathcal{P}_{\mathcal{G}} \times \mathcal{P}_{\mathcal{L}} \times \mathcal{Q}^{\text{PF}}$, be convex. The DC power flow model $\mathcal{Q}^{\text{DC}}$ satisfies this requirement, yielding a polyhedral set at the expense of an accurate representation of power flows. In contrast, the Jabr relaxation $\mathcal{Q}^{\text{Jabr}}$ provides a more accurate model but leads to a nonlinear formulation that faces scalability challenges, as documented in this chapter. Our cutting-plane pricing algorithm (CPPA) provides an intermediate option. This is backed by our results in Section 2.8.4 where we show that empirically the dual variables associated with active power balance constraints converge to some limit.

2.10.2 Cutting-Plane Pricing Algorithm

We extend the approach in Section 2.6 to derive accurate electricity prices for large power grids. First, the algorithm constructs a set $\mathcal{Q}^{\text{CP}}$ for (2.52) by dynamically generating linear cuts that outer-approximate $\mathcal{Q}^{\text{Jabr}}$. The cuts are implemented under the adequate cut management procedure discussed in Section 2.6.2. This yields a rapid and numerically stable algorithm that gives a tight relaxation with respect to $\mathcal{Q}^{\text{Jabr}}$. Since the welfare problem (2.52) under $\mathcal{Q}^{\text{CP}}$ is a mixed-integer linear program, it is amenable to be paired with either CH or IP pricing, see Algorithm 3. The set $\mathcal{Q}^{\text{Jabr}}_{\text{linear}}$ refers to the region described by solely linear inequalities of $\mathcal{Q}^{\text{Jabr}}$, and M_{relax} stands for relaxing binary constraints in M.

Relevant input parameters for our procedure are: a time limit $T > 0$; a number of admissible

Algorithm 3 Cutting-Plane Pricing Algorithm (CPPA)

procedure CUTTING-PLANE PRICING
    Initialize $r \leftarrow 0$, $z_0 \leftarrow +\infty$
    $M \leftarrow$ Problem (2.52) with $\mathcal{Q}^{\text{Jabr}}_{\text{linear}}$
    if CH pricing **then**
        $M \leftarrow M_{\text{relax}}$
    end if
    while $t < T$ and $r < T_{ftol}$ **do**
        $z \leftarrow \min M$ and $\bar{x} \leftarrow \operatorname{argmin} M$
        If $\overline{x} \notin \mathcal{Q}^{\text{Jabr}}$:
            Compute cuts for the most violated inequalities
            Add cuts if not too parallel to cuts in M
            Drop cuts of age $\geq T_{age}$ which are not tight
        if $z - z_0 < z_0\, \epsilon_{ftol}$ **then**
            $r \leftarrow r + 1$
        else if $z - z_0 \geq z_0\, \epsilon_{ftol}$ **then**
            $r \leftarrow 0$
        end if
    end while
    Solve M
    if IP pricing **then**
        Fix binary variables in M and re-solve
    end if
    Extract dual variables of constraints (2.52b)-(2.52c) as prices λ
end procedure

iterations without sufficient objective improvement $T_{ftol} \in \mathbb{N}$; and a threshold for relative objective improvement $\epsilon_{ftol} > 0$. Our measure of cut-quality is the amount by which a solution $\bar{x}$ to M violates some nonlinear inequality in $\mathcal{Q}^{\text{Jabr}}$, among which we add a top percentage of the most violated inequalities to M. We filter the cuts by age T_{age} and how parallel they are. See 2.6.2 for further details.

2.10.3 Computational Setup and Results

We report numerical experiments on scenarios for a 617-bus instance and the three largest instances from the GOC2 dataset [111]. As in [27], we abstract from contingencies, switched shunts, and line and transformer switching. All of our experiments were run on an Intel(R) Xeon(R) Linux64 machine CPU E5-2687W v3 3.10 GHZ with 20 physical cores, 40 logical processors and 256 GB RAM. We used three commercial solvers: Gurobi version 10.0.1 [89], Artelys Knitro version 13.2.0 [78], and Mosek 10.0.43 [95]. Gurobi was used to solve the LPs for CPPA, Mosek was used to solve the Jabr SOCs, and Knitro for finding AC-feasible solutions.

For the first case study (617), we pair the Jabr SOC, DC and CP formulations with the IP pricing rule to obtain an allocation $z_{\text{R}} = (P^g, P^d, y, P, Q, x)$ for the welfare problem (2.52), and prices λ^{PF} where PF in $\{\text{Jabr}, \text{DC}, \text{CP}\}$. We compare prices and running times for the above models, see Table 2.18. We then follow [27] and simulate a *redispatch* procedure by taking the optimal allocation z_{R} and feeding it as a warm-start solution to Knitro, so to find an adjusted AC-feasible allocation. To this end, we require that all variables remain fixed up to a tolerance of 10^{-5}, except for P^g and P^d, which are allowed to vary within a tolerance of 1 p.u. to restore feasibility of the power flow equations. We denote the resulting AC-feasible allocation obtained through this redispatch procedure by $\phi(z_{\text{R}})$.

Using this AC-feasible allocation, as in [27], we compute metrics of economic efficiency, see Table 2.19: (i) make-whole payments (MWPs), (ii) global and local lost opportunity costs (GLOCs/LLOCs), and (iii) re-dispatch costs (RDCs) which we compare against total welfare. Formally, for an allocation $z = (P^g, P^d, y, P, Q, x)$ and prices λ, consider the *direct utility* as

Table 2.18: Price Statistics and Run Time for Jabr SOCP, DC and CP with IP

	Jabr SOCP Pricing			DC Pricing				CP Pricing				
Scenario	Avg	Std	Time	Avg	Std	δ	Time	Avg	Std	δ	Time	#Cuts
617-005	95.67	1.17	13.44	94.81	1.4*e*-11	1.15	0.97	95.55	1.16	0.11	9.61	1,945
617-017	118.64	1.24	43.46	118.74	5.5*e*-12	0.84	0.94	118.58	1.23	0.07	12.07	2,052
617-024	61.32	1.10	5.18	59.94	4.8*e*-11	1.50	1.07	61.26	1.09	0.07	7.41	1,917
617-062	82.26	1.25	14.22	79.06	3.7*e*-11	3.24	1.02	82.23	1.24	0.04	7.35	1,558
617-073	106.61	1.34	10.47	106.48	1.1*e*-12	0.91	1.12	106.53	1.34	0.08	7.43	1,963
617-Avg	92.90	1.22	17.35	91.81	2.1*e*-11	1.53	1.03	92.83	1.21	0.07	8.77	1,887

Table 2.19: Economic Efficiency Metrics for Jabr SOCP, DC and CP Relaxations with IP after Redispatch

	Jabr SOCP Pricing					DC Pricing					CP Pricing				
Scenario	Welfare	MWPs	GLOCs	LLOCs	RDCs	Welfare	MWPs	GLOCs	LLOCs	RDCs	Welfare	MWPs	GLOCs	LLOCs	RDCs
617-005	646,240	37,429	230,322	196,674	331	645,158	36,974	231,608	198,404	6,866	646,226	37,432	230,357	196,709	635
617-017	573,962	47,221	299,280	254,546	335	573,447	47,209	299,824	255,091	5,915	573,967	47,222	299,287	254,553	1,235
617-024	741,538	9,739	141,700	138,485	288	739,697	9,749	143,844	140,628	6,876	741,533	9,742	141,719	138,504	468
617-062	701,515	19,339	177,561	163,122	318	700,209	19,397	179,414	164,975	7,035	701,508	19,340	177,572	163,133	560
617-073	610,392	48,218	265,102	220,130	366	609,556	48,202	266,001	221,029	6,844	610,388	48,220	265,119	220,146	643
617-Avg	654,729	32,389	222,793	194,591	327	653,631	32,306	224,138	196,026	6,707	654,724	32,391	222,811	194,609	712

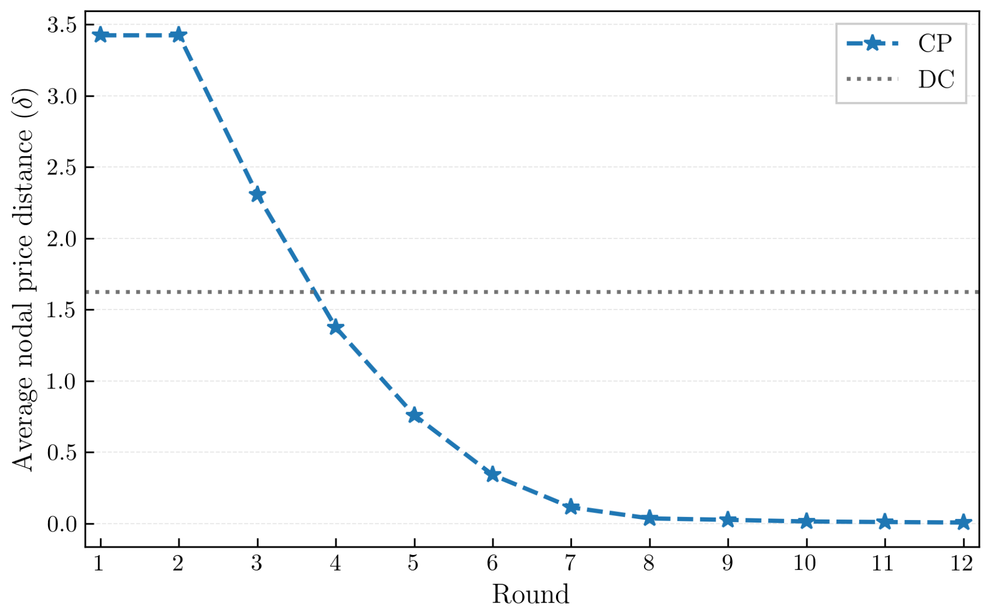

Figure 2.17: Average nodal price distance δ for the CP and DC relaxations relative to the Jabr SOCP relaxation.

$u^g_\ell(z,\lambda) := P^g_\ell \lambda_{k^g(\ell)} - (-\pi^g_\ell(P^g_\ell, x_\ell))$ or $u^d_\ell(z,\lambda) := \pi^d_\ell(z) - P^d_\ell \lambda_{k^d(\ell)}$ where $k^\tau(\ell)$ for $\tau \in \{d,g\}$ indicates the bus associated to the load or generator ℓ. The metrics are then defined as

$$MWP := \sum_{\tau\in\{d,g\}} \sum_{\ell\in\mathcal{L}\cup\mathcal{G}} [-u^\tau_\ell(\phi(z),\lambda)]^+ \tag{2.54a}$$

$$GLOC := \sum_{\tau\in\{d,g\}} \sum_{\ell\in\mathcal{L}\cup\mathcal{G}} [\sup_{z'}\{u^\tau_\ell(z',\lambda)\} - u^\tau_\ell(\phi(z),\lambda)] \tag{2.54b}$$

$$LLOC := \sum_{\tau\in\{d,g\}} \sum_{\ell\in\mathcal{L}\cup\mathcal{G}} [\sup_{P^\tau_\ell}\{u^\tau_\ell(P^\tau_\ell, P^{-\tau}_\ell,\lambda)\} - u^\tau_\ell(\phi(z),\lambda)] \tag{2.54c}$$

$$RDC := \sum_{\tau\in\{d,g\}} \sum_{\ell\in\mathcal{L}\cup\mathcal{G}} [u^\tau_\ell(z,\lambda) - u^\tau_\ell(\phi(z),\lambda)]^+ \tag{2.54d}$$

where $P^{-\tau}_\ell$ in (2.54c) denotes all variables other than P^τ_ℓ, i.e., unit τ controls only P^τ_ℓ.

For the larger instances, we pair the $\mathcal{Q}^{\text{Jabr}}$, $\mathcal{Q}^{\text{DC}}$, $\mathcal{Q}^{\text{CP}}$ formulations with the CH pricing rule. We compare prices and running times across all the above models, see Table 2.20. In Table 2.21 we report on scenarios under contingencies.

Table 2.20: Price Statistics and Run Time for Jabr SOCP, DC and CP with CH Policy

	Jabr SOCP Pricing				DC Pricing					CP Pricing						
Scenario	Avg	Std	Time	Obj	Avg	Std	δ	Time	Gap%	Avg	Std	δ	TimeLP	TimeCut	#Cuts	Gap%
17,700-019	1.01	1.07	66.06	7,789,436	15.12	0.61	13.38	19.73	9.40	14.82	1.05	13.09	19.33	285.40	69,364	9.659
17,700-020	1.22	1.62	73.55	7,757,103	14.34	1.12	12.45	21.96	9.260	14.72	1.30	12.80	23.88	285.48	70,744	9.535
17,700-021	1.26	1.58	70.68	7,705,752	13.59	1.21	11.69	26.43	8.874	14.08	1.33	12.15	34.67	290.93	71,346	9.163
17,700-089	0.80	0.71	64.22	7,792,189	14.71	0.73	13.18	24.43	9.290	14.71	1.06	13.20	22.83	304.61	70,043	9.518
17,700-094	1.09	1.25	66.70	7,778,704	13.99	1.01	12.24	20.59	8.390	14.28	1.21	12.52	21.62	287.68	70,223	8.659
17,700-Avg	1.08	1.25	68.24	7,764,637	14.35	0.94	12.59	22.63	9.043	14.52	1.19	12.75	24.47	290.82	70,344	9.307
19,402-006	1.69	0.036	150.22	233,510	1.66	0.078	0.049	28.24	0.92	1.68	0.04	0.002	48.30	331.12	147,868	0.053
19,402-010	1.33	0.030	153.81	280,531	1.32	0.11	0.055	28.31	0.45	1.33	0.03	0.002	48.13	345.71	147,857	0.047
19,402-069	1.64	0.038	155.89	478,515	1.62	0.084	0.052	26.09	0.36	1.64	0.04	0.003	44.16	307.01	143,397	0.060
19,402-077	1.35	0.022	137.88	202,145	1.34	0.043	0.029	27.48	0.67	1.35	0.02	0.003	45.89	323.92	143,657	0.095
19,402-095	1.48	0.023	144.32	225,801	1.47	0.027	0.024	27.00	0.57	1.47	0.02	0.003	51.19	331.81	141,628	0.031
19,402-Avg	1.50	0.030	148.42	284,100	1.48	0.069	0.042	27.43	0.60	1.49	0.03	0.002	47.54	327.91	144,881	0.057
31,777-011	3.95	0.92	146.40	64,740,242	4.53	11.14	1.15	37.13	0.14	3.91	0.34	0.039	56.51	299.09	137,764	0.002
31,777-012	3.10	0.30	127.06	52,756,414	3.55	7.92	1.08	36.35	0.12	3.09	0.30	0.012	40.39	262.79	137,764	0.002
31,777-015	3.36	1.82	151.64	62,000,434	3.92	12.27	1.39	36.45	0.15	3.34	1.49	0.026	40.03	262.42	136,885	0.002
31,777-019	3.12	0.87	139.62	55,916,108	3.61	8.80	1.15	37.06	0.11	3.10	0.48	0.039	40.19	275.03	136,536	0.002
31,777-051	3.42	0.29	132.43	56,825,370	4.17	11.24	1.40	36.01	0.15	3.41	0.29	0.013	33.66	298.17	120,857	0.002
31,777-Avg	3.39	0.84	139.43	58,447,713	3.96	10.27	1.23	36.60	0.14	3.37	0.58	0.026	42.16	279.50	133,991	0.002

Table 2.21: Price Statistics and Run Time for Jabr SOCP, DC and CP with CH Policy under Contingencies

	Jabr SOCP Pricing				DC Pricing					CP Pricing						
Scenario	Avg	Std	Time	Obj	Avg	Std	δ	Time	Gap%	Avg	Std	δ	TimeLP	TimeCut	#Cuts	Gap%
31,777-011 (C)	*	-	162.63	-	5.09	13.14	-	36.47	-	Infeas	-	-	78.97	-	129,689	-
31,777-012 (C)	*	-	138.43	-	3.55	7.92	-	35.93	-	Infeas	-	-	65.69	-	107,467	-
31,777-015 (C)	3.45	3.66	149.27	61,985,332	4.33	12.20	1.78	36.23	0.16	3.42	3.60	0.029	72.26	-	108,718	0.150
31,777-019 (C)	*	-	103.61	-	3.94	10.26	-	36.36	-	Infeas	-	-	80.05	-	107,606	-
31,777-051 (C)	3.50	2.79	147.85	56,810,808	4.26	11.95	1.61	36.44	0.176	3.47	2.41	0.039	37.56	-	109,477	0.003

Table 2.22: Reactive Power Prices for Jabr SOCP, DC and CP with CH

	Jabr SOCP Pricing		DC Pricing		CP Pricing		
Scenario	Avg	Std	Avg	Std	Avg	Std	δ
31,777-011	-0.0069	1.76e-5	-	-	-0.0079	0.67	0.14
31,777-012	-0.0159	2.02e-5	-	-	-0.0163	0.06	0.02
31,777-015	0.0168	1.6e-4	-	-	0.0269	1.32	0.60
31,777-019	-0.0120	8.2e-6	-	-	-0.0126	0.15	0.05
31,777-051	-0.0113	8.2e-7	-	-	-0.0124	0.08	0.10
31,777-Avg	-0.0059	4.1e-5	-	-	-0.0045	0.46	0.24

Case 617

Table 2.18 displays the average (“Avg") and standard deviation[26] (“Std") of nodal prices (\$/MWh) for each scenario, as well as the average across scenarios, for $\mathcal{Q}^{\text{Jabr}}$, $\mathcal{Q}^{\text{DC}}$, and $\mathcal{Q}^{\text{CP}}$.[27] DC produces a slightly lower average price across nodes and scenarios, while exhibiting zero variability across nodes, thereby failing to capture any network congestion. In contrast, the SOCP and CP relaxations yield prices with nonzero and very similar standard deviations. Moreover, CP prices are very close to those obtained from the Jabr SOC, as measured by average distance (“δ"), defined as $\delta(\text{PF}) := \|\lambda_{\text{PF}} - \lambda_{\text{SOC}}\|_1/|\mathcal{B}|$ for each power flow model and network $\mathcal{N} = (\mathcal{B}, \mathcal{E})$. The running times (“Time") for each model in Table 2.18 report the time per scenario, measured in seconds, as well as the average time across scenarios. The number of cuts (#Cuts) generated for each scenario is shown in the right-most column of the table.

Table 2.19 displays economic efficiency metrics (in \$) with respect to the allocation and prices produced by different formulations *after the redispatch process*. After AC-feasibility adjustments, welfare is slightly higher for the Jabr SOCP and CP relaxations than for the DC approximation. More importantly, redispatch costs are one order of magnitude smaller: in the case of DC, the RDCs are, on average, a 1.02% of total welfare, while Jabr SOCP and CPPA show relative values of 0.05% and 0.11%, respectively.

Large Cases ($\geq$ 15,000 Nodes)

Table 2.20 displays the average and standard deviation of nodal prices for all the considered scenarios. We ran the CPPA for 300 seconds and report the total time spent generating cuts (“TCut"), as well as the time it takes to solve the last LP (“TimeLP"). We report the objective (in \$) of (2.52) for Jabr SOCP (“Obj") and the % gap relative to the DC and CP objectives (“Gap%"). Overall, CPPA very quickly achieves prices that are close (small δ) to those attained by Jabr SOC; along

[26] We use the standard deviation purely as a measure of dispersion across nodes, without any statistical interpretation in terms of randomness or sampling.

[27] We conducted a thorough proof check of our experimental setting and code while deriving our results. Still, we acknowledge small discrepancies with results obtained for the 617-bus instances with [27].

the lines of what is reported in [27] for small networks. The running times of the last LPs are comparable to the those of the DC model and faster than Jabr SOC, which serves as an estimate of how long CPPA would take if warm-started. For the 17,700-bus scenarios, we find that neither DC nor CP resemble Jabr SOCP prices. Further investigation of these scenarios let us confirm that these instances are actually AC-infeasible, via a computation of a minimally infeasible system for the i2 relaxation for ACOPF, see Section 2.5.2, which is stronger than the Jabr SOC. Hence, in these scenarios, Jabr SOCP fails to capture the physics of the AC power flows.

Warm-Starts for Case 31,777

We tested the warm-starting capabilities for CPPA by considerably stressing the five scenarios of the 31,777-bus network (see Table 2.21). To this end, first we ran CPPA for 400 seconds and stored the cuts computed in the last round. We observe in Figure 2.17 the improvement, in terms of average distance δ, of every round of the algorithm for scenario 51. We created stressed instances by loading branch contingencies, at most 1% of the total branches of the network, provided in the GOC2 dataset for each scenario. Next, we warm-started CPPA with the previously computed cuts (“#Cuts"), and run CPPA for only one round. Mosek failed to converge for scenarios 11, 12 and 19 under contingencies (“$*$"), whereas CPPA quickly identified these scenarios as infeasible (“Infeas"), and DC failed to do so. In fact, since these scenarios are infeasible, the price signals reported by the DC model are entirely inaccurate, with average prices being significantly underestimated. For scenarios 15 and 51, CPPA attains, twice as fast, prices that are very close to Jabr SOC, with $\delta(\text{CP})$ values being at least an order of magnitude smaller than $\delta(\text{DC})$ values. In Table 2.22 we display reactive power prices for the SOCP and CP relaxations. Note that the average reactive power prices of the CP relaxation are very close to the Jabr SOCP, however, there is a larger nodal price variation. We note that the DC approximation does not produce reactive power prices.

2.11 AC-Feasibility is Weakly $\mathbb{NP}$-hard on Trees: A Corrected Proof

In this section, we revisit the proof of Theorem [21, Theorem 1], which establishes the weak $\mathbb{NP}$-hardness of AC-feasibility (see below for a formal statement of the problem). We identify a gap in the proposed reduction from the weakly $\mathbb{NP}$-hard problem SUBSET SUM, arising from the encoding of numerical quantities. Specifically, the argument implicitly assumes that *any* real number obtained from rational numbers through the sine and cosine functions admits a polynomial-size encoding in the input size. This assumption does not hold in general—for example, $\cos(1)$ is irrational. We restate the construction using their notation and provide a corrected proof that addresses this issue.

Consider a power system with undirected network $\mathcal{N} := (\mathcal{B}, \mathcal{E})$. We denote by $\mathcal{G} \subseteq \mathcal{B}$ the set of generators and by $\mathcal{L} \subseteq \mathcal{B}$ the set of loads (we assume that at each bus there is at most one generator). For each line $\{k, m\} \in \mathcal{E}$, we denote its conductance by $g_{\{k,m\}} \geq 0$ and its susceptance by $b_{\{k,m\}} \leq 0$. We assume there are no transformers nor shunt elements, which implies symmetric line admittance matrices.

Our simplified AC-FEASIBILITY problem (cf. [21]) has as input fixed demands for active (P^d) and reactive (Q^d) power, all voltages magnitudes are fixed to one, and it assumes that each line $\{k, m\}$ has a fixed capacity $U_{\{k,m\}} \geq 0$. We require non-negative active power generation. Moreover, our proof assumes that each line $\{k, m\}$ has susceptance $b_{\{k,m\}} < 0$ and conductance $g_{\{k,m\}} > 0$. Therefore, our AC-FEASIBILITY problem consists in finding the phase angles θ_k, active (P_k^g) and reactive (Q_k^g) power generation, active power flows P_{km}, P_{mk}, and reactive power flows

Q_{km}, Q_{mk} satisfying the following non-linear system:

$\forall$ bus $k \in \mathcal{B}$:

$$\sum_{\{k,m\} \in \delta(k)} P_{km} = P_k^g - P_k^d \tag{2.55}$$

$$\sum_{\{k,m\} \in \delta(k)} Q_{km} = Q_k^g - Q_k^d \tag{2.56}$$

$\forall$ branch $\{k, m\} \in \mathcal{E}$:

$$P_{km} = g_{\{k,m\}}(1 - \cos(\theta_k - \theta_m)) - b_{\{k,m\}} \sin(\theta_k - \theta_m) \tag{2.57}$$

$$P_{mk} = g_{\{k,m\}}(1 - \cos(\theta_k - \theta_m)) + b_{\{k,m\}} \sin(\theta_k - \theta_m) \tag{2.58}$$

$$Q_{km} = -b_{\{k,m\}}(1 - \cos(\theta_k - \theta_m)) - g_{\{k,m\}} \sin(\theta_k - \theta_m) \tag{2.59}$$

$$Q_{mk} = -b_{\{k,m\}}(1 - \cos(\theta_k - \theta_m)) + g_{\{k,m\}} \sin(\theta_k - \theta_m) \tag{2.60}$$

$$\max\left\{P_{km}^2 + Q_{km}^2, P_{mk}^2 + Q_{mk}^2\right\} \le U_{\{k,m\}} \tag{2.61}$$

$\forall$ generator $k \in \mathcal{G}$:

$$P_k^g \ge 0 \tag{2.62}$$

$\forall$ generator $k \notin \mathcal{G}$:

$$P_k^g = 0. \tag{2.63}$$

Remark 2.23. We note that imposing a thermal limit $U_{\{k,m\}} \le 2(g_{\{k,m\}}^2 + b_{\{k,m\}}^2)$ on each line is equivalent to imposing a maximum phase angle difference $\bar{\Delta}_{\{k,m\}} \in [0, \pi/2]$ (see also Appendix A.1). Indeed, let $A \coloneqq 1 - \cos(\theta_k - \theta_m)$ and $B \coloneqq \sin(\theta_k - \theta_m)$, then for any line $\{k, m\}$

we have that

$$\begin{aligned}
P_{km}^2 + Q_{km}^2 &= (g_{\{k,m\}}A - b_{\{k,m\}}B)^2 + (b_{\{k,m\}}A + g_{\{k,m\}}B)^2 \\
&= (g_{\{k,m\}}^2 + b_{\{k,m\}}^2)(A^2 + B^2) \\
&= (g_{\{k,m\}}^2 + b_{\{k,m\}}^2)(1 - 2\cos(\theta_k - \theta_m) + \cos^2(\theta_k - \theta_m) + \sin^2(\theta_k - \theta_m)) \\
&= 2(g_{\{k,m\}}^2 + b_{\{k,m\}}^2)(1 - \cos(\theta_k - \theta_m)).
\end{aligned}$$

Hence, $P_{km}^2 + Q_{km}^2 \le U_{\{k,m\}}$ if and only if $0 \le \cos(\theta_k - \theta_m)$, i.e., $|\theta_k - \theta_m| \le \pi/2$. The bijection is given by

$$y \mapsto \arccos(1 - (y/2(g_{\{k,m\}}^2 + b_{\{k,m\}}^2))) \tag{2.64}$$

for every $0 \le y \le 2(g_{\{k,m\}}^2 + b_{\{k,m\}}^2)$. Then we can define the maximum phase angle difference for branch $\{i,j\}$ as

$$\bar{\Delta}_{\{k,m\}} := \arccos(1 - (U_{\{k,m\}}/2(g_{\{k,m\}}^2 + b_{\{k,m\}}^2))).$$

Therefore, we can formally define our problem as follows:

AC-FEASIBILITY: Given a network $\mathcal{N} = (\mathcal{B}, \mathcal{E})$, sets of generators and loads $\mathcal{G}, \mathcal{L} \subseteq \mathcal{B}$ and rational vectors $P^d, Q^d \in \mathcal{Q}^{\lvert\mathcal{B}\rvert}$, $g, b, U \in \mathcal{Q}^{\lvert\mathcal{E}\rvert}$, is there a feasible solution to the system of constraints (2.55)-(2.63)?

Theorem 2.24. AC-FEASIBILITY *on trees is weakly* $\mathbb{NP}$*-hard.*

Proof. We prove that for star networks AC-FEASIBILITY is weakly $\mathbb{NP}$-hard by exhibiting a reduction from SUBSET-SUM. Given a set $M \subseteq \mathbf{N}$ and a number $w \in \mathbf{N}$, SUBSET-SUM consists in deciding whether there exists $V \subseteq M$ such that $\sum_{x \in V} x = w$. If such a set V exists, we call the problem instance (M, w) solvable. Before defining the injective mapping between SUBSET-SUM and AC-FEASIBILITY instances, we define some constants which will be key for our proof. Let $g := 1$, $b := -5/4$, and $\bar{U} := 41/20$, hence the bijection (2.64) implies $\bar{\Delta} := \arccos(3/5) \approx 0.9273$. Clearly $\bar{\Delta} \in (0, \pi/2)$ and $\sin(\bar{\Delta}) = 4/5$.

Given an arbitrary instance (M, v) of SUBSET-SUM, we define an AC-network $\mathcal{N}_{(M,w)}$ as follows. Let $\mathcal{G} := M$; $\mathcal{L} := \{\ell\}$; $\mathcal{E} := \{\{x, \ell\} : x \in M\}$, where each line $\{x, \ell\}$ has conductance $g_{\{x,\ell\}} := x > 0$ and susceptance $b_{\{x,\ell\}} := (-5/4)x < 0$ and capacity $U_{\{x,\ell\}} := \bar{U}x^2$; $P^d_\ell := (3/5)w$, $Q^d_\ell := (-13/10)w$ and $P^d_i := 0, Q^d_i := 0$ for all buses $i \neq \ell$. Therefore, this encoding $((\mathcal{G} \cup \mathcal{L}), \mathcal{E}, g, b, U, P^d, Q^d)$ is polynomial in the size of (M, w), thereby resolving the encoding issue identified above.

(M, w) **is solvable** $\implies \mathcal{N}_{(M,w)}$**.** Let V be a solution for (M, w). We define $\hat{\theta}_\ell := 0$; $\forall x \in V$, $\hat{\theta}_x := \bar{\Delta}$, $\hat{P}_{\ell x} := (-3/5)x$, $\hat{Q}_{\ell x} := (13/10)x$, $\hat{P}_{x\ell} := (7/5)x$, $\hat{Q}_{x\ell} := (-3/10)x$, $\hat{P}^g_{x\ell} := (7/5)x$, $\hat{Q}^g_{x\ell} := (-3/10)x$; and $\forall x \in M \setminus V$, $\hat{\theta}_x := 0$, $\hat{P}_{\ell x} := 0$, $\hat{Q}_{\ell x} := 0$, $\hat{P}_{x\ell} := 0$, $\hat{Q}_{x\ell} := 0$, $\hat{P}^g_{x\ell} := 0$, and $\hat{Q}^g_{x\ell} := 0$. It can be easily checked that this assignment $\hat{\theta}, \hat{P}^g, \hat{Q}^g, \hat{P}, \hat{Q}$ satisfies the active and reactive power flow definition equations (2.57)-(2.60). Line capacitiy constraints (2.61) are satisfied since

$$\hat{P}^2_{\ell x} + \hat{Q}^2_{\ell x} = (205/100)x^2 = U_{\{x,\ell\}},$$
$$\hat{P}^2_{x\ell} + \hat{Q}^2_{x\ell} = (205/100)x^2 = U_{\{x,\ell\}}.$$

Constraints (2.55)-(2.56) for active and reactive power balance at bus ℓ are satisfied since

$$\sum_{\{\ell,j\}\in\delta(\ell)} \hat{P}_{\ell x} = \sum_{j\in M}(-3/5)x = (-3/5)w = -P^d_\ell,$$
$$\sum_{\{\ell,j\}\in\delta(\ell)} \hat{Q}_{\ell x} = \sum_{j\in M}(13/10)x = (13/10)x = -Q^d_\ell.$$

Power balance constraints at buses $x \in M$ are clearly satisfied. Finally, non-negative generation constraints (2.62) are satisfied since $\hat{P}^g_x = (7/5)x > 0$ for all $x \in \mathcal{G}$. In consequence, our solution is feasible.

$\mathcal{N}_{(M,w)}$ **has a feasible solution** $\implies$ (M,w) **is solvable.** Let $\hat{\theta}, \hat{P}^g, \hat{P}, \hat{Q}$ be a feasible solution. Since $0 < \bar{\Delta}$, $|\hat{\theta}_x - \hat{\theta}_\ell| \le \bar{\Delta}$ and

$$\hat{P}_{x\ell} = (1 - \cos(\hat{\theta}_x - \hat{\theta}_\ell)) + (5/4)\sin(\hat{\theta}_x - \hat{\theta}_\ell) = \hat{P}^g_x \ge 0 \tag{2.65}$$

for all $x \in M$, then $\hat{\theta}_x - \hat{\theta}_\ell \ge 0$. Indeed, suppose on the contrary that $\hat{\theta}_x - \hat{\theta}_\ell < 0$. Then dividing (2.65) by $\sin(\hat{\theta}_x - \hat{\theta}_\ell) < 0$, we have that $\tan((\hat{\theta}_x - \hat{\theta}_\ell)/2) \le (-5/4)$. Since $-\bar{\Delta}/2 \le (\hat{\theta}_x - \hat{\theta}_\ell)/2$ and $\tan$ being increasing in $(-\pi/2, 0)$, then $\tan(-\hat{\Delta}/2) \le \tan((\hat{\theta}_x - \hat{\theta}_\ell)/2) \iff (1/3) \le (-5/4)$ which is a contradiction. We define $V := \{x \in M \ : \ \hat{\theta}_x - \hat{\theta}_\ell > 0\}$ which is non-empty since $\theta_x - \theta_\ell = 0$ for all $x \in \mathcal{G}$ would imply $\sum_{x \in \mathcal{G}} P^g_x = 0$ but $P^d_\ell > 0$. Moreover, active and reactive power balance constraints at ℓ become $\sum_{x \in M} \hat{P}_{\ell x} = (-3/5)w$ and $\sum_{x \in M} \hat{Q}_{\ell x} = (-13/10)w$. By rearranging the sums we have that

$$\begin{aligned}\sum_{x \in M} \hat{P}_{\ell x}(-3/5) = \sum_{x \in M} \hat{Q}_{\ell x}(10/13) &\iff 0 = \sum_{x \in M} \left(\hat{P}_{\ell x}(-3/5) - \hat{Q}_{\ell x}(10/13)\right) \\ &\iff 0 = \sum_{x \in M} \left(\hat{P}_{\ell x}(13/10) - \hat{Q}_{\ell x}(-3/5)\right).\end{aligned} \tag{2.66}$$

We claim that for every $x \in M$ the summand $\hat{P}_{\ell x}(13/10) - \hat{Q}_{\ell x}(-3/5)$ is non-positive. Indeed,

$$\begin{aligned}\hat{P}_{\ell x}(13/10) - \hat{Q}_{\ell x}(-3/5) &= \left((1 - \cos(\hat{\theta}_x - \hat{\theta}_\ell)) - (5/4)\sin(\hat{\theta}_x - \hat{\theta}_\ell)\right)(13/10) \\ &\quad - \left((5/4)(1 - \cos(\hat{\theta}_x - \hat{\theta}_\ell)) + \sin(\hat{\theta}_x - \hat{\theta}_\ell)\right)(-3/5) \\ &= (82/40)(1 - \cos(\hat{\theta}_x - \hat{\theta}_\ell)) - (41/40)\sin(\hat{\theta}_x - \hat{\theta}_\ell).\end{aligned}$$

Hence $\hat{P}_{\ell x}(13/10) - \hat{Q}_{\ell x}(-3/5) \le 0$ if and only if

$$2(1 - \cos(\hat{\theta}_x - \hat{\theta}_\ell)) - \sin(\hat{\theta}_x - \hat{\theta}_\ell) \le 0. \tag{2.67}$$

If $(\hat{\theta}_x - \hat{\theta}_\ell) = 0$ the latter clearly holds. If $(\hat{\theta}_x - \hat{\theta}_\ell) > 0$ then dividing the inequality (2.67) by $\sin(\hat{\theta}_x - \hat{\theta}_\ell) > 0$ gives us the equivalence

$$\hat{P}_{\ell x}(13/10) - \hat{Q}_{\ell x}(-3/5) \leq 0 \iff \tan((\hat{\theta}_x - \hat{\theta}_\ell)/2) \leq 1/2. \tag{2.68}$$

Since $\tan(\bar{\Delta}/2) = 1/2$, $(\hat{\theta}_x - \hat{\theta}_\ell) \leq \bar{\Delta} < \pi/2$ and the fact that $\tan$ is strictly increasing in $(0, \pi/2)$, the claim follows. Moreover, $\hat{P}_{\ell x}(13/10) - \hat{Q}_{\ell x}(-3/5) = 0 \iff \hat{\theta}_x - \hat{\theta}_\ell \in \{0, \bar{\Delta}\}$.

The previous argument implies that all summands in (2.66) must be 0, hence given our choice of V, we have that $\forall x \in V$, $\hat{\theta}_x - \hat{\theta}_\ell = \bar{\Delta}$. This implies that $\hat{P}_{\ell x} = (-3/5)x$. Therefore, from active power balance at ℓ we obtain $\sum_{x \in V} \hat{P}_{\ell x} = \sum_{x \in V} (-3/5)x = (-3/5)w$ which proves $\sum_{x \in V} x = w$. $\square$

2.12 Conclusions

In this chapter, we developed a fast linear cutting-plane method used to obtain tight relaxations for even the largest single and multi-period ACOPF instances, by appropriately outer-approximating the SOCP relaxations. The resulting formulations can be constructed and solved efficiently using mature linear programming technology, supported by careful cut management. We also provided theoretical justification for the strength of the underlying SOCP relaxations and for the use of their linear outer approximations, and showed that these can be further strengthened with SDP-based cuts.

The central focus on this paper concerns *reoptimization*. We showed that our procedure possesses efficient warm-starting capabilities – previously computed cuts, for some given instance, can be re-utilized and loaded into new runs of *related* instances, hence leveraging previous computational effort. As a main contribution we demonstrated, through extensive numerical testing in medium to (very) large multi-period instances, that the warm-start feature for our cutting-plane algorithm yields tight and accurate bounds far faster than otherwise possible. It is worth noting that this capability stands in contrast to what is possible using nonlinear (convex) solvers.

More broadly, our results highlighted the effectiveness of linearly constrained relaxations for ACOPF. In particular, we showed that linear relaxations with convex quadratic objectives provided robust and reliable lower bounds, offering both theoretical guarantees and strong numerical performance, in sharp contrast to what nonlinear relaxations such as SOCPs can offer.

Finally, we demonstrated the practical impact of our framework through a scalable pricing algorithm for electricity markets. Our experiments show that the proposed approach delivers accurate price signals and lower redispatch costs relative to DC approximations, while achieving accuracy comparable to SOCP relaxations at similar computational cost when warm-started.

Chapter 3: Symmetry and Relaxations in Integer Programming

3.1 Introduction

Understanding what makes an integer program (IP) challenging for state-of-the-art algorithms is a fundamental step toward developing more effective solvers.[1] While it is reasonable to assume that a larger number of variables and constraints would slow down any algorithm for solving an IP, there are other, more subtle features of IPs that are problematic due to the techniques we currently employ. One such feature is *symmetry*, informally defined as any transformation of a mathematical object that leaves its fundamental properties unchanged.

Mathematicians usually welcome a high degree of symmetry, as it often implies strong properties of the objects under consideration. For instance, systems of linear equalities $Ax = b$ can be solved faster in practice when $A = A^\top$, because factorizations of A are guaranteed to exist in this case, see, e.g., [119]. In contrast, symmetry is usually problematic in integer programming. Even small-sized symmetric problems are often hard to solve [120], [121], [122], [123], [124]. In his survey on the topic [121], Margot states that *the trouble comes from the fact that many subproblems in the enumeration tree are isomorphic, forcing a wasteful duplication of effort*–but as we will see, this is not the only problem caused by symmetry.

Understanding how to tackle symmetric IPs is essential, as some degree of symmetry is prevalent in both theory and practice. The structure of the problem itself sometimes causes symmetry – for instance, if we wish to schedule jobs on identical machines, or if we solve a combinatorial optimization problem on a highly symmetric graph.

[1]This chapter is based on joint work with Yuri Faenza, Víctor Verdugo, and José Verschae [117], with additional exposition and results. An earlier version of this work appeared in the proceedings of the 27th Conference on Integer Programming and Combinatorial Optimization (IPCO) [118]. I gratefully acknowledge support from the Air Force Office of Scientific Research (AFOSR) under grant FA9550-23-1-0697, as well as travel awards from the Mixed Integer Programming (MIP) Workshop, organized by the Mixed Integer Programming Society (MIPS), and the 2026 IOE–ISyE-MS&E Rising Stars Workshop.

Whichever the cause, more than 36% of the instances from MIPLIB 2010 admit nontrivial symmetries; see the work by Pfetsch and Rehn [124] for details. Moreover, symmetries have played an important role for theoretical work, both from a polyhedral perspective (e.g., [125], [126]), as a device to construct lower bounds of relaxations (e.g., [127], [128]), and a way of reducing the dimensionality of convex relaxations (e.g., [123], [129]), to mention a few.

In this chapter, we investigate how symmetry can harm state-of-the-art techniques for solving binary integer programs. Our work differs from the existing literature in three main aspects.

- First, rather than focusing on the role of symmetry in generating many "equivalent" integer solutions, we study the relationship between symmetry and the performance of LP- and SDP-based cut operators, broadly referred to as hierarchies. These methods iteratively strengthen LP relaxations of integer programs. Our work is motivated by the fundamental question: can one characterize the conditions under which a symmetric optimal face of a relaxation is eliminated at a fractional node of the branch-and-cut tree? We provide a positive answer; see Theorem 3.1.

- Second, most prior work has focused on specific problem classes, such as matchings [130], [131], knapsack [127], [132], max-cut and sparsest-cut [128], [133], [134], scheduling [135], [136], vertex cover [137], and planted clique [138], [139]. In contrast, with the notable exception of the work by Kurpisz et al. [140], our assumptions only rely on a well-established group-theoretic parametrization of the symmetry of the underlying polytope, namely k-*transitivity*.

- Third, with the exception of [131], the literature has largely focused on determining how many rounds are required to close the integrality gap. In contrast, we address the more fundamental question of how many rounds are needed to *improve the bound*.

Our main result is the following.

Theorem 3.1. *Let $k, n \in \mathbb{N}$ such that $k < n$, and let $P \subseteq [0,1]^n$ be a $(k+1)$-transitive integer-empty polytope. Then the following statements are equivalent:*

(A) *P intersects all of the $(n-k)$-dimensional faces of the hypercube $[0,1]^n$;*

(B) $\mathrm{SA}^k(P) \neq \emptyset$*;*

(C) $\mathrm{LS}^k(P) \neq \emptyset$*;*

(D) $\mathrm{LS}_0^k(P) \neq \emptyset$.

We defer to Section 3.2 the definition of the hierarchies SA^k, LS^k, and LS_0^k. A polytope P is G-invariant for a permutation group G over $[n] := \{1,\ldots,n\}$, if under the action of the elements of G, given by $\pi x := (x_{\pi^{-1}(1)},\ldots,x_{\pi^{-1}(n)})$, we have that $P = \pi(P)$ for all $\pi \in G$. A group G is k-transitive if it can map any ordered sequence of k indices in $[n]$ to any other ordered sequence of k indices in $[n]$ by some element in G. A polytope is k-transitive if it is G-invariant for a k-transitive group G; observe that $(k+1)$-transitivity implies k-transitivity. Thus, the larger the k, the *more symmetric* P is.

A high-level description of Theorem 3.1 is as follows. Assume that the polytope $P \subseteq [0,1]^n$ is integer-empty and "symmetric enough", i.e., $(k+1)$-transitive. Theorem 3.1 implies that we can decide whether k rounds of any of the three hierarchies certify that P is integer-empty just by looking at the intersection of P with the faces of the hypercube $[0,1]^n$. In particular, one of the hierarchies gives us such a certificate if and only if all of them do.

It is well-known that, for any $k \in \mathbb{N}$ and polytope $P \subseteq [0,1]^n$, $\mathrm{SA}^k(P) \subseteq \mathrm{LS}^k(P) \subseteq \mathrm{LS}_0^k(P)$, and in general they can be significantly different. In contrast, in our setting, we observe a *hierarchy collapse*: the three methods are equally effective in determining whether the polytope is integer-empty. This result may explain why the integer-emptiness of symmetric polytopes is hard to prove for LP-based hierarchies: the strongest of the three operators does not outperform the weakest one. Moreover, if the other features stay constant, a "more symmetric" polytope (i.e., k-transitive for larger k) will "resist" these operators for at least as many rounds as a "less symmetric" polytope.

Our next result establishes that the non-polyhedral hierarchy LS_+ is equally effective as the above polyhedral hierarchies in determining integer-emptiness, if the polytope P is not full-dimensional. The technical condition $\dim P \leq n-1$ is key in our argument. It is known that ${\mathrm{LS}^k}_+(P) \subseteq \mathrm{LS}^k(P)$

but SA^k and LS_+^k are incomparable in general.

Theorem 3.2. *Let $P \subseteq [0,1]^n$ be a $(k+1)$-transitive polytope with $\dim P \leq n-1$. Then $\mathrm{LS}_0^k(P) \neq \emptyset$ if and only if $\mathrm{LS}_+^k(P) \neq \emptyset$.*

The following corollary is immediate.

Corollary 3.3. *Let $k, n \in \mathbb{N}$ such that $k < n$, and let $P \subseteq [0,1]^n$ be a $(k+1)$-transitive integer-empty polytope with $\dim P \leq n-1$. Then the following statements are equivalent:*

(A) *P intersects all of the $(n-k)$-dimensional faces of the hypercube $[0,1]^n$;*

(B) $\mathrm{SA}^k(P) \neq \emptyset$*;*

(C) $\mathrm{LS}_+^k(P) \neq \emptyset$*;*

(D) $\mathrm{LS}_0^k(P) \neq \emptyset$.

Contributions

- We provide a unified perspective on several well-known examples from the literature where lift-and-project hierarchies exhibit poor performance [141], [142], [143], [144], [145]. This behavior is characterized by slow progress in closing the integrality gap, often requiring many iterations to improve upon weak LP bounds (Sections 3.7)

- We introduce the notion of k-transitivity – well-known in computational group theory (see, e.g., [146, Sec. 7]) – into integer programming, and use it to parametrize the degree of symmetry of polytopes (Sections 3.2.2 and 3.4.5).

- Motivated by weak LP relaxations in covering problems such as the Steiner Triple Covering problem (see Section 3.6.2), we study the conditions under which a symmetric optimal face of a relaxation is eliminated at a fractional node of the branch-and-cut tree. We identify this condition for symmetric integer programs under cuts derived from lift-and-project hierarchies (Section 3.5).

- We define the notion of k-resistance for polytopes, and show that even applying $\Theta(n)$ mixed-degree RLT factors to a class of Minimum Knapsack instances does not improve the LP bound (Section 3.5).

- Our main result (Theorem 3.1) establishes such a characterization, and we show that its assumptions are essentially tight (Section 3.7.1). We illustrate the applicability of our results through several applications to problems from the literature (Section 3.6).

- We develop a general procedure for analyzing Sherali–Adams rank by extending the proof technique underlying Theorem 3.1. The resulting framework provides a systematic way to construct fractional feasible solutions that certify lower bounds on the hierarchy (Section 3.9). We demonstrate its effectiveness on three different problems, obtaining bounds that match those in the literature (Sections 3.9.3, 3.9.4, and 3.6.3).

- We provide proofs of several folklore results relating symmetry, hierarchies, and polyhedral geometry (Section 3.4).

- We identify and correct an incorrect statement in the literature whose original proof relies on invalid numerical assumptions (Section 3.4.3).

3.2 Background

For $i, j \in \mathbb{Z}_{\geq 0}$ and $x \in \mathbb{R}^n$ we write $\binom{i}{j} := 0$ and $\prod_j^i x_j := 1$ if $j > i$. Moreover, we define $\binom{i}{r} := 0$, $\prod_j^r x_j := 1$ if $r \in \mathbb{Z}_{<0}$ and $[0] := \emptyset$. We denote by $\mathbf{1}$ and $\mathbf{0}$ the all-ones vector and the all-zeros vector, respectively. Moreover, we write $\binom{n}{\leq i} := \sum_{j=0}^{i} \binom{n}{j}$. For $\emptyset \neq S \subseteq [n]$ and subsets $J_0, J_1 \subseteq S$, we say that (J_0, J_1) is a 2-partition of S if $J_0 \cup J_1 = S$ and $J_0 \cap J_1 = \emptyset$, and we denote it as $J_0 \sqcup J_1 = S$. For any vectors $y \in \mathbb{R}^{\binom{n}{\leq j}}$ and $z \in \mathbb{R}^{\binom{n}{\leq i}}$, let (y, z) denote their vertical concatenation, i.e., $(y, z) := \binom{y}{z}$. Whenever $I = \{j\}$, we simply write $y_I = y_j$. Throughout this chapter, we let $m \in \mathbb{N}$ and $P := \{x \in \mathbb{R}^n \,:\, a_i^\top x \geq b_i, i \in [m]\} \subseteq [0, 1]^n$. We also write this system compactly as $Ax \geq b$, where $A \in \mathbb{R}^{m \times n}$ has rows $a_i^\top$ and $b \in \mathbb{R}^m$. We denote the integer hull of P by $P_I := \text{conv}(P \cap \{0, 1\}^n)$.

3.2.1 Integer Programming Hierarchies

We now present the Linear Programming hierarchies of interest for this work. For a polytope P as above, let

$$P^{i,1} \coloneqq P \cap \{x \in \mathbb{R}^n \,:\, x_i = 1\} \quad \text{and} \quad P^{i,0} \coloneqq P \cap \{x \in \mathbb{R}^n \,:\, x_i = 0\} \text{ for } i \in [n].$$

We denote by P_i the *Balas-Céria-Cornuéjols operator* [147] associated with the variable x_i, defined as $P_i(P) \coloneqq \text{conv}(P^{i,1} \cup P^{i,0})$. This operator serves as a fundamental building block of LP hierarchies.

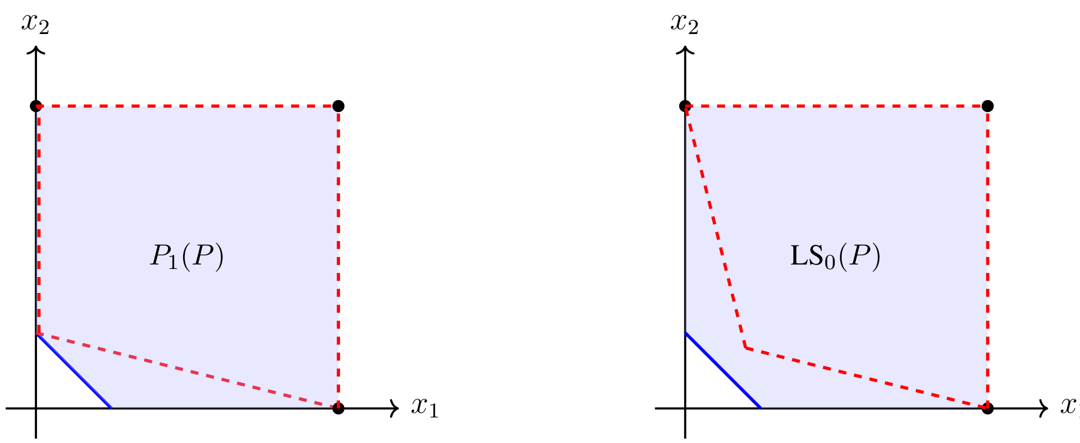

Figure 3.1: Polytope $P \coloneqq \{x \in \mathbb{R}^2 \,:\, x_1 + x_2 \geq \frac{1}{4},\, 0 \leq x_1, x_2 \leq 1\}$ (shaded in light blue) together with $P_1(P)$ (left) and $\text{LS}_0(P)$ (right), whose contours are shown in red dashed lines.

The *lift-and-project closure* of P can be written as $\text{LS}_0(P) \coloneqq \bigcap_{i\in[n]} P_i(P)$. We will employ a useful characterization of $\text{LS}_0(P)$ (see, e.g., [143], [148]):

$$\text{LS}_0(P) = \{x \in \mathbb{R}^n \,: \exists Y \in \mathbb{R}^{(n+1)\times(n+1)},\, Ye_i, Y(e_0 - e_i) \in K(P), \forall i \in [n],$$
$$Ye_0 = Y^\top e_0 = \text{diag}(Y) = (1, x)^\top\},$$

where the matrices Y have columns and rows indexed by $\{0, 1, \ldots, n\}$, e_i denotes the i-th canonical vector, and $K(P) \coloneqq \{(\lambda, \lambda x)^\top : \lambda \in \mathbb{R}_+,\, x \in P\}$. We denote by LS_0^k the k iterative applications of LS_0. That is, $\mathrm{LS}_0^2(P) = \mathrm{LS}_0(\mathrm{LS}_0(P))$, $\mathrm{LS}_0^3(P) = \mathrm{LS}_0(\mathrm{LS}_0^2(P))$, etc. The *Lovász-Schrijver* operator, LS, requires the lifting matrices Y to be symmetric, while its PSD variant, LS_+, further requires Y to be PSD. We similarly denote repeated applications of LS and LS_+ by LS^k and LS_+^k. Following the presentation in [149], we define the *Sherali-Adams level* k relaxation [150] $\mathrm{M}^k(P)$ of P as follows.

(i) **Extend**: For every row $a_i^\top x \geq b_i$, with $i \in [m]$, every set $S \subseteq [n]$, $|S| = k$ and every 2-partition $J_0 \sqcup J_1 = S$, introduce a new constraint

$$\prod_{i \in J_1} x_i \prod_{j \in J_0} (1 - x_j) \cdot (a_i^\top x - b_i \geq 0). \tag{3.1}$$

(ii) **Linearize**: multilinearize each of the constraints in the previous step by replacing every monomial $\prod_{i \in I} x_i$ with a new variable y_I, define $y_\emptyset := 1$, and set $x_i^2 = x_i$:

$$\mathrm{M}^k(P) := \left\{ y \in \mathbb{R}^{\binom{n}{\leq k+1}} \ : \ \text{multilinearized inequalities (3.1)} \right\}.$$

We denote by $\mathrm{SA}^k(P)$ the (coordinate) projection of $\mathrm{M}^k(P)$ onto the space of original variables, i.e.,

$$\mathrm{SA}^k(P) := \{x \in \mathbb{R}^n : \text{there exists } y \in \mathrm{M}^k(P),\ y_i = x_i,\ i \in [n]\}.$$

Remark 3.4. Standard definitions of the Sherali-Adams hierarchy either augment the defining system $Ax \geq b$ with the redundant inequality $1 \geq 0$, see [149], or augment $Ax \geq b$ with the box constraints $\mathbf{0} \leq x \leq \mathbf{1}$ and explicitly include the corresponding consistency constraints in the lifted space, see [150], [151]. Since we assume that $P \subseteq [0,1]^n$, these additional inequalities are valid for P. Throughout this document, we omit them. Lemma 3.24 shows that the Sherali-Adams relaxation is invariant under augmenting $Ax \geq b$ with valid inequalities. Consequently, our convention is equivalent to these standard definitions.

The multivariate polynomial factors in (3.1) are known as *juntas*, or *reformulation-linearization technique (RLT) factors*. Given a set $S \subseteq [n]$, each 2-partition $J_0 \sqcup J_1 = S$ induces a polynomial of degree $|S|$, which we refer to as a $|S|$-*junta*, and denote by

$$F_{|S|}(J_0, J_1)(x) := \prod_{i \in J_1} x_i \prod_{j \in J_0} (1 - x_j).$$

We denote its multilinearization by $f_{|S|}(J_0, J_1)(y)$. When the degree is clear from the context, we omit the subscript $|S|$ from both $F_{|S|}$ and $f_{|S|}$.

Remark 3.5. Throughout this document, a k-junta denotes a polynomial of degree exactly k, whereas [149] uses the term for polynomials of degree at most k. This distinction is inconsequential for the Sherali-Adams hierarchy: by [150, Lemma 1], the inequalities generated by j-juntas with $j < k$ are implied by those generated by k-juntas. Accordingly, we restrict the Sherali-Adams construction only to k-juntas. For completeness, we reprove this simple fact in Appendix A.8; see Lemma A.7.

It is routine to check that

$$f_{|S|}(J_0, J_1)(y) = \sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup H}.$$

Consequently, the multilinearization of $F_{|S|}(J_0, J_1)(x)\big(a_i^\top x - b_i\big) \geq 0$ is

$$\left(\sum_{j \in J_1} a_{i,j} - b_i\right) f_k(J_0, J_1) + \sum_{j \in [n] \setminus S} a_{i,j}\, f_{k+1}(J_1 \cup \{j\}, J_0) \geq 0. \tag{3.2}$$

Therefore, an explicit inequality description of $\mathrm{M}^k(P)$ is obtained by considering all $i \in [m]$, all sets $S \subseteq [n]$ of cardinality k, and all 2-partitions $J_0 \sqcup J_1 = S$.

The procedure obtained by applying the multiplication (Step (i)) and multilinearization (Step (ii)) steps to only a subset of the inequalities $Ax \geq b$, and by considering only a subset of the juntas of degree at most k, is known as the *reformulation–linearization technique (RLT)*; see [150], [151]. We occasionally refer to juntas as *RLT factors*.

Example 3.6. Consider P from Figure 3.1. In [152] it is shown that using the procedure described above we obtain an algebraic characterization of $P_1(P)$. Indeed, instead of multiplying the constraints of P by all of the variables, to obtain P_1 we just multiply by x_1 and $(1 - x_1)$. This gives us:

$$x_1 \cdot (x_1 + x_2 \geq 1/4) \rightarrow x_1^2 + x_1 x_2 \geq (1/4)x_1$$

$$x_1 \cdot (x_1 \geq 0) \rightarrow x_1^2 \geq 0$$

$$x_1 \cdot (x_2 \geq 0) \rightarrow x_1 x_2 \geq 0$$

$$x_1 \cdot (-x_1 \geq -1) \rightarrow -x_1 \geq -x_1^2$$

$$x_1 \cdot (-x_2 \geq -1) \rightarrow -x_1 x_2 \geq -x_1$$

$$(1 - x_1) \cdot (x_1 + x_2 \geq 1/4) \rightarrow x_1 + x_2 - x_1^2 - x_1 x_2 \geq (1/4) - (1/4)x_1$$

$$(1 - x_1) \cdot (x_1 \geq 0) \rightarrow x_1 - x_1^2 \geq 0$$

$$(1 - x_1) \cdot (x_2 \geq 0) \rightarrow x_2 - x_1 x_2 \geq 0$$

$$(1 - x_1) \cdot (-x_1 \geq -1) \rightarrow -x_1 + x_1^2 \geq -1 + x_1$$

$$(1 - x_1) \cdot (-x_2 \geq -1) \rightarrow -x_2 + x_1 x_2 \geq -1 + x_1$$

Next, we linearize $x_1 x_2$ as y_{12} and $x_1^2 = x_1$, which is a valid constraint since $x_1, x_2 \in \{0, 1\}$, and obtain the nonredundant inequalities:

$$x_{12} \geq 0$$

$$x_1 - y_{12} \geq 0$$

$$(1/4)x_1 + x_2 - y_{12} \geq (1/4)$$

$$x_2 - y_{12} \geq 0$$

$$-x_1 \geq -1$$

$$-x_1 - x_2 + y_{12} \geq -1$$

Finally, projecting out y_{12} yields the inequalities

$$(1/4)x_1 + x_2 \geq 1/4$$
$$1 \geq x_1 \geq 0$$
$$1 \geq x_2 \geq 0.$$

which describe the polytope $P_1(P)$.

Example 3.7. Consider again the polytope P from Figure 3.1. We next apply the RLT procedure to its inequalities using the 2-juntas x_1x_2 and $(1-x_1)(1-x_2)$. In particular,

$$x_1x_2 \cdot (x_1 \geq 0) \to y_{12} \geq 0$$
$$(1-x_1)(1-x_2) \cdot (x_1 + x_2 \geq 1/4) \to 0 \geq (1/4)(1 - x_1 - x_2 + y_{12}).$$

Since $x_1 + x_2 \geq 1 + y_{12}$ and $y_{12} \geq 0$, we have that $x_1 + x_2 \geq 1$, giving us the integer hull of P.

Let $t \in [n]$. Possibly by permuting the indices we define $P^{\leq t} := P_t(\cdots P_2((P_1(P)))\cdots)$. Below we state Balas' Sequential Convexification Theorem.

Theorem 3.8 ([153])**.** *Let $P \subseteq [0,1]^n$ be a polytope. Then*

$$P^{\leq t} = \mathit{conv}(\{x \in P \,:\, x_j \in \{0,1\},\, j \in [t]\}).$$

In particular, $P^{\leq n} = \mathit{conv}(P \cap \{0,1\}^n)$.

It follows that $P_I = \mathrm{LS}_0^n(P)$ by Theorem 3.8. Moreover, it is well known that for all $k \in [n]$ we have $P_I \subseteq \mathrm{SA}^k(P) \subseteq \mathrm{LS}^k(P) \subseteq \mathrm{LS}_0^k(P) \subseteq P$; see, e.g., [142], [145].

Given an objective vector $c \in \mathbb{R}^n$ and a polytope P, we define the *integrality gap* (of the relaxation P) as the ratio between the optimal value of c over P_I and its optimal value over P.

A Characterization of Sherali-Adams

Fix $d \le k \le n$ and a polytope $P \subseteq [0,1]^n$. The Sherali-Adams hierarchy has the following *locally consistent distribution property*: for every point $y \in \mathrm{M}^k(P)$ and every set of d coordinates $S \subseteq [n]$, we can write y as convex combination of certain points z^{J_0,J_1} (for any $J_0 \sqcup J_1 = S$) that belong to $\mathrm{M}^{k-d}(P)$ and are integer on coordinates from S. The Sherali-Adams hierarchy shares this property with any stronger hierarchy, such as SoS. However, it can be shown that this property *characterizes* the Sherali-Adams hierarchy (see, e.g., [149] and Appendix A.8 for a complete proof): y belongs to $\mathrm{M}^k(P)$ if and only if the vectors z^{J_0,J_1} belong to $\mathrm{M}^{k-d}(P)$. To make this property formal, we begin by introducing some notation.

Let $d, k \in \mathbb{Z}_{\geq 0}$ such that $d \le k$ and $y \in [0,1]^{\binom{n}{\le k+1}}$. Following [154], for $S \subseteq [n]$ of cardinality d and a 2-partition $J_0 \sqcup J_1 = S$, we define:

$$y_I^{J_0,J_1} := \sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup I \cup H}, \quad z^{J_0,J_1}(y)_I := \begin{cases} \frac{y_I^{J_0,J_1}}{y_\emptyset^{J_0,J_1}} & \text{if } y_\emptyset^{J_0,J_1} > 0, \\ y_I^{J_0,J_1} & \text{if } y_\emptyset^{J_0,J_1} = 0, \end{cases} \tag{3.3}$$

for every $I \subseteq [n]$ with $|I| \le k-d+1$.

The following result strengthens Theorem 2.25 of [149] by establishing an equivalence. In [149], the authors establish one implication, with some details omitted, and state that the converse follows by reversing their argument. For completeness, we provide a full proof of both directions in Appendix A.8.

Theorem 3.9. *Let $P \subseteq [0,1]^n$, $d, k \in \mathbb{Z}_{\geq 0}$ such that $d \le k$ and $y \in [0,1]^{\binom{n}{\le k+1}}$. Then $y \in \mathrm{M}^k(P)$ if and only if for every $S \subseteq [n]$ of cardinality d, the following hold:*

1. *For every 2-partition $J_0 \sqcup J_1 = S$, we have $z^{J_0,J_1}(y) \in \mathrm{M}^{k-d}(P)$ whenever $y_\emptyset^{J_0,J_1} > 0$; and*

2. *The restriction of y to its first $\binom{n}{\le k-d+1}$ coordinates can be expressed as the convex combina-*

tion of points $z^{J_0,J_1}(y)$, namely,

$$y_I = \sum_{J_0,J_1 \subseteq [n] \,:\, J_0 \sqcup J_1 = S} y_\emptyset^{J_0,J_1} z_I^{J_0,J_1}(y), \tag{3.4}$$

for every $I \subseteq [n]$ with $|I| \leq k - d + 1$.

It is worth emphasizing that Condition 2 above implies that for any $S \subseteq [n]$, $\sum_{J_0 \sqcup J_1 = S} y_\emptyset^{J_0,J_1} = 1$. Moreover, the vectors $\{z^{J_0,J_1}(y)\}_{J_0 \sqcup J_1 = S}$ are integer in the coordinates indexed by S. Throughout, expressions of the form $J_0 \sqcup J_1 = S$ in indices range all over $J_0, J_1 \subseteq [n]$ satisfying $J_0 \sqcup J_1 = S$.

Given $S \subseteq [n]$ and $y \in \mathrm{M}^k(P)$, we say that we *induce on the variables (indexed by) S* if we can write y (restricted to its first $\binom{n}{\leq k-d+1}$ coordinates) as in (3.4). See the proof of Proposition 3.20 for an application of the above result to a 2-dimensional instance with $k = d = 1$.

3.2.2 Symmetries

Let $\mathcal{S}_n \coloneqq \{\text{bijection } \pi : [n] \to [n]\}$ denote the set of all permutations on $[n]$. For $\pi, \pi' \in \mathcal{S}_n$, we denote their (function) composition $\pi' \circ \pi$ as $\pi'\pi$, and by π^{-1} the inverse function of π. We say that a subset $G \subseteq \mathcal{S}_n$ is a *(permutation) group* if G is non-empty and it is closed under compositions and inverses. We write $H \leq G \leq \mathcal{S}_n$ to denote that H, G are permutation groups with $H \subseteq G$. For any $x \in \mathbb{R}^n$, we define a *left permutation action on $\mathbb{R}^n$* as $\pi x \coloneqq (x_{\pi^{-1}(1)}, x_{\pi^{-1}(2)}, \ldots, x_{\pi^{-1}(n)})$ for all $\pi \in \mathcal{S}_n$. Then, the *set of (permutation) symmetries acting on P* can be defined as

$$G(P) \coloneqq \{\pi \in \mathcal{S}_n \,:\, \pi x \in P \text{ for all } x \in P\}.$$

Next, assume that P can be described by a system of linear inequalities $Ax \geq b$. Let $A_{\cdot,i}$ denote the ith column of A and $A_{j,\cdot}$ its jth row. We define a *right column action on $\mathbb{R}^{m\times n}$* as $A\pi \coloneqq (A_{\cdot,\pi(i)})_{i\in[n]}$, $\forall \pi \in \mathcal{S}_n$. Analogously, a *left row action on $\mathbb{R}^{m\times n}$* is defined by $\sigma A \coloneqq (A_{\sigma(j),\cdot})_{j\in[m]}$, $\forall \sigma \in \mathcal{S}_m$. Then, the *set of formulation symmetries acting on $Ax \geq b$* is given by

$$G(A, b) \coloneqq \{\pi \in \mathcal{S}_n \,:\, \exists \sigma \in \mathcal{S}_m,\, \sigma A \pi = A \text{ and } \sigma b = b\}.$$

It is well-known and straightforward to verify that $G(P)$ and $G(A, b)$ are permutation groups. In general, $G(A, b)$ is a subgroup of $G(P)$ since, for example, $Ax \geq b$ may contain redundant inequalities which do not preserve the symmetries of P. In fact, the two coincide if P is full-dimensional and has no redundant inequalities [155].

Consider the linear program $\mathcal{P} : \min\{c^\top x : Ax \geq b\}$, with $c \in \mathbb{R}^n$. Then, the *set of of formulation symmetries of* $\mathcal{P}(A, b, c)$ and the *set of symmetries acting on* $\mathcal{P}$ are defined, respectively, as:

$$G(\mathcal{P}) \coloneqq \{\pi \in \mathcal{S}_n : \pi x \in P \text{ and } c^\top \pi x = c^\top x \text{ for all } x \in P\},$$

$$G(A, b, c) \coloneqq \{\pi \in \mathcal{S}_n : \pi c = c \text{ and } \exists \sigma \in \mathcal{S}_m,\, \sigma A \pi = A \text{ and } \sigma b = b\}.$$

As above, the remaining sets of symmetries are also permutation groups, and $G(A, b, c) \leq G(\mathcal{P})$.

Remark 3.10. Throughout this chapter, we assume that all groups are subgroups of $G(P)$ or $G(\mathcal{P})$. In computational studies, these groups may not be directly available; accordingly, the literature typically considers the group of formulation symmetries $G(A, b, c)$. In [156], computing generators of $G(A, b, c)$ is reduced to the computation of generators of a graph automorphism group, and hence to Graph Isomorphism, for which fast practical quasi-polynomial time algorithms exist; see also [121], [124]. We also note that [155] describes a procedure to compute $G(\mathcal{P})$.

k-Transitivity and the Averaging Operator

Let $k \in [n]$ and $G \leq \mathcal{S}_n$. G is *k-transitive* if, for every pair of (ordered) k-tuples t^1 and t^2 of $[n]$, where the elements of each tuple are all-different, there exists a permutation $\pi \in G$ such that $\pi(t^1_j) = t^2_j$ for $j \in [k]$. We say that P is *G-invariant* if $G \leq G(P)$ and it is k-transitive if moreover G is k-transitive.

Given a subset $S \subseteq [n]$ and $G \leq \mathcal{S}_n$, we call the subgroup of G which fixes pointwise each element of S the *pointwise stabilizer of* S and we denote it by $G_S \coloneqq \{\pi \in G : \pi(i) = i,\, \forall \pi \in G, \forall i \in S\}$. Given any $x \in \mathbb{R}^n$, we define the *orbit of x under* G as $\mathrm{orb}_G(x) \coloneqq \{\pi x : \pi \in G\}$.

Moreover, we define the *Reynolds operator* [157] $\beta_G : \mathbb{R}^n \to \mathbb{R}^n$ as

$$\beta_G(x) \coloneqq \frac{1}{|G|} \sum_{\pi \in G} \pi x$$

for $x \in \mathbb{R}^n$. If $x \in P$ for some polytope P and $G \leq G(P)$, then the convexity of P implies that $\beta_G(x) \in P$ for every $x \in P$. Next, we describe the averaging induced by the Reynolds operator with respect to the pointwise stabilizer of a set $S \subseteq [n]$.

Proposition 3.11. *Let $P \subseteq [0,1]^n$ be a polytope, $G \leq G(P)$, and S a subset of $[n]$ of cardinality at most k, for some $k \in \mathbb{Z}_{\geq 0}$. Assume that G is $(k+1)$-transitive. Then for any $x \in P$, the point $\beta_{G_S}(x)$ belongs to P. Moreover, for $i \in [n]$, it has entries:*

$$\beta_{G_S}(x)_i = \begin{cases} x_i & \text{if } i \in S, \\ \frac{\sum_{j \in [n] \setminus S} x_j}{n - |S|} & \text{if } i \notin S. \end{cases}$$

Proof. The statement is obvious for $i \in S$. We next claim $\beta_{G_S}(x)_j = \beta_{G_S}(x)_i$ for any $i, j \in [n] \setminus S$. Thus, let $i, j \in [n] \setminus S$. Since G is $(k+1)$-transitive, there exists $h \in G$ such that $h(\ell) = \ell$ for $\ell \in S$ and $h(j) = i$. By construction, $\pi \in G_S$. Thus, using the well-known fact that G_S is a subgroup of G, we can write

$$\begin{aligned}
\beta_{G_S}(x)_j &= \frac{1}{|G_S|} \sum_{\pi \in G_S} (\pi x)_j \\
&= \frac{1}{|G_S|} \sum_{\pi \in G_S} x_{\pi^{-1}(j)} \\
&= \frac{1}{|G_S|} \sum_{\pi \in G_S} x_{\pi^{-1} h^{-1}(i)} \\
&= \frac{1}{|G_S|} \sum_{\pi \in G_S} x_{(h\pi)^{-1}(i)} = \frac{1}{|G_S|} \sum_{r \in G_S} x_{r^{-1}(i)} = \frac{1}{|G_S|} \sum_{r \in G_S} x_{r(i)} = \beta_{G_S}(x)_i.
\end{aligned}$$

Therefore, for $i \in [n] \setminus S$, we have

$$\sum_{j\in[n]\setminus S} x_j + \sum_{j\in S} x_j = \|x\|_1 = \|\beta_G(x)\|_1 = \sum_{j\in[n]} \beta_G(x)_j = (n - |S|) \cdot \beta_G(x)_i + \sum_{j\in S} x_j,$$

from which the statement follows. $\square$

In all the definitions given above, we omit the subscript if the group is clear from the context. For $G \leq \mathcal{S}_n$ we say that G is *transitive on* $T \subseteq [n]$ if for every $i, j \in T$, there exists $\pi \in G$ such that $\pi(i) = \pi(j)$. Thus, in particular, G is 1-transitive if and only if it is transitive on $[n]$. The following result follows from a fact in [158, Section 1.8].

Proposition 3.12. *A permutation group $G \leq \mathcal{S}_n$ is $(k+1)$-transitive if and only if G is k-transitive and for each $S \subseteq [n]$, with $|S| = k$, we have that G_S is transitive on $[n] \setminus S$.*

The permutation group action of $\mathcal{S}_n$ on singletons $\{i\} \subseteq [n]$ can be naturally extended to arbitrary subsets $S \subseteq [n]$ as: $\pi(S) \coloneqq \{\pi(i) \,:\, i \in S\}$ for all $\pi \in \mathcal{S}_n$. Hence we can define the following *action* on $\mathbb{R}^{\binom{n}{\leq k+1}}$:

$$(\pi x)_S \coloneqq x_{\pi^{-1}(S)} \tag{3.5}$$

for every $x \in \mathbb{R}^{\binom{n}{\leq k+1}}$, $S \subseteq [n]$ and $\pi \in \mathcal{S}_n$. See [146] for further background on permutation groups and their actions.

3.3 Known Results on Hierarchies and Integrality Gaps

We summarize known results on integrality gaps and hierarchies that are relevant for our analysis.

Linear and semidefinite hierarchies offer systematic procedures for strengthening an initial relaxation of an integer program [148], [150], [152], [159]; see, e.g., [145] for a detailed comparison. They are used as an automatic way to generate stronger relaxations, while keeping efficiency at low levels of the hierarchy. However, their success in improving integrality gaps is limited, and symmetries partially explain this phenomenon. Machine scheduling showcases this behavior: while

Sherali-Adams and Lasserre are not able to close the gap after linearly many rounds over the symmetric assignment relaxation, they yield approximation schemes after symmetry-breaking [135], [136]. [140] study the impact of symmetries for Lasserre and show a sufficient condition to prove rank lower bounds whenever the solutions and constraints are invariant under every permutation of $[n]$.

Several approaches for handling symmetries in IPs are found in the literature, including symmetry-breaking procedures in the formulation or in the algorithms employed; see, e.g., [160], [161], [162] and the survey [124] for a detailed exposition. Towards the goal of providing a unified perspective, [163] recently introduce a framework that captures many of the existing symmetry-handling methods. In the context of LP hierarchies, [123] shows how symmetry can be exploited to construct shrunken LP formulations that attain the same objective value as the corresponding Sherali-Adams relaxations.

Ranks and Integrality Gap

For any polytope $P \subseteq [0,1]^n$ and any well-defined operator T, the *T-rank of P* is defined as the smallest $k \in [n]$ such that $T^k(P) = P_I$. We next summarize known rank results for well-studied polytopes (see also the notion of resistance in Section 3.5). When discussing integrality gaps, we assume the objective is given by $\mathbf{1}$.

We begin with the Maximum Matching problem. Let K_{2n+1} be the complete graph on $2n+1$ nodes, with edge set $E = \{\{i,j\} : i,j \in [2n+1],\ i<j\}$, so that $|E| = \binom{2n+1}{2}$. We consider the edge formulation, whose LP relaxation is given by

$$P^{MM}(K_{2n+1}) = \left\{ x \in [0,1]^{|E|} : \sum_{j:\{i,j\}\in E} x_{ij} \le 1,\ \forall i \in [2n+1] \right\}.$$

It is shown in [141] that $P^{MM}(K_{2n+1})$ has BCC-rank n^2, while [144] prove that it has LS$_+$-rank n (note that $P^{MM}(K_n)$ has $\Theta(n^2)$ variables). Moreover, [131] characterize the integrality gap of $P^{MM}(K_n)$, showing in particular that it is $(n-1)$-SA-resistant, and that SA^{2n-1} is required to close the gap.

Next, for an arbitrary graph $G = (V, E)$, the vertex formulation of the Maximum Stable Set problem can be defined as

$$P^{SS}(G) := \left\{ x \in [0,1]^{|V|} \,:\, x_i + x_j \leq 1,\, \forall \{i,j\} \in E \right\}.$$

In [148] this polytope is studied in detail. Indeed, it is shown that the LS-rank of $P^{SS}(G)$ satisfies the inequality

$$\frac{n}{\alpha(G)} - 2 \leq t \leq n - \alpha(G) - 1$$

where the stability number of G defined as $\alpha(G) := \max \{ \mathbf{1}^\top x \,:\, x \in P^{SS}(G), x \in \{0,1\}^{|V|} \}$. For instance, if G corresponds to the complete graph on n nodes then the LS-rank is $n - 2$. Moreover, the point $(1/(t+2))\mathbf{1}$ belongs to $\mathrm{LS}^t(P^{SS}(G))$ and $\mathrm{SA}^t(P^{SS}(G))$ On the other hand, they show that the LS_+-rank t of $P^{SS}(G)$ satisfies $t \leq \alpha(G)$.

Another polytope well-studied polytope is

$$P_n^K := \left\{ x \in [0,1]^n \,:\, \sum_{i=1}^{n} x_i \geq 1/2 \right\}.$$

[142] show its LS_+ and SA-rank is n. Moreover, they show that the equalities $\mathrm{LS}_0^k(P_n^K) = \mathrm{LS}^k(P_n^K) = \mathrm{LS}_+^k(P_n^K)$ hold for $0 \leq k \leq n$. Then, [145] shows that $\frac{1}{n+1}\mathbf{1}$ belongs to $\mathrm{SA}^{n-1}(P_n^K)$.

Finally, we note that P_n^K and $P^{SS}(K_n)$ are n-transitive polytopes, whereas $P^{MM}(K_{2n+1})$ is only 1-transitive. The reason is that the symmetry group of $P^{MM}(K_{2n+1})$ acts on variables indexed by the edges of K_{2n+1}. If this group were 2-transitive, then any pair of adjacent edges could be mapped to any pair of disjoint edges. However, such a mapping would not preserve adjacency, and hence cannot correspond to a symmetry of K_{2n+1}. Therefore, $P^{MM}(K_{2n+1})$ is not 2-transitive.

See Sections 3.6.4 and 3.6.3 for two other classes of polytopes studied in the literature.

3.4 Symmetry, Hierarchies, and Polyhedral Structure

In this section, we present general geometric results on LP hierarchies, together with structural properties of 2-transitive polytopes. We further show that the LP hierarchies considered in this work are *symmetry-preserving*, in the sense that symmetries of a polytope P are inherited by the corresponding relaxation. Although these results will be used later, we believe they are of independent interest. In addition, we discuss a folklore result concerning the dominance of SA over LS_0.

Let $\mathcal{C}_n$ be the collection of all polytopes in $\mathbb{R}^n$. We say that an operator $T : \mathcal{C}_n \to \mathcal{C}_n$ is *well-defined* if for every polytope $P \subseteq [0,1]^n$: (1) T is integer valid, i.e., $T(P) \cap \{0,1\}^n = P \cap \{0,1\}^n$, and (2) and $T(P) \subseteq P$; and (3) T is monotone, i.e., for every pair of polytopes $P, Q \subseteq [0,1]$, if $P \subseteq Q$, then $T(P) \subseteq T(Q)$.

3.4.1 Geometric Properties of LP Hierarchies

For any $(c, \delta) \in \mathbb{R}^{n+1}$ we write $H_{c,\delta} := \{x \in \mathbb{R}^n \,:\, c^\top x = \delta\}$ for the hyperplane defined by (c, δ). We say that the inequality $c^\top x \geq \delta$ is *valid* for P if every x in P satisfies $c^\top x \geq \delta$. A *face* F of a polytope P is a set of the form $F := P \cap H_{c,\delta}$, where $c^\top x \geq \delta$ is a valid inequality for P. See [12, Chapter 3] for details.

We prove that the Sherali-Adams operator is *face-preserving*, i.e., given any face F of P, we have that $\text{SA}^k(F) = \text{SA}^k(P) \cap F$. To this end, we first present a geometric characterization of M^k.

A Disjunctive Characterization of M^k The operator LS_0 admits a disjunctive characterization in the space of original variables (see Section 3.2.1), while SA admits such a characterization in an extended space. Indeed, by [151, Theorem 4.1] the set $\text{M}^k(P)$ can be written as

$$\text{M}^k(P) = \bigcap_{S \subseteq [n],\, |S| = k} \text{conv}\left(\bigcup_{J_0 \sqcup J_1 = S} P_{J_0, J_1} \right)$$

where (J_0, J_1) is a 2-partition of S, and P_{J_0,J_1} is the polytope P in $\mathbb{R}^{\binom{n}{\leq k+1}}$ where variables $x_j = y_j$ are fixed to 0 and $x_{j'} = y_{j'}$ are fixed to 1 for each $j \in J_0$, and each $j' \in J_1$, as well as all the resulting consequences on the lifted variables y_I with $2 \leq |I| \leq k+1$, i.e.,

$$
\begin{aligned}
P_{J_0,J_1} := \{y \in [0,1]^{\binom{n}{\leq k+1}} : \ & \sum_{j \in [n] \setminus S} a_{i,j} y_j \geq \left(b_i - \sum_{j \in J_1} a_{i,j} \right) \quad \forall i \in [m], \\
& 0 \leq y_j \leq 1 \quad \forall j \in [n] \setminus S \\
& y_j = 1 \quad \forall j \in J_1, \quad y_j = 0 \quad \forall j \in J_0 \\
& y_I = 1 \quad \forall I \subseteq J_1, |I| \geq 2 \\
& y_I = 0 \quad \forall I \subseteq S, |I| \geq 2, I \cap J_0 \neq \emptyset \\
& y_{j \cup I} = y_j \quad \forall j \in [n] \setminus S, \forall I \subseteq S, \emptyset \subsetneq I \subseteq J_1 \\
& y_{j \cup I} = 0 \quad \forall j \in [n] \setminus S, I \cap J_0 \neq \emptyset \}.
\end{aligned}
$$

The disjunctive characterization of the operators LS_0, LS and SA can be used to prove a useful geometric property. Let F be a face of P. By monotonicity of a well-defined operator T, it holds that $T(F) \subseteq T(P) \cap F$. For the operators considered here, this inclusion holds at equality. This fact is known for LS and LS_0; see [143], [148]. The following lemma [12, Claim in Theorem 5.22, Chapter 5] will be used to extended this result to SA.

Lemma 3.13. *Let $A \subseteq \mathbb{R}^n$ be arbitrary and assume that $c^\top x \geq \delta$ is valid for A. Then conv$(H_{c,\delta} \cap A) = H_{c,\delta} \cap$ conv(A).*

Proof. Since $H_{c,\delta} \cap A \subseteq H_{c,\delta} \cap \text{conv}(A)$ and $\text{conv}(A)$ and $H_{c,\delta}$ are convex sets, then the left-to-right inclusion follows. To prove the converse, let $\overline{x} \in H_{c,\delta} \cap \text{conv}(A)$ arbitrary. By Caratheodory's Theorem, there exits at most $k \leq n+1$ positive weights $\{\lambda_i\}_{i=1}^k$ that add up to 1, $x^i \in A$ for $i \in [k]$, and such that $\overline{x} = \sum_{i=1}^k \lambda_i x^i$. Moreover, $c^\top \overline{x} = \delta$ and $c^\top x^i \geq \delta$ for $i \in [k]$ since the inequality is valid for A. Since $c^\top \overline{x} = b$ we have that $\sum_{i=1}^k \lambda_i c^\top x^i = \delta$. We conclude that $x^i \in H_{c,\delta}$ given that $\sum_{i=1}^k \lambda_i = 1$. □

Before proving the extension to SA, we first illustrate the argument revisiting the known case of the BCC operator P_j. For completeness, we provide a simple proof that $P_j(F) = P_j(P) \cap F$ using the lemma.

Proposition 3.14. *Let F be a face of P. Then $P_j(F) = P_j(P) \cap F$ for $j \in [n]$.*

Proof. Let F be a face of P defined by the some hyperplane $H_{a,b}$ with $(a, b) \in \mathbb{R}^{n+1}$. Then

$$
\begin{aligned}
P_j(F) &= P_j(P \cap H_{a,b}) \\
&= \text{conv}((P \cap H_{a,b} \cap \{x \in \mathbb{R}^n \,:\, x_j = 1\}) \cup (P \cap H_{a,b} \cap \{x \in \mathbb{R}^n \,:\, x_j = 0\}) \\
&= \text{conv}(H_{a,b} \cap (P \cap (\{x \in \mathbb{R}^n \,:\, x_j = 1\} \cup \{x \in \mathbb{R}^n \,:\, x_j = 0\}))) \\
&= H_{a,b} \cap \text{conv}(P \cap (\{x \in \mathbb{R}^n \,:\, x_j = 1\} \cup \{x \in \mathbb{R}^n \,:\, x_j = 0\})) \quad \text{by Lemma (3.13)} \\
&= H_{a,b} \cap P_j(P) \\
&= F \cap P_j(P) \quad \text{since } P \cap P_j(P) = P_j(P).
\end{aligned}
$$

□

We now turn to SA. For an arbitrary set $R \subseteq \mathbb{R}^n$, we denote its trivial extension to $\mathbb{R}^{\binom{n}{\leq k+1}}$ by

$$R^k := \{y \in \mathbb{R}^{\binom{n}{\leq k+1}} \,:\, \Pi_x(y) \in R\} = R \times \mathbb{R}^{\binom{n}{\leq k+1}-n}.$$

Proposition 3.15. *Let F be a face of P. Then it holds that*

$$\text{M}^k(F) = \text{M}^k(P) \cap F^k \quad \textit{and} \quad \text{SA}^k(F) = \text{SA}^k(P) \cap F.$$

Proof. Assume that F is induced by the inequality $c^\top x \geq \delta$, i.e., $F = P \cap H_{c,\delta}$. Consider an arbitrary $S \subseteq [n]$ with $|S| = k$, and a 2-partition (J_0, J_1) of S. We note that $c^\top x \geq \delta$ is valid for

P_{J_0,J_1}. Therefore,

$$\begin{aligned}\operatorname{conv}\left(\left(\bigcup_{J_0\sqcup J_1=S} P_{J_0,J_1}\right)\bigcap F^k\right) &= \operatorname{conv}\left(\left(\bigcup_{J_0\sqcup J_1=S} P_{J_0,J_1}\right)\bigcap H^k_{c,\delta}\right)\\ &= \operatorname{conv}\left(\bigcup_{J_0\sqcup J_1=S} P_{J_0,J_1}\right)\bigcap H^k_{c,\delta}\end{aligned}$$

where the last equality follows by Lemma 3.13. In consequence,

$$\begin{aligned}\mathrm{M}^k(F) &= \bigcap_{S\subseteq[n],|S|=k} \operatorname{conv}\left(\left(\bigcup_{J_0\sqcup J_1=S} P_{J_0,J_1}\right)\bigcap F^k\right)\\ &= \bigcap_{S\subseteq[n],|S|=k} \operatorname{conv}\left(\bigcup_{J_0\sqcup J_1=S} P_{J_0,J_1}\right)\bigcap H^k_{c,\delta}\\ &= \left(\bigcap_{S\subseteq[n],|S|=k} \operatorname{conv}\left(\bigcup_{J_0\sqcup J_1=S} P_{J_0,J_1}\right)\right)\bigcap H^k_{c,\delta}\\ &= \mathrm{M}^k(P)\cap F^k\end{aligned}$$

since $P^k \cap \mathrm{M}^k(P) = \mathrm{M}^k(P)$.

Finally, we show that

$$\Pi_x(\mathrm{M}^k(P)\cap F^k) = \mathrm{SA}^k(P)\cap F.$$

Indeed, let $x \in \Pi_x(\mathrm{M}^k(P)\cap F^k)$. Then there exists $y \in \mathbb{R}^{\binom{n}{\le k+1}-n}$ such that $(x,y) \in \mathrm{M}^k(P)\cap F^k$, i.e., $(x,y)\in \mathrm{M}^k(P)$ and $x \in F$. The former implies $x \in \mathrm{SA}^k(P)$, and together with $x\in F$, we obtain $x \in \mathrm{SA}^k(P)\cap F$. The converse follows readily. □

3.4.2 Geometric Properties of 2-transitive Polytopes

For a permutation group $G \le G(P)$, we define the *fixed set of P under G* as

$$\operatorname{fix}_G(P) := \{x \in P : \pi x = x \text{ for all } \pi \in G\}.$$

When $P = \mathbb{R}^n$, this set is a linear subspace, commonly referred to as the *fixed subspace of G*.

Proposition 3.16. *Let $P \subseteq [0,1]^n$ be a 2-transitive polytope. Then $\dim(P) \in \{0, 1, n-1, n\}$.*

Proof. Let G be a 2-transitive group under which P is invariant. By definition, it holds that $\text{fix}_G(P) \subseteq P$. If $\text{fix}_G(P) = P$, then $\dim(P) \in \{0, 1\}$ since G is transitive.

On the other hand, suppose that $P \setminus \text{fix}_G(P) \neq \emptyset$. Let $u \in P \setminus \text{fix}_G(P)$ and consider the orbit of u under G. Suppose, without loss of generality, that u_1 is the largest entry in u. Since G is transitive, u_1 can be mapped to each entry $j \in \{2, \ldots, n\}$. Next, pick any set of vectors $\{u^1, \ldots, u^n\}$ where $u^j = g_j(u)$ and $u^j_j = u_1$ for some $g_j \in G$, for each $j \in [n]$. We claim that the set $\{\beta_{G_j}(u^j)\}_{j=1}^n \subseteq P$ is affinely independent. Indeed, for each $j \in [n]$ the group G_j is transitive on $[n] \setminus \{j\}$ since G is 2-transitive, hence we can write

$$\begin{aligned}\beta_{G_j}(u^j) - \beta_{G_n}(u^n) &= \left(u_1 e_j + \left(\frac{\|u\|_1 - u_1}{n-1}\right)\sum_{i \neq j} e_i\right) - \left(u_1 e_n + \left(\frac{\|u\|_1 - u_1}{n-1}\right)\sum_{i \neq n} e_i\right) \\ &= \overline{u}\,(e_j - e_n) \quad \text{where} \quad \overline{u} := \left(u_1 - \frac{\|u\|_1 - u_1}{n-1}\right).\end{aligned}$$

Considering that $u \notin \text{fix}_G(P)$, the scalar $\overline{u}$ must be nonzero, implying that the $n-1$ vectors $\beta_{G_j}(u^j) - \beta_{G_n}(u^n)$ are linearly independent. In consequence $\dim(P)$ is at least $n-1$. $\square$

Proposition 3.17. *Let $P \subseteq [0,1]^n$ be a polytope and let G be a 2-transitive group under which P is invariant. If*

1. $\dim(P) \leq 1$, *then* $\text{fix}_G(P) = P$*;*
2. $\dim(P) = n-1$, *then* $\dim(\text{fix}_G(P)) = 0$*;*
3. $\dim(P) = n$, *then* $\dim(\text{fix}_G(P)) = 1$.

Proof. (1) If $\dim(P) \leq 1$, then $\text{fix}_G(P) = P$ since G is transitive. (2) Assume $\dim(P) = n-1 > 1$, and suppose, on the contrary, that $\text{fix}_G(P)$ is not a singleton. By Proposition 3.11 there exist some $t_0, t_1 \in [0,1]$, $t_0 < t_1$, $\text{fix}_G(P) = \{te \in \mathbb{R}^n \,:\, t \in [t_0, t_1]\}$. As in the proof of Proposition 3.16, pick some $u \in P \setminus \text{fix}_G(P)$, take its orbit under G, and consider the set $\{\beta_{G_j}(u^j)\}_{j=1}^n \subseteq P$. The latter set is affinely independent and lies on the hyperplane $H := \{x \in \mathbb{R}^n \,:\, e^\top x = \|u\|_1\}$.

Clearly $\beta_G(u) = (\|u\|_1/n)e$ belongs to $\text{fix}_G(P) \cap H$, hence $t_0 \leq \|u\|_1/n \leq t_1$. Pick any other $t' \in [t_0, t_1]$ such that $t' \neq \|u\|_1/n$. Then $x' = t'e \in \text{fix}_G(P) \subseteq P$ and it belongs to the orthogonal complement of H. This implies that P must be full-dimensional, which contradicts our assumption $\dim(P) = n - 1$. $\square$

Remark 3.18. The above results are tight with respect to transitivity. Consider $x \coloneqq (1, 1/2, 1, 1/2)$ and $z \coloneqq (1, 1/3, 1, 1/3)$ and let P be the convex hull of the orbits of x and y under C_4. It is straightforward to show that $\dim P = 2$. Moreover, its fixed set $\{te \in \mathbb{R}^4 : t \in [2/3, 3/4]\}$ is one-dimensional.

3.4.3 A Folklore Result: Sherali–Adams Dominates the Lift-and-Project Closure

It is a folklore result that one round of Sherali–Adams is strictly stronger than the lift-and-project closure. While a proof is claimed in [164], we first show that their argument is incorrect. To the best of our knowledge, a correct proof has not appeared in the literature. We then present an instance in $[0, 1]^2$ that illustrates this dominance.

In [164, Section 3, Subsection 14] the authors exhibit an instance of the Steiner Triple Covering problem (see Section 3.6.2) for which they claim that a particular SA level 1 cut is not valid for the split closure – which is known to be strictly stronger than the lift-and-project closure. Their argument relies on the claim that a particular point lies in the interior of the split closure of the relaxation, and hence in the interior of the lift-and-project closure. This claim is supported by verifying the infeasibility of a mixed-integer nonlinear system. However, such infeasibility conclusions should be interpreted with caution, as numerical solvers for mixed-integer nonlinear systems may be significantly affected by rounding errors, potentially leading to incorrect assessments. We show below that this claim is incorrect.

Claim 3.19. Consider STC(9) and denote by P its LP relaxation. The point $(3/7)e$, which is claimed in [164, Section 3, Subsection 14] to lie in the interior of $\text{LS}_0(P)$, does not in fact belong to it.

Proof. We show that $(3/7)e$ is not in the interior of $P_i(P)$ for any $i \in [9]$. Consider $i \in [9]$ and let $P^k := P \cap \{x \in [0,1]^9 \,:\, x_i = k\}$ with $k \in \{0,1\}$. Let $x^1 \in P^1$ and $x^0 \in P^0$ be arbitrary. We show that any $\lambda \in [0,1]$ such that $\hat{x} := \lambda x^1 + (1-\lambda)x^0 \in P_i(P) \cap \text{span}(e)$ satisfies $\lambda \geq 3/7$. To this end, first, we note that $\sum_{j \neq i} x_j \geq 8/3$ is valid for P^1 (summing up the 8 constraints such that x_i is not in their support), while $\sum_{j \neq i} x_j \geq 4$ is valid for P^0 (summing up the 4 constraints such that x_i belongs to their support). Since $\hat{x} = \lambda x^1 + (1-\lambda)x^0 \in \text{span}(e)$ for $j \in [9]$ and $\hat{x}_i = \lambda$, we have that $\lambda = \lambda x_j^1 + (1-\lambda)x_j^0$ for all $j \in [9]$. Summing up over $j \in [9]$ yields

$$\begin{aligned} 9\lambda &= \lambda + \lambda \left(\sum_{j \neq i} x_j^1 \right) + (1-\lambda) \left(\sum_{j \neq i} x_j^0 \right) \\ &\geq \lambda + (8/3)\lambda + (1-\lambda)4 \end{aligned}$$

which implies $\lambda \geq 3/7$. $\square$

We now present a correct proof that Sherali–Adams with $k = 1$ strictly dominates the lift-and-project closure.

Proposition 3.20. *For $0 < \epsilon < 1/2$ there exists a polytope $P_\epsilon \subseteq [0,1]^2$ such that* $\text{SA}(P) \subsetneq \text{LS}_0(P)$.

Proof. Let $0 < \epsilon < 1/2$ and $P_\epsilon \subseteq [0,1]^2$ be the polytope defined as the convex hull of the following points: $(0,1/2), (1/2,1), (1/2,0)$, and $(1, 1/2+\epsilon)$. Note that $P_1(P)$ is the segment with endpoints $\{(0,1/2), (1,1/2+\epsilon)\}$, while $P_2(P)$ is the segment with endpoints $\{(1/2,1), (1/2,0)\}$. It is straightforward to check that $\text{LS}_0(P_\epsilon) = P_1(P_\epsilon) \cap P_2(P_\epsilon) = \{(1/2, (1+\epsilon)1/2)\}$.

Next, we show that $\text{SA}(P_\epsilon) = \emptyset$ using Theorem 3.9 with $k = d = 1$. Suppose for the sake of contradiction that this is not the case. Let $y \in \text{M}^1(P)$ where $\Pi_x(y/y_\emptyset) \in \text{SA}(P_\epsilon)$. Let $v^i, w^i \in \text{M}^0(P_\epsilon)$ for $i \in \{1,2\}$ given by

$$\begin{aligned} &v_\emptyset^1 = y_1, \quad w_\emptyset^1 = y_\emptyset - y_1 \\ &v_1^1 = y_1, \quad w_1^1 = 0 \\ &v_2^1 = y_{12}, \quad w_2^1 = y_2 - y_{12}, \end{aligned}$$

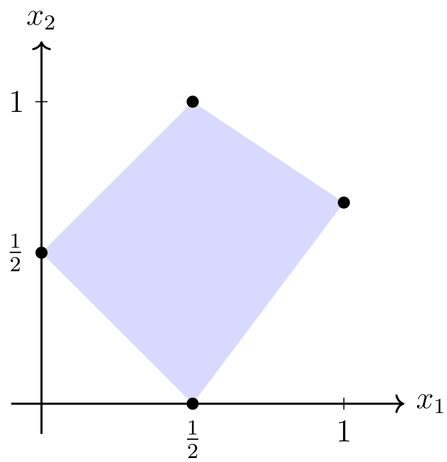

Figure 3.2: Polytope P_ϵ with $\epsilon = 1/6$.

and

$$\begin{aligned} v^2_\emptyset &= y_2, \quad w^2_\emptyset = y_\emptyset - y_2 \\ v^2_1 &= y_{12}, \quad w^2_1 = y_1 - y_{12} \\ v^2_2 &= y_2, \quad w^2_2 = 0. \end{aligned}$$

We note that we can induce on both variables x_1 and x_2 since the only points of P_ϵ with at least one integer entry are cut-off by SA level 1, i.e., $\Pi_x(y/y_\emptyset)$ must be fully fractional.

It holds that $\Pi_x(v^1/v^1_\emptyset) \in P_\epsilon$ and $\Pi_x(w^1/w^1_\emptyset) \in P_\epsilon$. Since $v^1_1/v^1_\emptyset = 1$, we have that $v^1_2/v^1_\emptyset$ must be equal to $1/2 + \epsilon$, i.e., $y_{12} = (1/2 + \epsilon)y_1$. On the other hand, since $w^1_1/w^1_\emptyset = 0$ we have that $\frac{y_2 - y_{12}}{y_\emptyset - y_1} = 1/2$. Replacing the former relation into the latter equation we have $y_2 = (1/2)y_\emptyset + \epsilon y_1$. Next, we have that $\Pi_x(v^2/v^2_\emptyset) \in P_\epsilon$ and $\Pi_x(w^2/w^2_\emptyset) \in P_\epsilon$. Since $v^2_2/v^2_\emptyset = 1$ we have that $v^2_1/v^2_\emptyset = 1/2$, i.e., $y_{12} = (1/2)y_2$. Moreover, $w^2_2/w^2_\emptyset = 0$ implies that $w^2_1/w^2_\emptyset = 1/2$, i.e., $\frac{y_1 - y_{12}}{y_\emptyset - y_2} = 1/2$.

Therefore, $y_1 = (1/2)y_\emptyset$. Summing up, we have

$$\begin{aligned} y_{12} &= (1/2+\epsilon)y_1 \\ y_2 &= (1/2)y_\emptyset + \epsilon y_1 \\ y_{12} &= (1/2)y_2 \\ y_1 &= (1/2)y_\emptyset \end{aligned}$$

The latter system reduces to $y_{12} = (1/2)(1+\epsilon)y_1$ and $y_{12} = (1/2+\epsilon)y_1$, i.e., $\epsilon = 0$, a contradiction. $\square$

3.4.4 Symmetric Operators

We say that T is *symmetry-preserving* if, for every polytope P and every subgroup $G \le G(P)$, the convex set $T(P)$ is also G-invariant. We show that the hierarchies LS_0, LS and SA are symmetry-preserving. This property is implicit in the literature, see e.g., Ostrowski [123]. The formal statements are given in Propositions 3.25 and 3.26.

We begin by introducing the auxiliary results needed for their proofs.

Proposition 3.21. *Let P be a G-invariant polytope. If T is a well-defined symmetry-preserving operator, then*

$$T(P) \neq \emptyset \iff \mathrm{fix}_G(P) \cap T(P) \neq \emptyset.$$

Proof. Necessity is straightforward. To prove sufficiency, pick any $x \in T(P)$. Given that T is symmetry-preserving we have that $\mathrm{orb}_G(x) \subseteq T(P)$. By convexity of $T(P)$ we have that $\mathrm{conv}(\mathrm{orb}_G(x)) \subseteq T(P)$. Finally, the barycenter of the orbit of x under G, $\beta_G(x)$, clearly belongs to $\mathrm{conv}(\mathrm{orb}_G(x)) \subseteq T(P)$. We observe that $\beta_G(x) \in \mathrm{fix}_G(P)$, as for every $\pi' \in G$,

$$\pi'\beta_G(x) = \pi'\frac{1}{|G|}\sum_{\pi\in G}\pi x = \frac{1}{|G|}\sum_{\pi\in G}\pi'\pi x = \frac{1}{|G|}\sum_{\pi''\in G}\pi'' x,$$

where in the last equality we used the changed of variables $\pi'' := \pi'\pi$. $\square$

We will use the following characterization of the symmetries of a polytope P.

Lemma 3.22. *Let P be a polytope. A permutation $\pi \in \mathcal{S}_n$ belongs to $G(P)$ if and only if for any valid inequality $\alpha^\top x \geq \beta$ of P we have that $(\pi\alpha)^\top x \geq \beta$ is also valid for P.*

Proof. Consider a permutation $\pi \in G(P)$ and a valid inequality $\alpha^\top x \geq \beta$ for P. We must show that $(\pi\alpha)^\top x \geq \beta$ is also valid. Since $G(P)$ is a group, then $\pi^{-1} \in G(P)$. If $x \in P$ then $\pi^{-1}x \in P$, and thus $\alpha^\top \pi^{-1} x \geq \beta$. Observing that $\alpha^\top(\pi^{-1}x) = (\pi\alpha)^\top x$, then $(\pi\alpha)^\top x \geq \beta$ for all $x \in P$. Hence $(\pi\alpha)^\top x \geq \beta$ is a valid inequality for P.

To show the other implication, consider a permutation π such that for any valid inequality $\alpha^\top x \geq \beta$ for P, we obtain that $(\pi\alpha)^\top x \geq \beta$ is also valid. It suffices to show that for any $x \in P$ we have that $\pi x \in P$. Consider any formulation of P given by $P = \{x : Ax \geq b\}$, for some matrix $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Let $x \in P$. Then $a_i^\top x \geq b_i$ for all $i \in [m]$, where a_i is the transpose of the i-th row of A. Then $(\pi a_i)^\top x \geq b_i$. Therefore $(\pi a_i)^\top x = a_i^\top(\pi^{-1}(x)) \geq b_i$ for all $i \in [m]$, and thus $\pi^{-1}(x) \in P$. We conclude that $\pi^{-1} \in G(P)$. As $G(P)$ is a group, it is closed under taking inverses, and thus $\pi \in G(P)$. $\square$

The next lemma shows how the action of a permutation can be transferred from a lifted point to the vector of coefficients defining an inequality.

Lemma 3.23. *Suppose that G acts on $\mathbb{R}^{\binom{n}{\leq k+1}}$ as in* (3.5)*. Let $a, y \in \mathbb{R}^{\binom{n}{\leq k+1}}$ and $\pi \in G$. Then $a^\top \pi y \geq 0$ if and only if $(\pi^{-1}a)^\top y \geq 0$.*

Proof. It is not hard to see that each $\pi \in G$ can be associated with a permutation matrix $R_\pi \in \{0,1\}^{\binom{n}{\leq k+1} \times \binom{n}{\leq k+1}}$, hence we can write the action of π on y, i.e., πy, as $R_\pi y$. Moreover, a permutation matrix is an orthogonal transformation, i.e., $(R_\pi)^{-1} = R_\pi^\top$, then we can write $a^\top \pi y = a^\top R_\pi y = (a^\top R_\pi) y = (R_\pi^\top a)^\top y = (R_\pi^{-1} a)^\top y = (\pi^{-1} a)^\top y$. $\square$

Finally, we show that the Sherali-Adams relaxation is invariant under augmenting the defining system with valid inequalities. This property will be used to prove that $\mathrm{M}^k(P)$ is G-invariant and, in particular, justifies the convention adopted in Section 3.2.1 of omitting certain redundant valid inequalities when constructing $\mathrm{M}^k(P)$.

Lemma 3.24. *Let $Ax \geq b$ be a system of linear inequalities defining a polytope P contained in $[0,1]^n$, and let $\alpha^\top x \geq \beta$ be an inequality implied by $Ax \geq b$. Then*

$$\mathrm{M}^k(\{Ax \geq b\}) = \mathrm{M}^k(\{Ax \geq b, \alpha^\top x \geq \beta\}). \tag{3.6}$$

Proof. Clearly $\mathrm{M}^k(\{Ax \geq b, \alpha^\top x \geq \beta\}) \subseteq \mathrm{M}^k(\{Ax \geq b\})$. We now turn to the reverse inclusion.

Consider the linearization of the inequality $\prod_{i \in J_1} x_i \prod_{j \in J_0}(1 - x_j)(\alpha^\top x - \beta) \geq 0$, i.e.,

$$\left(\sum_{j \in J_1} \alpha_j - \beta\right) f_k(J_0, J_1) + \sum_{j \in [n] \setminus S} \alpha_j\, f_{k+1}(J_1 \cup \{j\}, J_0) \geq 0. \tag{3.7}$$

We show that (3.7) is implied by the inequalities of $\mathrm{M}^k(\{Ax \geq b\})$. First, suppose that the polytope P induced by the system $Ax \geq b$ has positive dimension. Then, by Proposition A.5 in Appendix A.7 there exists a vector of duals $u \geq 0$ such that $u^\top A = \alpha$ and $u^\top b = \beta$. Therefore, if we denote by A^j the j-th column of A, we have

$$\begin{aligned}
&\left(\sum_{j \in J_1} \alpha_j - \beta\right) f_k(J_0, J_1) + \sum_{j \in [n] \setminus S} \alpha_j\, f_{k+1}(J_1 \cup \{j\}, J_0) \\
&= \left(\sum_{j \in J_1} (u^\top A^j) - u^\top b\right) f_k(J_0, J_1) + \sum_{j \in [n] \setminus S} (u^\top A^j)\, f_{k+1}(J_1 \cup \{j\}, J_0) \\
&= u^\top \left(\left(\sum_{j \in J_1} A^j - b\right) f_k(J_0, J_1) + \sum_{j \in [n] \setminus S} A^j\, f_{k+1}(J_1 \cup \{j\}, J_0) \right),
\end{aligned}$$

which is a conic combination of inequalities of $\mathrm{M}^k(\{Ax \geq b\})$.

It remains to consider the case where P is 0-dimensional, i.e., it consists of a single point. This point is either integral or fractional. If it is integral, then (3.6) holds because $\alpha^\top x \geq \beta$ is trivially valid for the integer hull of P, and the projection of M^k onto the original space of variables, i.e., SA^k, preserves integrality. If the point is fractional, then (3.6) holds because both sides are empty: SA is at least as strong as the lift-and-project closure LS_0, and therefore cuts off every fractional vertex of P. □

Proposition 3.25. *Let $G \leq G(P)$ and $k \in \mathbb{N}$. Then $\mathrm{M}^k(P)$ is G-invariant.*

Proof. Consider $P := \{x \in \mathbb{R}^n : a_i^\top x \geq b_i, i \in [m]\} \subseteq [0,1]^n$. Due to Lemma 3.22, if $\pi \in G(P)$ we have that $(\pi a_i)^\top x \geq b_i$ is valid for P. Hence, by Lemma 3.24, if we add $(\pi a_i)^\top x \geq b_i$ to the description of P, the set $\mathrm{M}^k(P)$ does not change. Therefore, we can assume that the set of inequalities describing P, i.e., $\{a_i^\top x \geq b_i : i \in [m]\}$, is closed under permutations in $G(P)$, in the sense that $\alpha^\top x \geq \beta$ belongs to $\{a_i^\top x \geq b_i : i \in [m]\}$ if and only if $(\pi\alpha)^\top x \geq \beta$ belongs to $\{a_i^\top x \geq b_i : i \in [m]\}$.

We now introduce notation for the inequalities defining $\mathrm{M}^k(P)$. For each $S \subseteq [n]$ with $|S| = k$, each 2-partition $J_0 \sqcup J_1 = S$, and each row $a_i^\top x \geq b_i$ of the system $Ax \geq b$, inequality (3.2) takes the form

$$\left(\sum_{j \in J_1} a_{i,j} - b_i\right) f_k(J_0, J_1) + \sum_{j \in [n] \setminus S} a_{i,j}\, f_{k+1}(J_1 \cup \{j\}, J_0) \geq 0.$$

We define the vectors $r^{J_0,J_1}(a_i, b_i) \in \mathbb{R}^{\binom{n}{\leq k+1}}$ by setting, for every subset $I \subseteq [n]$ with $|I| \leq k+1$,

$$r^{J_0,J_1}(a_i, b_i)_I := \begin{cases} (-1)^{|H|}\left(\sum_{j \in J_1} a_{i,j} - b_i\right) & \text{if } I = J_1 \sqcup H,\ H \subseteq J_0, \\ (-1)^{|H|} a_{i,j} & \text{if } I = J_1 \sqcup \{j\} \sqcup H,\ H \subseteq J_0,\ j \in [n] \setminus S, \\ 0 & \text{o.w.} \end{cases}$$

Next, we prove that for any $y \in \mathrm{M}^k(P)$ and any $\pi \in G$, we have that $\pi y \in \mathrm{M}^k(P)$. Indeed, for $J_0 \sqcup J_1 = S$ with $|S| = k$ observe that $(\pi r^{J_0,J_1}(a_i, b_i))_I = (r^{J_0,J_1}(a_i, b_i))_{\pi^{-1}(I)} = (r^{\pi(J_0),\pi(J_1)}(\pi^{-1} a_i, b_i))_I$ for $I \subseteq [n]$ with $|I| \leq k+1$. Therefore, by Lemma 3.23 we have that

$$\begin{aligned} r^{J_0,J_1}(a_i, b)^\top (\pi y) \geq 0 &\iff (\pi^{-1} r^{J_0,J_1}(a_i, b))^\top y \geq 0 \\ &\iff (r^{\pi^{-1}(J_0),\pi^{-1}(J_1)}(\pi^{-1} a_i, b_i))^\top y \geq 0. \end{aligned}$$

Since π acts as a permutation of $[n]$, we have that $\pi(S) \subseteq [n]$, $|\pi(S)| = k$ and $\pi(J_0) \sqcup \pi(J_1) = \pi(S)$ is a 2-partition. Moreover, by assumption, if the inequality $a_i^\top x \geq b$ is present in the defining system of P, then so is $(\pi^{-1}a_i)^\top x \geq b$. Hence, $(r^{\pi^{-1}(J_0),\pi^{-1}(J_1)}(\pi^{-1}a_i, b_i))^\top y \geq 0$ is a valid inequality of $\mathrm{M}^k(P)$. We conclude that $\pi y \in \mathrm{M}^k(P)$ whenever $y \in \mathrm{M}^k(P)$. In consequence, $\mathrm{M}^k(P)$ is G-invariant. $\square$

A similar statement holds for LS_0 and LS.

Let $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ be a matrix with row and column indices in $\{0,1,\ldots,n\}$. For a permutation $\pi \in \mathcal{S}_n$, we extend it to $\{0,1\ldots,n\}$ by simply defining $\pi(0) = 0$. We define the *action of* G *on any lifting matrix* as

$$(\pi Y)_{i,j} \coloneqq Y_{\pi^{-1}(i),\pi^{-1}(j)} \tag{3.8}$$

for every $i,j \in \{0,1,\ldots,n\}$.

Proposition 3.26. *Let* $G \leq G(P)$*,* $k \in \mathbb{N}$*. Then* $\mathrm{LS}_0^k(P)$ *and* $\mathrm{LS}^k(P)$ *are* G*-invariant.*

Proof. It suffices to prove the statement for $k = 1$. Let us show first that LS_0 is symmetry-preserving. Assume that P is G-invariant. We aim to show that $\mathrm{LS}_0(P)$ is also G-invariant, that is, if $x \in \mathrm{LS}_0(P)$ and $\pi \in G$ then $\pi x \in \mathrm{LS}_0(P)$. Let Y be the lifting matrix associated to x, where $Ye_0 = (1,x)^\top$ and $Ye_i, Y(e_0 - e_i) \in K(P)$. We show that πY is a lifting matrix for πx. Let us extend the action of π to vectors in $\mathbb{R}^{n+1}$ as $\pi(x_0,\ldots,x_n) = (x_0, x_{\pi^{-1}(1)},\ldots,x_{\pi^{-1}(n)})$. Then, $(\pi Y)e_i = \pi(Ye_{\pi^{-1}(i)}) \in \pi K(P) = \{\pi x : x \in K(P)\}$. Using the straightforward fact that $K(P)$ is G-invariant (for the action of G extended to $\mathbb{R}^{n+1}$), we obtain that $\pi K(P) = K(P)$, and hence $(Ye_{\pi^{-1}(i)}) \in K(P)$ for every $i \in [n]$. With this we conclude that $(\pi Y)e_i \in K(P)$. The exact same analysis shows that $(\pi Y)(e_0 - e_i) \in K(P)$. Finally, we observe that due to the definition of the action of π on Y it holds that $(\pi Y)e_0 = (1,\pi x) = (\pi Y)^\top e_0 = \mathrm{diag}(\pi Y)$. We conclude that $\pi x \in \mathrm{LS}_0$.

Finally, assume that $x \in \mathrm{LS}(P)$. We follow the same proof as for LS_0 and show that if Y is the lifting matrix for $x \in \mathrm{LS}$, then πY is the corresponding lifting matrix for πx. Indeed, we have that Y is symmetric, $Y = Y^\top$. By using the definition of πY, we obtain that πY is also symmetric, as for all $i, j \in \{0, 1 \ldots, n\}$,

$$(\pi Y)_{i,j} = Y_{\pi^{-1}(i),\pi^{-1}(j)} = Y_{\pi^{-1}(j),\pi^{-1}(i)} = (\pi Y)_{j,i}.$$

We conclude that $\pi x \in \mathrm{LS}(P)$. $\square$

3.4.5 On k-Transitive Groups

We present a structural result on k-transitivity. This result follows from one of the major breakthroughs in modern mathematics, the classification of finite simple groups [165].

Theorem 3.27. *[158, Theorem 4.11] The finite* 2*-transitive permutation groups are explicitly known. In particular, the only finite* 6*-transitive groups are the full symmetric group* $\mathcal{S}_n$ *and the alternating group*[2] $\mathcal{A}_n$.

Computing a set of generators S for permutation group can be performed in quasipolynomial time [166], [167]. In practical computations $|S|$ is usually small, and any $G \leq \mathcal{S}_n$ can be generated by at most $n/2$ permutations; see [168]. On the other hand, if we are given a set of generators for a group, we show that (possibly known fact) we can identify its degree of transitivity in polynomial time.

Proposition 3.28. *Given* $G \leq S_n$ *by a set of generators* $S \subseteq G$*, for* $k \leq n$*, we can check in polynomial time (in* $|S|$ *and* n*), whether* G *is* k*-transitive or not.*

Proof. In [168, Page 49, Chapter 3] there is a list of tasks that can be computed in polynomial-time. Consider the following tasks: 1) Task (a), Find orbits of G; 2) Task (e), Given Δ subset of $[n]$, find the pointwise stabilizer of Δ in G; 3) Task (b), Given h in $\mathcal{S}_n$, test whether h belongs to G. Therefore, since we have that G is k-transitive iff for any i in $[n]$ the subgroup G_i is $(k-1)$-transitive, we can decide in polynomial time whether G is k-transitive.

[2] This is the group of all the permutations in $\mathcal{S}_n$ which can be written as product of an even number of transpositions $(i\,j)$.

Indeed, given G by a set of generators, by 1) we can check whether G is transitive or not (we get only one orbit). If it is not, we are done. If it is transitive, then for every i in $[n]$, by 2) we can compute the subgroup G_i (which is the pointwise stabilizer of the set $\{i\}$. Then by 1) again we can check whether it is transitive or not. If it is not for some i, then we have that G is 1-transitive but not 2-transitive. Otherwise, it is at least 2-transitive and we continue. For every $j > i$, by 2) we find the pointwise stabilizer of $\{j\}$ in G_i (we computed G_i before). Then by 1) we can find the orbits of $G_{i,j}$ etc. Since for $k \geq 6$ there are only two possible k-transitive groups ($\mathcal{S}_n$ and $\mathcal{A}_n$), then for $k \geq 6$ it suffices to know whether G is 6-transitive or not. If it is, then we just need to distinguish between $\mathcal{S}_n$ and $\mathcal{A}_n$ by 3) we can check whether any odd permutation belongs to G or not. ($\mathcal{S}_n$ is n-transitive while $\mathcal{A}_n$ is $(n-2)$-transitive). □

See [168] for further background on computational group theory.

3.5 On the Improvement of Integrality Gaps

Consider the following integer programs.

$$\mathcal{P}_1 : \quad \min_{x\in\{0,1\}^3} \quad x_1+x_2+x_3$$

$$x_1 - 2x_2 - 2x_3 \geq -3 \tag{3.9a}$$

$$-2x_1 + x_2 - 2x_3 \geq -3 \tag{3.9b}$$

$$-2x_1 - 2x_2 + x_3 \geq -3 \tag{3.9c}$$

$$2x_1 + 2x_2 + 2x_3 \geq 3 \tag{3.9d}$$

$$\mathcal{P}_2 : \quad \min_{x\in\{0,1\}^3} \quad x_1+x_2+x_3$$

$$2x_1 - x_2 - 4x_3 \geq -3 \tag{3.10a}$$

$$-4x_1 + 2x_2 - x_3 \geq -3 \tag{3.10b}$$

$$-x_1 - 4x_2 + 2x_3 \geq -3 \tag{3.10c}$$

$$2x_1 + 2x_2 + 2x_3 \geq 3 \tag{3.10d}$$

We denote by F_1^* and F_2^* the optimal LP faces of the formulations (3.9) and (3.10), respectively; see Figure 3.3. It is straightforward to verify that both integer programs admit a unique feasible integer solution, namely $(1, 1, 1)$, and are therefore as integer programs. However, their linear programming relaxations are geometrically distinct. Despite this, they are equivalent from a bounding perspective, as both optimal LP faces attain the same objective value of $3/2$.

Interestingly, applying one round of the LS_0 operator yields significantly different behavior. By Proposition 3.14, $\text{LS}_0(F) = \text{LS}_0(P) \cap F$, so it suffices to analyze the operator on the optimal LP face. Because this face contain no integer points, the relevant question is whether LS_0 eliminates it, i.e., whether LS_0 *improves the LP bound*.

From Figure 3.3, we readily deduce that $\text{LS}_0(F_2^*) = \emptyset$, since for $i \neq j$, the polytopes $P_i(F_2^*) \cap P_j(F_2^*)$ correspond to distinct vertices of F_2^*. In constrast, $\text{LS}_0(F_1^*) \neq \emptyset$, as it containts the point $(1/2, 1/2, 1/2)$.

This discrepancy is not incidental. Rather, it reflects an intrinsic structural difference between the two polytopes, which can be understood through their symmetries. Geometrically, F_2^* lacks certain *supporting* points (highlighted in blue in Figure 3.3) that would allow the formation of convex combinations *resistant* to the action of LS_0.

This distinction is captured by the symmetry groups acting on their vertex sets. The polytope F_1^* is the convex hull of all permutations of $(1, 1/2, 0)$, and thus admits the full symmetric group $\mathcal{S}_3$. In contrast, F_2^* admits only the cyclic group C_3. Since both groups act transitively on the vertices, this alone is insufficient to distinguish their behavior. A finer notion is therefore required, namely k-transitivity. In particular, the symmetry group of F_1^* is 3-transitive, whereas that of F_2^* is only 1-transitive.

Remark 3.29. Let P be a linear relaxation of $\mathcal{P}$, and let F be a face of P. Then $G(P) \leq G(F)$. Moreover, if F is induced by an inequality of the form $\mathbf{1}^\top x \geq \beta$, as is often the case when studying integrality gaps in symmetric problems, then $G(P) = G(F)$.

As Theorem 3.1 suggests, symmetry alone is not sufficient for a polytope to resist cut operators such as LS_0. Indeed, let $F' \coloneqq \text{conv}(\{\pi((1+t)/2, 1/2, (1-t)/2) : \pi \in \mathcal{S}_3\})$ for $0 < t < 1$. By definition F' is 3-transitive polytope, and it is straightforward to verify that it is strictly contained in F_1^*. Morever, $\text{LS}_0(F') = \emptyset$, in agreement with the theorem.

Therefore, Theorem 3.1 ultimately hinges on an additional condition beyond symmetry. Returning to the guiding question posed in the introduction – *can we characterize the conditions under which a symmetric optimal face of a relaxation is eliminated at a fractional node of the*

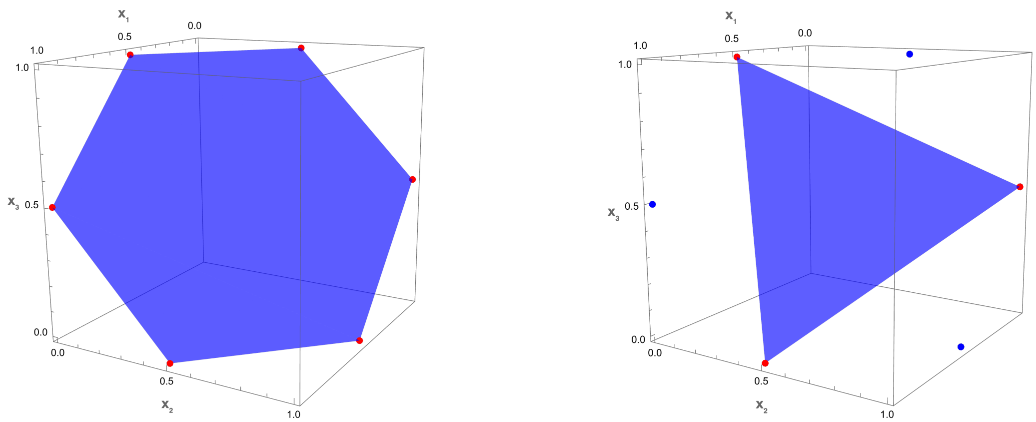

Figure 3.3: A depiction of the optimal LP faces of (3.9) (left) and (3.10) (right).

branch-and-cut tree? – the theorem provides a partial positive answer in the presence of sufficient symmetry in the underlying binary integer problem, when the cuts considered arise from the LP hierarchies discussed throughout this chapter. See Section 3.6.2 for an example of LP bound resistance after adding a non-optimal objective cut.

3.5.1 k-Resistance

We formalize the notion of resistance. In general, we aim to understand which properties an integer program (not necessarily symmetric) must satisfy so that its integrality gap does not improve after a large (typically linear in n) number of rounds of a cut operator T. We say that P is k-T *resistant* if there exists $c \in \mathbb{R}^n$ such that

$$\min_{x \in P} c^T x = \min_{x \in T^k(P)} c^T x < \min_{x \in P \cap \{0,1\}^n} c^T x, \tag{3.11}$$

where any of the values above is defined to be ∞ if the corresponding problem is infeasible.

Consider now a class of polytopes $P(n) \subseteq [0,1]^n$ for $n \in \mathbb{N}$, and let $f : \mathbb{N} \to \mathbb{N}$. We say that $P(n)$ is $\Theta(f)$-T resistant if there exists a function $g : \mathbb{N} \to \mathbb{N}$ with $g = \Theta(f)$ such that, for each $n \in \mathbb{N}$, $P(n)$ is $g(n)$-T resistant.

A direct application of Theorem 3.1 provides sufficient conditions toward answering this question. In particular, let $\{P(n)\}_{n \in \mathbb{N}}$ be a family of polytopes with $P(n) \subseteq [0,1]^n$. If, for each

n, $P(n)$ is integer-empty, its optimal LP face is $(k+1)$-transitive, and it intersects all $(n-k)$-dimensional faces of $[0,1]^n$, then $P(n)$ is k-T resistant for $T \in \{\mathrm{LS}_0, \mathrm{LS}_+, \mathrm{SA}\}$.

Resistance under Mixed-Order RLT Factors

We next consider the use of RLT factors of different orders and exhibit an instance in which, even after adding $\Theta(n)$ such factors, there is still no improvement in the LP bound. This shows that the inclusion of higher-order factors does not suffice to strengthen the relaxation.

Consider a polytope $P \subseteq [0,1]^n$ described by $a_i^\top x \geq b_i$, $i \in [m]$. Given a set K of RLT factors, we say that the ***RLT procedure*** is applied to P using K if every factor in K is used in the Sherali-Adams extend-and-linearize construction, namely, steps (i) and (ii) of Section 3.2.1. We say that we ***RLT by*** a d-junta $F_d(J_0, J_1)(x)$ if we multiply every inequality defining P by this factor and then multilinearize. We denote by $[\,\cdot\,]_L$ the multilinearization operator acting on polynomial expressions by first applying the identities $x_i^2 = x_i$ for every $i \in [n]$, then replacing each monomial $\prod_{i\in I} x_i$, where $I \subseteq [n]$, with y_I, setting $y_{\{i\}} = x_i$, and $y_\emptyset = 1$.

Using this notation, the relaxation $\mathrm{M}^k(P)$ can be written, following the original definition of Sherali and Adams [150], as follows; see Remark 3.4.

(SA-0) $y_\emptyset = 1$;

(SA-1) For each 2-partition $J_0 \sqcup J_1 = S \subseteq [n]$ with $|S| = \min\{k+1, n\}$,

$$\left[F_{|S|}(J_0, J_1)(x)\right]_L \geq 0;$$

(SA-2) For each $i \in [m]$ and each 2-partition $J_0 \sqcup J_1 = S \subseteq [n]$ with $|S| = k$,

$$\left[F_k(J_0, J_1)(x)\,(a_i^\top x - b_i)\right]_L \geq 0.$$

Given a set K of polynomial RLT factors, let $\deg(K)$ denote the maximum degree of a factor in K. We define $\mathrm{R}_K(P)$ as the linear system obtained by applying the RLT procedure to P using the factors in K and intersecting the resulting system with $P \times \mathbb{R}^{\binom{n}{\leq \deg(K)+1}-n}$. In particular, $\mathrm{R}_K(P)$ implies all the original inequalities defining P, and we assume that the equation $y_\emptyset = 1$ is included.

Now we exhibit an instance that is resistant to $\Theta(n)$ RLT factors of arbitrary degree. Let

$$P := \left\{ x \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} x_i \geq \delta \right\}$$

for some $0 < \delta < 1$, and consider the binary integer program

$$\min \left\{ \mathbf{1}^\top x \;\middle|\; x \in P \cap \{0,1\}^n \right\}.$$

We show that if the RLT procedure is applied to P using any collection of $(n-1)$ juntas, of arbitrary degrees $d_i \in [n]$, then the optimal value of the resulting LP relaxation is unchanged.

Recall that the multilinearization of any d-junta $F_d(J_0, J_1)(x)$ is given by

$$[F_d(J_0, J_1)(x)]_L = \sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup H}. \tag{3.12}$$

Moreover, for $k \in J_0 \sqcup J_1$, we have the following *absorption properties*

$$x_k \cdot F_d(J_0, J_1)(x) = \begin{cases} F_d(J_1, J_0)(x) & \text{if } k \in J_1 \\ 0 & \text{otherwise, i.e., } k \in J_0, \end{cases} \tag{3.13}$$

and

$$(1 - x_k) \cdot F_d(J_0, J_1)(x) = \begin{cases} F_d(J_0, J_1)(x) & \text{if } k \in J_0 \\ 0 & \text{otherwise, i.e., } k \in J_1. \end{cases} \tag{3.14}$$

Lemma 3.30. *Let $\mathcal{D}$ be any set of distinct d_i-juntas, where $1 \leq d_i \leq n$ and $i \in [|\mathcal{D}|]$. If $|\mathcal{D}| \leq \lfloor n/2 \rfloor - 1$, then there exists some pair of indices $i^*, j^* \in [n]$, such that, for every junta*

$F_{d_i}(J_0, J_1)(x) \in \mathcal{D}$ *whose corresponding* 2*-partition satisfies* $|J_1| \leq 2$*, we have*

$$\{i^*, j^*\} \cap J_1 = \emptyset.$$

Proof. Notice that if $|\mathcal{D}| \leq \lfloor n/2 \rfloor - 1$ at most $\lfloor n/2 \rfloor - 1$ disjoint pairs $\{i, j\}$ can be chosen. Hence, there exists some pair $\{i^*, j^*\}$ which is disjoint to all of them. $\square$

Proposition 3.31. *Let K be any set of d_i-juntas. Suppose that $|K| \leq \lfloor n/2 \rfloor - 1$. Then*

$$\min \{\mathbf{1}^\top x \,:\, (x, y) \in \mathrm{R}_K(P)\} \leq \delta.$$

Proof. Since $|K| \leq \lfloor n/2 \rfloor - 1$, let $\{i^*, j^*\}$ be a pair of indices as in Lemma 3.30. Consider the following vector: $x'_{i^*} = x'_{j^*} = \frac{\delta}{2}$, $x'_i = 0$ for all $i \in [n] \setminus \{i^*, j^*\}$, $y'_\emptyset = 1$, and $y'_I = -(1-\delta)$ if $I = \{i^*, j^*\}$, and $y'_I = 0$ if I is a non-empty subset of variables. Notice that $\mathbf{1}^\top x' = x'_{i^*} + x'_{j^*} = \delta$, therefore it suffices to prove that $(x', y') \in \mathrm{R}_K(P)$.

(SA-1) **Constraints** First, we show that (x', y') satisfies all the SA-1 constraints. Let $F_d(J_0, J_1)$ be in K. Notice that any consistency constraint can be written as

$$[F_{d+1}(J_0, J_1 \cup \{k\})(x)]_L \geq 0 \quad \text{or} \quad [F_{d+1}(J_0 \cup \{k\}, J_1)(x)]_L \geq 0$$

if $k \notin J_1 \cup J_0$.

First, observe that $[F_d(J_0, J_1)(x')]_L \geq 0$ if $J_1 \neq \emptyset$. Suppose $J_1 \neq \emptyset$. Either $|J_1| \geq 3$, or $|J_1| \leq 2$ and $\{i^*, j^*\} \cap J_1 = \emptyset$. In both cases, $k \in J_1$, $k \neq i^*, j^*$ hence all variables are of the kind $y_{I \cup \{k\}}$. Moreover, notice that if $J_1 \neq \emptyset$, then (x', y') satisfies the corresponding SA-1 constraints since the only nonzero entries of (x', y') are $x'_{i^*}, x'_{j^*}, y'_{i^*, j^*}$ and they do not participate in this constraint.

Hence, the non-trivial cases are when $J_1 = \emptyset$, which we enumerate below.

(i) Suppose $k \in [n] \setminus J_0$. Then by (3.12), we have that $[F_{d+1}(J_0 \cup \{k\}, \emptyset)(x)]_L$ equals

$$\sum_{H \subseteq J_0 \cup \{k\}} (-1)^{|H|} y'_H = \begin{cases} 1 - x'_{i^*} & \text{if } i^* \in J_0 \cup \{k\} \text{ and } j^* \notin J_0 \\ 1 - x'_{j^*} & \text{if } j^* \in J_0 \cup \{k\} \text{ and } i^* \notin J_0 \\ 1 - x'_{i^*} - x'_{j^*} + y_{i^*,j^*} & \text{if } i^*, j^* \in J_0 \cup \{k\} \\ 1 & \text{otherwise.} \end{cases}$$

Hence, in any case the latter sum is nonnegative.

To study $[F_{d+1}(J_0, \{k\})(x)]_L$ we consider two sub-cases:

(a) Let $k \in \{i^*, j^*\}$ and assume without loss of generality that $k = i^*$. Then by (3.12), we have

$$[F_{d+1}(J_0, \{k\})(x)]_L = \sum_{H \subseteq J_0} (-1)^{|H|} y'_{\{i^*\} \cup H} = \begin{cases} x'_{i^*} & \text{if } j^* \notin J_0 \\ x'_{i^*} - y_{i^*,j^*} & \text{if } j^* \in J_0. \end{cases}$$

Clearly any case is positive.

(b) Assume $k \neq i^*, j^*$. Then

$$[F_{d+1}(J_0, \{k\})(x)]_L = \sum_{H \subseteq J_0} (-1)^{|H|} y'_{\{k\} \cup H} = 0,$$

Since $y'_I = 0$ whenever $I \setminus \{i^*, j^*\} \neq \emptyset$.

(ii) Suppose $k \in J_0$. Then by (3.12) and (3.14), we have

$$
\begin{aligned}
[(1-x_k)\cdot F_d(J_0,\emptyset)(x)]_L &= [F_d(J_0,\emptyset)(x)]_L \\
&= \textstyle\sum_{H\subseteq J_0}(-1)^{|H|}y'_{\{k\}\cup H} \\
&= \begin{cases} 1-x'_{i^*} & \text{if } i^*\in J_0 \text{ and } j^*\in[n]\setminus J_0 \\ 1-x'_{j^*} & \text{if } j^*\in J_0 \text{ and } i^*\in[n]\setminus J_0 \\ 1-x'_{i^*}-x'_{j^*}+y_{i^*,j^*} & \text{if } i^*,j^*\in J_0 \\ 1 & \text{otherwise.} \end{cases}
\end{aligned}
$$

Moreover, the absorption property (3.13) and $k \in J_0$ implies $x_1 \cdot F_d(J_0,\emptyset)(x) = 0$. Therefore (x', y') satisfies any SA-1 constraint.

(SA-2) **Constraints** We show that (x', y') satisfies any SA-2 constraint. Using (3.14), we need to verify that

$$\left[\sum_{k\in[n]\setminus J_0} x_k\left(\prod_{j\in J_1} x_j\prod_{i\in J_0}(1-x_i)\right)-\delta\left(\prod_{j\in J_1} x_j\prod_{i\in J_0}(1-x_i)\right)\right]_L \geq 0.$$

for every $F_d(J_0, J_1)(x) \in K$. Using (3.12), the previous inequality can be rewritten as:

$$\underbrace{\sum_{k\in[n]\setminus J_0}\sum_{H\subseteq J_0}(-1)^{|H|}y'_{J_1\cup\{k\}\cup H}}_{A}-\delta\underbrace{\left(\sum_{H\subseteq J_0}(-1)^{|H|}y'_{J_1\cup H}\right)}_{B}\geq 0 \tag{3.15}$$

As with the SA-1 constraints, we can assume that $J_1 = \emptyset$. In fact, if $J_1 \neq \emptyset$, then $J_1\setminus\{i^*, j^*\} \neq \emptyset$, hence $y'_{J_1\cup\{k\}\cup I} = y'_{J_1\cup I} = 0$ for every I. Hence, the inequality (3.15) is satisfied by (x', y'). We enumerate the cases where $J_1 = \emptyset$.

(i) Suppose $|\{i^*, j^*\} \cap J_0| = 2$. Notice that

$$B = \sum_{H\subseteq J_0}(-1)^{|H|}y'_H = 1 - x'_{i^*} - x'_{j^*} + y'_{i^*,j^*} = 0.$$

If $[n] \setminus J_0 \neq \emptyset$ then for $k \in [n] \setminus J_0$, we have $y'_{\{k\} \cup H} = 0$ for all $H \subseteq J_0$. If conversely $J_0 = [n]$, then the first summation in A is empty. In both cases, $A = 0$ and (x', y') satisfies the inequality (3.15).

(ii) Suppose $|\{i^*, j^*\} \cap J_0| = 1$. Without loss of generality, suppose $i^* \in J_0$ and $j^* \in [n] \setminus J_0$. Then

$$A = \sum_{k \in [n] \setminus J_0} \sum_{H \subseteq J_0} (-1)^{|H|} y'_{\{k\} \cup H} = x'_{j^*} - y'_{i^*, j^*} = 1 - \frac{\delta}{2}.$$

On the other hand,

$$B = \delta \left(\sum_{H \subseteq J_0} (-1)^{|H|} y'_{J_1 \cup H} \right) = \delta(1 - x'_{i^*}) = \delta - \frac{\delta^2}{2}.$$

Since $1 - \frac{3\delta}{2} + \frac{\delta^2}{2} = \frac{1}{2}(\delta - 1)(\delta - 2) \geq 0$ for any $0 < \delta \leq 1$, then (x', y') satisfies the inequality (3.15).

(iii) Assume $|\{i^*, j^*\} \cap J_0| = 0$. Then

$$A = \sum_{k \in [n] \setminus J_0} \sum_{H \subseteq J_0} (-1)^{|H|} y'_{\{k\} \cup H} = x_{i^*} + x_{j^*} = \delta,$$

while

$$B = \delta \left(\sum_{H \subseteq J_0} (-1)^{|H|} y'_H \right) = \delta.$$

Therefore (x', y') satisfies the inequality (3.15).

Therefore (x', y') is feasible for $\mathrm{R}_K(P)$ and we conclude the proof. □

3.6 Examples and Applications

3.6.1 A 3-Dimensional Example

We illustrate Theorem 3.1 on a 3-dimensional polytope. In Figure 3.4 we consider the polytope $P = \mathrm{conv}(\{\pi(0, 1/2, 1) : \pi \in \mathcal{S}_3\})$, that is, the convex hull of all points obtained by permuting the coordinates of $(0, 1/2, 1)$ in every possible way. This polytope coincides with the LP optimal face of (3.9).

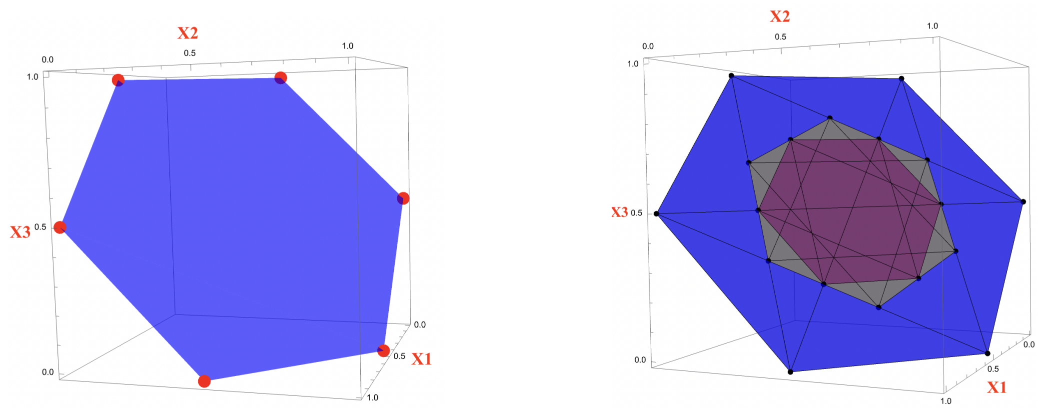

Figure 3.4: A depiction of $P = \text{conv}(\{\pi(0, 1/2, 1) : \pi \in \mathcal{S}_3\})$ (left), $\text{SA}^1(P) = \text{LS}^1(P)$ (right, violet), and $\text{LS}_0^1(P)$ (right, gray).

By construction, P is 3-transitive (and hence 2-transitive). In the left figure, one can see the original polytope. We note that P intersects every 2-dimensional face of $[0, 1]^3$. On the right, we illustrate $\text{SA}^1(P)$, $\text{LS}^1(P)$, and $\text{LS}_0^1(P)$. As predicted by Theorem 3.1, when taking $k = 1$ these three polytopes are non-empty; however they are not necessarily equal as $\text{SA}^1(P) = \text{LS}^1(P) \subsetneq \text{LS}_0^1(P)$. In Proposition 3.37, we prove that inclusion is strict by exhibiting a point in $\text{LS}_0(P)$ that does not belong to $\text{SA}(P)$.

On the other hand, since P does not intersect every 1-dimensional face (e.g., the face defined by $x_1 = 1$ and $x_2 = 1$ does not intersect P), Theorem 3.1 implies that $\text{SA}^2(P) = \text{LS}^2(P) = \text{LS}_0^2(P) = \emptyset$.

3.6.2 Steiner Triple Covering Problems

Starting with the work of Fulkerson, Nemhauser, and Trotter [120], Steiner Triple Covering Problems (STSs) have been studied computationally by many researchers in Integer Programming [122], [123], [169], [170]. These covering problems are based on a combinatorial structure known as a *Steiner Triple System* on a ground set Ω. Formally, given a finite set Ω, a Steiner Triple System a collection $\mathcal{B} \subseteq 2^\Omega$ whose elements, called *triples*, have cardinality 3, and such that every pair of elements of Ω is contained in exactly one triple T of $\mathcal{B}$.

In the associated covering formulation, we introduce a binary variable x_i for each element $i \in \Omega$. Without loss of generality, we identify the ground set Ω with $[|\Omega|]$, so that the number of variables is $|\Omega|$. Most of the work in the area has focused on two classes of STSs in which $|\Omega|$ is either a power of 3 or 5 times a power of 3. For $n \in \mathbb{N}$, we let STS_n denote the corresponding covering instances with $|\Omega| = 3^n$ variables. We now provide a formal description; see [171] for background.

STS_1 is defined as the following set covering instance

$$\min \quad \{\mathbf{1}^T x \,:\, x_1 + x_2 + x_3 \geq 1,\, x \in \{0,1\}^3\}.$$

For $n \in \mathbb{N}_{\geq 2}$, we define STS_n as a function of STS_{n-1} as follows. We let A_{n-1} be the constraint matrix for STS_{n-1}. We let I_{n-1} be the identity matrix with 3^{n-1} rows and columns. We let B_n be the matrix with 3^n columns and whose rows are all and only the vectors that satisfy the following: (a) three entries i_0, i_1, i_2 are equal to 1, and the remaining entries are equal to 0; (b) $3^j + 1 \leq i_j \leq 3^{j+1}$ for $j = 0, 1, 2$; (c) $i_j \mod 3^n$ assume three different values for $j \in \{0, 1, 2\}$. STS_n is then defined as the following set covering problem (for the sake of readability, we omit the 0s):

$$\begin{array}{lll}
\displaystyle\min_{x \in \{0,1\}^n} & \mathbf{1}^T x & \\
\text{s.t.} & \begin{pmatrix} A_{n-1} & & \\ & A_{n-1} & \\ & & A_{n-1} \\ I_{n-1} & I_{n-1} & I_{n-1} \\ & B_n & \end{pmatrix} x & \geq \mathbf{1}
\end{array}$$

We note that the matrix A_n is sparse, with exactly three ones per row. In the construction above, each row corresponds to a triple of the underlying Steiner Triple System on $\Omega = [3^n]$. Moreover, it is well known, via a simple counting argument, that A_n has $3^n(3^n - 1)/6$ rows [171].

Lemma 3.32. *Every Steiner Triple System over a set* Ω *with triple set* $\mathcal{B}$ *satisfies* $|\mathcal{B}| = \frac{|\Omega|(|\Omega|-1)}{6}$. *In particular,* A_n *has* $3^n(3^n-1)/6$ *rows.*

Proof. We count in two different ways the total number of pairs (i, j) over the ground set Ω. We know that every $T \in \mathcal{B}$ contains 3 elements, so each T has exactly 3 different pairs and these pairs do not belong to any other T' of $\mathcal{B}$. Therefore, $3 \cdot |\mathcal{B}| = \#$pairs which implies $|B| = \frac{|\Omega|(|\Omega|-1)}{6}$. $\square$

Example 3.33. STS_2 is given by the following binary integer program.

$$
\begin{array}{lccccccccccl}
\min\limits_{x\in\{0,1\}^9} & x_1 & +x_2 & +x_3 & +x_4 & +x_5 & +x_6 & +x_7 & +x_8 & +x_9 & \\
 & x_1 & +x_2 & +x_3 & & & & & & & \geq 1 \\
 & & & & x_4 & +x_5 & +x_6 & & & & \geq 1 \\
 & & & & & & & x_7 & +x_8 & +x_9 & \geq 1 \\
 & x_1 & & & +x_4 & & & +x_7 & & & \geq 1 \\
 & & x_2 & & & +x_5 & & & +x_8 & & \geq 1 \\
 & & & x_3 & & & +x_6 & & & +x_9 & \geq 1 \\
 & x_1 & & & & +x_5 & & & & +x_9 & \geq 1 \\
 & x_1 & & & & & +x_6 & & +x_8 & & \geq 1 \\
 & & x_2 & & +x_4 & & & & & +x_9 & \geq 1 \\
 & & x_2 & & & & +x_6 & +x_7 & & & \geq 1 \\
 & & & x_3 & +x_4 & & & & +x_8 & & \geq 1 \\
 & & & x_3 & & +x_5 & & +x_7 & & & \geq 1
\end{array}
$$

An optimal solution to this instance is, e.g., $x^* = (1, 1, 0, 1, 0, 1, 0, 1, 0)$ with objective value 5.

For $n \in \mathbb{N}_{\geq 2}$, the feasible region of the LP relaxation of STS_n, which we denote by P_n, is known to be 2-transitive [172]. It is straightforward to see that it is not 3-transitive. Fix $n \in \mathbb{N}_{\geq 2}$.

By construction, $x_1 + x_2 + x_3 \geq 1$ is a valid constraint of STS_n. It is not hard to see that there exists a feasible solution $\overline{x}$ for STS_n that sets at least 3 variables to 0. If the feasible region of the linear programming relaxation of STS_n were 3-transitive, then there would exist a feasible solution $\hat{x}$ such that $\hat{x}_1 = \hat{x}_2 = \hat{x}_3 = 0$, violating $x_1 + x_2 + x_3 \geq 1$.

On the 2-Resistance of STS

We first show that its optimal of the linear programming relaxation of STS_n is a fractional vertex.

Proposition 3.34. *For $n \in \mathbb{N}$, the point $\frac{1}{3}\mathbf{1}_{3^n}$ is the unique minimizer of the LP relaxation of STS_n.*

Proof. By induction on $n \in \mathbb{N}$. Base case $n = 2$. By summing up all the rows of A_9 we obtain the valid inequality $4\mathbf{1}_9^\top x \geq 12$ for the LP, i.e., $\mathbf{1}_9^\top x \geq 3$, hence the LP bound is lower bounded by 3. Notice that $\overline{x} := \frac{1}{3}\mathbf{1}_9$ is feasible for the LP and it lies on $F := \{x \in P \,:\, \mathbf{1}_9^\top x = 3\}$. We show that $\dim(F) = 0$. To this end, we first show that every row of $A_9 x \geq \mathbf{1}_9$ is an implicit equality in F. Indeed, it straightforward to verify that, for any $x \in F$, if $x_{i_1} + x_{i_2} + x_{i_3} > 1$ for some for triple $\{i_1, i_2, i_3\}$, then adding it up with its associated two other triples with disjoint support, i.e., on $[9] \setminus \{i_1, i_2, i_3\}$, we obtain $\mathbf{1}_9^\top x > 3$. Since the rank of A_9 is 9, we conclude the base case. For $n > 2$ arbitrary, we can readily apply the induction hypothesis since $A_{3^{n-1}}$ is embedded by construction on A_{3^n}. Clearly, the rows associated to the three diagonal blocks $A_{3^{n-1}}$ of A_{3^n} are linearly independent and there are $3 \cdot \frac{3^{n-1}(3^{n-1}-1)}{6}$ many of them by Lemma 3.32. Therefore, since $3 \cdot \frac{3^{n-1}(3^{n-1}-1)}{6} > 3^n$ for $n > 2$, we conclude that the optimal LP face of the linear programming relaxation of STS_n is 0-dimensional given by the point $\frac{1}{3}\mathbf{1}_{3^n}$. $\square$

By Proposition 3.34, applying any hierarchy considered in this work to P_n of any hierarchy improves the optimality gap (even the BCC operator P_i for any variable $i \in [n]$, since the optimal LP face is a fully fractional vertex). Consequently, applying Theorem 3.1 to the face $P_n \cap \{x \in \mathbb{R}^n \,:\, \mathbf{1}^\top x = 3^n/3\}$ is not particularly informative.

We therefore construct a 2-transitive, integer-empty face of a tighter relaxation of STS_2 that satisfies property (A) with $k = 1$ of Theorem 3.1. This is achieved by adding a non-optimal objective function cut (i.e., a valid inequality) to the LP relaxation of STS_2. A similar argument applies for larger values of n.

Consider the cut $\sum_{i=1}^{n} x_i \geq 4$, whose intersection with the linear programming relaxation of STS_2 defines a face F that is integer-empty. Since the cut is parallel to the objective, it preserves the symmetries of P_n; hence, by [172], we conclude that F is 2-transitive. Moreover, because the cut is non-optimal, F does not contain integer feasible points.

Thus, we are left to show that F intersects all $(n-1)$-dimensional faces of the hypercube. By symmetry, we only need to consider $F_i = \{x \in [0,1]^9 : x_1 = i\}$ for $i = 0, 1$. We then conclude the argument by observing that

$$\left(0, 1, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \in F \cap F_0$$

and

$$\left(1, 0, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \in F \cap F_1.$$

Therefore, one round of any of the LP hierarchies considered in this work does not suffice to improve the bound given by the LP relaxation where the cut has been added. STS_n illustrates how Theorem 3.1 can be employed to give a bound on the number of rounds we need to apply the hierarchies for, in order to improve the bound given by some LP relaxation of a (symmetric) integer program.

3.6.3 Parity and Knapsack-Cover

Understanding the strength of hierarchies to prove approximate bounds has attracted significant attention. In particular, constructing lower bounds is a technically challenging but important task in order to understand the limitation of such techniques. The Sum-of-Squares (SoS) hierarchy, a stronger construction than SA, LS, and LS_0, has been studied extensively in this setting. The starting point of several of these constructions [128], [139], [173] was given in [127], who showed that

$k = \lfloor n/2 \rfloor$ rounds of the SoS hierarchy are needed to prove that $P = \{x \in [0,1]^n : \sum_{i=1}^n x_i = n/2\}$ is integer-empty when n is odd. Noting that P is $\mathcal{S}_n$-invariant, and thus n-transitive, our main theorem implies directly that the (weaker) Sherali-Adams hierarchy needs exactly $k+1 = \lfloor n/2 \rfloor + 1$ rounds to prove that P is integer-empty. Although this is not a new result, it gives a very simple proof for this fact and exemplifies the usefulness of our techniques: note that it suffices to find a vector in the intersection of P with each $(n-k)$-dimensional face of $[0,1]^n$, a straightforward exercise for this polytope. Moreover, we note that the construction of a feasible point of $\text{SA}^k(P)$ we use in the proof of (A)⇒(B) in Theorem 3.1 is a strict generalization of the ideas in [127] to full-dimensional polytopes.

Another application considers the Knapsack-Cover problem (also known as Min-Knapsack). While the basic linear relaxation of this problem has an unbounded integrality gap, the introduction of knapsack-cover inequalities reduces this gap to 2. Despite efforts to tighten this relaxation using hierarchies like Sherali-Adams, Lovász-Schrijver, and SoS, the integrality gap remains at 2 even after several rounds. We show that our construction generalizes the approach in [174], [175], offering a simpler method to generate feasible vectors in the Sherali-Adams hierarchy. For a specific instance, the knapsack-cover linear relaxation yields

$$P = \min\left\{\sum_{i=1}^n x_i : \sum_{i=1}^n x_i \geq 1 + \tfrac{1}{n-1}, \sum_{j\in[n]\setminus\{i\}} x_j \geq 1, x \in [0,1]^n\right\}.$$

In Section 3.9.5, we show that our techniques of Section 3.7 give a direct method for constructing a solution in $x \in \text{SA}^k(P)$ with value $\sum_{i=1}^n x_i = 1 + O(k/n)$, while the optimal integral solution has value 2, which implies an integrality gap of 2 as long as k is sublinear, $k = o(n)$.

We emphasize that in both constructions our explicit solution is guaranteed to belong to the Sherali-Adams hierarchy; we leave as an open problem if the same solution belongs to the tighter SoS hierarchy at the same level.

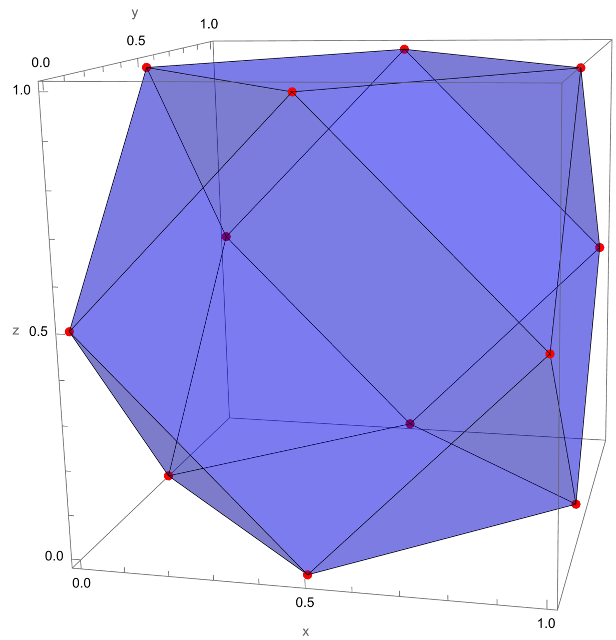

Figure 3.5: A geometric representation of P_3^{CC} in $[0,1]^3$.

3.6.4 The Cropped Cube

The cropped cubed is a polytope obtained by cropping each vertex of the 0-1 hypercube, given by

$$P_n^{CC} := \left\{ x \in [0,1]^n \ : \ \textstyle\sum_{i \in I} x_i + \sum_{i \notin I} (1 - x_i) \geq 1/2, \ \text{ for all } I \subseteq [n] \right\}.$$

See Figure 3.5. It is straightforward to verify that this polytope is integer-empty and n-transitive. In [142], [143] is shown that it resists $n-1$ rounds of the Lovász-Schrijver operator with PSD constraint, meaning that $\mathrm{LS}_+^{n-1}(P_n^{CC}) = \mathrm{LS}^{n-1}(P_n^{CC}) \neq \emptyset$. Moreover, in [128] it is shown that $\mathrm{SA}^{n-1}(P_n^{CC}) \neq \emptyset$. Theorem 3.1 implies directly these results for the operators LS and SA: we only need to observe that P_n^{CC} intersects every $n-1$ dimensional face of $[0,1]^n$. Namely, we need to verify that for any two disjoint sets $J_0, J_1 \subseteq [n]$ with $|J_0 \cup J_1| = n-1$, there is a vector $x \in P_n^{CC}$ with $x_i = 0$ if $i \in J_0$ and $x_i = 1$ for $i \in J_1$. To do so, simply define the only free variable as $x_i = 1/2$, for $i \in [n] \setminus (J_0 \cup J_1)$.

3.7 Main Theorem: LP Hierarchies Collapse under Symmetry

We restate the main theorem of this chapter. We refer the reader to Section 3.6 for examples and applications.

Theorem 3.35 (Restatement of Theorem 3.1)**.** *Let $k, n \in \mathbb{N}$ such that $k < n$, and let $P \subseteq [0,1]^n$ be a $(k+1)$-transitive integer-empty polytope. Then the following statements are equivalent:*

(A) *P intersects all of the $(n-k)$-dimensional faces of the hypercube $[0,1]^n$;*

(B) $\mathrm{SA}^k(P) \neq \emptyset$*;*

(C) $\mathrm{LS}^k(P) \neq \emptyset$*;*

(D) $\mathrm{LS}_0^k(P) \neq \emptyset$*.*

In what follows, we first discuss the assumptions and a consequence of the theorem. Since the most technically involved part of the proof is the implication (A)⇒(B), we then provide a high-level description of the argument. This overview is intended to improve readability and will also be useful for establishing other claims, as well as for illustrating a general proof technique. We next prove the remaining implications. We conclude with an elementary proof of Theorem 3.1 for $k = 1$, based on a different argument.

3.7.1 Assumptions and Consequences of Theorem 3.1

We illustrate how Theorem 3.1 is tight in two distinct ways and discuss the consequence that the relaxations $\mathrm{SA}^k(P), \mathrm{LS}^k(P), \mathrm{LS}_0^k(P)$ are either all nonempty or all empty. This equivalence, however, does not imply that the relaxations define the same feasible region. The distinction is made explicit through a concrete example, followed by a discussion of applicability of Theorem 3.1 under k-transitivity.

Tightness with respect to k-Transitivity First, we claim that we cannot weaken the $(k+1)$-transitivity assumption to k-transitivity. Consider the polytope

$$P \coloneqq \text{conv}\left(\{\pi(1,1/2,0), \pi(1,1,1/10) \,:\, \pi \in C_3\}\right) \tag{3.16}$$

where C_3 is the cyclic group acting on the set $[3]$; see Fig. 3.6. This polytope is 1-transitive but not 2-transitive since $(1,1/2,0) \in P$ while $(1/2,1,0) \notin P$.

Proposition 3.36. *The polytope P (3.16) satisfies $\mathrm{LS}_0(P) \neq \emptyset$ while $\mathrm{SA}(P) = \mathrm{LS}(P) = \emptyset$.*

Proof. A description by inequalities of P computed with POLYMAKE [176] is given by

$$10x_1 + 4x_2 - 2x_3 \geq 3 \tag{3.17a}$$

$$-2x_1 + 10x_2 + 4x_3 \geq 3 \tag{3.17b}$$

$$4x_1 - 2x_2 + 10x_3 \geq 3 \tag{3.17c}$$

$$2x_1 + 2x_2 + 2x_3 \geq 3 \tag{3.17d}$$

$$-10x_1 - 10x_2 - 10x_3 \geq -21 \tag{3.17e}$$

$$x_1, x_2, x_3 \leq 1 \tag{3.17f}$$

We claim that $\mathrm{SA}(P) = \emptyset$ while $\mathrm{LS}_0(P) \neq \emptyset$. To this end, first we show that there exists a uniform point $(t,t,t) \in P$ that belongs to $P_1(P)$, where t must lie in $[1/2, 7/10]$ by inequalities (3.17d) and (3.17e). Indeed, $P_1(P)$ consists of the convex hull of the following vertices: $(1,1/2,0)$, $(1,1,1/10)$, $(1,1/10,1)$ and $(0,1,1/2)$. We want to show that there exists $t \in [1/2, 7/10]$ and $\sum_{i=1}^{4} \lambda_i = 1$ with $\lambda_i \geq 0$ for $i \in [4]$ such that

$$\begin{pmatrix} t \\ t \\ t \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1/2 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 1/10 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 1/10 \\ 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 1 \\ 1/2 \end{pmatrix}. \tag{3.18}$$

It is straightforward to verify that the solution $\bar{t} = 3/5$, $\bar{\lambda}_1 = 1/10$, $\bar{\lambda}_2 = 1/9$, $\bar{\lambda}_3 = 7/18$ and

$\bar{\lambda}_4 = 2/5$ satisfies the above constraints. Since the points that define $P_2(P)$ and $P_3(P)$ correspond to a cyclic permutation of the vertices of $P_1(P)$, and the system (3.18) is invariant under cyclic row permutations, we can conclude that $(\bar{t}, \bar{t}, \bar{t})$ belongs to $\text{LS}_0(P)$.

Finally, we show that $\text{SA}(P) = \emptyset$. Consider the following system of inequalities:

$$\begin{aligned} -x_1 - x_2 + x_{1,2} &\geq -1 \\ -x_1 - x_3 + x_{1,3} &\geq -1 \\ -x_2 - x_3 + x_{2,3} &\geq -1 \\ 11x_1 - 10x_{1,2} - 10x_{1,3} &\geq 0 \\ 3x_1 + 4x_2 - 2x_3 - 4x_{1,2} + 2x_{1,3} &\geq 3 \\ -2x_1 + 3x_2 + 4x_3 + 2x_{1,2} - 4x_{2,3} &\geq 3 \\ 4x_1 - 2x_2 + 3x_3 - 4x_{1,3} + 2x_{2,3} &\geq 3 \end{aligned}$$

These inequalities are obtained by multiplying the original constraints (3.17) by the following terms and then linearizing: inequalities (3.17f) (with bounds $x_1 \leq 1$ and $x_3 \leq 1$) by $1 - x_2$ and $1 - x_3$ for the former, and by $1 - x_2$ for the latter; inequality (3.17e) by x_1; inequality (3.17a) by $1 - x_1$; inequality (3.17b) by $1 - x_2$; and inequality (3.17c) by $1 - x_3$. A Farkas' certificate of infeasibility is given by $(26, 32, 20, 1, 9, 10, 10)$. $\square$

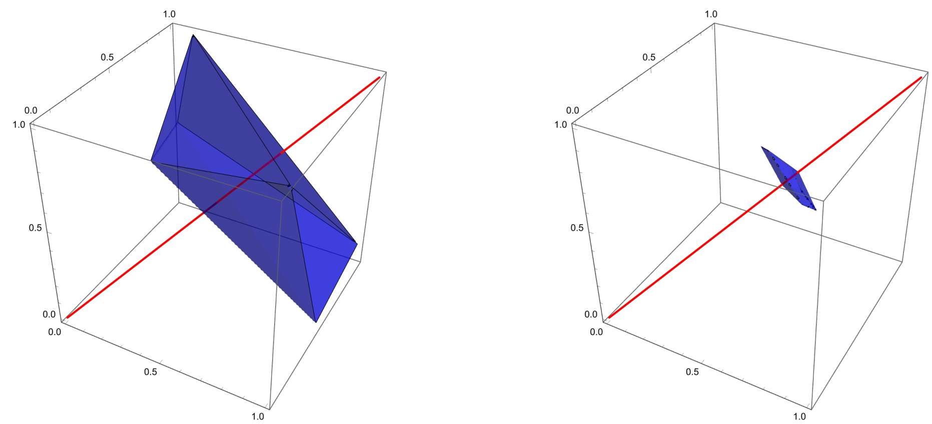

Figure 3.6: On the left, the polytope $P := \operatorname{conv}(\{\pi(1,1/2,0), \pi(1,1,1/10) : \pi \in C_3\})$ from Section 3.7.1. On the right, a nonempty subset of $\mathrm{LS}_0(P)$.

Thus, (B) does not hold, which implies that P is integer-empty, while (D) holds. Therefore, our assumption on the degree of transitivity cannot be relaxed.

Tightness with respect to Integer-emptiness Observe that we cannot drop the integer-emptyness assumption, since trivially (B), (C), (D) hold for any polytope with an integer point, irrespectively of whether (A) holds.

Integer-emptiness Equivalence We emphasize that Theorem 3.1 establishes an equivalence only with respect to integer-emptiness, and does not imply that the relaxations themselves coincide. In particular, $\mathrm{SA}^1(P)$ and $\mathrm{LS}_0^1(P)$ may differ (while both being nonempty), even when the hypotheses of the theorem hold for $k = 1$; see Example 3.6.1. We provide a proof of this claim below.

Proposition 3.37. *Let* $P := $ *conv*$(\{\pi(0,1/2,1) : \pi \in \mathcal{S}_3\})$. *Then* $\mathrm{SA}(P)$ *is nonempty and it is strictly contained in* $\mathrm{LS}_0(P)$.

Proof. Since P is a 3-transitive integer-empty polytope which intersects all of the 2-dimensional faces of the cube $[0,1]^3$, Theorem (3.1) implies that $\mathrm{SA}(P)$ and $\mathrm{LS}_0(P)$ are nonempty. Next, we exhibit a point in $\mathrm{LS}_0(P)$ which does not belong to $\mathrm{SA}(P)$. We claim that $x := (5/6, 1/3, 1/3)$ belongs to $P_2(P) = \operatorname{conv}(P^{2,1} \cup P^{2,0})$. This follows since $x = (1/3)(1/2,1,0) + (2/3)(1,0,1/2) \in$

$P^2(P)$, and $x = (1/3)(1/2, 0, 1) + (2/3)(1, 1/2, 0) \in P^3(P)$. Moreover,

$$x = \frac{1}{3}\begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 1/2 \\ 0 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 0 \\ 1/2 \\ 1 \end{pmatrix} \in P^1(P)$$

On the other hand, we have x cannot be lifted to a solution to $\mathrm{M}(P)$ since the following subsystem is infeasible

$$\begin{aligned} -x_1 + 2y_{12} + 2y_{13} &\geq 0 \\ x_2 - 2y_{12} + y_{23} &\geq 0 \\ x_3 - 2y_{13} + y_{23} &\geq 0 \\ 2x_1 + 2x_2 + 3x_3 - 2y_{13} - 2y_{23} &\geq 0 \end{aligned}$$

These inequalities are obtained by multiplying the constraints (3.9) by the following terms and then linearizing: inequalities (3.9d) by x_1; inequality (3.9c) by x_2; inequality (3.9b) by x_3; and inequality (3.9d) by $1 - x_3$. Therefore, $x \notin \mathrm{SA}(P)$. □

Applicability under k-Transitivity From a worst-case performance perspective of the LP hierarchies considered, Theorem 3.1 explains the poor behavior (i.e., $\Theta(n)$-resistance) observed in most instances studied in the literature, as these typically admit a full symmetric group acting on the variables; see Section 3.6.

A notable exception is [131], where the permutation action on the edge formulation of the Maximum Matching problem is not even 2-transitive (it is only 1-transitive), and therefore our framework does not apply directly; see Section 3.3. This suggests that additional mechanisms are responsible for the observed behavior in that setting, and motivates to interesting open questions; see Section 3.10. The approach in [131] relies on problem-specific symmetry arguments, whereas our approach is structural and general.

Furthermore, Theorem 3.27 shows that k-transitive permutation groups are highly restricted for $k \geq 6$. Consequently, the assumptions of Theorem 3.1 hold only in a limited class of highly symmetric settings. This highlights both the strength and the limitations of our result: it provides a unifying explanation for a broad class of symmetric instances, while inherently applying only to problems exhibiting sufficiently strong symmetry.

3.7.2 A High-Level Description of the Proof of (A) $\Rightarrow$ (B)

Assume that (A) holds. Let P be a polytope with $G \leq G(P)$ and G $(k+1)$-transitive. We apply Theorem 3.9 with $d = k$. Thus, we want to exhibit a point $\overline{y}$ in $\mathrm{M}^k(P)$, and show that, for every $S \subseteq [n]$ with $|S| = k$, it can be supported by the points $z^{J_0,J_1}(\overline{y}) \in \mathrm{M}^0(P)$ as in Theorem 3.9. However, our construction will proceed in reverse: we build the supporting points first and then show that they satisfy (3.3) and (3.4) for some $\overline{y}$. In particular, we perform the following steps.

1. Construction of points from $\mathrm{M}^0(P)$. By leveraging the $(k+1)$-transitivity of P, for every $S \subseteq [n]$ with $|S| = k$ and $J_0 \sqcup J_1 = S$, we first construct "highly symmetric" points, i.e., points $x^{J_0,J_1} \in P$ such that, for any $S' \subseteq [n]$ with $|S'| = k$ and $J_0' \sqcup J_1'$ with $|J_1| = |J_1'|$, we have $x^{J_0,J_1} = \pi x^{J_0',J_1'}$ for some $\pi \in G$. By definition, $(1, x^{J_0,J_1}) \in \mathrm{M}^0(P)$.

2. System of constraints defining $\overline{y}$. Note that any point $\overline{y}$ that satisfies (3.3) and (3.4) with respect to our constructed vectors $z^{J_0,J_1} := (1, x^{J_0,J_1})$ is guaranteed to belong to $\mathrm{M}^k(P)$. Imposing these conditions leads to a system of equations whose variables are the coordinates of $\overline{y}$. This system appears to be highly complex and non-tractable. However, because of the symmetric structure of z^{J_0,J_1}, we look for solutions $\overline{y}$ that are highly symmetric. Indeed, for every $I, J \subseteq [n]$ with $|I| = |J| \leq k+1$ and every $J_0 \sqcup J_1 = S$, $J_0' \sqcup J_1' = S'$ with $|J_1| = |J_1'|$ and $|S| = |S'| = k$, we impose the following conditions:

(C-1) $\overline{y}_I = \overline{y}_J =: \overline{\gamma}_{|I|}$;

(C-2) $\overline{y}_{\emptyset}^{J_0,J_1} = \overline{y}_{\emptyset}^{J_0',J_1'} =: \overline{\lambda}_{|J_1|}$.

Conditions (C-1) and (C-2) above lead to a (much smaller) system in the variables γ and λ, which have $k+1$ components each. Thus, we are left to show that this system has a solution. This is the most technical part of the proof, divided into two steps.

3. Solution to the system defining $\overline{\gamma}$. We first consider a subsystem of linear equalities in the variables γ only, and show that the corresponding constraint matrix is invertible.

4. Solution to the general system. Based on the previous step, we give an explicit formula for $\overline{\gamma}$ and $\overline{\lambda}$, and proceed to show that they satisfy all constraints of the system in γ and λ defined in Step 2.

3.7.3 Proof of (A) $\Rightarrow$ (B)

We now present full details of the proof (A) $\Rightarrow$ (B) by following the steps outlined in the previous section.

1. Construction of points from $\mathrm{M}^0(P)$**.** For $\ell \in \{0, \dots, k\}$, define

$$P_\ell := P \cap \{x \in [0,1]^n \,:\, x_i = 1 \text{ for all } i \in [\ell], x_j = 0 \text{ for all } j \in [k] \setminus [\ell]\}, \tag{3.20}$$

and observe that it is nonempty because of Assumption (A). Thus, let $x^\ell \in P_\ell$. Recall that $G_{[k]}$ is the pointwise stabilizer of $[k]$. Observe that P_ℓ is $G_{[k]}$-invariant. Moreover, since G is $(k+1)$-transitive, by Proposition 3.12, $G_{[k]}$ is transitive on $[n] \setminus [k]$. Therefore, by applying Proposition 3.11 to x^ℓ with $G_{[k]}$ and P_ℓ we obtain

$$\overline{x}^\ell := \beta_{G_{[k]}}(x^\ell) = (\underbrace{1, \dots, 1}_{\ell}, \underbrace{0, \dots, 0}_{k-\ell}, \underbrace{\Delta_\ell, \dots, \Delta_\ell}_{n-k}) \in P_\ell, \tag{3.21}$$

for some $\Delta_\ell \in [0,1]$. Define $\Delta := (\Delta_0, \dots, \Delta_k) \in [0,1]^{k+1}$ and observe that $\Delta \in (0,1)^{k+1}$ since P is integer-empty.

Next, let $S \subseteq [n]$ with $|S| = k$ and consider an arbitrary 2-partition $J_0 \sqcup J_1 = S$ such that $|J_1| = \ell$. By definition, the point $\overline{x}^\ell$ has ℓ ones in entries indexed by $[\ell]$ and zeros in entries indexed by $[k] \setminus [\ell]$. Since G is $(k+1)$-transitive and $G \leq G(P)$, there exists some permutation $\pi \in G$ and some point $\overline{x}^{J_0,J_1} \in P$ such that π maps: all of the ones in positions $[\ell]$ of $\overline{x}^\ell$ to entries indexed by

J_1 of $\overline{x}^{J_0,J_1}$; all of the zeros in positions $[k] \setminus [\ell]$ to entries indexed by J_0; and the fractional Δ_ℓ's in positions $[n] \setminus [k]$ to entries indexed by $[n] \setminus S$. In other words,

$$\overline{x}_i^{J_0,J_1} := (\pi \overline{x}^\ell)_i = \begin{cases} 0 & \text{if } i \in J_0, \\ 1 & \text{if } i \in J_1, \\ \Delta_\ell & \text{if } i \in [n] \setminus S. \end{cases} \tag{3.22}$$

By definition of $\mathrm{M}^0(P)$, we immediately deduce the following.

Lemma 3.38. *For each $S \subseteq [n]$ with $|S| = k$ and 2-partition $J_0 \sqcup J_1 = S$, we have $(1, \overline{x}^{J_0,J_1}) \in \mathrm{M}^0(P)$.*

2. System of constraints defining $\overline{y}$. We next give sufficient conditions for the existence of $\overline{y} \in \mathrm{M}^d(P)$.

For $t \in \mathbb{N}$ and $\Theta = (\Theta_0, \dots, \Theta_t) \in \mathbb{R}^{t+1}$, we let $A_\Theta^t \in \mathbb{R}^{(t+1)\times(t+1)}$ with rows indices in $\{0, \dots, t\}$ and column indices in $\{1, \dots, t+1\}$ be defined as follows:

$$(A_\Theta^t)_{\ell,r} := \begin{cases} (-1)^{\ell+r-1} \left[\binom{t-\ell}{r-\ell-1} + \binom{t-\ell}{r-\ell}\Theta_\ell\right] & \text{if } \ell < r, \\ -\Theta_\ell & \text{if } \ell = r, \\ 0 & \text{if } \ell > r. \end{cases} \tag{3.23}$$

See Example 3.42 for an illustration of the matrix A_Θ^t.

Lemma 3.39. *Consider the following system in the variables $\gamma_0, \ldots, \gamma_{k+1} \in \mathbb{R}$, $\lambda_0, \ldots, \lambda_k \in \mathbb{R}$*

$$A_\Delta^k \gamma = \Delta_0(1, 0, \ldots, 0), \tag{3.24a}$$

$$1 = \sum_{\ell=0}^{k} \binom{k}{\ell} \lambda_\ell, \tag{3.24b}$$

$$\gamma_1 = \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \lambda_\ell, \tag{3.24c}$$

$$\gamma_1 = \sum_{\ell=0}^{k} \binom{k}{\ell} \lambda_\ell \Delta_\ell, \tag{3.24d}$$

$$\lambda_\ell = \sum_{r=0}^{k-\ell} (-1)^r \binom{k-\ell}{r} \gamma_{r+\ell} \quad \forall \ell \in \{0, 1, \ldots, k\}, \tag{3.24e}$$

$$\mathbf{0} \le \lambda, \gamma \le \mathbf{1}, \quad \gamma_0 = 1, \tag{3.24f}$$

where we abbreviate $\lambda \coloneqq (\lambda_0, \ldots, \lambda_k)$ and $\gamma \coloneqq (\gamma_1, \ldots, \gamma_{k+1})$. Suppose the system admits a solution $(\overline{\gamma}_0, \overline{\gamma}, \overline{\lambda})$ strictly positive in all components.

Then, $\overline{y} \in \mathrm{M}^k(P)$, where

$$\overline{y}_I = \overline{\gamma}_{|I|} \quad \textit{for all } I \subseteq [n], \textit{ with } |I| \le k+1. \tag{3.25}$$

The following lemma is an auxiliary result for the proof of Lemma 3.39. We then proceed with that proof.

Lemma 3.40. *Let $d \in \mathbb{Z}_{\ge 0}$ and $y \in [0,1]^{\binom{n}{\le d}}$. For every $I \subseteq [n]$ with $|I| \le 1$, $S \subseteq [n]$ of cardinality d, every 2-partition $J_0 \sqcup J_1 = S$, and every $j \in J_0$, we have*

$$\sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup \{j\} \cup H} = 0.$$

Proof. We have

$$\begin{aligned}
&\sum_{H\subseteq J_0}(-1)^{|H|}y_{J_1\cup\{j\}\cup H}\\
&=\sum_{H\subseteq J_0: j\in H}(-1)^{|H|}y_{J_1\cup\{j\}\cup H}+\sum_{H\subseteq J_0: j\notin H}(-1)^{|H|}y_{J_1\cup\{j\}\cup H}\\
&=\sum_{H\subseteq J_0: j\notin H}(-1)^{|H|+1}y_{J_1\cup\{j\}\cup H}+\sum_{H\subseteq J_0: j\notin H}(-1)^{|H|}y_{J_1\cup\{j\}\cup H}=0. \qquad \square
\end{aligned}$$

Proof of Lemma 3.39. Let $(\overline{\gamma}_0,\overline{\gamma},\overline{\lambda})$ be a solution to the system (3.24a)–(3.24f) that is strictly positive in all components and set $\overline{y}_I=\overline{\gamma}_{|I|}$ for all $I\subseteq[n]$, $|I|\le k+1$. Observe that $\overline{y}\in[0,1]^{\binom{n}{\le k+1}}$.

For a 2-partition $J_0\sqcup J_1$ of S with $|S|=k$, let $\overline{y}^{J_0,J_1}$ and $z^{J_0,J_1}(\overline{y})$ be as in (3.3). To show that $\overline{y}\in \mathrm{M}^k(P)$, it suffices to verify that the conditions from Theorem 3.9 hold. In other words, (i) $z^{J_0,J_1}(\overline{y})\in \mathrm{M}^0(P)$ for every 2-partition $J_0\sqcup J_1=S$, and (ii) $\overline{y}_I=\sum_{J_0\sqcup J_1=S}\overline{y}_{\emptyset}^{J_0,J_1}z_I^{J_0,J_1}(\overline{y})$ for every $I\subseteq[n]$ with $|I|\le 1$.

Observe that, for $|J_1|=\ell$, we have

$$\overline{y}_{\emptyset}^{J_0,J_1}=\sum_{H\subseteq J_0}(-1)^{|H|}\overline{y}_{J_1\cup H}=\sum_{r=0}^{k-\ell}(-1)^r\binom{k-\ell}{r}\overline{\gamma}_{r+\ell}=\overline{\lambda}_\ell>0, \tag{3.26}$$

where the first equality holds by definition of $y_{\emptyset}^{J_0,J_1}$, the second by definition of $\overline{y}$, the third by (3.24e), and the inequality by the fact that $\overline{\lambda}_\ell>0$ by hypothesis. Thus, $z_I^{J_0,J_1}(\overline{y})=\overline{y}_I^{J_0,J_1}/\overline{y}_{\emptyset}^{J_0,J_1}$ for all $I\subseteq[n]$, $|I|\le 1$.

We first show that condition (i) holds for 2-partitions of the form $J_0=[k]\setminus[\ell]$, $J_1=[\ell]$ for $\ell\in\{0,\dots,k\}$. We abbreviate $z^\ell(\overline{y}):=z^{[k]\setminus[\ell],[\ell]}(\overline{y})$ and $\overline{y}^\ell:=\overline{y}^{[k]\setminus[\ell],[\ell]}$.

We show (i) by verifying that $z^\ell(\overline{y})=(1,\overline{x}^\ell)$ for every $\ell\in\{0,\dots,k\}$. The statement follows by Lemma 3.38. Let $I\subseteq[n]$ with $|I|\le 1$. For $I=\emptyset$, then $z_I^\ell(\overline{y})=1$ by definition. Thus, let $I=\{j\}$. To compute $z_j^\ell(\overline{y})$, observe that for $j\in[\ell]$ we can write

$$\overline{y}_j^\ell=\sum_{H\subseteq[k]\setminus[\ell]}(-1)^{|H|}\overline{y}_{[\ell]\cup\{j\}\cup H}=\sum_{H\subseteq[k]\setminus[\ell]}(-1)^{|H|}\overline{y}_{[\ell]\cup H}=\overline{y}_{\emptyset}^\ell,$$

where the last equality follows by definition of $\overline{y}_\emptyset^\ell$ in (3.3). Thus, $z_j^\ell(\overline{y}) = \overline{y}_\emptyset^\ell / \overline{y}_\emptyset^\ell = 1 = \overline{x}_j^\ell$, as required. Suppose now $j \in [k] \setminus [\ell]$. Then $j \in J_0$. Since we know that $\overline{y} \in [0,1]^{\binom{n}{\leq k+1}}$, we can apply Lemma 3.40 and obtain

$$\overline{y}_j^{J_0,J_1} = \sum_{H \subseteq J_0} (-1)^{|H|} \overline{y}_{J_1 \cup \{j\} \cup H} = 0.$$

Hence, we conclude that $z_j^\ell(\overline{y}) = 0 = \overline{x}_j^\ell$, as required.

Last, let $j \notin [k]$. We want to show that $\overline{z}_j^\ell = \overline{x}_j^\ell$. Notice that $\overline{z}_j^\ell = \overline{y}_j^\ell / \overline{y}_\emptyset^\ell = \overline{y}_j^\ell / \overline{\lambda}_\ell$, where the last equation follows by (3.26). In turn, $\overline{x}_j^\ell = \Delta_\ell$ by definition. Since

$$\overline{y}_j^\ell = \sum_{H \subseteq [k] \setminus [\ell]} (-1)^{|H|} \overline{y}_{[\ell] \cup \{j\} \cup H} = \sum_{r=0}^{k-\ell} (-1)^r \binom{k-\ell}{r} \overline{\gamma}_{r+\ell+1},$$

we have that $\frac{\overline{y}_j^\ell}{\overline{\lambda}_\ell} = \overline{x}_j^\ell$ is equivalent to

$$\begin{aligned}
& \sum_{r=0}^{k-\ell} \binom{k-\ell}{r} (-1)^r \overline{\gamma}_{\ell+1+r} = \Delta_\ell \sum_{r=0}^{k-\ell} \binom{k-\ell}{r} (-1)^r \overline{\gamma}_{\ell+r} \\
\iff & \sum_{r=0}^{k-\ell} \binom{k-\ell}{r} (-1)^r \overline{\gamma}_{\ell+1+r} + \sum_{r=1}^{k-\ell} \binom{k-\ell}{r} \Delta_\ell (-1)^{r+1} \overline{\gamma}_{\ell+r} = \Delta_\ell \overline{\gamma}_\ell \\
\iff & \sum_{r=0}^{k-\ell} (-1)^r \left[\binom{k-\ell}{r} + \binom{k-\ell}{r+1} \Delta_\ell \right] \overline{\gamma}_{\ell+1+r} = \Delta_\ell \overline{\gamma}_\ell \\
\iff & -\Delta_\ell \overline{\gamma}_\ell + \sum_{r=\ell+1}^{k+1} (-1)^{\ell+r-1} \left[\binom{k-\ell}{r-\ell-1} + \binom{k-\ell}{r-\ell} \Delta_\ell \right] \overline{\gamma}_r = 0
\end{aligned} \tag{3.27}$$

where the penultimate equivalence follows by a change of index variables $r' = r+1$ in the second sum at the LHS. Equation (3.27) is then implied by (3.24a), (3.24e), (3.23) and $\overline{\gamma}_0 = 1$. This concludes the proof that $\overline{z}^\ell = (1, \overline{x}^\ell)$. Since P is G-invariant, we deduce that $\overline{z}^{J_0,J_1} = (1, x^{J_0,J_1})$ for all $S \subseteq [n]$, $|S| = k$, and $J_0 \sqcup J_1 = S$. Hence, (i) holds.

To show (ii), recall the we observed in (3.26) that $\overline{y}_\emptyset^{J_0,J_1} = \overline{\lambda}_{|J_1|}$ for every 2-partition $J_0 \sqcup J_1$ of $S \subseteq [n]$ with $|S| = k$. Let $I = \{j\} \in [n]$. Recall that we also just observed that $z_j^{J_0,J_1}(\overline{y}) = \overline{x}_j^{J_0,J_1}$.

If $j \notin S$, then

$$\begin{aligned}
\sum_{J_0 \sqcup J_1 = S} \overline{y}_\emptyset^{J_0,J_1} z_j^{J_0,J_1}(\overline{y}) &= \sum_{J_0 \sqcup J_1 = S} \overline{y}_\emptyset^{J_0,J_1} \overline{x}_j^{J_0,J_1} \\
&= \sum_{J_0 \sqcup J_1 = S} \overline{\lambda}_{|J_1|} \Delta_{|J_1|} && \text{by (3.21) and (3.26)} \\
&= \sum_{\ell=0}^{k} \binom{k}{\ell} \overline{\lambda}_\ell \Delta_\ell \\
&= \overline{\gamma}_1 = \overline{y}_j && \text{by (3.24d) and (3.25).}
\end{aligned}$$

If conversely $j \in S$, we have

$$\begin{aligned}
\sum_{J_0 \sqcup J_1 = S} \overline{y}_\emptyset^{J_0,J_1} z_j^{J_0,J_1}(\overline{y}) &= \sum_{J_0 \sqcup J_1 = S} \overline{y}_\emptyset^{J_0,J_1} \overline{x}_j^{J_0,J_1} \\
&= \sum_{J_0 \sqcup J_1 = S : j \in J_0} \overline{y}_\emptyset^{J_0,J_1} \overline{x}_j^{J_0,J_1} + \sum_{J_0 \sqcup J_1 = S : j \in J_1} \overline{y}_\emptyset^{J_0,J_1} \overline{x}_j^{J_0,J_1} \\
&= \sum_{J_0 \sqcup J_1 = S : j \in J_1} \overline{\lambda}_{|J_1|} && \text{by (3.21)} \\
&= \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \overline{\lambda}_\ell \\
&= \overline{\gamma}_1 = \overline{y}_1 && \text{by (3.24c) and (3.25).}
\end{aligned}$$

To conclude that (ii) holds, let $I = \emptyset$ and observe

$$\begin{aligned}
\sum_{J_0 \sqcup J_1 = S} y_\emptyset^{J_0,J_1} z_\emptyset^{J_0,J_1}(y) &= \sum_{J_0 \sqcup J_1 = S} \overline{y}_\emptyset^{J_0,J_1}, \\
&= \sum_{J_0 \sqcup J_1 = S} \overline{\lambda}_{|J_1|} && \text{by (3.26)} \\
&= \sum_{\ell=0}^{k} \binom{k}{\ell} \overline{\lambda}_\ell \\
&= 1 && \text{by (3.24b).}
\end{aligned}$$

$\square$

3. Solution to the system defining $\overline{\gamma}$. We are left to show the existence of a solution to the system from Lemma 3.39. To achieve that, we give in Lemma 3.44 an explicit formula for the determinant of A^t_Θ. Recall that this is the constraint matrix of system (3.24a) that uniquely defines $\overline{\gamma}$. To prove this formula, we start with some intermediate steps. For a vector $\Theta = (\Theta_0, \dots, \Theta_t) \in (0,1)^{t+1}$, we sometimes make the dependency on the components of Θ more explicit and write $A^t_{\Theta_0,\dots,\Theta_t} = A^t_\Theta$.

Lemma 3.41. *Let $t \in \mathbb{N}$ and $\Theta = (\Theta_0, \dots, \Theta_t) \in (0,1)^{t+1}$. The determinant of A^t_Θ can be decomposed as*

$$\det A^t_\Theta = \Theta_0 \det A^{t-1}_{\Theta_1,\dots,\Theta_t} + (1-\Theta_t) \det A^{t-1}_{\Theta_0,\dots,\Theta_{t-1}}. \tag{3.28}$$

The idea of the proof of Lemma 3.41 is to transform A^t_Θ via elementary column operations into another matrix B^t_Θ for which we deduce the relation (3.28). As $\det A^{t'}_{\Theta'} = \det B^{t'}_{\Theta'}$ for all t', Θ', the relation (3.28) also holds for A^t_Θ. For $t \in \mathbb{N}$ and $\Theta = (\Theta_0, \dots, \Theta_t) \in \mathbb{R}^t$, we define $B^t_{\Theta_0,\dots,\Theta_t} \in \mathbb{R}^{(t+1)\times(t+1)}$ with rows indexed in $\{0, \dots, t\}$ and columns indexed in $\{1, \dots, t+1\}$ as follows: its r-th column is given by

$$(B^t_\Theta)_{\cdot,r} := \sum_{j=r}^{t} (A^t_\Theta)_{\cdot,j}$$

Example 3.42. We next illustrate matrices A^t_Θ and B^t_Θ with some examples. We can write $A^1_{\Theta_0,\Theta_1}$ and transform it into $B^1_{\Theta_0,\Theta_1}$ as follows:

$$A^1_{\Theta_0,\Theta_1} = \begin{pmatrix} 1+\Theta_0 & -1 \\ -\Theta_1 & 1 \end{pmatrix} \underset{\text{adding 2 to 1}}{\rightsquigarrow} \begin{pmatrix} \Theta_0 & -1 \\ 1-\Theta_1 & 1 \end{pmatrix} = B^1_{\Theta_0,\Theta_1}.$$

Similarly, letting $\Theta = (\Theta_0, \Theta_1, \Theta_2)$, we have,

$$A_\Theta^2 = \begin{pmatrix} 1+2\Theta_0 & -2-\Theta_0 & 1 \\ -\Theta_1 & 1+\Theta_1 & -1 \\ 0 & -\Theta_2 & 1 \end{pmatrix} \underset{\text{adding 2,3 to 1}}{\rightsquigarrow} \begin{pmatrix} \Theta_0 & -2-\Theta_0 & 1 \\ 0 & 1+\Theta_1 & -1 \\ 1-\Theta_2 & -\Theta_2 & 1 \end{pmatrix}$$

$$\underset{\text{adding 3 to 2}}{\rightsquigarrow} \begin{pmatrix} \Theta_0 & -1-\Theta_0 & 1 \\ 0 & \Theta_1 & -1 \\ 1-\Theta_2 & 1-\Theta_2 & 1 \end{pmatrix} = B_\Theta^2.$$

Claim 3.43. For $\ell \in \{0, \dots, t\}$ and $r \in \{1, \dots, t+1\}$, entry (ℓ, r) of B_Δ^t is given by

$$(B_\Delta^t)_{\ell,r} = \begin{cases} (-1)^{\ell+r-1}\left[\binom{t-\ell-1}{r-\ell-2} + \binom{t-\ell-1}{r-\ell-1}\Theta_\ell\right] & \text{if } \ell < r \le t, \\ 1-\Theta_t & \text{if } \ell = t \text{ and } r \le t, \\ (-1)^{\ell+t} & \text{if } \ell \le t \text{ and } r = t+1, \\ 0 & \text{o.w. (if } \ell \ge r,\ \ell < t) \end{cases} \tag{3.29}$$

Proof. We will use the following binomial identities (see e.g. [177]):

$$\sum_{j=0}^{t} (-1)^j \binom{t}{j} = 0, \tag{3.30}$$

$$\sum_{j=r}^{t} (-1)^j \binom{t}{j} = (-1)^r \binom{t-1}{r-1}. \tag{3.31}$$

Row $\ell = t$ and column $r = t+1$. By definition $(B_\Theta^t)_{\cdot,t+1} = (A_\Theta^t)_{\cdot,t+1}$, and this agrees with formula (3.29) since $(A_\Theta^t)_{\ell,t+1} = (-1)^{\ell+t}$ for $\ell \in \{0, \dots, t\}$. Next, we check the entries of row $\ell = t$ of B_Θ^t for columns $r \in \{0, \dots, t\}$. Since the only nonzeros in row $\ell = t$ of A_Θ^t are $-\Theta_t$ and $(-1)^{t+t}$, at entries (t,t) and $(t, t+1)$, respectively, it holds that $(B_\Theta^t)_{t,r} = 1 - \Theta_t$ for $r \in \{1, \dots, t\}$.

Lower triangular part of B_Θ^t. We check $(B_\Theta^t)_{\ell,r} = 0$ for $1 \leq r \leq \ell < t$. Indeed,

$$\begin{aligned}
(B_\Theta^t)_{\ell,r} &= -\Theta_\ell + \sum_{j=\ell+1}^{t+1} (-1)^{\ell+j-1}\left[\binom{t-\ell}{j-\ell-1} + \binom{t-\ell}{j-\ell}\Theta_\ell\right] \\
&= -\Theta_\ell + \sum_{j=\ell+1}^{t+1} (-1)^{\ell+j-1}\binom{t-\ell}{j-\ell-1} + \sum_{j=\ell+1}^{t} (-1)^{\ell+j-1}\binom{t-\ell}{j-\ell}\Theta_\ell \\
&= -\Theta_\ell + \sum_{j=0}^{t-\ell}(-1)^{2\ell+j}\binom{t-\ell}{j} + \sum_{j=1}^{t-\ell}(-1)^{2\ell+j-1}\binom{t-\ell}{j}\Theta_\ell \\
&= -\Theta_\ell + (-1)^{2\ell}\sum_{j=0}^{t-\ell}(-1)^{j}\binom{t-\ell}{j} + (-1)^{2\ell-1}\sum_{j=1}^{t-\ell}(-1)^{j}\binom{t-\ell}{j}\Theta_\ell \\
&= -\Theta_\ell + (-1)\left(-\Theta_\ell + \sum_{j=0}^{t-\ell}(-1)^{j}\binom{t-\ell}{j}\Theta_\ell\right) = 0,
\end{aligned}$$

where in the last two equalities we used the identity (3.30).

Diagonal and upper triangular part of B_Θ^t. Let $0 \leq \ell < r \leq t$. Using (3.31), we have:

$$\begin{aligned}
(B_\Theta^t)_{\ell,r} &= \sum_{j=r}^{t+1}(-1)^{\ell+j-1}\left[\binom{t-\ell}{j-\ell-1} + \binom{t-\ell}{j-\ell}\Theta_\ell\right] \\
&= (-1)^{2\ell}\sum_{j=r-\ell-1}^{t-\ell}(-1)^{j}\binom{t-\ell}{j} + (-1)^{2\ell-1}\sum_{j=r-\ell}^{t-\ell}(-1)^{j}\binom{t-\ell}{j}\Theta_\ell \\
&= (-1)^{2\ell}(-1)^{r-\ell-1}\binom{t-\ell-1}{r-\ell-2} + (-1)^{2\ell-1}(-1)^{r-\ell}\binom{t-\ell-1}{r-\ell-1}\Theta_\ell \\
&= (-1)^{\ell+r-1}\left[\binom{t-\ell-1}{r-\ell-2} + \binom{t-\ell-1}{r-\ell-1}\Theta_\ell\right].
\end{aligned}$$

$\square$

Notice, in particular, that $(B_\Theta^t)_{\ell,\ell+1} = (-1)^{2\ell}\left[\binom{t-\ell-1}{-1} + \binom{t-\ell-1}{0}\Theta_\ell\right] = \Theta_\ell$.

Proof of Lemma 3.41. It suffices to prove (3.28) for B_Θ^t. To this end, we expand the determinant along the first column of B_Θ^t. This column has exactly two nonzeros, Θ_0 and $1-\Theta_t$ in row 0 and t

respectively, so we obtain

$$\det B^t_\Theta = \Theta_0 \det B^{t-1}_{\Theta_1,\dots,\Theta_t} + (-1)^t (1-\Theta_t) \det C,$$

where C is the matrix obtained by removing the first column and the last row of B^t_Θ. We claim that $\det C = (-1)^t \det A^{t-1}_{\Theta_0,\dots,\Theta_{t-1}}$. Indeed, observe that C is the $t \times t$ matrix obtained by removing the first column and the last row of B^t_Θ, and every nonzero entry in this submatrix has the opposite sign of the corresponding entries in $A^{t-1}_{\Theta_0,\dots,\Theta_{t-1}}$. Since $\det A^{t-1}_{\Theta_0,\dots,\Theta_{t-1}} = \det B^{t-1}_{\Theta_0,\dots,\Theta_{t-1}}$, by the multilinearity of the determinant, we can conclude that

$$\det B^t_\Theta = \Theta_0 \det B^{t-1}_{\Theta_1,\dots,\Theta_t} + (1-\Theta_t) \det B^{t-1}_{\Theta_0,\dots,\Theta_{t-1}}.$$

□

For $t \in \mathbb{N}$ and a vector $\Theta = (\Theta_0, \dots, \Theta_t) \in (0,1)^{t+1}$, we define

$$\omega^t_\Theta \coloneqq \sum_{r=0}^{t} \binom{t}{r} \prod_{j=0}^{r-1} \Theta_j \prod_{j=r+1}^{t} (1-\Theta_j), \tag{3.32}$$

and observe that ω^t_Θ is strictly positive. Equivalently, we write $\omega^t_\Theta = \omega^t_{\Theta_0,\dots,\Theta_t}$.

Lemma 3.44. *Let $t \in \mathbb{N}$ and $\Theta = (\Theta_0, \dots, \Theta_t) \in (0,1)^{t+1}$. The determinant of A^t_Θ is equal to ω^t_Θ.*

Proof. We use the decomposition (3.28) and show the claim by induction on t. The case $t = 1$ easily follows since

$$\omega^1_{\Theta_0,\Theta_1} = (1-\Theta_1) + \Theta_0 = \det A^1_{\Theta_0,\Theta_1}.$$

Next, suppose that $\omega^{t-1}_{\Theta'} = \det A^{t-1}_{\Theta'}$ for $\Theta' \in (0,1)^t$ arbitrary. Using (3.28) and our induction hypothesis, we have

$$
\begin{aligned}
\det A_\Theta^t &= \Theta_0 \det A^{t-1}_{\Theta_1,\dots,\Theta_t} + (1-\Theta_t)\det A^{t-1}_{\Theta_0,\dots,\Theta_{t-1}} \\
&= \Theta_0 \omega^{t-1}_{\Theta_1,\dots,\Theta_t} + (1-\Theta_t)\omega^{t-1}_{\Theta_0,\dots,\Theta_{t-1}} \\
&= \Theta_0 \omega^{t-1}_{\Theta_1,\dots,\Theta_t} + (1-\Theta_t)\left(\prod_{j=1}^{t-1}(1-\Theta_j) + \sum_{r=1}^{t-1}\binom{t-1}{r}\prod_{j=0}^{r-1}\Theta_j \prod_{j=r+1}^{t-1}(1-\Theta_j)\right) \\
&= \Theta_0 \omega^{t-1}_{\Theta_1,\dots,\Theta_t} + \prod_{j=1}^{t}(1-\Theta_j) + \Theta_0\sum_{r=1}^{t-1}\binom{t-1}{r}\prod_{j=1}^{r-1}\Theta_j \prod_{j=r+1}^{t}(1-\Theta_j) \\
&= \prod_{j=1}^{t}(1-\Theta_j) \\
&\quad + \Theta_0\left(\sum_{r=1}^{t-1}\left[\binom{t-1}{r-1}+\binom{t-1}{r}\right]\prod_{j=1}^{r-1}\Theta_j \prod_{j=r+1}^{t}(1-\Theta_j) + \binom{t-1}{t-1}\prod_{j=1}^{t-1}\Theta_j\right) \\
(*) &= \prod_{j=1}^{t}(1-\Theta_j) + \Theta_0\left(\sum_{r=1}^{t}\binom{t}{r}\prod_{j=1}^{r-1}\Theta_j \prod_{j=r+1}^{t}(1-\Theta_j)\right) \\
&= \sum_{r=0}^{t}\binom{t}{r}\prod_{j=0}^{r-1}\Theta_j \prod_{j=r+1}^{t}(1-\Theta_j) \\
&= \omega_\Theta^t,
\end{aligned}
$$

where in $(*)$ we used $\binom{t-1}{r-1}+\binom{t-1}{r}=\binom{t}{r}$. $\square$

4. Solution to the general system. We next define a vector $(\overline{\gamma}_0, \overline{\gamma}, \overline{\lambda}) \in \mathbb{R}^{2k+3}$ and show that it satisfies all constraints of Lemma 3.39.

For $\ell \in \{0,\dots,k\}$, define

$$\overline{\lambda}_\ell \coloneqq \frac{1}{\omega_\Delta^k}\prod_{i=0}^{\ell-1}\Delta_i \prod_{i=\ell+1}^{k}(1-\Delta_i). \tag{3.33}$$

For $i \in [k+1]$, we define

$$\rho_\Delta^i \coloneqq \begin{cases} \Delta_{i-1}\left(\frac{\omega^{k-i}_{\Delta_i,\ldots,\Delta_k}}{\omega^{k-(i-1)}_{\Delta_{i-1},\ldots,\Delta_k}}\right) & \text{if } i \leq k, \\ \Delta_k & \text{if } i = k+1. \end{cases} \tag{3.34}$$

These scalars are well-defined since $\omega^{k-i}_{\Delta_i,\ldots,\Delta_k} > 0$. Then define $\overline{\gamma} \in \mathbb{R}^{k+1}$ as

$$\overline{\gamma}_i \coloneqq \rho_\Delta^i \overline{\gamma}_{i-1} \tag{3.35}$$

for $i \in [k+1]$ and $\overline{\gamma}_0 \coloneqq 1$.

Constraints (3.24f)**.** $\overline{\lambda} \in (0,1)^{k+1}$ immediately from $\Delta \in (0,1)^{k+1}$ and the definition of $\overline{\lambda}$. $\overline{\gamma}_0 = 1$ follows by definition. Last, we prove that $\overline{\gamma} \in (0,1]^{k+1}$. For $i \in \{0,\ldots,k\}$ we can rewrite $\overline{\gamma}_i$ as

$$\overline{\gamma}_i = \left(\prod_{j=0}^{i-1} \Delta_j\right)\left(\frac{\omega^{k-i}_{\Delta_i,\ldots,\Delta_k}}{\omega^{k}_{\Delta_0,\ldots,\Delta_k}}\right). \tag{3.36}$$

Using (3.32), we write

$$
\begin{aligned}
\omega^{k-(i-1)}_{\Delta_{i-1},\ldots,\Delta_k} &= \sum_{r=0}^{k-(i-1)} \binom{k-(i-1)}{r} \prod_{j=0}^{r-1} \Delta_{j+i-1} \prod_{j=r+1}^{k} (1-\Delta_{j+i-1}) \\
&= \sum_{r=i-1}^{k} \binom{k-(i-1)}{r-(i-1)} \prod_{j=i-1}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j) \\
&= \prod_{j=i}^{k}(1-\Delta_j) + \sum_{r=i}^{k} \binom{k-(i-1)}{r-(i-1)} \prod_{j=i-1}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j) \\
&= \prod_{j=i}^{k}(1-\Delta_j) + \sum_{r=i}^{k} \left(\binom{k-i}{r-i} + \binom{k-i}{r-(i-1)} \right) \prod_{j=i-1}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j) \\
&= \sum_{r=i}^{k} \binom{k-i}{r-i} \prod_{j=i-1}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j) + E(\Delta) \\
&= \Delta_{i-1} \sum_{r=i}^{k} \binom{k-i}{r-i} \prod_{j=i}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j) + E(\Delta) \\
&= \Delta_{i-1}\omega^{k-i}_{\Delta_i,\ldots,\Delta_k} + E(\Delta),
\end{aligned}
$$

where $E(\Delta)$ is strictly positive. Thus, $0 < \Delta_{i-1}\omega^{k-i}_{\Delta_i,\ldots,\Delta_k} < \omega^{k-(i-1)}_{\Delta_{i-1},\ldots,\Delta_k}$ and therefore $\rho^i_\Delta \in (0,1)$. Consequently, in view of (3.35), and $\overline{\gamma}_0 = 1$, we deduce that $\overline{\gamma} \in (0,1)^{k+1}$.

Constraints (3.24a)**.** We first show the following.

Lemma 3.45. *For $i \in \{0, \ldots, k\}$, the determinant of $A^{k-i}_{\Delta_i,\ldots,\Delta_k}$ can be written as*

$$\sum_{r=i}^{k} (-1)^{r+i} \left[\binom{k-i}{r-i} + \binom{k-i}{r-i+1} \Delta_i \right] \prod_{j=i}^{r-1} \Delta_{j+1} \det A^{k-(r+1)}_{\Delta_{r+1},\ldots,\Delta_k}. \tag{3.37}$$

where $A^{k-(r+1)}_{\Delta_{r+1},\ldots,\Delta_k} := 1$ if $r = k$.

Proof. Since $A^{k-i}_{\Delta_i,\ldots,\Delta_k}$ is the $(k+1-i) \times (k+1-i)$ principal submatrix of A^k_Δ indexed by rows $\{i, \ldots, k\}$ and columns $\{i+1, \ldots, k+1\}$, it suffices, without loss of generality, to consider the case $i = 0$. Let $r \in \{1, \ldots, k\}$ and let $\tilde{A}^{k-r}$ be the $(k+1-r) \times (k+1-r)$ submatrix of A^k_Δ obtained by deleting rows 1 through r and columns 1 through r (recall that $\{1, \ldots, k+1\}$ is the set of column indices of A^k_Δ and $\{0, \ldots, k\}$ is the set of row indices of A^k_Δ). We claim that the matrix $\tilde{A}^{k-r}$ has only two nonzero entries in its first column (as does A^k_Δ). Indeed, first, observe that A^k_Δ has the nonzero subdiagonal $(-\Delta_1, \ldots, -\Delta_k)$ lying below its main diagonal. Moroeover, after deleting columns 1 through r, the first column of the resulting submatrix contains $r+2$ consecutive nonzeros, followed by zeros. Removing rows 1 through r then leaves exactly two nonzero entries in the first column of $\tilde{A}^{k-r}$ (see Figure 3.7).

$$\left(\begin{array}{cc|ccc} \times & \times & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \hline \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ \hline 0 & 0 & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ 0 & 0 & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \end{array} \right)$$

Figure 3.7: Schematic representation of A^4_Δ, with the entries in bold corresponding to the surviving $\tilde{A}^2$ submatrix after deleting rows 1-2 and columns 1-2.

Therefore, the Laplace expansion along the first column of A_Δ^k and $\tilde{A}^{k-r}$ yields:

$$\det A_\Delta^k = \left[\binom{k}{0} + \binom{k}{1}\Delta_0\right] \det A_{\Delta_1,\ldots,\Delta_k}^{k-1} + \Delta_1 \det \tilde{A}^{k-1} \tag{3.38}$$

$$\begin{aligned}\det \tilde{A}^{k-r} = (-1)^r &\left[\binom{k}{r} + \binom{k}{r+1}\Delta_0\right] \det A_{\Delta_{r+1},\ldots,\Delta_k}^{k-(r+1)} \\ &+ \Delta_{r+1} \det \tilde{A}^{k-(r+1)}.\end{aligned} \tag{3.39}$$

By repeated substitution of $\det \tilde{A}^{k-r}$ in (3.38), the claimed formula follows. □

Let $i \in \{0, \ldots, k\}$. Equate the formula for the determinant of $A_{\Delta_i,\ldots,\Delta_k}^{k-i}$ from Lemma 3.44 with that from Lemma 3.45. Then multiply both sides by $\prod_{j=0}^{i} \Delta_i / \omega_\Delta^k$ to obtain

$$\frac{\prod_{j=0}^{i} \Delta_i}{\omega_\Delta^k} \omega_{\Delta_i,\ldots,\Delta_k}^{k-i} = \frac{\prod_{j=0}^{i} \Delta_i}{\omega_\Delta^k} \sum_{r=i}^{k} (-1)^{r+i} \left[\binom{k-i}{r-i} + \binom{k-i}{r-i+1}\Delta_i\right] \prod_{j=i}^{r-1} \Delta_{j+1} \omega_{\Delta_{r+1},\ldots,\Delta_k}^{k-(r+1)}.$$

By applying (3.36) to both sides, we deduce,

$$\Delta_i \overline{\gamma}_i = \sum_{r=i}^{k} (-1)^{r+i} \left[\binom{k-i}{r-i} + \binom{k-i}{r-i+1}\Delta_i\right] \overline{\gamma}_{r+1},$$

thus showing that the i-th constraint from (3.24a) is satisfied.

Constraints (3.24d). We have

$$\omega_\Delta^k \sum_{r=0}^{k} \binom{k}{r} \overline{\lambda}_r \Delta_r \tag{3.40}$$

$$= \sum_{r=0}^{k} \binom{k}{r} \prod_{j=0}^{r} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j)$$

$$= \Delta_0 \sum_{r=0}^{k} \binom{k}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+1}^{k} (1-\Delta_j)$$

$$= \Delta_0 \sum_{r=1}^{k} \binom{k}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+1}^{k} (1-\Delta_j) + \Delta_0 \binom{k}{0} \prod_{j=1}^{k} (1-\Delta_j) + \Delta_0 \binom{k}{k} \prod_{j=0}^{k-1} \Delta_{j+1}$$

$$= \Delta_0 \sum_{r=1}^{k-1} \left[\binom{k-1}{r-1} + \binom{k-1}{r} \right] \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+1}^{k} (1-\Delta_j) + \Delta_0 \binom{k}{0} \prod_{j=1}^{k} (1-\Delta_j)$$

$$+ \Delta_0 \binom{k}{k} \prod_{j=0}^{k-1} \Delta_{j+1}$$

$$= \Delta_0 \sum_{r=1}^{k-1} \binom{k-1}{r-1} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+1}^{k} (1-\Delta_j) \tag{3.41}$$

$$+ \Delta_0 \sum_{r=1}^{k-1} \binom{k-1}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+1}^{k} (1-\Delta_j) \tag{3.42}$$

$$+ \Delta_0 \binom{k-1}{0} \prod_{j=2}^{k} (1-\Delta_j) - \Delta_0 \Delta_1 \binom{k-1}{0} \prod_{j=2}^{k} (1-\Delta_j) \tag{3.43}$$

$$+ \Delta_0 \binom{k-1}{k-1} \prod_{j=0}^{k-1} \Delta_{j+1}. \tag{3.44}$$

(3.42) can be expanded as

$$\Delta_0 \sum_{r=1}^{k-1} \left(\binom{k-1}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+2}^{k} (1-\Delta_j) \right.$$
$$\left. - \Delta_{r+1} \binom{k-1}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+2}^{k} (1-\Delta_j) \right). \tag{3.45}$$

Then (3.40) equals (3.43)+ (3.41)+ (3.45)+ (3.44), which can be written as

$$
\begin{aligned}
&\Delta_0 \Bigg(\underbrace{\binom{k-1}{0}\prod_{j=2}^{k}(1-\Delta_j) - \Delta_1\binom{k-1}{0}\prod_{j=2}^{k}(1-\Delta_j)}_{(3.43)/\Delta_0} + \underbrace{\sum_{r=1}^{k-1}\binom{k-1}{r-1}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+1}^{k}(1-\Delta_j)}_{(3.41)/\Delta_0} \\
&+ \underbrace{\sum_{r=1}^{k-1}\binom{k-1}{r}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+2}^{k}(1-\Delta_j) - \sum_{r=1}^{k-1}\Delta_{r+1}\binom{k-1}{r}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+2}^{k}(1-\Delta_j)}_{(3.45)/\Delta_0} \\
&+ \underbrace{\binom{k-1}{k-1}\prod_{j=0}^{k-1}\Delta_{j+1}}_{(3.44)/\Delta_0} \Bigg) \\
&= \Delta_0 \Bigg(\sum_{r=0}^{k-1}\binom{k-1}{r}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+2}^{k}(1-\Delta_j) + \sum_{r=1}^{k}\binom{k-1}{r-1}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+1}^{k}(1-\Delta_j) \\
&\quad - \sum_{r=0}^{k-1}\binom{k-1}{r}\Delta_{r+1}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+2}^{k}(1-\Delta_j) \Bigg),
\end{aligned} \tag{3.46}
$$

where in the last equation we used

$$
\begin{aligned}
\binom{k-1}{0}\prod_{j=2}^{k}(1-\Delta_j) &+ \sum_{r=1}^{k-1}\binom{k-1}{r}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+2}^{k}(1-\Delta_j) \\
&= \sum_{r=0}^{k-1}\binom{k-1}{r}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+2}^{k}(1-\Delta_j),
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{r=1}^{k-1}\binom{k-1}{r-1}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+1}^{k}(1-\Delta_j) &+ \binom{k-1}{k-1}\prod_{j=0}^{k-1}\Delta_{j+1} \\
&= \sum_{r=1}^{k}\binom{k-1}{r-1}\prod_{j=0}^{r-1}\Delta_{j+1}\prod_{j=r+1}^{k}(1-\Delta_j),
\end{aligned}
$$

and

$$
\begin{aligned}
-\Delta_1 \binom{k-1}{0} \prod_{j=2}^{k}(1-\Delta_j) - \sum_{r=1}^{k-1} \Delta_{r+1} \binom{k-1}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+2}^{k} (1-\Delta_j) \\
= -\sum_{r=0}^{k-1} \binom{k-1}{r} \Delta_{r+1} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+2}^{k} (1-\Delta_j).
\end{aligned}
$$

By doing a change of index variables $r' = r - 1$ in the second summand from (3.46), we have that the second and third summands cancel out. Therefore (3.40) equals to

$$
\begin{aligned}
&\Delta_0 \left(\sum_{r=0}^{k-1} \binom{k-1}{r} \prod_{j=0}^{r-1} \Delta_{j+1} \prod_{j=r+2}^{k} (1-\Delta_j) \right) \\
&= \Delta_0 \left(\sum_{r'=1}^{k} \binom{k-1}{r'-1} \prod_{j=0}^{r'-2} \Delta_{j+1} \prod_{j=r'+1}^{k} (1-\Delta_j) \right) \\
&= \Delta_0 \left(\sum_{r'=1}^{k} \binom{k-1}{r'-1} \prod_{j'=1}^{r'-1} \Delta_{j'} \prod_{j=r'+1}^{k} (1-\Delta_j) \right) \\
&= \Delta_0 \left(\sum_{r'=0}^{k-1} \binom{k-1}{r'} \prod_{j'=1}^{r'} \Delta_{j'} \prod_{j=r'+2}^{k} (1-\Delta_j) \right) \\
&= \Delta_0 \omega^{k-1}_{\Delta_1,\ldots,\Delta_k} = \overline{\gamma}_1 \omega^k_{\Delta},
\end{aligned}
$$

where last equation follows by definition (3.36) of $\overline{\gamma}$.

Constraints (3.24b)**,** (3.24c)**, and** (3.24e)**.** These constraints can be written compactly as follows:

$$
\begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = Q^k \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_k \end{pmatrix} \tag{3.47}
$$

where Q^k is the $(k+1) \times (k+1)$ *inverse Pascal upper-triangular matrix* defined as

$$Q^k_{\ell,r} := \begin{cases} (-1)^{\ell+r}\binom{k-\ell}{r-\ell} & \text{if } \ell \leq r \\ 0 & \text{o.w.} \end{cases} \tag{3.48}$$

for $\ell, r \in \{0, \ldots, k\}$. Indeed, row ℓ at the RHS of (3.47) corresponds to

$$\begin{aligned} \sum_{r=0}^{k} Q^k_{\ell,r}\gamma_r &= \sum_{r=0}^{k}(-1)^{\ell+r}\binom{k-\ell}{r-\ell}\gamma_r \\ &= \sum_{r=\ell}^{k-\ell}(-1)^r\binom{k-\ell}{r}\gamma_{\ell+r} \\ &= \lambda_\ell. \end{aligned}$$

where the last equality follows by (3.24e). We make use of the following lemma.

Proposition 3.46. *The matrix Q^k is invertible and its inverse is given by the matrix P^k defined as*

$$P^k_{\ell,r} = \begin{cases} \binom{k-\ell}{r-\ell} & \textit{if } \ell \leq r \\ 0 & \textit{o.w.} \end{cases} \tag{3.49}$$

for $\ell, r \in \{0, \ldots, k\}$.

Proof. This fact is implicit in the literature, but we give a proof here for completeness. Let $\ell, r \in \{0, \ldots, k\}$. We show that

$$(P^k Q^k)_{\ell,r} = \begin{cases} 1 & \text{if } \ell = r, \\ 0 & \text{if } \ell \neq r. \end{cases}$$

First, notice that the (ℓ, r) entry of $P^k Q^k$ is given by

$$\begin{aligned}\sum_{j=0}^{k} P^k_{\ell,j} Q^k_{j,r} &= \sum_{j=0}^{k} \binom{k-\ell}{j-\ell} (-1)^{j+r} \binom{k-j}{r-j} \\ &= (-1)^r \sum_{j=\ell}^{r} (-1)^j \binom{k-\ell}{j-\ell}\binom{k-j}{r-j} \\ &= (-1)^r \sum_{j=0}^{r-\ell} (-1)^{j+\ell} \binom{k-\ell}{j}\binom{(k-\ell)-j}{(r-\ell)-j} \\ &= (-1)^{r+\ell} \sum_{j=0}^{r-\ell} (-1)^j \binom{k-\ell}{r-\ell}\binom{r-\ell}{j} \\ &= (-1)^{r+\ell} \binom{k-\ell}{r-\ell} \sum_{j=0}^{r-\ell} \binom{r-\ell}{j}\end{aligned}$$

where the next to last equality follows by the binomial identity $\binom{a}{b}\binom{b}{c} = \binom{a}{c}\binom{a-c}{b-c}$ for integers a, b, c, see e.g. [177]. If $\ell = r$, then we have that $(P^k Q^k)_{\ell,r}$ equals $(-1)^{\ell+\ell}\binom{k-\ell}{0}\binom{0}{0} = 1$. On the other hand, if $\ell < r$, then the identity (3.31) implies that $(P^k Q^k)_{\ell,r}$ equals 0, and if $\ell > r$, the same holds since $\binom{k-\ell}{r-\ell} = 0$. $\square$

The latter proposition and (3.47) imply that there is a biunivocal relation between $(\lambda_0, \dots, \lambda_k)$ and $(\gamma_0, \dots, \gamma_k)$. Therefore, the following system is equivalent to (3.47):

$$\begin{pmatrix}\gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_k\end{pmatrix} = P^k \begin{pmatrix}\lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_k\end{pmatrix} \tag{3.50}$$

Next, we verify that $(\overline{\gamma}_0, \overline{\gamma}_1, \dots, \overline{\gamma}_k)$ and $(\overline{\lambda}_0, \overline{\lambda}_1, \dots, \overline{\lambda}_k)$ defined as in (3.36) and (3.33) satisfy (3.50). Indeed, for row $\ell \in \{0, \ldots, k\}$ we have

$$\begin{aligned}
&\overline{\gamma}_\ell = \sum_{r=0}^{k} P^k_{\ell,r}\overline{\lambda}_r = \sum_{r=\ell}^{k} \binom{k-\ell}{r-\ell}\overline{\lambda}_r \qquad (3.51)\\
\iff &\left(\prod_{j=0}^{\ell-1}\Delta_j\right)\frac{\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k}}{\omega^k_\Delta} = \sum_{r=\ell}^{k}\binom{k-\ell}{r-\ell}\frac{1}{\omega^k_\Delta}\prod_{j=0}^{r-1}\Delta_j\prod_{j=r+1}^{k}(1-\Delta_j)\\
\iff &\left(\prod_{j=0}^{\ell-1}\Delta_j\right)\frac{\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k}}{\omega^k_\Delta} = \left(\frac{\prod_{j=0}^{\ell-1}\Delta_j}{\omega^k_\Delta}\right)\sum_{r=\ell}^{k}\binom{k-\ell}{r-\ell}\prod_{j=\ell}^{r-1}\Delta_j\prod_{j=r+1}^{k}(1-\Delta_j)\\
\iff &\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k} = \sum_{r=\ell}^{k}\binom{k-\ell}{r-\ell}\prod_{j=\ell}^{r-1}\Delta_j\prod_{j=r+1}^{k}(1-\Delta_j)
\end{aligned}$$

where the last statement follows by Definition (3.32). In consequence, (3.50) is satisfied by $(\overline{\gamma}_0, \overline{\gamma}_1, \ldots, \overline{\gamma}_k)$ and $(\overline{\lambda}_0, \overline{\lambda}_1, \ldots, \overline{\lambda}_k)$. This verifies (3.24b) and (3.24c) by looking at rows $\ell = 0$ and $\ell = 1$ in (3.50), respectively. At last, (3.24e) is verified by (3.47).

3.7.4 Remaining Proofs for Theorem 3.1

We note that (B) implies (C) and (C) implies (D) are well-known facts, see, e.g., [145], [148]. We show that (D) implies (A).

Suppose therefore that $\mathrm{LS}^k_0(P)$ is nonempty and let G be a $(k+1)$-transitive subgroup of $G(P)$. We make use of the following auxiliary lemmas.

Lemma 3.47. *Let $j \in [k]$ and $i \in [n]$. Then*

$$\mathrm{LS}^j_0(P) \cap \{x \in \mathbb{R}^n \,:\, x_i \in \{0,1\}\} = \mathrm{LS}_0\left(\mathrm{LS}^{j-1}_0(P) \cap \{x \in \mathbb{R}^n \,:\, x_i \in \{0,1\}\}\right). \qquad (3.52)$$

Moreover, if $x \in \mathrm{LS}^j_0(P)$ and $x_i \in (0,1)$, then there exists $x^{i,1}, x^{i,0} \in \mathrm{LS}^{j-1}_0(P)$ such that $x^{i,1}_i = 1$ and $x^{i,0} = 0$.

Proof. See, e.g., [142, Lemma 2.2] for a proof of (3.52). The second statement follows by the definition of membership in $\mathrm{LS}^j_0(P)$. $\square$

Lemma 3.48. *Let $j \in [k]$ and $i \in [n]$. Suppose that P is $(k+1)$-transitive and that there exists $x \in \mathrm{LS}_0^{k-j+1}(P)$ such that $x_i \in (0,1)$. Then for every subset $J \subseteq [n]$ with $i \in J$ and $|J| = j$, there exists $x^{i,1}, x^{i,0} \in \mathrm{LS}_0^{k-j}(P)$ such that: a) $x_i^{i,1} = 1$ and $x_i^{i,0} = 0$; b) $x_\ell^{i,1} = x_r^{i,1}$ and $x_\ell^{i,0} = x_r^{i,0}$ for all $\ell, r \in [n] \setminus J$.*

Proof. By Lemma 3.47 we have that there exists $\tilde{x}^{i,1}, \tilde{x}^{i,0} \in \mathrm{LS}^{k-j}(P)$ such that a) holds. Next, since P is $(k+1)$-transitive, there exists a $(k+1)$-transitive group G under which P is G-invariant. By Proposition 3.26 the polytope $\mathrm{LS}_0^{k-j}(P)$ is G-invariant as well. This implies that the pointwise stabilizer of J, G_J, is transitive on the set of indices $[n] \setminus J$, by Proposition 3.12. Therefore, we have that $x^{i,1} := \beta_{G_J}(\tilde{x}^{i,1})$ and $x^{i,0} := \beta_{G_J}(\tilde{x}^{i,1})$ belong to $\mathrm{LS}_0^{k-j}(P)$ by the first statement of Proposition 3.11. Since G_J fixes the indices in J, we deduce that a) holds for $x^{i,1}, x^{i,0}$. The second statement of Proposition 3.11 implies that $x^{i,1}, x^{i,0}$ also satisfy b). $\square$

Let $x \in \mathrm{LS}_0^k(P)$ and consider $\overline{x} := \beta_G(x)$. Since P is integer-empty, we deduce that $\overline{x}$ is fractional in all components. By Proposition 3.26, $\mathrm{LS}_0^k(P)$ is G-invariant, hence $\overline{x} \in \mathrm{LS}_0^k(P)$. Pick $i = 1$ and apply Lemma 3.48 as to obtain points $x^{1,1}$ and $x^{1,0} \in \mathrm{LS}_0^{k-1}(P)$. Let G_1 be the subgroup of G that fixes the first coordinate. Using again the fact that P is integer-empty, we deduce that all coordinates of $\overline{x}^{1,1} := \beta_{G_1}(x^{1,,1})$ are fractional, except the first that is equal to 1. Similarly, all coordinates of $\overline{x}^{1,0} := \beta_{G_1}(x^{1,0})$ are fractional, except the first that is equal to 0. Since $\mathrm{LS}^{k-1}(P)$ is G-invariant by Proposition 3.26, we have $\overline{x}^{1,0}, \overline{x}^{1,1} \in \mathrm{LS}^{k-1}(P)$. Iterating, for all $j \in [k]$ and vectors $q^j \in \{0,1\}^j$ we obtain points $\overline{x}^{q^2}, \overline{x}^{q^3} \ldots, \overline{x}^{q^k}$ such that

- $\overline{x}^{q^j} \in \mathrm{LS}_0^{k-j}(P)$;
- $\overline{x}_\ell^{q^j} = q_\ell^j$ for all $\ell \in [j]$.

In particular, for all binary k-tuples q, we obtain a point $\overline{x}^q \in \mathrm{LS}_0^0(P) = P$ with $\overline{x}_\ell^q = q_\ell$ for all $\ell \in [k]$. Now fix a $(n-k)$-dimensional face F of $[0,1]^n$, defined as:

$$F = \{x \in [0,1]^n : \, x_i = 1 \, \forall i \in S_1; \, x_i = 0 \, \forall i \in S_0\}.$$

Let $q \in \{0,1\}^k$ such that $q_1 = q_2 = \cdots = q_{|S_1|} = 1$ and $q_{|S_1|+1} = q_{|S_1|+2} = \cdots = q_k = 0$. Since G is $(k+1)$-transitive, there exists $\pi \in G$ that maps $\overline{x}^q$ to a point of $\pi\overline{x}^q \in F$. In particular, $\pi\overline{x}^q \in F \cap P$, as required.

3.7.5 An Elementary Proof for 2-Transitive Polytopes

We present an elementary proof of Theorem 3.1 for $k = 1$.

Proposition 3.49. *Let $P \subseteq [0,1]^n$ be 2-transitive polytope. If there exists some $i \in [n]$ for which the faces $P^{i,1}$ and $P^{i,0}$ are nonempty, then $\mathrm{LS}_0(P) \neq \emptyset$.*

Proof. We assume that P does not have integral points, else the statement is trivially true. Additionally, we assume, without loss of generality, that the points x^1 and x^0 of P are such that $x^1_1 = 1$ and $x^0_1 = 0$. Let $F_1 := \operatorname{conv}(\{gx^1 \,:\, g \in G_1\})$ and $F_0 := \operatorname{conv}(\{gx^0 \,:\, g \in G_1\})$. These sets are contained in $P^{1,1} = P \cap \{x \in \mathbb{R}^n \,:\, x_1 = 1\}$ and $P^{1,0} = P \cap \{x \in \mathbb{R}^n \,:\, x_1 = 0\}$, respectively, since G_1 is the subgroup of G which fixes the first coordinate. Therefore, F_1 and F_0 are contained in $P_1(P)$.

Since G_1 is transitive on the coordinates $\{2,\dots,n\}$ and F_1 and F_0 are $(n-1)$-dimensional, the corresponding fixed points under G_1 are given by

$$\overline{x}^1_j = \begin{cases} 1 & \text{if } j = 1 \\ \frac{\|x^1\|_1 - 1}{n-1} & \text{if } j \neq 1 \end{cases}, \qquad \overline{x}^0_j = \begin{cases} 0 & \text{if } j = 1 \\ \frac{\|x^0\|_1}{n-1} & \text{if } j \neq 1 \end{cases}.$$

Clearly $x^1_j \geq 0$ for $j \neq 1$ since $P^{1,1} \neq \emptyset$.

Next, we show that there exists $\lambda^* \in [0,1]$ such that $\lambda^*\overline{x}^1 + (1-\lambda^*)\overline{x}^0 = \lambda^* e$. Indeed, let $\lambda^* := \frac{\|x^0\|_1}{n - (\|x^1\|_1 - \|x^0\|_1)}$. First, we note that λ^* is a well-defined, and non-negative, real number since $n - (\|x^1\|_1 - \|x^0\|_1) \geq n - \|x^1\|_1 > 0$ given that $x^1 \in [0,1]^n$ and $e \notin P$. Moreover, λ^* is less or equal than 1 since

$$\frac{\|x^0\|_1}{n-(\|x^1\|_1-\|x^0\|_1)} \leq 1 \iff \|x^0\|_1 \leq n-(\|x^1\|_1-\|x^0\|_1)$$
$$\iff \|x^1\|_1 \leq n.$$

We check whether $\lambda^*\overline{x}^1_j + (1-\lambda^*)\overline{x}^0_j = \lambda^*$ holds for every $j \in [n]$. If $j = 1$ this is trivially true. If $j \neq 1$, then

$$\begin{aligned}\lambda^*\overline{x}^1_j + (1-\lambda^*)\overline{x}^0_j = \lambda^* &\iff \overline{x}^0_j = \lambda^*(1-\overline{x}^1_j+\overline{x}^0_j)\\ &\iff \frac{\|x^0\|_1}{n-1} = \lambda^*\left(\frac{(n-1)-\|x^1\|_1+1+\|x^0\|_1}{n-1}\right)\\ &\iff \|x^0\|_1 = \lambda^*(n-(\|x^1\|_1-\|x^0\|_1)).\end{aligned}$$

In other words, $\lambda^* e \in P_1(P)$. Since our previous argument holds analogously for any $i \in [n]$ because: (1) G acts transitively on $[n]$ therefore the faces $P^{i,1}$ and $P^{i,0}$ are nonempty for every $i \in [n]$; and (2) G_i acts transitively on the set $[n] \setminus \{i\}$ of $n-1$ coordinates, i.e., our choice of λ^* is independent of which coordinate of x^1 and x^0 we picked, we have that $\lambda^* e$ belongs to $P_i(P)$ for every $i \in [n]$. □

Proposition 3.50. *Let $P \subseteq [0,1]^n$ be a 2-transitive polytope such that $0, e \notin P$. Then $\mathrm{LS}_0(P) \neq \emptyset$ if and only if there exists $i \in [n]$ such that $P^{i,1}, P^{i,0} \neq \emptyset$.*

Proof. Sufficiency follows by Proposition 3.49. We prove necessity by showing that the contrapositive is true. Suppose that for every $i \in [n]$ it holds that either (inclusive) $P^{i,1} = \emptyset$ or $P^{i,0} = \emptyset$. Given that G is transitive there are three possible cases. Case (1): If $P^{\tilde{i},1} = \emptyset$ for some $\tilde{i} \in [n]$, then it must be empty for every $i \in [n]$. Indeed, if this were not the case and for some $i' \in [n]$, $P^{i',1} \neq \emptyset$, there would exist a $g \in G$ and $x \in P$ such that $x_{i'} = 1$ and $(gx)_{\tilde{i}} = 1$. Therefore, $\mathrm{LS}_0(P) \subseteq \cap_{i=1}^n P^{i,0} \subseteq \{0\}$. Since we assumed that $0 \notin P$, then $\mathrm{LS}_0(P) = \emptyset$. Case (2): By an analogous argument, if $P^{\tilde{i},0} = \emptyset$ for some $\tilde{i} \in [n]$, then it must be empty for every $i \in [n]$, hence $\mathrm{LS}_0(P) \subseteq \cap_{i=1}^n P^{i,1} \subseteq \{e\}$. Given that $e \notin P$, we have that $\mathrm{LS}_0(P) = \emptyset$. Case (3): for every

$i \in [n]$, $P^{i,1} = P^{i,0} = \emptyset$. In this case it is clear that $\mathrm{LS}_0(P)$ must be empty.

□

Theorem 3.51. *Let $P \subseteq [0,1]^n$ be a 2-transitive polytope such that $0, e \notin P$. The following are equivalent:*

(a) *P intersects all the facets of the hypercube.*

(b) $\mathrm{SA}(P) \neq \emptyset$*;*

(c) $\mathrm{LS}(P) \neq \emptyset$*;*

(d) $\mathrm{LS}_0(P) \neq \emptyset$*.*

Proof. The statement (b) if and only (c) follows by definition. The equivalence between (c) and (d) follows by Proposition 3.52. We next show that if P is transitive, then the statement that there exists some $i \in [n]$ for which the faces $P^{i,1}$ and $P^{i,0}$ are nonempty is equivalent to P intersects all the facets of P. Indeed, if there exists $x^1, x^0 \in P$ such that $x^1_i = 1$ and $x^0_i = 0$, then for any $j \in [n]$ there exists some permutations $g_1, g_2 \in G(P)$ such that $(g_1(x^1))_j = 1$ and $(g_1(x^0))_j = 0$. In other words, $P^{j,1}$ and $P^{j,0}$ are nonempty. Therefore, by Propositions 3.49 and 3.50 (a) is equivalent to (d). □

3.8 Extending the Main Theorem: Beyond Polyhedral Hierarchies

We extend Theorem 3.1 to a non-polyhedral hierarchy. This extension requires an stronger assumption on P, namely that P is not full-dimensional.

3.8.1 Lovász–Schrijver Operator LS_+ with PSD constraint

In this subsection we show that if $P \subseteq [0,1]^n$ is an integer-empty $(k+1)$-transitive polytope with $\dim P \leq n-1$, then $\mathrm{LS}^k(P) \neq \emptyset$ if and only if $\mathrm{LS}^k_+(P) \neq \emptyset$. We start with an alternative proof of Theorem 3.1 for LS_0 and LS with $k = 1$ to illustrate that lifting matrices can be symmetrized, and then proceed with a technical lemma.

Let $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ be a matrix with row and column indices in $\{0,1,\ldots,n\}$. Recall the action (3.8) of G on any lifting matrix Y defined by $(\pi Y)_{i,j} := Y_{\pi^{-1}(i),\pi^{-1}(j)}$ for $i,j \in \{0,1,\ldots,n\}$ and $\pi \in G$ with $\pi(0) := 0$.

Proposition 3.52. *Let $P \subseteq [0,1]^n$ be a 2-transitive polytope. Then*

$$LS_0(P) \neq \emptyset \iff LS(P) \neq \emptyset. \tag{3.53}$$

Proof. Sufficiency follows by definition, therefore we prove necessity. Let G be a 2-transitive group under which P is invariant. Let $x \in \text{fix}_G(P) \cap \text{LS}_0(P)$ and let $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ such that $Ye_i, Y(e_0 - e_i) \in K(P)$, and $Ye_0 = Y^\top e_0 = \text{diag}(Y) = (1,x)^\top$. We show that $\overline{x}$ belongs to $\text{LS}(P)$ by exhibiting a symmetric matrix $\overline{Y}$ that satisfies the above properties satisfied by Y.

We define $\overline{Y} \in \mathbb{R}^{(n+1)\times(n+1)}$ by $\overline{Y}_{0,\ell} = Y_{0,\ell}$ and $\overline{Y}_{\ell,0} = Y_{\ell,0}$ for $\ell \in [n]$ and

$$\overline{Y}_{i,j} := \frac{1}{|G|}\sum_{g\in G}(g \cdot Y) = \frac{1}{|G|}\sum_{g\in G} Y_{g^{-1}(i),g^{-1}(j)} \tag{3.54}$$

for all $i,j \in [n]$. We show that $\overline{Y}_{i,j} = \overline{Y}_{j,i}$ for every pair $i,j \in [n]$. Indeed, since G is 2-transitive, then there exists $h \in G$ such that $h(i) = j$ and $h(j) = i$, hence

$$\begin{aligned}
\overline{Y}_{i,j} &= \frac{1}{|G|}\sum_{g\in h^{-1}G} Y_{(h^{-1}hg)^{-1}(i),(h^{-1}hg)^{-1}(j)} \\
&= \frac{1}{|G|}\sum_{r\in G} Y_{(h^{-1}r)^{-1}(i),(h^{-1}r)^{-1}(j)} \qquad r := hg \\
&= \frac{1}{|G|}\sum_{r\in G} Y_{r^{-1}(j),r^{-1}(i)} \qquad \text{since } (h^{-1}r)^{-1}(i) = r^{-1}h(i),\, (h^{-1}r)^{-1}(j) = r^{-1}h(j) \\
&= \overline{Y}_{j,i}.
\end{aligned}$$

Next, we have that $\overline{Y}e_0 = \overline{Y}^\top e_0 = \text{diag}(\overline{Y}) = (1,x)$ clearly holds since $gx = x$ for all $g \in G$. Finally, we show that $\overline{Y}$ satisfies $\overline{Y}e_i, \overline{Y}(e_0 - e_i) \in K(P)$ for all $i \in [n]$. Indeed, for any $g \in G$, $g \cdot Y$ is the lifting associated to gx, hence $(gY)e_i, (gY)(e_0 - e_i) \in K(P)$. Since $K(P)$ is a cone, we conclude that $x \in \mathrm{fix}_G(P) \cap \mathrm{LS}(P)$. Therefore, by propositions 3.26 and 3.21 the claim follows. $\square$

Lemma 3.53. *Let $P \subseteq [0,1]^n$ be a $(k+1)$-transitive polytope with $\dim(P) \leq n-1$. Let $\overline{x} \in \mathrm{LS}_0^k(P)$ and integers k_1, k_2 such that $k_1 + k_2 \leq k$ and $\overline{n} := n - (k_1 + k_2)$. By inducing the first k_1 variables to 0 and the next k_2 variables to 1, we obtain the following lifting matrix Y associated to some $x' \in \mathrm{LS}_0^{\overline{n}}(P)$ given by*

$$Y := \begin{pmatrix} 1 & \mathbf{0}_{1\times k_1} & \mathbf{1}^\top_{1\times k_2} & \gamma e_{1\times \overline{n}} \\ \mathbf{0}_{k_1\times 1} & \mathbf{0}_{k_1\times k_1} & \mathbf{0}_{k_1\times k_2} & \mathbf{0}_{k_1\times \overline{n}} \\ \mathbf{1}_{k_2\times 1} & \mathbf{0}_{k_2\times k_1} & \mathbf{1}_{k_2\times k_2} & \gamma e_{k_2\times \overline{n}} \\ \gamma e_{\overline{n}\times 1} & \mathbf{0}_{\overline{n}\times k_1} & \gamma e_{\overline{n}\times k_2} & Z \end{pmatrix}, \quad Z := \gamma I_{\overline{n}} + \sigma(e_{\overline{n}\times\overline{n}} - I_{\overline{n}}).$$

for some $\gamma, \sigma \in [0,1]$.

Proof. By Propositions 3.21 and 3.26 we can take $\overline{x} = \alpha e$, for some $\alpha \in [0,1]$, in $\mathrm{LS}_0^k(P) \cap \mathrm{fix}_G(P)$. Given that P is 2-transitive, we claim that there exists a lifting matrix $\overline{Y} \in \mathbb{R}^{(n+1)\times(n+1)}$ for $\overline{x}$ defined as

$$\overline{Y} := \begin{pmatrix} 1 & \alpha & \alpha & \cdots & \alpha \\ \alpha & \alpha & \beta & \cdots & \beta \\ \alpha & \beta & \alpha & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \beta \\ \alpha & \beta & \cdots & \beta & \alpha \end{pmatrix}$$

for some $\beta \in [0,1]$. This fact follows by a similar argument as the one used in the proof of Proposition 3.52. Indeed, 2-transitivity implies that for every pair $i, j \in \{1, \ldots, n+1\}$ such that $i - j = 1$ and $\ell \leq n - i$, there exists some $h^\ell_{i,j} \in G$ such that $h^\ell_{i,j}(i) = j$ and $h^\ell_{i,j}(j) = i + \ell$ hence $\overline{Y}_{i,j} = \overline{Y}_{j,i+\ell}$.

Next, we exhibit a lifting matrix for some $x^0 \in \mathrm{LS}_0^{n-k_1}(P)$ obtained by, starting from $\overline{x}$, inducing the first k_1 variables to 0. We proceed by induction over the finite set $\{1, \ldots, k_1\}$. The base case follows by noting that, by the definition of $\overline{x}$ being in $\mathrm{LS}_0^k(P)$ and by construction of $\overline{Y}$, there exists a point $x^{0,1} \in \text{LS}_0^{k-1}(P)$ whose lifting matrix can be described as

$$Y^{0,1} := \begin{pmatrix} 1 & 0 & \alpha' \mathbf{1}_{1\times(n-1)}^\top \\ 0 & 0 & \mathbf{0}_{1\times(n-1)} \\ \alpha' \mathbf{1}_{(n-1)\times 1} & \mathbf{0}_{(n-1)\times 1} & Z^{0,1} \end{pmatrix}$$

where $\alpha' = (\alpha - \beta)/(1-\alpha) \in [0,1]$, by definition when inducing a variable to 0, and $Z^{0,1}$ is some matrix in $[0,1]^{(n-1)\times(n-1)}$ whose diagonal is given by $\alpha' e_{n-1}$. We symmetrize $Y^{0,1}$ with respect to the group G_1, i.e., we consider $\frac{1}{|G_1|}\sum_{g\in G_1} gY^{0,1}$. In other words, this reduces to fixing the first row and column of $Y^{0,1}$ and symmetrizing the $(n-1)\times(n-1)$ matrix $Z^{0,1}$. Since G_1 is 2-transitive on $\{2,\ldots,n\}$, following the proof of Proposition 3.52, for every pair $i,j \in \{2,\ldots,n+1\}$ such that $i-j=1$ and $\ell \leq n-i$, there exists some $h_{i,j}^\ell \in G_1$ such that $h_{i,j}^\ell(i) = j$ and $h_{i,j}^\ell(j) = i+\ell$, hence $\overline{Z}_{i,j}^{0,1} = \overline{Z}_{j,i+\ell}^{0,1}$. This implies that $\overline{Z}^{0,1} = \alpha' I_{n-1} + \beta'(e_{(n-1)\times(n-1)} - I_{n-1})$. In consequence, the symmetrization of $Y^{0,1}$ with respect to the group G_1 is given by

$$\overline{Y}^{0,1} := \begin{pmatrix} 1 & 0 & \alpha' \mathbf{1}_{1\times(n-1)}^\top \\ 0 & 0 & \mathbf{0}_{1\times(n-1)} \\ \alpha' \mathbf{1}_{(n-1)\times 1} & \mathbf{0}_{(n-1)\times 1} & \overline{Z}^{0,1} \end{pmatrix}.$$

Next, suppose that we have induced the first $1 \leq k' < k_1$ variables to 0, so there exists some $x^{0;1,\ldots,k'} \in \text{LS}_0^{k-k'}(P) \cap \text{fix}_{G_{[k']}}(P)$ whose lifting matrix, which is $G_{[k']}$-invariant, can be written as

$$\overline{Y}^{0;1,\ldots,k'} := \begin{pmatrix} 1 & \mathbf{0}_{1\times k'} & \alpha'' \mathbf{1}_{1\times(n-k')}^\top \\ \mathbf{0}_{k'\times 1} & \mathbf{0}_{k'\times k'} & \mathbf{0}_{k'\times(n-k')} \\ \alpha'' \mathbf{1}_{(n-k')\times 1} & \mathbf{0}_{(n-k')\times k'} & \overline{Z}^{0,1,\ldots,k'} \end{pmatrix},$$

where $\overline{Z}^{0;1,\ldots,k'} = \alpha'' I_{n-k'} + \beta''(e_{(n-k')\times(n-k')} - I_{n-k'})$ for some $\alpha'', \beta'' \in [0,1]$. By definition, we have that $\overline{Y}^{0;1,\ldots,k'}(e_0 - e_i) \in K(\text{LS}_0^{n-(k'+1)}(P))$ for all $i \in [n]$ and $\text{diag}(\overline{Y}^{0;1,\ldots,k'}) = (1, 0_{k'}, \alpha'' 1_{n-k'})^\top$. By inducing $i = k'+1$ to 0 we have that $\hat{x} := (0_{k'+1}, 0, \hat{\alpha})^\top \in \text{LS}_0^{n-(k'+1)}(P)$, where $\hat{\alpha} := (\alpha'' - \beta'')/(1-\alpha'')$, and with a lifting matrix given by

$$Y^{0;1,\dots,k',k'+1} := \begin{pmatrix} 1 & \mathbf{0}_{1\times(k'+1)} & \alpha''\mathbf{1}^\top_{1\times(n-k')} \\ \mathbf{0}_{(k'+1)\times 1} & \mathbf{0}_{(k'+1)\times(k'+1)} & \mathbf{0}_{(k'+1)\times(n-(k'+1))} \\ \alpha''\mathbf{1}_{(n-(k'+1))\times 1} & \mathbf{0}_{(n-(k'+1))\times(k'+1)} & Z^{0,1,\dots,(k'+1)} \end{pmatrix},$$

Since $G_{1,\dots,k'+1}$ is 2-transitive on the set of variables $\{k'+1,\dots,n\}$, we can fix the first $k'+1$ rows and columns, and by the arguments given before we can symmetrize $Z^{0,1,\dots,k'+1}$ obtaining $\overline{Z}^{0;1,\dots,k'+1} = \hat{\alpha} I_{n-(k'+1)} + \hat{\beta}(e_{(n-(k'+1))\times(n-(k'+1))} - I_{n-(k'+1)})$.

In consequence, it is straightforward to verify that after k_1+k_2 rounds of LS_0, where where we induced k_1 variables to 0 and then k_2 variables to 1, we end up with a lifting matrix $\tilde{Y}$ associated to a point $x' \in \mathrm{LS}_0^{\overline{n}}(P)$ described as

$$\tilde{Y} := \begin{pmatrix} 1 & \mathbf{0}_{1\times k_1} & \mathbf{1}^\top_{1\times k_2} & \tilde{\alpha} e_{1\times\overline{n}} \\ \mathbf{0}_{k_1\times 1} & \mathbf{0}_{k_1\times k_1} & \mathbf{0}_{k_1\times k_2} & \mathbf{0}_{k_1\times\overline{n}} \\ \mathbf{1}_{k_2\times 1} & \mathbf{0}_{k_2\times k_1} & \mathbf{1}_{k_2\times k_2} & \tilde{\alpha} e_{k_2\times\overline{n}} \\ \tilde{\alpha} e_{\overline{n}\times 1} & \mathbf{0}_{\overline{n}\times k_1} & \tilde{\alpha} e_{\overline{n}\times k_2} & Z \end{pmatrix}$$

where $\tilde{\alpha} \in [0,1]$ and Z is some $G_{[k_1+k_2]}$-invariant matrix in $[0,1]^{\overline{n}\times\overline{n}}$ whose diagonal is given by $\tilde{\alpha} e_{\overline{n}}$. □

Proof of Theorem 3.2. Let $1 \le k \le n$ and let $\overline{x} \in \mathrm{LS}_0^k(P)$. For $k = n$ the claim follows by results in [148], [152]. Assume that $k < n$. Starting from $\overline{x}$, suppose we induce the first k_1 variables to 0, and the next k_2 variables to 1 such that $k_1 + k_2 \le k$. Let $\overline{n} := n - (k_1+k_2)$. By taking symmetrized lifting matrix from Lemma 3.53 and using the fact that $|\mathrm{fix}_G(P)| = 1$, by Proposition 3.17 since P is not full-dimensional, we obtain a point in $\mathrm{LS}_0^k(P)$ given by $\hat{x} = \sum_{i\in K_2} e_i + \sum_{i\in R} \tilde{\alpha} e_i$ where

$K_2 := \{k_1 + 1, \ldots, k_1 + k_2\}$ and $R := [n] \setminus [k_1 + k_2]$, and lifting matrix associated to $\hat{x}$:

$$
\begin{aligned}
Y :&= \begin{pmatrix} 1 & \mathbf{0}_{1\times k_1} & \mathbf{1}_{1\times k_2}^\top & \tilde{\alpha} e_{1\times \overline{n}} \\ \mathbf{0}_{k_1\times 1} & \mathbf{0}_{k_1\times k_1} & \mathbf{0}_{k_1\times k_2} & \mathbf{0}_{k_1\times \overline{n}} \\ \mathbf{1}_{k_2\times 1} & \mathbf{0}_{k_2\times k_1} & \mathbf{1}_{k_2\times k_2} & \tilde{\alpha} e_{k_2\times \overline{n}} \\ \tilde{\alpha} e_{\overline{n}\times 1} & \mathbf{0}_{\overline{n}\times k_1} & \tilde{\alpha} e_{\overline{n}\times k_2} & Z \end{pmatrix} \\
Z :&= \tilde{\alpha} I_{\overline{n}} + \tilde{\beta}(e_{\overline{n}\times\overline{n}} - I_{\overline{n}})
\end{aligned}
$$

where $\tilde{\alpha} := \frac{(n\alpha - k_2)}{\overline{n}}$ and $\tilde{\beta} := \frac{\overline{n}\tilde{\alpha}^2 - \tilde{\alpha}}{\overline{n}-1}$ since

$$
k_2 + 1 + (\overline{n} - 1)\frac{\tilde{\beta}}{\tilde{\alpha}} = k_2 + \overline{n}\tilde{\alpha}. \tag{3.55}
$$

Indeed, this equality follows since $|\text{fix}_G(P)| = 1$, which implies that the Reynolds operator applied to the point obtained after inducing on new variable coincides with that applied to $\tilde{x}$. In particular, the sums of their components must agree, yielding (3.55).

We claim that Y is positive semidefinite. Indeed, by taking Schur's complement of Y, psdness of Y can be reduced to

$$
\begin{aligned}
Y \succeq 0 \iff & \begin{pmatrix} \mathbf{0}_{k_1\times k_1} & \mathbf{0}_{k_1\times k_2} & \mathbf{0}_{k_1\times \overline{n}} \\ \mathbf{0}_{k_2\times k_1} & \mathbf{1}_{k_2\times k_2} & \tilde{\alpha} e_{k_2\times \overline{n}} \\ \mathbf{0}_{\overline{n}\times k_1} & \tilde{\alpha} e_{\overline{n}\times k_2} & Z \end{pmatrix} - \begin{pmatrix} \mathbf{0}_{k_1\times 1} \\ \mathbf{1}_{k_2\times 1} \\ \tilde{\alpha} e_{k_2\times \overline{n}} \end{pmatrix} \begin{pmatrix} \mathbf{0}_{k_1\times 1} & \mathbf{1}_{k_2\times 1} & \tilde{\alpha} e_{k_2\times \overline{n}} \end{pmatrix} \succeq 0 \\
\iff & W := (\tilde{\alpha} - \tilde{\alpha}^2) I_{\overline{n}} + (\tilde{\beta} - \tilde{\alpha}^2)(e_{\overline{n}\times\overline{n}} - I_{\overline{n}}) \succeq 0
\end{aligned}
$$

We show that W is diagonally dominant, i.e., that for $i \in [\overline{n}]$ we have that $|W_{i,i}| - \sum_{j\neq i} |W_{i,j}| \geq 0$.

Since $\tilde{\beta} - \tilde{\alpha}^2 \leq 0$ it follows that

$$
\begin{aligned}
|W_{i,i}| - \sum_{j \neq i} |W_{i,j}| &= \tilde{\alpha} - \tilde{\alpha}^2 - (\overline{n} - 1)(\tilde{\alpha}^2 - \tilde{\beta}) \\
&= \tilde{\alpha} - \overline{n}\tilde{\alpha}^2 + (\overline{n} - 1)\tilde{\beta} \\
&= \tilde{\alpha} - \overline{n}\tilde{\alpha}^2 + \overline{n}\tilde{\alpha}^2 - \tilde{\alpha} \\
&= 0.
\end{aligned}
$$

In consequence, all the symmetrized lifting matrices are positive semidefinite. □

3.9 Extension: A Procedure for Analyzing the SA Rank

By now, we have seen the strength of Theorem 3.9. As noted in the introduction, its power lies in the "single-step" characterization of SA, in contrast to the iterative construction of the LS operators. In particular, this proposition allows us to reason about high levels of the hierarchy by examining specific points in P that lie on faces of $[0, 1]^n$. In other words, understanding the intersections of P with the faces of $[0, 1]^n$ suffices to characterize the hierarchy – this is independent of whether P is symmetric or not. If P is $(k + 1)$-transitive, these conditions can be verified efficiently, namely in time polynomial in the encoding length of P and k.

In this section, we leverage Theorem 3.9 to extend the proof technique from Section 3.7.3 to construct fractional solutions to the Sherali-Adams relaxation in settings where P is highly symmetric but not necessarily integer-empty. The question here is different: we know that $\mathrm{SA}^k(P) \neq \emptyset$ for all k, so the goal is no longer to certify integer-emptiness, we would like to establish how large k needs to be for $\mathrm{SA}^k(P) = P_I$. Previously, the approach was as follows: finding a point in $\mathrm{M}^k(P)$ immediately implied that the k-th level fails to capture the integer structure of P. In the present setting, however, the existence of $y \in \mathrm{M}^k(P)$ does not by itself imply that the k-th level has not reached the integer hull.

Instead, the goal is to construct a fractional point $y \in \mathrm{M}^k(P)$ whose projection $\mathrm{proj}_x(y) \in$

$\mathrm{SA}^k(P)$ lies outside P_I and which persists at high levels (large k) of the hierarchy. Such points can then be used to derive lower bounds on either the Sherali–Adams rank of P or the rank of valid inequalities for P_I.

If we were interested in a minimization problem whose group of formulation symmetries $G(A, b, c)$, see Section 3.2.2, is transitive, so that the objective lies in the span of the all-ones vector, then a natural approach is to construct fractional points $y \in \mathrm{M}^k(P)$ supported on points in P with *small objective value* and, for simplicity, *small support*. While small support is not essential, it simplifies the computations and aligns with our goal of a parsimonious technique for deriving bounds in highly symmetric problems.

To certify that $\mathrm{proj}_x(y) \notin P_I$, it suffices to exhibit a valid inequality for P_I that is violated by $\mathrm{proj}_x(y)$. In that case, we can conclude that the Sherali–Adams rank of P is strictly greater than k. In this sense, the construction aims to produce points that "fool" the relaxation while remaining feasible for $\mathrm{M}^k(P)$.

We proceed by describing the assumptions needed to apply this generalization to P, which are weaker than those used in our previous analysis. We then establish the validity of the approach by revisiting the arguments in Section 3.7. While this construction does not guarantee the existence of a point in $\mathrm{SA}^k(P) \setminus P_I$, Sections 3.9.3, 3.9.4, and 3.9.5 demonstrate that the strategy can nevertheless be effective.

3.9.1 Proof of the Generalization by Revisiting Section 3.7

Suppose that $P \subseteq [0, 1]^n$ is a $(k + 1)$-transitive polytope, not necessarily integer-empty. Let $G \leq G(P)$.

1. Construction of points from $\mathrm{M}^0(P)$. Let $i \in \mathbb{N}$ fixed with $i \leq k$. We assume that, for $\ell \in \{0, \ldots, i\}$, the following sets are nonempty

$$P_\ell := P \cap \{x \in [0, 1]^n \,:\, x_i = 1 \text{ for all } i \in [\ell], x_j = 0 \text{ for all } j \in [k] \setminus [\ell]\}. \tag{3.56}$$

For $i < k$, we note that this assumption is weaker than assuming that P intersects all the $(n-k)$-dimensional faces of $[0,1]^n$. Moreover, we assume that $\dim(P_0) > 0$.

Since P_ℓ is nonempty, for each ℓ, there exists some $x^\ell \in P_\ell$. Then, as in Section 3.7, by $(k+1)$-transitivity of G we can construct the points

$$\overline{x}^\ell := \beta_{G_{[k]}}(x^\ell) = (\underbrace{1,\ldots,1}_{\ell}, \underbrace{0,\ldots,0}_{k-\ell}, \underbrace{\Delta_\ell,\ldots,\Delta_\ell}_{n-k}) \in P_\ell, \tag{3.57}$$

for some $\Delta_\ell \in [0,1]$. Notice that in Section 3.7.3 we deduced that $\Delta_\ell \in (0,1)$ for $\ell \in \{0,\ldots,k\}$; however, this need not hold if P contains integer points. Let $i^* \le i$ denote the smallest integer such that: $\Delta_\ell \in (0,1)$ for all $\ell \in \{0,\ldots,i^*-1\}$ and $\Delta_{i^*} = 0$. We note that $i^* \neq 0$ since $\dim(P_0) > 0$.

From this point on, for the sake of simplicity, we write i^* by simply i, and define $\Delta := (\Delta_0,\ldots,\Delta_i,\Delta_{i+1},\ldots,\Delta_k)^{k+1}$ where $\Delta_j := 0$ for $i+1 \le j \le k$.

Next, for a fixed 2-partition $J_0 \sqcup J_1 = S \subseteq [n]$ with $|S| = k$ and $|J_1| = \ell \in \{0,\ldots,i\}$ we define, as in (3.22), the points

$$\overline{x}_j^{J_0,J_1} := \begin{cases} 0 & \text{if } j \in J_0, \\ 1 & \text{if } j \in J_1, \\ \Delta_\ell & \text{if } j \in [n] \setminus S. \end{cases} \tag{3.58}$$

Following an argument similar to the proof of Lemma 3.38, we have that the points $(1, \overline{x}^{J_0,J_1})$ belong to $\mathrm{M}^0(P)$. We stress that this holds for 2-partitions with $|J_1| \le i$.

2. System of constraints defining $\overline{y}$. We give sufficient conditions for the existence of $\overline{y} \in \mathrm{M}^k(P)$ supported on the points previously defined, i.e., points $\overline{x}^{J_0,J_1}$ with $|J_0 \sqcup J_1| = k$ such that $|J_1| \le i$. To this end, we modify Lemma 3.39 by imposing some extra conditions on γ and λ.

Lemma 3.54. *Consider the following system in the variables $\gamma_0, \ldots, \gamma_{k+1} \in \mathbb{R}$, $\lambda_0, \ldots, \lambda_k \in \mathbb{R}$*

$$A_\Delta^k \gamma = \Delta_0(1, 0, \ldots, 0), \tag{3.59a}$$

$$1 = \sum_{\ell=0}^{k} \binom{k}{\ell} \lambda_\ell, \tag{3.59b}$$

$$\gamma_1 = \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \lambda_\ell, \tag{3.59c}$$

$$\gamma_1 = \sum_{\ell=0}^{k} \binom{k}{\ell} \lambda_\ell \Delta_\ell, \tag{3.59d}$$

$$\lambda_\ell = \sum_{r=0}^{k-\ell} (-1)^r \binom{k-\ell}{r} \gamma_{r+\ell} \quad \forall \ell \in \{0, 1, \ldots, k\}, \tag{3.59e}$$

$$\lambda_\ell = 0 \quad \forall \ell \in \{i+1, \ldots, k\}, \tag{3.59f}$$

$$\gamma_\ell = 0 \quad \forall \ell \in \{i+1, \ldots, k+1\}, \tag{3.59g}$$

$$\mathbf{0} \leq \lambda, \gamma \leq \mathbf{1}, \quad \gamma_0 = 1, \tag{3.59h}$$

where we abbreviate $\lambda \coloneqq (\lambda_0, \ldots, \lambda_k)$ and $\gamma \coloneqq (\gamma_1, \ldots, \gamma_{k+1})$. Suppose the system admits a solution $(\overline{\gamma}_0, \overline{\gamma}, \overline{\lambda})$ with $\overline{\lambda}_\ell > 0$ for all $\ell \in \{0, \ldots, i\}$.

Then $\overline{y} \in \mathrm{M}^k(P)$ where

$$\overline{y}_I = \overline{\gamma}_{|I|} \quad \textit{for all } I \subseteq [n], \textit{ with } |I| \leq k+1. \tag{3.60}$$

To prove Lemma 3.54, we follow the steps in our proof of Lemma 3.39 and show that the argument carries through. Assume that $(\overline{\gamma}_0, \overline{\gamma}, \overline{\lambda})$ is a solution to the system (3.54). For each subset $S \subseteq [n]$, and every 2-partition $J_0 \sqcup J_1 = S$ with $|S| = k$, we let $\overline{y}^{J_0,J_1}$ and $z^{J_0,J_1}(\overline{y})$ be as in (3.3). To show that $\overline{y} \in \mathrm{M}^k(P)$, we verify that the conditions (i) and (ii) from Theorem 3.9 hold.

We begin with (i). To this end, observe that for $\ell = |J_1| \in \{0, \ldots, i\}$, we have by assumption

$$\overline{y}_\emptyset^{J_0,J_1} = \sum_{H \subseteq J_0} (-1)^{|H|} \overline{y}_{J_1 \cup H} = \sum_{r=0}^{k-\ell} (-1)^r \binom{k-\ell}{r} \overline{\gamma}_{r+\ell} = \overline{\lambda}_\ell > 0. \tag{3.61}$$

Thus, $z^{J_0,J_1}(\overline{y}) = \overline{y}^{J_0,J_1}/\overline{y}_\emptyset^{J_0,J_1}$. If $i+1 \leq |J_1| \leq k$, then we have $\overline{y}_\emptyset^{J_0,J_1} = \overline{\lambda}_{|J_1|} = 0$ and $z^{J_0,J_1}(\overline{y}) = y^{J_0,J_1}$, hence $z^{J_0,J_1}(\overline{y}) \notin \mathrm{M}^0(P)$.

Next, we show that condition (i) holds for 2-partitions of the form $J_0 = [k] \setminus [\ell]$, $J_1 = [\ell]$ for $\ell \in \{0, \dots, i\}$. As in Section 3.7.3, we abbreviate $z^\ell(\overline{y}) \coloneqq z^{[k]\setminus[\ell],[\ell]}(\overline{y})$ and $\overline{y}^\ell \coloneqq \overline{y}^{[k]\setminus[\ell],[\ell]}$. This condition is readily verified since $z^\ell(\overline{y}) = (1, \overline{x}^\ell)$ holds for $\ell \in \{0, \dots, i\}$ by the exact same arguments as in Section 3.7.3. We remark that the equality $\frac{\overline{y}_j^\ell}{\overline{\lambda}_\ell} = \overline{x}_j^\ell$, for $j \notin [k]$, is captured by the ℓ-th row of $A_\Delta^k \gamma = \Delta_0(1, 0, \dots, 0)$, i.e., constraints (3.59a). As for the remaining rows with indices $i+1 \leq \ell \leq k$, since $\Delta_\ell = 0$ for $i \leq \ell \leq k$ and constraints (3.59g) enforce $\gamma_\ell = 0$ for $i \leq \ell \leq k+1$, all terms in

$$-\Delta_\ell \overline{\gamma}_\ell + \sum_{r=\ell+1}^{k+1} (-1)^{\ell+r-1} \left[\binom{k-\ell}{r-\ell-1} + \binom{k-\ell}{r-\ell} \Delta_\ell \right] \overline{\gamma}_r = 0$$

vanish, and hence these rows are trivially satisfied.

The verification of condition (ii), i.e., $\overline{y}_I = \sum_{J_0 \sqcup J_1 = S} \overline{y}_\emptyset^{J_0,J_1} z_I^{J_0,J_1}(\overline{y})$ for every $I \subseteq [n]$ with $|I| \leq 1$, is straightforward. Indeed, the only difference with Section 3.7 is that in this case 2-partitions with $i+1 \leq |J_1| \leq k$ have $y_\emptyset^{J_0,J_1} = 0$. Equations (3.59b), (3.59c) and (3.59d) readily hold.

3. Solution to the system defining $\overline{\gamma}$. Our characterization of the $\det A_\Theta^t$ only depends on the entries of $\Theta \in \mathbb{R}^{k+1}$ and its length $t \in \mathbb{N}$ (see Lemmas 3.41 and 3.44). Hence, for our choice of $\Theta = \Delta$ and $t = k$ we have

$$\omega_\Delta^k = \sum_{r=0}^{k} \binom{k}{r} \prod_{j=0}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1 - \Delta_j) \tag{3.62}$$

$$= \sum_{r=0}^{i} \binom{k}{r} \prod_{j=0}^{r-1} \Delta_j \prod_{j=r+1}^{i} (1 - \Delta_j) \tag{3.63}$$

since $\prod_{j=0}^{r-1} \Delta_j = 0$ for $i+1 \leq r \leq k$ and $\prod_{j=i+1}^{k} (1 - \Delta_j) = 1$. We note that $\omega_\Delta^k > 0$ since every summand is nonnegative and the first one, i.e., for $r = 0$, is strictly positive. By a similar argument we have that

$$\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k} = \sum_{r=\ell}^{k} \binom{k-\ell}{r-\ell} \prod_{j=\ell}^{r-1} \Delta_j \prod_{j=r+1}^{k} (1-\Delta_j) \tag{3.64}$$

is strictly positive for $1 \le \ell \le i$. In particular, $\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k} = 1$ for $i+1 \le \ell \le k$.

4. Solution to the general system. Finally, we define a vector $(\overline{\gamma}_0, \overline{\gamma}, \overline{\lambda}) \in \mathbb{R}^{2k+3}$ and show that it satisfies all constraints of Lemma 3.54. In particular, for $\ell \in \{0, \ldots, k\}$, we define

$$\overline{\lambda}_\ell \coloneqq \frac{1}{\omega^k_\Delta} \prod_{j=0}^{\ell-1} \Delta_j \prod_{j=\ell+1}^{i} (1-\Delta_j). \tag{3.65}$$

For $\ell \in [i+1]$, we define

$$\rho^\ell_\Delta \coloneqq \begin{cases} \Delta_{\ell-1} \left(\frac{\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k}}{\omega^{k-(\ell-1)}_{\Delta_{\ell-1},\ldots,\Delta_k}} \right) & \text{if } \ell \le i, \\ \Delta_i & \text{if } \ell = i+1. \end{cases}$$

These scalars are well-defined since $\omega^{k-\ell}_{\Delta_\ell,\ldots,\Delta_k} > 0$. Then define $\overline{\gamma} \in \mathbb{R}^{k+1}$ as

$$\overline{\gamma}_\ell \coloneqq \begin{cases} \rho^\ell_\Delta \overline{\gamma}_{\ell-1} & \text{if } \ell \in [i+1] \\ 0 & \text{if } \ell \in [k+1] \setminus [i+1] \end{cases} \tag{3.66}$$

and $\overline{\gamma}_0 \coloneqq 1$.

Notice that the solutions defined above share the same structure as those in Section 3.7.3. The only modification is that the entries of γ and λ are set to zero from index $i+1$ onward. Consequently, constraints (3.59h) are immediately satisfied: $\overline{\lambda} \in [0,1]^{k+1}$ follows directly from $\Delta \in [0,1]^{k+1}$ and the definition of $\overline{\lambda}$; and $(\overline{\gamma}_0, \overline{\gamma}) \in [0,1]^{k+2}$ since $\overline{\gamma}_0 = 1$ and $\rho^\ell_\Delta \in (0,1)$ for all $\ell \in [i]$ and $\rho^{i+1}_\Delta = 0$. Moreover, constraints (3.59f) hold by definition (3.65) together with our definition of Δ. Similarly, constraints (3.59f) follow from the construction of $\overline{\gamma}$ in (3.66).

As noted earlier, the determinant of A^t_Θ depends only on the entries of Θ (and hence on its length t), and not on any structural properties of Θ. Therefore, Lemma 3.45 implies that the ℓ-th constraint in (3.59a), for $\ell \in \{0, \dots, i\}$, is satisfied. Rows $i+1$ onwards in (3.59a) are trivially satisfied. Moreover, constraint (3.59d) is satisfied by the definition of $\overline{\lambda}_\ell$, ω^k_Δ and $\omega^{k-1}_{\Delta_1,\dots,\Delta_k}$. See in Section 3.7.3 the verification of constraint (3.24d).

Next, recall that constraints (3.59e) correspond to system (3.47). The matrix Q_k in (3.47), which defines the bijective relationship between γ and λ, is an invertible upper-triangular matrix. Hence, we readily verify that our solution $\overline{\gamma}_\ell = \overline{\lambda}_\ell = 0$ for $\ell \in \{i+1, \dots, k\}$ is consistent with these constraints. In turn, system (3.47) is equivalent to (3.50). Since, for each $\ell \in \{0, \dots, k\}$, equation (3.51) holds by the identity (3.64), we conclude that constraints (3.59e) are satisfied by $(\overline{\gamma}_0, \overline{\gamma}, \overline{\lambda})$.

Finally, observe that the first two rows of (3.50) correspond to (3.24b) and (3.24c), completing the proof.

3.9.2 The Procedure

Let $P \subseteq [0,1]^n$ be a $(k+1)$-transitive polytope. We describe next a general technique, which relies on Lemma 3.54, for constructing fractional points in $y \in \mathrm{M}^k(P)$ with *small objective value*. The goal is to obtain a point $\mathrm{proj}_x(y) \notin P_I$. To certify this, it suffices to exhibit a valid inequality for P_I that is violated by $\mathrm{proj}_x(y)$.

The key idea is to use supporting points associated with 2-partitions of subsets of $[n]$ of cardinality k containing at most i ones. In particular, we focus on points with few ones and many zeros (or, equivalently, by applying the affine transformation $\mathbf{1} - x$ to P, few zeros and many ones). As discussed earlier, for problems whose objective is in the span of $\mathbf{1}$, the goal is to identify parameters i and k such that the resulting point has *small total mass* $\sum_{j=1}^n x_j$. If we are minimizing, this increases the chance of violating a known valid inequality for P_I, such as an objective (or cardinality) constraint.

Step 0) Consider the affine transformation $Q : \mathbb{R}^n \to \mathbb{R}^n$ defined by $Q(x) := \mathbf{1} - x$ to P and obtain $P' := Q(P)$. We note that affine transformations are (permutation) symmetry-preserving, since they preserve the face lattice, and SA rank preserving. We note that although the affine map Q preserves the face lattice of P, it reverses the roles of the 0- and 1-faces of $[0,1]^n$. Thus, the hypercube-face intersection pattern of P is mapped to the complementary hypercube faces. We use this as a heuristic to obtain a larger k.

Step 1) Identify the largest k such that P intersects all of the $(n-k)$-dimensional faces $[0,1]^n$. Apply the same procedure to P', and denote the resulting value by k'. Among these, choose the polytope corresponding to $\max\{k,k'\}$. From this point onward, we fix this polytope and denote by P, with parameter k.

Step 2) For $\ell \in \{0,\dots,k\}$, compute points

$$\overline{x}^{\ell} := (\underbrace{1,\dots,1}_{\ell},\underbrace{0,\dots,0}_{k-\ell},\underbrace{\Delta_\ell,\dots,\Delta_\ell}_{n-k}) \in P_\ell. \tag{3.67}$$

Find $i \leq k$ as in Section 3.9.1. If $i = 0$, return to Step 1) and replace P by $Q(P)$.

Step 3) Use (3.59b) to compute ω_Δ^k, i.e.,

$$\omega_\Delta^k = \sum_{\ell=0}^{i} \binom{k}{\ell} \prod_{j=0}^{\ell-1} \Delta_j \prod_{j=\ell+1}^{i} (1-\Delta_j)$$

Step 4) Use equation (3.65) to find weights $\overline{\lambda}$,

$$\begin{aligned}
\overline{\lambda}_0 &= \frac{1}{\omega_\Delta^k} \prod_{j=1}^{i} (1-\Delta_j), \\
\overline{\lambda}_1 &= \frac{1}{\omega_\Delta^k} \Delta_0 \prod_{i=2}^{i} (1-\Delta_i), \\
&\vdots \\
\overline{\lambda}_i &= \frac{1}{\omega_\Delta^k} \prod_{j=0}^{i-1} \Delta_j
\end{aligned}$$

Step 5) Finally, compute $\overline{\gamma}_1$ with (3.59d)

$$\overline{\gamma}_1 = \sum_{\ell=0}^{i} \binom{k}{\ell} \overline{\lambda}_\ell \Delta_\ell.$$

Step 6) Then, $x^* := \overline{\gamma}_1 \mathbf{1}$ is a solution to system (3.54), by Lemma 3.54 we have $x^* \in \mathrm{SA}^k(P)$.

Step 7) Given a known valid inequality for P_I, say $ax \geq b$, we check whether x^* violates it. If the inequality is satisfied and $k > 1$, we either return to Step 2) and select a smaller value of k, or return to Step 1) and replace P with $Q(P)$.

In the following sections, we present three applications of this template. Our procedure recovers the known ranks for these problems. Since these bounds are known (value of k), we do not necessarily follow the bound computed in Step 1); accordingly, those instructions are omitted.

3.9.3 Maximum Stable Set

We show how we can easily derive the the SA bound in [145] using our procedure outlined in 3.9.2. We note that this bound matches the LS bound in [148].

Consider the vertex formulation for the LP relaxation of the Maximum Stable Set polytope on the complete graph K_n given by

$$P^{SS} := P^{SS}(K_n) = \{x \in \mathbb{R}^n \,:\, \forall i, j \in [n],\, i < j,\, x_i + x_j \leq 1, \quad \mathbf{0} \leq x \leq \mathbf{1}\}.$$

This polytope is n-transitive but it is not integer-empty. Indeed, its integer points are $\mathbf{0}$ and e_i, for $i \in [n]$, where e_i denotes the i-th vector of the canonical basis of $\mathbb{R}^n$. It is known that $P_I^{SS} = \mathrm{SA}^{n-2}(P^{SS}) \subsetneq \mathrm{SA}^{n-3}(P^{SS})$. Clearly, $\mathrm{OPT} := \max\{\mathbf{1}^\top x \,:\, x \in P^{SS},\, x \in \{0,1\}^n\}$ is 1, hence $\sum_{i=1}^n x_i \leq 1$ is valid for P_I^{SS}.

We can readily see that the largest k in Step 1) for P^{SS} is $k = n-1$ (since for $Q(P^{SS})$ we have $k' = 1$). Since the rank is known to be $n-2$, and our goal is to certify this bound, we set $k = n-3$.

Then, we compute $i^* = 1$ as in Step 2) by consider the following points which belong to P^{SS}:

$$x^0 := \left(\underbrace{0, \ldots, 0}_{n-3}, \underbrace{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}}_{3}\right)$$

$$x^1 := \left(\underbrace{1, 0 \ldots, 0}_{n-3}, \underbrace{0, 0, 0}_{3}\right)$$

We define $\Delta_j := 0$ for $2 \leq j \leq n-3$ since we notice that there are no more points in P^{SS} with nonzero integer coordinates. Moreover, we pick $\Delta_0 := \frac{1}{2}$ and $\Delta_1 = 0$.

For Step 3), we compute ω_{Δ}^{n-3}. We have,

$$\begin{aligned}
\omega_{\Delta}^{n-3} &= \sum_{\ell=0}^{i} \binom{n-3}{\ell} \prod_{j=0}^{\ell-1} \Delta_j \prod_{j=\ell+1}^{i} (1-\Delta_j) \\
&= \binom{n-3}{0}(1-\Delta_1) + \binom{n-3}{1}\Delta_0 \\
&= 1 + \frac{n-3}{2} \\
&= \frac{n-1}{2}.
\end{aligned}$$

Next, for Step 4) we compute the weights $\overline{\lambda}_0$ and $\overline{\lambda}_1$ as

$$\begin{aligned}
\overline{\lambda}_0 &= \frac{2}{n-1}(1-\Delta_1) = \frac{2}{n-1} \\
\overline{\lambda}_1 &= \frac{2}{n-1}\Delta_0(1-\Delta_2) = \frac{1}{n-1}.
\end{aligned}$$

Then, following Step 5) we compute $\overline{\gamma}_1$ as

$$\begin{aligned}
\overline{\gamma}_1 &= \binom{n-3}{0}\overline{\lambda}_0\Delta_0 + \binom{n-3}{1}\overline{\lambda}_1\Delta_1 \\
&= \frac{1}{n-1}.
\end{aligned}$$

Hence, $x^* = \frac{1}{n-1}\mathbf{1} \in \mathrm{SA}^{n-3}(P^{SS})$ and its objective is $\frac{n}{n-1} = 1 + \frac{1}{n-1}$. We know that $\sum_{i=1}^{n} x_i \leq 1$

is valid for P_I^{SS}. Since x^* violates this inequality, it follows that P_I^{SS} is strictly contained in $\mathrm{SA}^{n-3}(P^{SS})$. Thus, we recover the bound in [145].

It is straightforward to verify that if we had chosen $k = n-2$, then the resulting point would belong to P_I^{SS}, as expected since $\mathrm{SA}^{n-2}(P^{SS}) = P_I^{SS}$. Indeed, $\omega_\Delta^{n-2} = \frac{n}{2}$, hence $\overline{\lambda}_0 = \frac{2}{n}$ and $\overline{\lambda}_1 = \frac{1}{n}$, hence $\overline{\gamma}_1 = \frac{1}{n}$. Notice that $\overline{\gamma}_1 \mathbf{1} \in \mathrm{conv}(\{\pi x_1 \,:\, \pi \in G(P^{SS})\}) \subseteq P_I^{SS}$.

3.9.4 Minimum Knapsack

Consider the minimum knapsack polytope P_n^K defined in Section 3.3:

$$P_n^K = \left\{ x \in \mathbb{R}^n \,:\, \sum_{i=1}^n x_i \geq \frac{1}{2}, \quad \mathbf{0} \leq x \leq \mathbf{1} \right\}. \tag{3.68}$$

We apply our template 3.9.2 with $k = n-1$. As in Section 3.9.3 consider the following points which belong to P_n^K:

$$x^0 := \left(\underbrace{0, \dots, 0}_{n-1}, \underbrace{\frac{1}{2}}_{1} \right)$$
$$x^1 := \left(\underbrace{1, 0 \dots, 0}_{n-1}, \underbrace{0}_{1} \right)$$

We define $\Delta_j := 0$ for $2 \leq j \leq n-1$, and notice that $\Delta_0 = \frac{1}{2}$ and $\Delta_1 = 0$.

With (3.64) we compute ω_Δ^{k-1}

$$\begin{aligned}
\omega_\Delta^{n-1} &= \binom{n-1}{0} \prod_{j=0}^{n-2} (1 - \Delta_j) + \binom{n-1}{1} \Delta_0 \\
&= 1 + \frac{n-1}{2} \\
&= \frac{n+1}{2}
\end{aligned}$$

Next, we compute the weights $\overline{\lambda}_0$ and $\overline{\lambda}_1$ using (3.65). Indeed,

$$\begin{aligned}\overline{\lambda}_0 &= \frac{1}{\omega_\Delta^{k-1}}(1-\Delta_1) = \frac{2}{n+1}\\ \overline{\lambda}_1 &= \frac{1}{\omega_\Delta^{k-1}}\Delta_0 = \frac{1}{n+1}.\end{aligned}$$

Finally, $\overline{\gamma}_1$ can be computed using (3.59d),

$$\begin{aligned}\overline{\gamma}_1 &= \binom{n-1}{0}\overline{\lambda}_0\Delta_0 + \binom{n-1}{1}\overline{\lambda}_1\Delta_1\\ &= \frac{1}{n+1}.\end{aligned}$$

Thus, $x^* := \frac{1}{n+1}\mathbf{1} \in \mathrm{SA}^{n-1}(P_n^K)$. Since $\sum_{i=1}^n x_i \geq 1$ is valid for the integer hull of P_n^K and it is violated by x^*, it follows that the SA rank of P_n^K is n. In other words, we recover the bound in [145].

3.9.5 Knapsack-Cover Inequalities

In this section, we apply our procedure to the knapsack-cover relaxation of the Minimum Knapsack problem, a stronger alternative to the standard LP relaxation discussed in the previous section. We begin by recalling the problem and introducing the knapsack-cover inequalities.

The knapsack cover problem consists of n items, each with a covering capacity $u_i \geq 0$ and a cost $c_i \geq 0$. Our aim is to select a subset of items of minimum cost whose total covering capacity is at least a given demand $D \geq 0$, i.e.,

$$\min\left\{\sum_{i=1}^n c_i x_i : \sum_{i=1}^n x_i u_i \geq D, x \in \{0,1\}^n\right\}.$$

The straightforward linear relaxation, where we replace the constraint $x \in \{0,1\}^n$ by $x \in [0,1]^n$, has an unbounded integrality gap. The knapsack-cover inequalities, introduced in [178] and later popularized by [179], correspond to a set of valid inequalities that reduced the integrality gap to 2, and are given by

$$\sum_{i \notin S} x_i \bar{u}_i(S) \geq D(S) \qquad \text{for all } S \subseteq [n],$$

where $D(S) := \max\{D - \sum_{i \in S} u_i, 0\}$ and $\bar{u}_i(S) := \min\{D(S), u_i\}$.

An important open problem, particularly relevant for approximation algorithms, is how to obtain relaxations with smaller integrality gaps. [174], [175] show that even after a sublinear number of rounds $k = o(n)$ of the Sherali-Adams hierarchy (as well as for the semidefinite strengthening LS_+ of the Lovász-Schrijver hierarchy), the integrality gap still remains 2. Later, [140] proves that even after $\Omega(\log(n)^{1-\varepsilon})$ (for any $\varepsilon > 0$) rounds of the stronger Sum-of-Squares (Lasserre) SDP hierarchy, the integrality gap remains 2.

In what follows, we apply the template developed in Section 3.9.2 and recover the same result for the Sherali-Adams relaxation as [174], [175]. Unlike their approach, which requires proposing the full vector $y \in \text{M}^k(P)$ from scratch, we just need to provide vectors that belong to P intersected with each of the $(n-k)$-dimensional faces of the hypercube; a significantly simpler task.

Consider the instance of knapsack cover where $u_i = 1$ and $c_i = 1$ for all $i \in [n]$, and $D = 1 + \frac{1}{n-1}$. The linear relaxation with knapsack-cover inequalities[3] for this instance is

$$\min\left\{\sum_{i=1}^{n} x_i : \sum_{i=1}^{n} x_i \geq 1 + \frac{1}{n-1}, \sum_{j \in [n] \setminus \{i\}} x_j \geq 1 \text{ and } 0 \leq x_i \leq 1 \text{ for all } i \in [n]\right\}.$$

Let P denote the feasible region of this linear program. We will show $\text{SA}^k(P)$ contains a solution x with $\sum_{i=1}^{n} x_i = 1 + O(\frac{k}{n})$. As the optimal integral solution of this LP has value 2, this shows that if $k = k(n)$ grows sublinearly on n (for example, $k = n^{1-\delta}$ for some $\delta > 0$) then $\sum_{i=1}^{n} x_i$ converges to 1, hence implying that the integrality gap remains 2.

Proposition 3.55. *For any $k \leq n/2$, there exists a vector $x \in \text{SA}^k(P)$ with value $\sum_{i=1}^{n} x_i = 1 + O(\frac{k}{n})$.*

Proof. Consider any number $k \leq n/2$. We obtain an explicit vector in $\text{SA}^k(P)$ by using the

[3] As Kurpisz et al. [140] observed, the inequality $\sum_{i=1}^{n} x_i \geq 1 + \frac{1}{n-1}$ is redundant in this formulation and thus it can be omitted, which yields the polytope studied by [174], [175].

construction from Section 3.9.2. We follow the notation from that section. Observe first that the polytope P is $\mathcal{S}_n$-invariant. Hence P is n-transitive, and therefore $(k+1)$-transitive. Observe that the following vectors belong to P:

$$\overline{x}^0 := \left(\underbrace{0,\ldots,0}_{k}, \underbrace{\frac{1}{n-k-1},\ldots,\frac{1}{n-k-1}}_{n-k}\right),$$
$$\overline{x}^1 := \left(\underbrace{1,0,\ldots,0}_{k}, \underbrace{\frac{1}{n-k},\ldots,\frac{1}{n-k}}_{n-k}\right),$$
$$\overline{x}^2 := (\underbrace{1,1,0,\ldots,0}_{k}, \underbrace{0,\ldots,0}_{n-k}).$$

Hence $\Delta_0 = \frac{1}{n-k-1}$, $\Delta_1 = \frac{1}{n-k}$ and $\Delta_\ell = 0$ for all $\ell = 2,\ldots,k$. Notice that $i = 2$.

Next, we compute ω_Δ^k following Step 3). Indeed,

$$\begin{aligned}
\omega_\Delta^k &= \sum_{\ell=0}^{2} \binom{k}{\ell} \prod_{j=0}^{\ell-1} \Delta_j \prod_{j=\ell+1}^{2} (1-\Delta_j) \\
&= \left(1 - \frac{1}{n-k}\right) + k\left(\frac{1}{n-k-1}\right) + \left(\frac{k(k-1)}{2}\right)\left(\frac{1}{n-k-1}\right)\left(\frac{1}{n-k}\right) \\
&= 1 - \frac{1}{n-k} + \frac{k}{n-k-1} + \frac{1}{2}\left(\frac{k}{n-k-1}\right)\left(\frac{k-1}{n-k}\right) \\
&= \frac{2(n-k-1)(n-k) - 2(n-k-1) + 2k(n-k) + k(k-1)}{2(n-k-1)(n-k)} \\
&= \frac{2n^2 - 2nk - 4n + k^2 + 3k + 2}{2(n-k-1)(n-k)}.
\end{aligned}$$

Then

$$\overline{\lambda}_0 = \frac{1}{\omega_\Delta^k}\prod_{i=1}^{k}(1-\Delta_i) = \frac{1}{\omega_\Delta^k}\left(1-\frac{1}{n-k}\right),$$
$$\overline{\lambda}_1 = \frac{1}{\omega_\Delta^k}\Delta_0\prod_{i=2}^{k}(1-\Delta_i) = \frac{1}{\omega_\Delta^k}\left(\frac{1}{n-k-1}\right),$$
$$\overline{\lambda}_2 = \frac{1}{\omega_\Delta^k}\Delta_0\Delta_1\prod_{i=3}^{k}(1-\Delta_i) = \frac{1}{\omega_\Delta^k}\left(\frac{1}{n-k-1}\right)\left(\frac{1}{n-k}\right),$$

and $\overline{\lambda_\ell} = 0$ for all $\ell \geq 3$.

Finally, the vector $x^* := \overline{\gamma}_1\mathbf{1}$ belongs to $\mathrm{SA}^k(P)$, where $\overline{\gamma}_1$ can be computed using (3.59c):

$$\begin{aligned}\overline{\gamma}_1 &= \sum_{\ell=1}^{2}\binom{k-1}{\ell-1}\overline{\lambda}_\ell \\ &= \frac{1}{\omega_\Delta^k}\left(\frac{1}{n-k-1}\right) + \frac{1}{\omega_\Delta^k}\left(\frac{1}{n-k-1}\right)\left(\frac{1}{n-k}\right)(k-1) \\ &= \frac{n-1}{\omega_\Delta^k(n-k-1)(n-k)}.\end{aligned}$$

Consequently, its objective value is

$$\begin{aligned}\sum_{i=1}^{n} x_i^* &= \frac{n(n-1)}{\omega_\Delta^k(n-k-1)(n-k)} \\ &= \frac{2n(n-1)}{2n^2-2nk-4n+k^2+3k+2} \\ &= \frac{1-(1/n)}{1+(1/2)(k/n)^2-(k/n)+(3/2)(k/n^2)-(2/n)+(1/n^2)} \\ &= \frac{1}{1-\Theta(k/n)} = 1+\Theta\left(k/n\right).\end{aligned}$$

$\square$

3.10 Conclusions

In this chapter, we studied the behavior of LP- and SDP-based hierarchies under symmetry, focusing on their ability to certify integer-emptiness of polytopes contained in the binary cube. We developed a unified framework that provides bounds on the number of iterations required by these hierarchies.

Our main result shows that symmetry imposes intrinsic limitations: the "more symmetric" P is, the larger is the number of rounds for which the hierarchies "collapse" – that is, they can either all certify integer-emptiness, or none do. This phenomenon provides a general explanation for the observed difficulty of strengthening relaxations in highly symmetric settings, where more sophisticated operators can perform no better than simpler ones.

A notable feature of our approach is that our results do not depend on any specific class of polytopes, relying only on symmetry. As a consequence, our framework recovers several results from the literature and yields new insights as direct corollaries. In addition, we established structural properties of symmetric operators and polytopes that are of independent interest.

Overall, our results identify a structural condition under which lift-and-project hierarchies are provably ineffective: high symmetry together with the property of intersecting every $(n-k)$-dimensional face of the hypercube.

Chapter 4: Conclusions and Future Work

In this dissertation, we have designed and analyzed convex relaxations for a fundamental problem in power systems, as well as for structured integer programs.

On the power systems side, we developed linear relaxations that are accurate, numerically stable, warm-startable, and scalable. These methods provide valuable information to system operators, enabling rapid assessment of the quality of operating points. Moreover, they push the limits of algorithmic development for power systems by making it possible to solve large-scale day-ahead problems, at fine temporal granularity, involving millions of variables and constraints. In addition, these relaxations show strong potential for computing nodal energy prices and for anticipating network congestion in multi-day-ahead settings. Their effectiveness stems from leveraging the underlying structure of the problem, with tightness achieved by preserving essential properties of AC-feasible solutions.

On the integer programming side, we provide structural contributions to the study of symmetry in optimization and new insights into the behavior of convex relaxations in symmetric settings. By parametrizing symmetry through the concept of k-transitivity, we show that symmetry is not an all-or-nothing property but rather admits different levels. This new perspective allows us to characterize when lift-and-project operators are effective or ineffective as proof systems for certifying integer-emptiness. Our work also introduces the notion of k-resistance, which provides a finer lens for analyzing integrality gaps and better reflects the behavior of modern solution methods such as branch-and-cut, namely, whether a given fractional face can be eliminated at a node of the search tree. Beyond advancing theoretical understanding, these results inform the design of stronger formulations and have direct implications for practitioners and solver developers.

By focusing on scalability and structure, this dissertation advances both domain-specific applications in power systems and the broader development of optimization methodology, highlighting the central role of carefully designed relaxations in large-scale optimization.

Future Directions: Optimization for the Next-Generatior Power Grid On the power systems side, key directions include the design of fast heuristics for large-scale multi-period ACOPF, the deployment of these methods in collaboration with system operators, and the extension of linearly constrained relaxations to more complex settings such as unit commitment, security constraints, and risk-aware models. On the theoretical and algorithmic fronts, important challenges include understanding the computational complexity of ACOPF, refining SDP-based techniques, strengthening relaxations through improved modeling choices – specially the representation of power losses – and better understanding why large-scale SOCP relaxations of ACOPF remain challenging for modern conic solvers. Despite the sparse and highly structured nature of the conic constraints considered in this dissertation, these relaxations can be substantially more difficult to solve in practice than their LP-based outer-approximations. Determining the extent to which this phenomenon is driven by numerical issues, algorithmic design, or solver implementation remains an important open question.

In electricity markets, a central direction is to move beyond empirical evaluation toward a principled understanding of *true prices*, potentially via higher-order convex relaxations, and how well computed prices approximate them. Another key direction is the study of price nonuniqueness, where different solvers may yield different prices for the same instance. Understanding and quantifying this phenomenon is essential for assessing the robustness of market outcomes. More broadly, these ideas can be extended to other nonconvex markets.

Future Directions: Symmetries in Optimization Extending the analysis to stronger hierarchies, such as the Lasserre (SoS) hierarchy, and to other classes of cutting-plane methods and closures is a natural next step. Such an extension will likely require new techniques. For example, for the cropped cube, Lasserre certifies infeasibility in at most $n - 1$ rounds [180], whereas SA and LS$_+$ require n rounds. This suggests that the relationship between symmetry and hierarchy rank is sensitive to hierarchy-specific features beyond those exhibited by SA and LS_+.

Understanding the role of *subsymmetries* and their impact on relaxations is particularly relevant in practice. From a structural perspective, further characterizing k-transitive polytopes and the evolution of the integrality gap in symmetric settings remains an important open problem. More broadly, these directions point toward the development of new algorithms that are robust to symmetry or capable of exploiting it constructively. Given the limitations of k-transitive symmetries, an important direction is to move beyond permutation actions on variables and capture more general forms of symmetry.

References

[1] National Academies of Sciences, Engineering, and Medicine, *Analytic Research Foundations for the Next-Generation Electric Grid*. Washington, DC: The National Academies Press, 2016.

[2] International Energy Agency, "World energy outlook 2024," IEA, Paris, Tech. Rep., 2024.

[3] M. B. Cain, R. P. O'Neill, and A. Castillo, "History of Optimal Power Flow and Formulations," *Federal Energy Regulatory Commission*, 2012.

[4] New York Independent System Operator, "Day-ahead scheduling manual," New York Independent System Operator (NYISO), Tech. Rep., 2023.

[5] F. Margot, "Symmetry in integer linear programming," in *50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art*, M. Jünger et al., Eds., Springer, 2009, pp. 647–686.

[6] M. Pfetsch and T. Rehn, "A computational comparison of symmetry handling methods for mixed integer programs," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 37–93, 2019.

[7] A. Schrijver, *Theory of Linear and Integer Programming*. New York: John Wiley & Sons, 1986.

[8] J. Lee and S. Leyffer, Eds., *Mixed Integer Nonlinear Programming* (The IMA Volumes in Mathematics and its Applications). New York, NY: Springer, 2012, vol. 154.

[9] P. M. Pardalos and S. A. Vavasis, "Quadratic programming with one negative eigenvalue is NP-hard," *Journal of Global Optimization*, vol. 1, no. 1, pp. 15–22, 1991.

[10] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[11] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2009.

[12] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming* (Graduate Texts in Mathematics). Springer International Publishing, 2014, vol. 271.

[13] D. Bienstock and M. Villagra, "Accurate and Warm-Startable Linear Cutting-Plane Relaxations for ACOPF," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 5024–5031.

[14] D. Bienstock and M. Villagra, “Accurate linear cutting-plane relaxations for ACOPF,” *Mathematical Programming Computation*, vol. 18, pp. 79–133, 2025.

[15] D. Bienstock and M. Villagra, *Advanced Cutting-Plane Algorithms for ACOPF*, arXiv:2510.21698, 2025.

[16] M. Romero, F. Verástegui, and M. Villagra, “A linear and scalable cutting-plane algorithm for electricity pricing,” in *2025 IEEE Power & Energy Society General Meeting (PESGM)*, 2025, pp. 1–5.

[17] I. Hiskens and R. Davy, “Exploring the Power Flow Solution Space Boundary,” *IEEE Transactions on Power Systems*, vol. 16, pp. 389–395, 2001.

[18] W. A. Bukhsh, A. Grothey, K. I. M. McKinnon, and P. A. Trodden, “Local Solutions of the Optimal Power Flow Problem,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4780–4788, 2013.

[19] A. Verma, “Power Grid Security Analysis : An Optimization Approach,” Ph.D. Thesis, Columbia University, 2009.

[20] D. Bienstock and A. Verma, “Strong NP-hardness of AC power flows feasibility,” *Operations Research Letters*, vol. 47, pp. 494–501, 2019.

[21] K. Lehmann, A. Grastien, and P. Van Hentenryck, “AC-Feasibility on Tree Networks is NP-Hard,” *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 798–801, 2016.

[22] R. A. Jabr, “Radial Distribution Load Flow using Conic Programming,” *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, 2006.

[23] B. Kocuk, S. Dey, and X. Sun, “Strong SOCP Relaxations for the Optimal Power Flow Problem,” *Operations Research*, vol. 64, no. 6, pp. 1177–1196, 2016.

[24] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, “The QC Relaxation: A Theoretical and Computational Study on Optimal Power Flow,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3008–3018, 2016.

[25] B. Kocuk, S. S. Dey, and X. A. Sun, “Matrix Minor Reformulation and SOCP-based Spatial Branch-and-Cut Method for the AC Optimal Power Flow Problem,” *Mathematical Programming Computation*, vol. 10, no. 4, pp. 557–596, 2018.

[26] A. Oustry, C. D’Ambrosio, L. Liberti, and M. Ruiz, “Certified and accurate SDP bounds for the ACOPF problem,” *Electric Power Systems Research*, vol. 212, p. 108 278, 2022.

[27] M. Bichler and J. Knorr, “Getting prices right on electricity spot markets: On the economic impact of advanced power flow models,” *Energy Economics*, vol. 126, 2023.

[28] R. O'Neill, P. Sotkiewicz, B. Hobbs, M. Rothkopf, and W. Stewart, "Efficient Market-Clearing Prices in Markets with Nonconvexities," *European Journal of Operational Research*, pp. 269–285, 2005.

[29] P. Gribik, W. Hogan, and S. Pope, "Market-clearing electricity prices and energy uplift," Harvard Electricity Policy Group, Cambridge, MA, Tech. Rep., 2007.

[30] M. Bichler and J. Knörr, "Getting Prices right on Electricity Spot Markets: On the Economic Impact of Advanced Power Flow models," *Energy Economics*, vol. 126, 2023.

[31] D. Bienstock and G. Muñoz, "On Linear Relaxations of OPF Problems," *arXiv:1411.1120*, pp. 1–15, 2014.

[32] D. Bienstock, *Electrical Transmission System Cascades and Vulnerability, an Operations Research viewpoint*. Society for Industrial and Applied Mathematics, 2015.

[33] A. R. Bergen and V. Vittal, *Power Systems Analysis*. Prentice-Hall, 1999.

[34] B. Stott, J Jardim, and O. Alsac, "DC Power Flow Revisited," *IEEE Transactions on Power Systems*, vol. 24, 2009.

[35] D. Molzahn and I. Hiskens, "A Survey of Relaxations and Approximations of the Power Flow Equations," *Foundations and Trends in Electric Energy Systems*, vol. 4, pp. 1–221, 2019.

[36] D. Bienstock, M. Escobar, C. Gentile, and L. Liberti, "Mathematical Programming Formulations for the Alternating Current Optimal Power Flow Problem," *Annals of Operations Research*, vol. 314, no. 1, pp. 277–315, 2022.

[37] A. Castillo, P. Lipka, J. Watson, S. Oren, and R. O'Neill, "A Successive Linear Programming approach to solving the IV-ACOPF," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2752–2763, Jul. 2016.

[38] K. Hedman, S. Oren, and R. O'Neill, "A Review of Transmission Switching and Network Topology Optimization," *IEEE Power and Energy Society General Meeting*, pp. 1–7, 2011.

[39] J. Holzer et al., "Grid Optimization Competition Challenge 3 Problem Formulation," Pacific Northwest National Laboratory, Tech. Rep., 2023.

[40] G. McCormick, "Computability of Global Solutions to Factorable Nonconvex programs: Part I—Convex underestimating problems," *Math. Program.*, vol. 10, pp. 147–175, 1976.

[41] D. Bienstock and G. Muñoz, "Approximate method for AC Transmission Switching based on a simple relaxation for ACOPF problems," in *2015 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2015, pp. 1–5.

[42] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "Network Flow and Copper Plate Relaxations for AC Transmission Systems," in *Power Systems Computation Conference (PSCC)*, 2016, pp. 1–8.

[43] J. A. Taylor and F. S. Hover, "Linear Relaxations for Transmission System Planning," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2533–2538, 2011.

[44] F. Glover, "Improved Linear Integer Programming Formulations of Nonlinear Integer problems," *Mgt. Science*, vol. 22, pp. 455–460, 1975.

[45] S. Mhanna, G. Verbic, and A. C. Chapman, "Tight LP Approximations for the Optimal Power Flow problem," in *2016 Power Systems Computation Conference (PSCC)*, IEEE, 2016, pp. 1–7.

[46] A. Ben-Tal and A. Nemirovski, "On Polyhedral Approximations of the Second-Order Cone," *Mathematics of Operations Research*, vol. 26, no. 2, pp. 193–205, 2001.

[47] S. Mhanna and P. Mancarella, "An Exact Sequential Linear Programming Algorithm for the Optimal Power Flow Problem," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 666–679, 2022.

[48] J. E. J. Kelley, "The Cutting-Plane Method for Solving Convex Programs," *Journal of the Society for Industrial and Applied Mathematics*, pp. 703–712, 1960.

[49] K. Baker, "Solutions of DC OPF Are Never AC Feasible," in *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, Association for Computing Machinery, 2021, 264–268.

[50] M. Farivar, C. R. Clarke, S. H. Low, and K. M. Chandy, "Inverter VAR Control for Distribution Systems with Renewables," *IEEE International Conference on Smart Grid Communications*, pp. 457–462, 2011.

[51] G. Qiu, M. Tanneau, and P. Van Hentenryck, "Dual Conic Proxies For Ac Optimal Power Flow," *Power Systems Computation Conference (PSCC)*, 2024.

[52] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite Programming for Optimal Power Flow Problems," *International Journal of Electrical Power and Energy Systems*, vol. 30, pp. 383–392, 2008.

[53] N. Shor, "Quadratic Optimization Problems," *Soviet Journal of Computer and Systems Sciences*, vol. 25, pp. 1–11, 1987.

[54] L. Lavaei and S. H. Low, "Zero Duality Gap in Optimal Power Flow Problem," *IEEE Transactions on Power Systems*, vol. 27, pp. 92–107, 2012.

[55] S. Bose, D. F. Gayme, S. H. Low, and K. M. Chandy, "Optimal power flow over tree networks," in *Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2011, pp. 1342–1348.

[56] S. Bose, S. H. Low, and K. M. Chandy, "Equivalence of Branch Flow and Bus Injection Models," in *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Allerton House, University of Illinois at Urbana–Champaign, USA, 2012, pp. 1–7.

[57] 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*, vol. 28, no. 4, pp. 3987–3998, 2013.

[58] M. Andersen, A. Hansson, and L. Vandenberghe, "Reduced-Complexity Semidefinite Relaxations of Optimal Power Flow Problems," *IEEE Transactions on Power Systems*, vol. 29, pp. 1855–1863, 2014.

[59] A. Eltved, J. Dahl, and M. S. Andersen, "On the Robustness and Scalability of Semidefinite Relaxation for Optimal Power Flow Problems," *Optimization and Engineering*, vol. 21, pp. 375–392, 2020.

[60] J. Sliwak, E. D. Andersen, M. Anjos, L. L., and E. Traversi, "A Clique Merging Algorithm to Solve Semidefinite Relaxations of Optimal Power Flow Problems," *IEEE Transactions on Power Systems*, vol. 36, pp. 1641–1644, 2021.

[61] D. K. Molzahn and I. A. Hiskens, "Sparsity-exploiting moment-based relaxations of the optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3168–3180, 2015.

[62] C. Josz, J. Maeght, P. Panciatici, and J. C. Gilbert, "Application of the moment-SOS approach to global optimization of the OPF problem," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 463–470, 2015.

[63] C. Josz and D. K. Molzahn, "Lasserre hierarchy for large scale polynomial optimization in real and complex variables," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1017–1048, 2018.

[64] H. Hijazi, C. Coffrin, and P. Van Hentenryck, "Polynomial SDP Cuts for Optimal Power Flow," in *Proceedings of the 2016 Power Systems Computation Conference (PSCC)*, Genoa, Italy, 2016, pp. 1–7.

[65] F. Zohrizadeh, C. Josz, M. Jin, R. Madani, J. Lavaei, and S. Sojoudi, "A Survey on Conic Relaxations of Optimal Power Flow problem," *European Journal of Operational Research*, vol. 287, no. 2, pp. 391–409, 2020.

[66] A. Lorca and X. Sun, "The Adaptive Robust Multi-Period Alternating Current Optimal Power Flow Problem," *IEEE Transactions on Power Systems*, vol. 33, pp. 1993–2003, 2017.

[67] G. E. Constante-Flores, A. J. Conejo, and F. Qiu, "Ac Network-Constrained Unit Commitment via Relaxation and Decomposition," *IEEE Transactions on Power Systems*, vol. 37, pp. 2187–2196, 2022.

[68] Y. Bai, H. Zhong, Q. Xia, C. Kang, and L. Xie, "A Decomposition Method for Network-constrained Unit Commitment with AC Power Flow Constraints," *Energy*, vol. 88, 2015.

[69] R. Quan, J. Jian, and Y. Mu, "Tighter Relaxation Method for Unit Commitment based on Second-order Cone Programming and Valid Inequalities," *Electrical Power and Energy Systems*, vol. 55, 2014.

[70] J. Liu, C. Laird, J. Scott, J. Watson, and A. Castillo, "Global Solution Strategies for the Network-Constrained Unit Commitment Problem with AC Transmission Constraints," *IEEE Transactions on Power Systems*, vol. 34, 2018.

[71] F. Zohrizadeh, M. Kheirandishfard, A. Nasir, and R. Madani, "Sequential Relaxation of Unit Commitment with AC Transmission Constraints," *IEEE Conference on Decision and Control*, pp. 2408–2413, 2018.

[72] Z. Yang, H. Zhong, A. Bose, T. Zheng, Q. Xia, and C. Kang, "A Linearized OPF model with Reactive Power and Voltage Magnitude: A Pathway to Improve the MW-Only DC OPF," *IEEE Transactions on Power Systems*, vol. 33, pp. 1734–1745, 2018.

[73] M. Alizadeh and F. Capitanescu, "A Tractable Linearization-Based Approximated Solution Methodology to Stochastic Multi-Period AC Security-Constrained Optimal Power Flow," *IEEE Transactions on Power Systems*, vol. 38, pp. 5896–5908, 2023.

[74] C. Coffrin and P. Van Hentenryck, "A Linear-Programming Approximation of AC Power Flows," *INFORMS Journal on Computing*, vol. 26, pp. 718–734, 2014.

[75] A. Castillo, C. Laird, J. Silva-Monroy C.A. Watson, and R. O'Neill, "The unit commitment problem with ac optimal power flow constraints," *IEEE Transactions on Power Systems*, vol. 31, 2016.

[76] A. Castillo, P. Lipka, J. Watson, S. Oren, and R. O'Neill, "A Successive Linear Programming Approach to Solving the IV-ACOPF," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2752–2763, 2016.

[77] N. Alguacil and A. Conejo, "Multiperiod Optimal Power Flow Using Benders Decomposition," *IEEE Transactions on Power Systems*, vol. 15, pp. 196–201, 2000.

[78] R. Byrd, J. Nocedal, and R. Waltz, "Knitro: An Integrated Package for Nonlinear Optimization," in *Large-Scale Nonlinear Optimization*, vol. 83, Springer, 2006, pp. 35–59.

[79] A. Wächter and L. T. Biegler, "On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming," *Mathematical Programming*, vol. 106, pp. 25–57, 2006.

[80] B. Subhonmesh, S. Low, and K. Chandy, "Equivalence of Branch Flow and Bus Injection models," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012, pp. 1893–1899.

[81] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "Network Flow and Copper Plate Relaxations for AC Transmission Systems," *arXiv:1506.05202*, 2015.

[82] R. Zimmerman, C. Murillo-Sanchez, and D. Gan, "Matpower, A MATLAB Power System Simulation Package," *IEE Trans. Power Sys.*, vol. 26, pp. 12–19, 2011.

[83] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Pearson, 1993.

[84] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "Distflow Extensions for AC Transmission Systems," *arXiv:1411.1120*, 2015.

[85] E. Klotz, "Identification, Assessment, and Correction of Ill-Conditioning and Numerical Instability in Linear and Integer Programs," *Bridging Data and Decisions*, pp. 54–108, 2014.

[86] A. Schrijver, *Theory of Linear and Integer Programming*. Wiley, 1998.

[87] F. Serrano, R. Schwarz, and A. Gleixner, "On the relation between the extended supporting hyperplane algorithm and Kelley's cutting plane algorithm," *J. Global Optimization*, pp. 161–179, 2020.

[88] N. Higham, *Accuracy and Stability of Numerical Algorithms*. SIAM, 1996.

[89] L. Gurobi Optimization, *Gurobi Optimizer Reference Manual*, 2023.

[90] D. L. Applegate, W. Cook, S. Dash, and D. Espinoza, "Exact Solutions to Linear Programming problems," *Operations Research Letters*, vol. 35, 693–699, 2007.

[91] A. Gleixner, D. Steffy, and K. Wolter, "Iterative Refinement for Linear Programming," *INFORMS Journal on Computing*, vol. 28, pp. 449–464, 2016.

[92] S. Vavasis, "Quadratic Programming is in NP," *Information Processing Letters*, vol. 36, pp. 73–77, 1990.

[93] D. Bienstock, A. Del Pia, and R. Hildebrand, “Complexity, Exactness, and Rationality in Polynomial Optimization,” *Mathematical Programming*, vol. 197, 2023.

[94] M. Farivar and S. Low, “Branch Flow Model: Relaxations and Convexification—Part I,” *IEEE Transactions on Power Systems*, vol. 28, pp. 2554–2564, 2013.

[95] M. ApS, *Mosek for AMPL User’s Guide*, 2024.

[96] R. Fourer, D. Gay, and B. Kernighan, “AMPL: A Mathematical Programing Language,” *Algorithms and Model Formulations in Mathematical Programming*, vol. 51, 1989.

[97] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, “AC Power Flow Data in MATPOWER and QCQP Format: iTesla, RTE Snapshots, and PEGASE,” *arXiv:1603.01533*, 2016.

[98] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, “Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive and corrective actions,” *IEEE Trans. on Power Systems*, pp. 4909–4917, 2013.

[99] A. Birchfield, T. Xu, K. Gegner, K. Shetye, and T. Overbye, “Grid structural characteristics as validation criteria for synthetic networks,” *IEEE Transactions on Power Systems*, vol. 32, pp. 3258–3265, 2017.

[100] A. Birchfield, T. Xu, and T. Overbye, “Power flow convergence and reactive power planning in the creation of large synthetic grids,” *IEEE Transactions on Power Systems*, vol. 33, pp. 6667–6674, 2018.

[101] S. Babaeinejadsarookolaee et al., “The power grid library for benchmarking ac optimal power flow algorithms,” *arXiv:1908.02788v2*, 2021.

[102] A. S. Xavier, A. M. Kazachkov, O. Yurdakul, J. He, and F. Qiu, *UnitCommitment.jl: A Julia/JuMP Optimization Package for Security-Constrained Unit Commitment (Version 0.4)*, Zenodo, 2024.

[103] H. Li, J. Yeo, A. Bornsheuer, and T. Overbye, “The Creation and Validation of Load Time Series for Synthetic Electric Power Systems,” *IEEE Transactions on Power Systems*, vol. 36, pp. 961–969, 2021.

[104] M. X. Goemans and D. P. Williamson, “Approximation Algorithms for Max-3-Cut and Other Problems via Complex Semidefinite Programming,” *Journal of Computer and System Sciences*, vol. 68, no. 2, pp. 442–470, 2004.

[105] J. A. Taylor, *Convex Optimization of Power Systems*. Cambridge, UK: Cambridge University Press, 2015.

[106] C. Josz and D. Molzahn, “Moment/Sum-of-Squares Hierarchy for Complex Polynomial Optimization,” *arxiv:1508.02068v2*, 2015.

[107] M. Fukuda, M. Kojima, K. Murota, and K. Nakata, “Exploiting Sparsity in Semidefinite Programming via Matrix Completion I: General Framework,” *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 647–674, 2000.

[108] A. Qualizza, P. Belotti, and F. Margot, “Linear Programming Relaxations of Quadratically Constrained Quadratic Programs,” in *Mixed Integer Nonlinear Programming*, ser. The IMA Volumes in Mathematics and its Applications, vol. 154, New York: Springer, 2012, pp. 407–426.

[109] S. S. Dey, A. Kazachkov, A. Lodi, and G. Muñoz, “Cutting Plane Generation through Sparse Principal Component Analysis,” *SIAM Journal on Optimization*, vol. 32, no. 2, pp. 1319–1347, 2022.

[110] P. J. Goulart, Y. Nakatsukasa, and N. Rontsis, “Accuracy of Approximate Projection to the Semidefinite Cone,” *Linear Algebra and its Applications*, vol. 594, pp. 177–192, 2020.

[111] J. Holzer et al., “Grid Optimization Competition Challenge 2 Problem Formulation,” Pacific Northwest National Laboratory, Tech. Rep., 2021.

[112] S. Bikhchandani and J. Ostroy, “The Package Assignment Model,” *Journal of Economic Theory*, pp. 377–406, 2002.

[113] J. Mays, “Sequential Pricing of Electricity,” *Energy Economics*, vol. 37, 2024.

[114] G. Liberopoulos and P. Andrianesis, “Critical Review of Pricing Schemes in Markets with Non-Convex Costs,” *Operations Research*, vol. 64, no. 1, pp. 17–31, 2016.

[115] W. W. Hogan and B. J. Ring, “On Minimum-Uplift Pricing for Electricity Markets,” Electricity Policy Group, Tech. Rep., 2003.

[116] B. Hua and R. Baldick, “A convex primal formulation for convex hull pricing,” *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3814–3823, 2017.

[117] Y. Faenza, V. Verdugo, J. Verschae, and M. Villagra, *Linear Programming Hierarchies Collapse under Symmetry*, arXiv:2511.07766, 2025.

[118] Y. Faenza, V. Verdugo, J. Verschae, and M. Villagra, “Linear programming hierarchies collapse under symmetry,” in *Integer Programming and Combinatorial Optimization*, S. S. Dey, M. Di Summa, and D. Salvagnin, Eds., Cham: Springer Nature Switzerland, 2026.

[119] J. R. Bunch and L. Kaufman, “Some Stable Methods for Calculating Inertia and Solving Symmetric Linear Systems,” *Mathematics of Computation*, vol. 31, no. 137, pp. 163–179, 1977.

[120] D. Fulkerson, G. L. Nemhauser, and L. E. Trotter, “Two computationally difficult set covering problems that arise in computing the 1-width of incidence matrices of steiner triple systems,” in *Approaches to Integer Programming*, Springer, 2009, pp. 72–81.

[121] F. Margot, “Symmetry in integer linear programming,” in *50 Years of Integer Programming 1958-2008*, Springer Berlin Heidelberg, 2010, pp. 647–686.

[122] J. Ostrowski, J. Linderoth, F. Rossi, and S. Smriglio, “Solving large steiner triple covering problems,” *Operations Research Letters*, vol. 39, no. 2, pp. 127–131, 2011.

[123] J. Ostrowski, “Using symmetry to optimize over the Sherali-Adams relaxation,” *Mathematical Programming Computation*, vol. 6, no. 4, pp. 405–428, 2014.

[124] M. Pfetsch and T. Rehn, “A computational comparison of symmetry handling methods for mixed integer programs,” *Mathematical Programming Computation*, vol. 11, no. 1, pp. 37–93, 2019.

[125] M. Yannakakis, “Expressing combinatorial optimization problems by linear programs,” *Journal of Computer and System Sciences*, vol. 43, pp. 441–466, 1991.

[126] A. Ben-Tal and A. Nemirovski, “On polyhedral approximations of the second-order cone,” *Mathematics of Operations Research*, vol. 26, no. 2, pp. 193–205, 2001.

[127] D. Grigoriev, “Complexity of positivstellensatz proofs for the knapsack,” *Computational Complexity*, vol. 10, pp. 139–154, 2001.

[128] M. Laurent, “Lower Bound for the Number of Iterations in Semidefinite Hierarchies for the Cut Polytope,” *Mathematics of Operations Research*, vol. 28, no. 4, pp. 871–883, 2003.

[129] K. Gatermann and P. A. Parrilo, “Symmetry groups, semidefinite programs, and sums of squares,” *Journal of Pure and Applied Algebra*, vol. 192, no. 1–3, pp. 95–128, 2004.

[130] A. Kurpisz, S. Leppänen, and M. Mastrolilli, “Sum-of-squares rank upper bounds for matching problems,” *Journal of Combinatorial Optimization*, vol. 36, no. 3, pp. 831–844, 2018.

[131] C. Mathieu and A. Sinclair, “Sherali-Adams relaxations of the matching polytope,” in *ACM Symposium on Theory of Computing (STOC)*, 2009, pp. 293–302.

[132] A. R. Karlin, C. Mathieu, and C. T. Nguyen, "Integrality gaps of linear and semi-definite programming relaxations for knapsack," in *Integer Programming and Combinatorial Optimization (IPCO)*, 2011, pp. 301–314.

[133] V. Cohen-Addad, T. Mömke, and V. Verdugo, "A 2-approximation for the bounded treewidth sparsest cut problem in FPT time," *Mathematical Programming*, vol. 206, no. 1, pp. 479–495, 2024.

[134] A. Gupta, K. Talwar, and D. Witmer, "Sparsest cut on bounded treewidth graphs: Algorithms and hardness results," in *ACM Symposium on Theory of Computing (STOC)*, 2013, pp. 281–290.

[135] A. Kurpisz, M. Mastrolilli, C. Mathieu, T. Mömke, V. Verdugo, and A. Wiese, "Semidefinite and linear programming integrality gaps for scheduling identical machines," *Mathematical Programming*, vol. 172, no. 1, pp. 231–248, 2018.

[136] V. Verdugo, J. Verschae, and A. Wiese, "Breaking symmetries to rescue sum of squares in the case of makespan scheduling," *Mathematical Programming*, vol. 183, no. 1, pp. 583–618, 2020.

[137] K. Georgiou, A. Magen, T. Pitassi, and I. Tourlakis, "Integrality gaps of 2-o(1) for vertex cover SDPs in the Lovász–Schrijver hierarchy," *SIAM Journal on Computing*, vol. 39, no. 8, pp. 3553–3570, 2010.

[138] B. Barak, S. Hopkins, J. Kelner, P. K. Kothari, A. Moitra, and A. Potechin, "A nearly tight sum-of-squares lower bound for the planted clique problem," *SIAM Journal on Computing*, vol. 48, no. 2, pp. 687–735, 2019.

[139] R. Meka, A. Potechin, and A. Wigderson, "Sum-of-squares lower bounds for Planted Clique," in *ACM Symposium on Theory of Computing (STOC)*, 2015, pp. 87–96.

[140] A. Kurpisz, S. Leppänen, and M. Mastrolilli, "Sum-of-squares hierarchy lower bounds for symmetric formulations," *Mathematical Programming*, vol. 182, no. 1, pp. 369–397, 2020.

[141] N. E. Aguilera, S. M. Bianchi, and G. L. Nasini, "Lift and project relaxations for the matching and related polytopes," *Discrete Applied Mathematics*, vol. 134, no. 1, pp. 193–212, 2004.

[142] W. Cook and S. Dash, "On the matrix-cut rank of polyhedra," *Mathematics of Operations Research*, vol. 26, no. 1, pp. 19–30, 2001.

[143] M. Goemans and L. Tunçel, "When Does the Positive Semidefiniteness Constraint Help in Lifting Procedures," *Mathematics of Operations Research*, vol. 26, no. 4, pp. 769–815, 2001.

[144] T. Stephen and L. Tunçel, "On a Representation of the Matching Polytope Via Semidefinite Liftings," *Mathematics of Operations Research*, pp. 1–7, 1999.

[145] M. Laurent, "A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming," *Mathematics of Operations Research*, vol. 28, no. 3, pp. 470–496, 2003.

[146] J. Dixon and B. Mortimer, *Permutation Groups*. Springer, 1996.

[147] E. Balas, S. Ceria, and G. Cornuéjols, "A lift-and-project cutting plane algorithm for mixed 0–1 programs," *Mathematical Programming*, vol. 58, no. 1, pp. 295–324, 1993.

[148] L. Lovász and A. Schrijver, "Cones of matrices and set-functions and 0-1 optimization," *SIAM Journal on Optimization*, vol. 1, no. 2, pp. 166–190, 1991.

[149] N. Fleming, P. Kothari, and T. Pitassi, "Semialgebraic Proofs and Efficient Algorithm Design," *Foundations and Trends in Theoretical Computer Science*, vol. 14, no. 1-2, pp. 1–221, 2019.

[150] H. Sherali and W. Adams, "A hierarchy of relaxations between the continuous and convex hull representations for zero one programming problems," *SIAM Journal on Discrete Mathematics*, vol. 3, no. 3, pp. 411–430, 1990.

[151] H. Sherali and W. Adams, "A hierarchy of relaxations and convex hull characterizations for mixed-integer zero—one programming problem," *Discrete Applied Mathematics*, vol. 52, no. 1, pp. 83–106, 1994.

[152] B. Balas, S. Ceria, and G. Cornuéjols, "A lift-and-project cutting plane algorithm for mixed 0–1 programs," *Mathematical Programming*, 295–324, 1993.

[153] E. Balas, "Disjunctive programming: Properties of the convex hull of feasible points," *Discrete Applied Mathematics*, vol. 89, no. 1–3, pp. 1–44, 1998, Originally GSIA Management Science Research Report MSRR 348, Carnegie Mellon University, 1974.

[154] T Rothvoss, "The Lasserre Hierarchy in Approximation Algorithms," Lecture Notes for the MAPSP 2013 Tutorial (preliminary version), 1–25, 2013.

[155] A. J. Geyer, D. A. Bulutoglu, and K. J. Ryan, "Finding the symmetry group of an LP with equality constraints and its application to classifying orthogonal arrays," *Discrete Optimization*, vol. 28, pp. 89–102, 2018.

[156] L. Liberti, "Reformulations in Mathematical Programming: Automatic Symmetry Detection and Exploitation," *Mathematical Programming Series A*, vol. 131, pp. 273 –304, 2012.

[157] H. Derksen and G. Kemper, *Computational Invariant Theory* (Encyclopaedia of Mathematical Sciences). Berlin: Springer-Verlag, 2002, vol. 130.

[158] P. J. Cameron, *Permutation Groups* (London Mathematical Society Student Texts). Cambridge University Press, 1999, vol. 45.

[159] J. B. Lasserre, “Global optimization with polynomials and the problem of moments,” *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[160] Y. Faenza and V. Kaibel, “Extended formulations for packing and partitioning orbitopes,” *Mathematics of Operations Research*, vol. 34, no. 3, pp. 686–697, 2009.

[161] V. Kaibel, M. Peinhardt, and M. E. Pfetsch, “Orbitopal fixing,” *Discrete Optimization*, vol. 8, no. 4, pp. 595–610, 2011.

[162] V. Kaibel and M. Pfetsch, “Packing and partitioning orbitopes,” *Mathematical Programming*, vol. 114, no. 1, pp. 1–36, 2008.

[163] J. van Doornmalen and C. Hojny, “A unified framework for symmetry handling,” *Mathematical Programming*, vol. 212, no. 1, pp. 217–271, 2025.

[164] G. Cornuéjols and Y. Li, “Elementary closures for integer programs,” *Operations Research Letters*, vol. 28, no. 1, pp. 1–8, 2001.

[165] D. Gorenstein, R. Lyons, and R. Solomon, *The Classification of the Finite Simple Groups*. Providence, RI: American Mathematical Society, 1994–, Mathematical Surveys and Monographs, multiple volumes.

[166] L. Babai, “Graph Isomorphism in Quasipolynomial Time,” *arXiv preprint arXiv:1512.03547v2*, 2016.

[167] L. Babai, “Graph Isomorphism in Quasipolynomial Time,” in *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, 2016, pp. 684–697.

[168] A. Seress, *Permutation Group Algorithms* (Cambridge Tracts in Mathematics). Cambridge University Press, 2003, vol. 152.

[169] T. A. Feo and M. G. Resende, “A probabilistic heuristic for a computationally difficult set covering problem,” *Operations Research Letters*, vol. 8, no. 2, pp. 67–71, 1989.

[170] C. Mannino and A. Sassano, “Solving hard set covering problems,” *Operations Research Letters*, vol. 18, no. 1, pp. 1–6, 1995.

[171] C. J. Colbourn and A. Rosa, *Triple Systems* (Oxford Mathematical Monographs). Oxford: Clarendon Press, 1999.

[172] M. Hall, "Automorphisms of steiner triple systems," *IBM Journal of Research and Development*, vol. 4, no. 5, pp. 460–472, 1960.

[173] A. Kurpisz, S. Leppänen, and M. Mastrolilli, "Sum-of-Squares Hierarchy Lower Bounds for Symmetric Formulations," in *Integer Programming and Combinatorial Optimization (IPCO)*, ser. Lecture Notes in Computer Science, 2016, pp. 362–374.

[174] D. Bienstock and M. Zuckerberg, "Approximate fixed-rank closures of covering problems," *Mathematical Programming*, vol. 105, no. 1, pp. 9–27, 2006.

[175] D. Bienstock and M. Zuckerberg, "Subset Algebra Lift Operators for 0-1 Integer Programming," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 63–95, 2004.

[176] E. Gawrilow and M. Joswig, "Polymake: A framework for analyzing convex polytopes," in *Polytopes—Combinatorics and Computation*, ser. DMV Seminar, G. Kalai and G. M. Ziegler, Eds., vol. 29, Basel: Birkhäuser, 2000, pp. 43–73.

[177] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Professional, 1994.

[178] L. A. Wolsey, "Faces for a linear inequality in 0-1 variables," *Mathematical Programming*, vol. 8, no. 1, pp. 165–178, 1975.

[179] R. D. Carr, L. K. Fleischer, V. J. Leung, and C. A. Phillips, "Strengthening integrality gaps for capacitated network design and covering problems," in *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2000, pp. 106–115.

[180] Y. H. Au and L. Tunçel, "Elementary polytopes with high lift-and-project ranks for strong positive semidefinite operators," *Discrete Optimization*, vol. 27, pp. 103–129, 2018.

Appendix A: Additional Proofs and Computational Results

A.1 Reformulating Maximum Angle Difference Constraints as Thermal Limits

We show that maximum angle-difference constraints can be modeled as thermal limit constraints. Consider an arbitrary branch $\{k, m\} \in \mathcal{E}$ and let its admittance matrix be given by

$$Y_{km} := \begin{pmatrix} \left(y + \frac{y^{sh}}{2}\right)\frac{1}{\tau^2} & -y\frac{1}{\tau e^{-j\sigma}} \\ -y\frac{1}{\tau e^{j\sigma}} & y + \frac{y^{sh}}{2} \end{pmatrix}$$

where $y = g + jb$ denotes its series admittance, $y^{sh} = g^{sh} + jb^{sh}$ its shunt admittance, $\tau > 0$ the magnitude of the transformer ratio, and σ its phase shift angle. Recall that complex power injected at bus k into the branch $\{k, m\}$ is given by $S_{km} = V_k I_{km}^*$. Since $S_{km} = P_{km} + jQ_{km}$, we have that $P_{km}^2 + Q_{km}^2 = (V_k I_{km}^*)(V_k I_{km}^*)^* = |V_k|^2 |I_{km}|^2$. In what follows, we will isolate the dependence of $|I_{km}|^2$ on θ_{km}. First, we define

$$f(|V_k|, |V_m|) := \frac{(g^2 + b^2 + gg^{sh} + bb^{sh} + (g^{sh2} + b^{sh2})/4)}{\tau^4}|V_k|^2 + \frac{(g^2 + b^2)}{\tau^2}|V_m|^2.$$

Notice that f is a convex quadratic function since $g, -b, \tau > 0$, $g^{sh}, b^{sh} \geq 0$, and $b^{sh} \ll |b|$, see Section 2.2. Using (A.10), we can write

$$\begin{aligned} |I_{km}|^2 &= f(|V_k|, |V_m|) + \psi|V_k||V_m|\cos(\theta_{km} - \sigma) + \omega|V_k|V_m|\sin(\theta_{km} - \sigma) \\ &= f(|V_k|, |V_m|) + |V_k||V_m|\left(\psi\cos(\theta_{km} - \sigma) + \omega\sin(\theta_{km} - \sigma)\right) \end{aligned}$$

where $\psi := \frac{1}{\tau^3}(-2g^2 - 2b^2 - gg^{sh} - bb^{sh})$ and $\omega := (gb^{sh} - bg^{sh})$. Given the assumptions on the branch parameters, $\psi < 0$ and $\omega \geq 0$. Therefore, assuming $|\theta_{km} - \sigma| \leq \frac{\pi}{2}$, which is typically satisfied in practice, as both phase shifts σ and angle differences are small (often a few degrees), $|I_{km}|^2$ is maximized over θ_{km} at $\theta_{km} = \Delta_{\{k,m\}} \leq \frac{\pi}{2}$. This implies that

$$\begin{aligned} P_{km}^2 + Q_{km}^2 &\leq |V_k|^2 f(|V_k|, |V_m|) + |V_k|^3 |V_m| \left(\psi \cos(\Delta_{\{k,m\}} - \sigma) + \omega \sin(\Delta_{\{k,m\}} - \sigma)\right) \\ &=: S(|V_k|, |V_m|) \end{aligned}$$

Since S is monotone increasing in $|V_k|$ and $|V_m|$, defining $\overline{S} := \min\{U_{\{k,m\}}, S(V_k^{\max}, V_m^{\max})\}$, yields the valid thermal limit constraint $P_{km}^2 + Q_{km}^2 \leq \overline{S}$.

A.2 i2 SOCP Stronger than Jabr SOCP

In what follows, we exhibit an example of a 2-bus network with reasonable transmission-line parameters for which a feasible point of the Jabr SOCP relaxation is not feasible for the i2 SOCP.[1]

Let k denote the 'from' bus and m the 'to' bus. Recall that the linearized power flow definitions (2.13) for a line (k, m) are given by:

$$P_{km} = G_{kk} v_k^{(2)} + G_{km} c_{km} + B_{km} s_{km} \tag{A.1a}$$

$$P_{mk} = G_{mm} v_m^{(2)} + G_{mk} c_{km} - B_{mk} s_{km} \tag{A.1b}$$

$$Q_{km} = -B_{kk} v_k^{(2)} - B_{km} c_{km} + G_{km} s_{km} \tag{A.1c}$$

$$Q_{mk} = -B_{mm} v_m^{(2)} - B_{mk} c_{km} - G_{mk} s_{km}. \tag{A.1d}$$

Moreover, the voltage and line capacity constraints are:

$$(V_k^{\min})^2 \leq v_k^{(2)} \leq (V_k^{\max})^2 \tag{A.2a}$$

$$(V_m^{\min})^2 \leq v_m^{(2)} \leq (V_m^{\max})^2 \tag{A.2b}$$

$$\max\{P_{km}^2 + Q_{km}^2, P_{mk}^2 + Q_{mk}^2\} \leq U_{\{k,m\}}. \tag{A.2c}$$

[1]A generative AI tool assisted in finding a smaller network exhibiting this phenomenon. The example was verified by the author.

We will consider the following subsystems:

$$\text{[Jabr]} \qquad (\text{A.1}),\ (\text{A.2}) \tag{A.3a}$$

$$c_{km}^2 + s_{km}^2 \leq v_k^{(2)} v_m^{(2)} \tag{A.3b}$$

and

$$\text{[i2]} \qquad (\text{A.1}),\ (\text{A.2}) \tag{A.4a}$$

$$i_{km}^{(2)} = \alpha_{km} v_k^{(2)} + \beta_{km} v_m^{(2)} + \gamma_{km} c_{km} + \zeta_{km} s_{km} \tag{A.4b}$$

$$P_{km}^2 + Q_{km}^2 \leq v_k^{(2)} i_{km}^{(2)} \tag{A.4c}$$

$$0 \leq i_{km}^{(2)} \leq \frac{U_{\{k,m\}}}{(V_k^{\min})^2}. \tag{A.4d}$$

Without loss of generality, we assume that power generation is unconstrained.

Next, we fix the line parameters. Indeed, consider

$$r_{\{k,m\}} = x_{\{k,m\}} = 10^{-3}, \quad g_{\{k,m\}}^{sh} = b_{\{k,m\}}^{sh} = 0, \quad \tau_{km} = 1, \quad \sigma_{km} = 0, \quad U_{\{k,m\}} = 10.$$

In particular, this corresponds to a simple transmission line with neither shunt admittance nor transformer. This choice of parameters implies that

$$g_{\{k,m\}} = \frac{r_{\{k,m\}}}{r_{\{k,m\}}^2 + x_{\{k,m\}}^2} = 500, \quad b_{\{k,m\}} = \frac{-x_{\{k,m\}}}{r_{\{k,m\}}^2 + x_{\{k,m\}}^2} = -500. \tag{A.5}$$

Therefore, $y_{\{k,m\}} = 500 - 500j$, and the admittance matrix is given by

$$Y_{km} = \begin{pmatrix} y & -y \\ -y & y \end{pmatrix} = \begin{pmatrix} 500 - j500 & -500 + j500 \\ -500 + j500 & 500 - j500 \end{pmatrix}. \tag{A.6}$$

Using (A.13) we compute the coefficients

$$\alpha_{km} = \beta_{km} = g^2_{\{k,m\}} + b^2_{\{k,m\}} = 5 \cdot 10^5, \quad \gamma_{km} = -2(g^2_{\{k,m\}} + b^2_{\{k,m\}}) = -10^6, \quad \zeta_{km} = 0. \tag{A.7}$$

Finally, assume

$$V_k^{\min} = V_m^{\min} = 0.9, \quad V_k^{\max} = V_m^{\max} = 1.1.$$

Next, we check that the solution $\bar{v}_k^{(2)} = \bar{v}_m^{(2)} = 1$, $\bar{c}_{km} = 0.9995$, $\bar{s}_{km} = 0$ is feasible for the subsystem (A.3). Indeed, by (A.6) we can write the power flow equations as:

$$\begin{aligned}
P_{km} &= 500v_k^{(2)} - 500c_{km} + 500s_{km} \\
P_{mk} &= 500v_m^{(2)} - 500c_{km} - 500s_{km} \\
Q_{km} &= 500v_k^{(2)} - 500c_{km} - 500s_{km} \\
Q_{mk} &= 500v_m^{(2)} - 500c_{km} + 500s_{km}.
\end{aligned}$$

Therefore, we have that $\bar{P}_{km} = \bar{P}_{mk} = \bar{Q}_{km} = \bar{Q}_{mk} = 500(1 - 0.9995) = 0.25$. Moreover, the Jabr inequality is satisfied since $\bar{c}^2_{km} + \bar{s}^2_{km} = 0.9995^2 < 1 = \bar{v}_k^{(2)}\bar{v}_m^{(2)}$. The line capacity constraint is satisfied since $\bar{P}^2_{km} + \bar{Q}^2_{km} = 0.0625 < 10$. In consequence, this solution is feasible for the subsystem (A.3).

On the other hand, using (A.7) we can write the constraint which defines $i_{km}^{(2)}$ as

$$i_{km}^{(2)} = 5 \cdot 10^5 (v_k^{(2)} + v_m^{(2)} - 2c_{km}), \tag{A.8}$$

while an upper bound on $i_{km}^{(2)}$ is given by

$$i_{km}^{(2)} \leq \frac{U_{\{k,m\}}}{(V_k^{\min})^2} \approx 12.3457.$$

By substituing our solution $\bar{v}_k^{(2)} = \bar{v}_m^{(2)} = 1$, $\bar{c}_{km} = 0.9995$ in (A.8) we obtain

$$i_{km}^{(}2) = 500 > 12.3457,$$

yielding a contradiction. Therefore, this candidate solution is not feasible for the subsystem (A.4).

A.3 Definition of the i2 Variable

Let the admittance matrix for line $\{km\}$ be

$$Y := \begin{pmatrix} \left(y + \frac{y^{sh}}{2}\right)\frac{1}{\tau^2} & -y\frac{1}{\tau e^{-j\sigma}} \\ -y\frac{1}{\tau e^{j\sigma}} & y + \frac{y^{sh}}{2} \end{pmatrix} = G + jB$$

where $y = g + jb$ denotes it's series admittance, shunt admittance is denoted by $y^{sh} = g^{sh} + jb^{sh}$, and $N := \tau e^{j\sigma}$ denotes the transformer ratio of magnitude $\tau > 0$ and phase shift angle σ.

Recall that Ohm's Law states that the current flowing from bus k to m through the transmission line $\{km\}$ is given by $\begin{pmatrix} I_{km} \\ I_{mk} \end{pmatrix} = Y \begin{pmatrix} V_k \\ V_m \end{pmatrix}$, hence $I_{km} = \frac{y}{\tau}\left(\frac{1}{\tau}V_k - e^{j\sigma}V_m\right) + V_k\frac{y^{sh}}{2\tau^2}$. Therefore, the modulus of I_{km} squared is

$$\begin{aligned}
|I_{km}|^2 &= I_{km}I_{km}^* \\
&= \left(\frac{y}{\tau}\left(\frac{1}{\tau}V_k - e^{j\sigma}V_m\right) + V_k\frac{y^{sh}}{2\tau^2}\right)\left(\frac{y^*}{\tau}\left(\frac{1}{\tau}V_k^* - e^{-j\sigma}V_m^*\right) + V_k^*\frac{y^{sh*}}{2\tau^2}\right) \\
&= \frac{|y|^2}{\tau^2}\left|\frac{1}{\tau}V_k - e^{j\sigma}V_m\right|^2 + 2\mathrm{Re}\left[\frac{y}{\tau}\left(\frac{1}{\tau}V_k - e^{j\sigma}V_m\right)V_k^*\frac{y^{sh*}}{2\tau^2}\right] \\
&+ \frac{1}{4\tau^4}|y^{sh}|^2|V_k|^2. \qquad \text{(A.9)}
\end{aligned}$$

Notice that

$$\begin{aligned}
\mathrm{Re}\left[\frac{y}{\tau}\left(\frac{1}{\tau}V_k - e^{j\sigma}V_m\right)V_k^*\frac{y^{sh*}}{\tau^2}\right] &= \frac{1}{\tau^3}\mathrm{Re}\left[yy^{sh*}\left(\frac{1}{\tau}V_k - e^{j\sigma}V_m\right)V_k^*\right] \\
&= \frac{1}{\tau^3}\mathrm{Re}\left[yy^{sh*}\left(\frac{1}{\tau}|V_k|^2 - e^{j\sigma}V_mV_k^*\right)\right] \\
&= \frac{1}{\tau^3}[(gg^{sh} + bb^{sh})\left((1/\tau)|V_k|^2\right. \\
&\left.-|V_k||V_m|\cos(\theta_{km} - \sigma)\right) \\
&- (bg^{sh} - gb^{sh})|V_k||V_m|\sin(\theta_{km} - \sigma)]
\end{aligned}$$

since $yy^{sh*} = (gg^{sh} + bb^{sh}) + j(bg^{sh} - gb^{sh})$. Moreover, the first term of the RHS of (A.9) can be written as

$$\frac{|y|^2}{\tau^2}\left|\frac{1}{\tau}V_k - e^{j\sigma}V_m\right|^2 = \frac{|y|^2}{\tau^2}\left(\frac{1}{\tau^2}|V_k|^2 + |V_m|^2 - \frac{2}{\tau}|V_k||V_m|\cos(\theta_{km} - \sigma)\right)$$

Putting everything together yields

$$\begin{aligned}
|I_{km}|^2 = \frac{(g^2 + b^2)}{\tau^2}&\left(\frac{1}{\tau^2}|V_k|^2 + |V_m|^2 - \frac{2}{\tau}|V_k||V_m|\cos(\theta_{km} - \sigma)\right) && \text{(A.10a)} \\
&+ \frac{1}{\tau^3}\left[(gg^{sh} + bb^{sh})\left(\frac{1}{\tau}|V_k|^2 - |V_k||V_m|\cos(\theta_{km} - \sigma)\right)\right. && \text{(A.10b)} \\
&\left.- (bg^{sh} - gb^{sh})|V_k||V_m|\sin(\theta_{km} - \sigma)\right] && \text{(A.10c)} \\
&+ \frac{1}{4\tau^4}(g^{sh2} + b^{sh2})|V_k|^2. && \text{(A.10d)}
\end{aligned}$$

Since $|V_k||V_m|\cos(\theta_{km} - \sigma) = c_{km}\cos(\sigma) + s_{km}\sin(\sigma)$ and $|V_k||V_m|\sin(\theta_{km} - \sigma) = s_{km}\cos(\sigma) - c_{km}\sin(\sigma)$ we can represent $i_{km}^{(2)} := |I_{km}|^2$, linearly, in terms of the fundamental variables $v_k^{(2)}$,$v_m^{(2)}$,

c_{km}, and s_{km} as

$$
\begin{aligned}
i_{km}^2 &= \frac{(g^2+b^2)}{\tau^2}\left(\frac{1}{\tau^2}v_k^{(2)} + v_m^{(2)} - \frac{2}{\tau}(c_{km}\cos(\sigma) + s_{km}\sin(\sigma))\right) \\
&+ \frac{1}{\tau^3}\Big[(gg^{sh}+bb^{sh})\left(\frac{1}{\tau}v_k^{(2)} - (c_{km}\cos(\sigma)+s_{km}\sin(\sigma))\right) \\
&- (bg^{sh}-gb^{sh})(s_{km}\cos(\sigma)-c_{km}\sin(\sigma))\Big] \\
&+ \frac{1}{4\tau^4}(g^{sh2}+b^{sh2})v_k^{(2)} && \text{(A.11)} \\
&= \alpha_{km}v_k^{(2)} + \beta_{km}v_m^{(2)} + \gamma_{km}c_{km} + \zeta_{km}s_{km}, && \text{(A.12)}
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_{km} &:= \frac{1}{\tau^4}\left((g^2+b^2) + (gg^{sh}+bb^{sh}) + \frac{(g^{sh2}+b^{sh2})}{4}\right) && \text{(A.13a)} \\
\beta_{km} &:= \frac{(g^2+b^2)}{\tau^2} && \text{(A.13b)} \\
\gamma_{km} &:= \frac{1}{\tau^3}\big(\cos(\sigma)(-2(g^2+b^2)-(gg^{sh}+bb^{sh})) + \sin(\sigma)(bg^{sh}-gb^{sh})\big) && \text{(A.13c)} \\
\zeta_{km} &:= \frac{1}{\tau^3}\big(\sin(\sigma)(-2(g^2+b^2)-(gg^{sh}+bb^{sh})) - \cos(\sigma)(bg^{sh}-gb^{sh})\big). && \text{(A.13d)}
\end{aligned}
$$

A.4 Superoptimality and SOCPs

Consider the following SOCP

$$
[1] \qquad \overline{p} \coloneqq \quad \min \qquad c^\top x
$$

subject to:

$$
\begin{aligned}
\|y_i\|_2 &\le t_i \quad i \in [m] \\
y_i &= A_i x + b_i, \quad i \in [m] \\
t_i &= c_i^\top x + d_i \quad i \in [m] \\
Ax &\ge b
\end{aligned}
$$

Its dual is given by

$$
\begin{aligned}
[2] \qquad \overline{d} \coloneqq \quad \min \qquad & \lambda^\top b - \sum_{i=1}^m (u_i d_i + v_i^\top b_i) \\
\text{subject to:} \qquad & \\
& \|v_i\|_2 \le u_i \quad i \in [m] \\
& c^\top = \lambda^\top A + \sum_{i=1}^m (v_i^\top A_i + u_i c_i^\top) \\
& \lambda, u_i \ge 0
\end{aligned}
$$

Lemma A.1. *Suppose that [1] is strictly feasible and $\overline{p} > -\infty$. Then strong duality holds, i.e., $\overline{p} = \overline{d}$, and there exists an optimal primal-dual pair $(\overline{x}, (\overline{y}_i, \overline{t}_i)_{i\in[m]})$, $(\overline{\lambda}, (\overline{v}_i, \overline{u}_i)_{i\in[m]})$. Let $(\tilde{x}, (\tilde{y}_i, \tilde{t}_i)_{i\in[m]})$ be an ϵ-feasible solution to the primal problem [1], i.e., it satisfies*

$$\epsilon \coloneqq \max\left\{\max\{b_i - A_i^\top \tilde{x} \,:\, i \in [m]\}, \max\{\|\tilde{y}_i\|_2 - \tilde{t}_i \,:\, i \in [m]\}\right\}.$$

Then,

$$c^\top \tilde{x} \ge \overline{p} - \epsilon(\|\overline{\lambda}\|_1 + \|\overline{u}\|_1). \tag{A.14}$$

Proof. Let e_m the all-ones vector in $\mathbb{R}^m$. Then,

$$\begin{aligned}
c^\top \tilde{x} &= \overline{\lambda}^\top A\tilde{x} + \sum_{i=1}^m (\overline{v}_i^\top A_i \tilde{x} + \overline{u}_i c_i^\top \tilde{x}) \\
&\geq \overline{\lambda}^\top (b - \epsilon e_m) + \sum_{i=1}^m (\overline{v}_i^\top (\tilde{y}_i - b_i) + \overline{u}_i(\tilde{t}_i - d_i)) \\
&= \overline{\lambda}^\top b - \epsilon\|\overline{\lambda}\|_1 + \sum_{i=1}^m (\overline{v}_i^\top \tilde{y}_i + \overline{u}_i \tilde{t}_i) - \sum_{i=1}^m (\overline{v}_i^\top b_i + \overline{u}_i d_i) \\
&\geq \overline{p} - \epsilon\|\overline{\lambda}\|_1 + \sum_{i=1}^m (\overline{v}_i^\top \tilde{y}_i + \overline{u}_i(\|\tilde{y}_i\|_2 - \epsilon)) \\
&= \overline{p} - \epsilon\|\overline{\lambda}\|_1 - \epsilon\|\overline{u}\|_1 + \sum_{i=1}^m (\overline{v}_i^\top \tilde{y}_i + \overline{u}_i\|\tilde{y}_i\|_2) \\
&\geq \overline{p} - \epsilon(\|\overline{\lambda}\|_1 + \|\overline{u}\|_1),
\end{aligned}$$

where the last inequality follows since if $\|\overline{v}_i\|_2 \leq \overline{u}_i$ then $-\overline{v}_i^\top \tilde{y}_i \leq \|\overline{v}_i\|_2 \|\tilde{y}_i\|_2 \leq \overline{u}_i \|\tilde{y}_i\|_2$ for all $i \in [m]$. $\square$

A.5 Warm-Started Relaxation for $T = 12, 24$

Table A.1: Warm-Started Relaxation versus i2 SOCP+, $T = 12$ hr active power load curve, ramping rates 50%

	Cutting-Plane													i2 SOCP+								ACOPF
				First Iteration				Last Iteration				Total				Obj			Time			
Case	#Vars	#Cons	FTime	Obj	Time	#Cuts	DInfs	Obj	Time	#Cuts	DInfs	Time	Rnds	#Vars	#Cons	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek	PBound
1354pegase	201209	219772	4.25	855287.79	9.75	18000	1.70e-08	860664.27	6.16	30236	1.06e-08	92.63	12	168974	257483	-	-	-	245.39	177.08	35.72	861473.50
9241pegase	1574972	1626484	34.46	3554525.89	163.27	205416	5.04e-07	3565252.49	271.12	354209	4.32e-04	3811.42	8	1354634	1969307	-	-	-	1439.33	TLim	352.61	-
ACTIVSg10k	1375020	1490960	30.21	28360735.35	106.03	84552	7.50e-09	28690135.28	94.69	148893	3.38e-07	1308.80	11	1151066	1714202	-	-	-	1543.11	TLim	190.39	-
10000goc-api	1401204	1503620	32.03	27353308.25	96.60	107568	3.27e-07	28459989.11	74.63	183308	2.40e-08	1066.39	11	1170122	1736383	-	-	-	TLim	TLim	256.71	-
10000goc-sad	1401204	1503620	30.42	16363341.55	133.30	105672	2.59e-06	16474320.51	105.74	144064	3.36e-06	1450.12	10	1170122	1736383	-	-	-	1918.80	TLim	247.71	-
ACTIVSg25k	3447015	3675848	74.34	68451480.40	357.88	206028	2.68e-07	69534077.06	972.97	796784	9.75e-07	4073.97	5	2885558	4247241	-	-	-	3539.20	TLim	523.18	-
30000goc-api	3768147	4127000	83.88	15513235.10	367.19	179196	7.00e-07	17079448.81	291.10	357763	1.06e-05	3667.40	9	3056138	4657273	-	-	-	TLim	TLim	756.77	-
30000goc-sad	3768147	4127000	86.66	13129167.74	521.03	539112	1.68e-06	13190466.44	329.66	246646	2.45e-06	3901.76	8	3056138	4657273	-	-	-	1694.63	TLim	888.11	-
ACTIVSg70k	9368583	10034036	207.22	184410148.60	1199.80	774684	4.07e-08	188736994.5	1276.64	1879714	3.66e-08	4702.30	3	7752122	11515211	-	-	-	3596.20	TLim	1124.19	-
78484epigrids-api	12551704	12730080	263.06	INF	137.79	1877340	-	INF	137.79	1877340	-	447.39	0	10674890	15270737	-	-	-	2875.97	TLim	1254.95	-
78484epigrids-sad	12551704	12730080	273.21	176961641.05	2697.25	2666748	1.36e-04	177335668.05	2828.46	3455898	1.36e-04	5465.67	1	10674890	15270737	-	-	-	TLim	TLim	1241.86	-

Table A.2: Warm-Started Relaxation versus i2 SOCP+, $T = 24$ hr active power load curve, ramping rates 50%

	Cutting-Plane													i2 SOCP+								ACOPF
				First Iteration				Last Iteration				Total				Obj			Time			
Case	#Vars	#Cons	FTime	Obj	Time	#Cuts	DInfs	Obj	Time	#Cuts	DInfs	Time	Rnds	#Vars	#Cons	Gurobi	Knitro	Mosek	Gurobi	Knitro	Mosek	PBound
1354pegase	402677	440584	8.66	1656321.34	20.38	36000	2.46e-08	1666609.97	12.74	63049	2.97e-08	180.61	11	337946	517475	-	-	-	345.44	TLim	53.32	-
9241pegase	3151388	3258748	66.05	6865741.86	368.23	410832	4.33e-08	6925029.02	2427.02	1354277	1.25e-03	8419.66	4	2709266	3957551	-	-	-	1972.91	TLim	525.04	-
ACTIVSg10k	2752524	2991860	66.13	54565119.63	226.82	169104	7.5e-09	55184650.09	211.53	298812	7.18e-08	2869.25	11	2302130	3446078	-	-	-	3918.14	TLim	396.91	-
10000goc-api	2804496	3015596	62.15	51667906.99	208.71	215136	5.822e-07	53800966.46	175.46	388226	3.18e-07	2273.58	10	2340242	3490135	-	-	-	1332.84	TLim	542.88	-
10000goc-sad	2804496	3015596	60.63	32512183.58	284.34	211344	3.46e-06	32678842.78	276.70	246720	7.97e-06	3099.33	9	2340242	3490135	-	-	-	2994.53	TLim	714.14	-
ACTIVSg25k	6898863	7371032	150.16	132201219.46	816.18	412056	4.74e-07	134142620.00	2042.59	1385811	2.08e-06	6991.16	3	5771114	8536377	-	-	-	4451.39	TLim	963.91	-
30000goc-api	7539819	8268104	167.43	28586461.36	755.74	358392	1.7313e-06	31616476.13	884.66	1064587	2.06e-05	6799.92	7	6112274	9356989	-	-	-	TLim	TLim	1139.21	-
30000goc-sad	7539819	8268104	169.75	25717079.58	1091.27	1078224	2.52e-06	25811075.96	1021.24	1393213	1.23e-05	6249.53	5	6112274	9356989	-	-	-	3213.23	TLim	1125.14	-
ACTIVSg70k	18747555	20109632	410.49	350310736.78	2641.18	1549368	2.407e-07	357083446.24	2687.31	3359811	2.98e-08	8496.71	2	1550424	23139407	-	-	-	5061.29	TLim	1331.11	-
78484epigrids-api	25110280	24855336	533.07	-	1521.64	3754680	-	-	1521.64	3754680	-	2138.78	0	21349922	30681233	-	-	-	TLim	287.62	1434.04	-
78484epigrids-sad	25110280	24855336	532.93	344805634.00	4749.24	5333496	1.34e-04	-	TLim	6915253	-	11561.02	1	21349922	30681233	-	-	-	TLim	284.28	1448.87	-

A.6 On the Inequivalence of the Matrices $L(X)$ and W

Proposition A.2. *Let $X \in \mathbb{C}^{n\times n}$ and $W \in \mathbb{R}^{2n\times 2n}$ be arbitrary rank-one matrices. Then, there does not exist an orthogonal matrix $Q \in \mathbb{R}^{2n\times 2n}$ such that $W = QL(X)Q^\top$.*

Proof. Assume, to the contrary, that there exists some orthogonal matrix $Q \in \mathbb{R}^{2n\times 2n}$ such that $W = QL(X)Q^\top$. Since orthogonal similarity transformations preserve rank, and Lemma A.3 implies that $L(X)$ is rank-two, we have that rank $W = 2$, which is a contradiction. $\square$

Lemma A.3. *Let $X \in \mathbb{C}^{n\times n}$. Then rank $L(X) = 2 \cdot$ rank X.*

Proof. Let $z = x + jy \in \mathbb{C}^n$ where $x, y \in \mathbb{R}^n$. Consider the $\mathbb{R}$-linear map $\phi : \mathbb{C}^n \to \mathbb{R}^{2n}$ given by $\phi(z) = (x, y)^\top$. We note that

$$\begin{aligned}
Xz = 0 &\iff (\Re X + j\Im X)(x + jy) = 0 \\
&\iff (\Re Xx - \Im Xy) + j(\Im Xx + \Re Xy) = 0 \\
&\iff \begin{pmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \\
&\iff L(X) \begin{pmatrix} x \\ y \end{pmatrix} = 0.
\end{aligned}$$

Next, let B be a basis for $\ker V$. By the above and since ϕ is injective, we have that $\dim\ker L(X) \geq |B|$. We show that $\dim\ker L(X) = 2 \cdot |B|$. For an arbitrary $z = x + jy \in \mathbb{C}^n$, we remark that clearly z and jz $\mathbb{C}$-linearly dependent, while $\phi(z)$ and $\phi(jz)$ are $\mathbb{R}$-linearly independent. Indeed, $\phi(jz) = (-y, x)^\top$ hence $\phi(z)$ is orthogonal to $\phi(jz)$. Since for different $z, z' \in B$, we have that z and jz' are $\mathbb{C}$-linearly independent, hence $\dim\ker L(X) \geq 2 \cdot |B|$. Finally, this inequality cannot be strict since if it were, there would be some $(\hat{x}, \hat{y}) \in \ker L(X)$ such that $\phi^{-1}(\hat{x}, \hat{y}) \in \ker V$ is not spanned by B. $\square$

A.7 Proof of a Result on Valid Inequalities for Polytopes

The following proposition is based on Exercise 3.14 of [12, Chapter 3]. For completeness, we provide a proof, which relies on the following auxiliary lemma.

Lemma A.4. *Let* $P := \{x \in \mathbb{R}^n \,:\, Ax \geq b\}$ *be a polytope such that* $\dim(P) > 0$*. Then the system of inequalities* $Ax \leq b$ *is infeasible.*

Proof. Let P' denote the polyhedron described by $Ax \leq b$. Suppose, for contradiction, that there exists $y \in P'$. Then, for any $x \in P$ we have that

$$Ay \leq b \leq Ax.$$

Hence, $A(x - y) \geq 0$. In particular, the latter holds for any point x_0 in the relative interior of P, which is nonempty since $\dim(P) > 0$. As x_0 is in the relative interior of P, at least one inequality $a_i^\top x \geq b_i$ is inactive at x_0, i.e., $a_i^\top x_0 > b_i$. Let $r := x_0 - y$. Then

$$a_i^\top r = a_i^\top x_0 - a_i^\top y > b_i - a_i^\top y \geq 0,$$

and therefore $r \neq 0$. Together with $Ar \geq 0$, this implies that r is a nonzero recession direction of P, contradicting the fact that P is bounded. □

Proposition A.5. *Let* $P := \{x \in \mathbb{R}^n \,:\, Ax \geq b\}$ *be a polytope such that* $\dim(P) > 0$*. Then an inequality* $\alpha^\top x \geq \beta$ *is valid for* P *if and only there exists* $u \geq 0$ *such that* $A^\top u = \alpha$ *and* $b^\top u = \beta$*.*

Proof. We first prove sufficiency. Let $x \in \mathbb{R}^n$ that belongs to P, i.e., $Ax \geq b$. Since there exists $u \geq 0$ such that $A^\top u = \alpha$ and $b^\top u = \beta$, we have

$$\alpha^\top x = u^\top Ax \geq u^\top b = \beta.$$

Since x was arbitrary, the inequality $\alpha^\top x \geq \beta$ is valid for P.

Next, we prove necessity. Since P is a polytope the linear program $\beta^* := \min\{\alpha^\top x \,:\, x \in P\}$ has a finite optimal value, i.e., $\beta^* \in \mathbb{R}$. By LP strong duality, there exists an optimal dual solution $u^* \geq 0$ which satisfies $A^\top u^* = \alpha$ and $b^\top u^* = \beta^*$. If $\beta^* = \beta$ the claim follows immediately. Otherwise, if $\beta^* > \beta$, it remains to construct a dual-feasible (but non-optimal) solution $\overline{u} \geq 0$ with objective value $b^\top \overline{u} = \beta < \beta^*$. To complete the proof, it suffices to show that there exists a vector $d \geq 0$ such that $A^\top d = 0$ with $b^\top d = -1$. Indeed, let $\Delta := \beta^* - \beta$ and define $\overline{u} := u^* + \Delta d$. Since $u^*, d \geq 0$ and $\Delta > 0$ we have $\overline{u} \geq 0$. Moreover, $A^\top \overline{u} = A^\top u^* + \Delta A^\top d = \alpha$ and $b^\top \overline{u} = b^\top u^* + \Delta b^\top d = \beta^* - \Delta = \beta$. Hence, $\overline{u}$ is dual feasible and satisfies $b^\top \overline{u} = \beta$, proving the claim. It therefore remains to establish the existence of such a vector d.

To this end, observe that the existence of d is equivalent to the linear program

$$\min\{b^\top d \,:\, A^\top d = 0,\, d \geq 0\}$$

being unbounded. By Farkas' Lemma, the latter is equivalent to the infeasibility of the system $Ax \leq b$, which follows from Lemma A.4. This completes the proof. □

A.8 Sherali-Adams Decomposition Lemma: A Complete Proof

To make this appendix self-contained, we recall the notation and definitions from previous sections that are needed in the proof.

A.8.1 Definitions

For $\emptyset \neq S \subseteq [n]$ and subsets $J_0, J_1 \subseteq S$, we say that (J_0, J_1) is a 2-partition of S if $J_0 \cup J_1 = S$ and $J_0 \cap J_1 = \emptyset$, and we denote it as $J_0 \sqcup J_1 = S$. Let $d, k, n \in \mathbb{N}_{\geq 0}$ with $k + 1 \leq n$ and $d \leq k$. Given $y \in [0,1]^{\binom{n}{\leq k+1}}$, for $S \subseteq [n]$ with $|S| = d$ and $J_0 \sqcup J_1 = S$, we define:

$$y_I^{J_0,J_1} := \sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup I \cup H}, \quad z_I^{J_0,J_1}(y) := \begin{cases} \frac{y_I^{J_0,J_1}}{y_\emptyset^{J_0,J_1}} & \text{if } y_\emptyset^{J_0,J_1} > 0, \\ y_I^{J_0,J_1} & \text{if } y_\emptyset^{J_0,J_1} = 0, \end{cases} \tag{A.15}$$

for every $I \subseteq [n]$ with $|I| \leq k+1-d$.

For any $S \subseteq [n]$ with $|S| = k$, each 2-partition $J_0 \sqcup J_1 = S$ induces the k-junta

$$F_k(J_0, J_1)(x) \coloneqq \prod_{j \in J_0} (1 - x_j) \prod_{j \in J_1} x_j.$$

When no confusion can arise, we omit the subscript and simply write $F(J_0, J_1)(x)$. Its multilinearization in the y-space can be written as

$$[F(J_0, J_1)(x)]_L = \sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup H},$$

where $[\,\cdot\,]_L$ denotes the multilinearization operator acting on polynomial expressions by first applying the identities $x_i^2 = x_i$ for every $i \in [n]$, then replacing each monomial $\prod_{i \in I} x_i$, where $I \subseteq [n]$, with y_I, setting $y_{\{i\}} = x_i$, and $y_\emptyset = 1$.

Let $P \subseteq [0,1]^n$ be a polytope described by $\sum_{j=1}^n a_{i,j} x_j - b_i \geq 0$, $i \in [m]$. Using this notation, we follow the original definition of Sherali and Adams [150] and write $\mathrm{M}^k(P)$ as follows:

(SA-0) $y_\emptyset = 1$;

(SA-1) For each 2-partition $J_0 \sqcup J_1 = S \subseteq [n]$ with $|S| \leq \min\{k+1, n\}$,

$$\sum_{H \subseteq J_0} (-1)^{|H|} y_{J_1 \cup H} \geq 0;$$

(SA-2) For each $i \in [m]$ and each 2-partition $J_0 \sqcup J_1 = S \subseteq [n]$ with $|S| \leq k$,

$$\sum_{H \subseteq J_0} (-1)^{|H|} \left(\sum_{j=1}^n a_{i,j} y_{J_1 \cup \{j\} \cup H} - b_i y_{J_1 \cup H} \right) \geq 0.$$

A.8.2 Simple Facts

Remark A.6. Recall that for any $y \in [0,1]^{\binom{n}{\leq k+1}}$ and $S \subseteq [n]$, we have that

$$y_\emptyset = \sum_{J_0, J_1 \subseteq [n]\,:\, J_0 \sqcup J_1 = S} y_\emptyset^{J_0,J_1}.$$

This is a well-known fact. In particular, it implies that if $y \in \mathrm{M}^k(P)$, for any $0 \leq k \leq n$, this sum is strictly positive and it equals 1.

We first establish some simple, well-known facts.

Lemma A.7. *Let $y \in [0,1]^{\binom{n}{\leq k+1}}$, $J_0 \sqcup J_1 = S \subseteq [n]$ be a 2-partition with $|S| \leq d$, and $I \subseteq [n]$ be an arbitrary subset with $|I| \leq k+1-d$. Then the following statements hold.*

i) $\forall R, T \subseteq [n]$, *such that* $|R \cup T| \leq d$ *and* $R \cap T \neq \emptyset$, *we have* $\sum_{H \subseteq R} (-1)^{|H|} y_{T \cup I \cup H} = 0$.

ii) If $I \cap J_0 \neq \emptyset$, *then* $y_I^{J_0,J_1} = 0$;

iii) $y_I^{J_0,J_1} = y_{I \setminus J_1}^{J_0,J_1} = y_I^{J_0, J_1 \setminus I}$. *In particular, if* $I \subseteq J_1$, *then* $y_I^{J_0,J_1} = y_\emptyset^{J_0,J_1}$;

iv) For $i \in [n] \setminus S$, *it holds that* $y_I^{J_0, J_1 \cup \{i\}} = y_{I \cup \{i\}}^{J_0,J_1}$ *and* $y_I^{J_0 \cup \{i\}, J_1} = y_I^{J_0,J_1} - y_{I \cup \{i\}}^{J_0,J_1}$; *and*

v) For $i \in [n] \setminus S$, *we have* $F(J_0, J_1)(x) = F(J_0 \cup \{i\}, J_1)(x) + F(J_0, J_1 \cup \{i\})(x)$. *This implies that to describe* M^k *it suffices to consider SA-1 constraints induced by* 2*-partitions of cardinality* $k+1$, *and SA-2 constraints induced by* 2*-partitions of cardinality* k.

Proof. i) Let $i \in R \cap T$. Then for any $I \subseteq [n]$ with $|I| \leq k+1-d$ we have

$$\begin{aligned}
\sum_{H \subseteq R} (-1)^{|H|} y_{T \cup I \cup H} &= \sum_{H \subseteq R\,:\, H \not\ni i} (-1)^{|H|} y_{T \cup I \cup H} + \sum_{H \subseteq R\,:\, H \ni i} (-1)^{|H|} y_{T \cup I \cup H} \\
&= \sum_{H \subseteq R \setminus \{i\}} (-1)^{|H|} y_{T \cup I \cup H} + \sum_{H \subseteq R \setminus \{i\}} (-1)^{|H|+1} y_{T \cup I \cup H \cup \{i\}} \\
&= \sum_{H \subseteq R \setminus \{i\}} (-1)^{|H|} y_{T \cup I \cup H} - \sum_{H \subseteq R \setminus \{i\}} (-1)^{|H|} y_{T \cup I \cup H} \\
&= 0.
\end{aligned}$$

ii) Let $i \in I \cap J_0$. Then $y_I^{J_0,J_1} = y_{I\setminus\{i\}}^{J_0,J_1\cup\{i\}} = 0$ where the last equality follows by i).

iii) We note that $y_I^{J_0,J_1}$ can be written as

$$\sum_{H\subseteq J_0}(-1)^{|H|}y_{J_1\cup I\cup H} = \sum_{H\subseteq J_0}(-1)^{|H|}y_{J_1\cup(I\setminus J_1)\cup H} = y_{I\setminus J_1}^{J_0,J_1} = \sum_{H\subseteq J_0}(-1)^{|H|}y_{(J_1\setminus I)\cup I\cup H} = y_I^{J_0,J_1\setminus I}.$$

iv) Let $i \in [n]$. Then it follows that $y_I^{J_0,J_1\cup\{i\}} = \sum_{H\subseteq J_0}(-1)^{|H|}y_{J_1\cup(\{i\}\cup I)} = y_{I\cup\{i\}}^{J_0,J_1}$. Next, we show $y_I^{J_0\cup\{i\},J_1} = y_I^{J_0,J_1} - y_{I\cup\{i\}}^{J_0,J_1}$. Indeed,

$$\begin{aligned}
y_I^{J_0\cup\{i\},J_1} &= \sum_{H\subseteq J_0\cup\{i\}}(-1)^{|H|}y_{J_1\cup H\cup I} \\
&= \sum_{H\subseteq J_0\cup\{i\}\,:\,H\not\ni i}(-1)^{|H|}y_{J_1\cup H\cup I} + \sum_{H\subseteq J_0\cup\{i\}\,:\,H\ni i}(-1)^{|H|}y_{J_1\cup H\cup I} \\
&= \sum_{H\subseteq J_0}(-1)^{|H|}y_{J_1\cup H\cup I} - \sum_{H\subseteq J_0}(-1)^{|H|}y_{J_1\cup H\cup I\cup\{i\}} \\
&= y_I^{J_0,J_1} - y_{I\cup\{i\}}^{J_0,J_1}.
\end{aligned}$$

v) For $i \in [n]\setminus S$, we have

$$F(J_0,J_1)(x) = (1-x_i+x_i)F(J_0,J_1)(x) = F(J_0\cup\{i\},J_1)(x) + F(J_0,J_1\cup\{i\})(x).$$

Hence, for every inequality $a^\top x - b \geq 0$ of P,

$$\begin{aligned}
\left[F(J_0,J_1)(x)(a^\top x-b)\right]_L &= \left[F(J_0\cup\{i\},J_1)(x)(a^\top x-b)\right]_L \\
&\quad + \left[F(J_0,J_1\cup\{i\})(x)(a^\top x-b)\right]_L.
\end{aligned}$$

□

Remark A.8. For $y \in [0,1]^{\binom{n}{\leq k+1}}$ and any 2-partition $J_0 \sqcup J_1 = S \subseteq [n]$ with $|S| \leq k+1$, we note that statement ii) of Lemma A.7 implies that $z_i^{J_0,J_1}(y) = 0$ for $i \in J_0$, while statement iii) implies that $z_i^{J_0,J_1}(y) = 1$ for $i \in J_1$, unless $y_\emptyset^{J_0,J_1} = 0$. Indeed, by iii), we have $y_i^{J_0,J_1} = y_\emptyset^{J_0,J_1}$ if $i \in J_1$;

hence $z_i^{J_0,J_1}(y) = \frac{y_i^{J_0,J_1}}{y_\emptyset^{J_0,J_1}} = 1$.

The following fact should be known but we are not aware of a formal proof. This is key when we use the decomposition lemma to induce on a variable y_i which is integral.

Lemma A.9. *Let $y \in \mathrm{M}^k(P)$ and let $i \in [n]$ such that $y_i \in \{0,1\}$. Then, for every $I \subseteq [n]$ with $|I| \leq k$, we have that $y_{I\cup\{i\}} = y_I$ if $y_i = 1$, whereas $y_{I\cup\{i\}} = 0$ if $y_i = 0$.*

Proof. We show that the above relations are implied by SA-1 constraints.

If $y_i = 0$, then $y_{I\cup\{i\}} = 0$. We prove that the inequality $y_i - y_{I\cup\{i\}} \geq 0$ holds. We proceed by induction on the cardinality of I. The base case $I = \{j\} \subseteq [n]$ follows since $y_i - y_{ij} \geq 0$ corresponds to the SA-1 constraint induced by the 2-partition $J_0 \sqcup J_1 = \{j\} \sqcup \{i\}$. Next, suppose that the statement holds for any I of cardinality $0 \leq \ell \leq k-1$. Consider $I \subseteq [n]$ such that $|I| = \ell + 1$. For some $j \in I$ we write $I = (I \setminus \{j\}) \sqcup \{j\}$. Then, the SA-1 constraint induced by the 2-partition $J_0 \sqcup J_1 = \{j\} \sqcup ((I \setminus \{j\}) \cup \{i\})$ is given by $y_{(I\setminus\{j\})\cup\{i\}} - y_{I\cup\{i\}} \geq 0$. By induction hypothesis we have that $y_i - y_{(I\setminus\{j\})\cup\{i\}} \geq 0$, hence $y_i - y_{I\cup\{i\}} \geq 0$. Since for every $S \subseteq [n]$ with $|S| \leq k+1$, the SA-1 constraint $y_S \geq 0$ is induced by the 2-partition $\emptyset \sqcup S = S$, and we assumed that $y_i = 0$, we conclude that $y_{I\cup\{i\}} = 0$.

If $y_i = 1$, then $y_{I\cup\{i\}} = y_I$. First, notice that the constraint $y_I - y_{I\cup\{i\}} \geq 0$ is induced by the 2-partition $J_0 \sqcup J_1 = \{i\} \sqcup I$. We show by induction on the cardinality of I that $y_{I\cup\{i\}} - y_I \geq 0$ holds if $y_i = 1$. We assume that $i \notin I$ else it the inequality trivially holds. The base case $I = \{j\} \subseteq [n]$ follows by noticing that the 2-partition $J_0 \sqcup J_1 = \{i,j\} \sqcup \emptyset$ induces $y_\emptyset - y_i - y_j + y_{ij} \geq 0$. Since $y \in \mathrm{M}^k(P)$ we have that $y_\emptyset = 1$, hence $y_i = 1$ implies $y_{ij} - y_j \geq 0$. Next, suppose that the inequality $y_{I\cup\{i\}} - y_I \geq 0$ holds for any I of cardinality $0 \leq \ell \leq k-1$. We show that it holds for $|I| = \ell + 1$. We do this by cases. First, assume that $|I| = \ell + 1$ is odd and consider the inequality induced by the 2-partition $J_0 \sqcup J_1 = (\{i\} \sqcup I) \sqcup \emptyset$. Indeed,

$$\begin{aligned}
0 \leq \sum_{H \subseteq \{i\} \sqcup I} (-1)^{|H|} y_H &= \sum_{H \subseteq \{i\} \sqcup I \,:\, H \not\ni i} (-1)^{|H|} y_H + \sum_{H \subseteq \{i\} \sqcup I \,:\, H \ni i} (-1)^{|H|} y_H \\
&= \sum_{H \subseteq I} (-1)^{|H|} y_H + \sum_{H \subseteq I} (-1)^{|H|+1} y_{H \cup \{i\}} \\
&= \sum_{H \subsetneq I} (-1)^{|H|} y_H + (-1)^{|I|} y_I - \sum_{H \subsetneq I} (-1)^{|H|} y_{H \cup \{i\}} + (-1)^{|I|+1} y_{I \cup \{i\}} \\
&= -y_I + y_{I \cup \{i\}}
\end{aligned}$$

where in the last equality we used the induction hypothesis $y_{H \cup \{i\}} = y_H$ for all $|H| < |I| = \ell + 1$, and that $|I|$ is odd.

Next, we address the case where $|I| = \ell + 1$ is even. Since $\ell \geq 1$, we can pick an arbitrary $j \in I$ and write $I' \coloneqq I \setminus \{j\}$. Consider the 2-partition $J_0 \sqcup J_1 = (\{i\} \sqcup I') \sqcup \{j\}$. The inequality induced by this 2-partition, $\sum_{H \subseteq \{i\} \sqcup I'} (-1)^{|H|} y_{H \cup \{j\}} \geq 0$, can be written as

$$\begin{aligned}
0 &\leq \sum_{H \subseteq \{i\} \sqcup I' \,:\, H \not\ni i} (-1)^{|H|} y_{H \cup \{j\}} + \sum_{H \subseteq \{i\} \sqcup I' \,:\, H \ni i} (-1)^{|H|} y_{H \cup \{j\}} \\
&= \sum_{H \subseteq I'} (-1)^{|H|} y_{H \cup \{j\}} + \sum_{H \subseteq I'} (-1)^{|H|+1} y_{H \cup \{j\} \cup \{i\}} \\
&= \sum_{H \subsetneq I'} (-1)^{|H|} y_{H \cup \{j\}} + (-1)^{|I'|} y_{I' \cup \{j\}} - \sum_{H \subsetneq I'} (-1)^{|H|} y_{H \cup \{j\} \cup \{i\}} + (-1)^{|I'|+1} y_{I' \cup \{j\} \cup \{i\}} \\
&= - y_I + y_{I \cup \{i\}}
\end{aligned}$$

where in the last equality we used the induction hypothesis $y_{H \cup \{j\} \cup \{i\}} = y_{H \cup \{j\}}$ for all $|H \cup \{j\}| < \ell + 1$, and that $|I'|$ is odd.

Since $y_I - y_{I \cup \{i\}} \geq 0$, we conclude that $y_{I \cup \{i\}} = y_I$ if $y_i = 1$. □

A.8.3 The Proof

We present a complete proof of a one-round version of the decomposition lemma for Sherali–Adams. This is a well-known result, and a partial proof of the implication $y \in \mathrm{M}^k(P) \Rightarrow$ 1) and 2) below appears in [149]. For completeness, we provide a full proof, including the converse implication and the case in which the variables on which we induce are integral, neither of which is addressed in [149].

Lemma A.10. *Let $y \in [0,1]^{\binom{n}{\leq k+1}}$. Then $y \in \mathrm{M}^k(P)$ if and only if for every $i \in [n]$ the following hold:*

1. *$z^{\{i\},\emptyset}(y) \in \mathrm{M}^{k-1}(P)$ if $y_\emptyset^{\{i\},\emptyset} > 0$, and $z^{\emptyset,\{i\}}(y) \in \mathrm{M}^{k-1}(P)$ if $y_\emptyset^{\emptyset,\{i\}} > 0$; and*
2. *The restriction of y to its first $\binom{n}{\leq k-d+1}$ coordinates can be expressed as the convex combination of points $z^{J_0,J_1}(y)$, namely,*

$$y_I = y_\emptyset^{\{i\},\emptyset} z_I^{\{i\},\emptyset}(y) + y_\emptyset^{\emptyset,\{i\}} z_I^{\emptyset,\{i\}}(y) \quad \textit{for } I \subseteq [n],\ |I| \leq k. \tag{A.16}$$

Proof. Let $y \in [0,1]^{\binom{n}{\leq k+1}}$.

$y \in \mathrm{M}^k(P)$ **implies 1) and 2).** We fix an arbitrary $i \in [n]$. First, assume that $y_\emptyset^{\{i\},\emptyset} > 0$ and $y_\emptyset^{\{i\},\emptyset} > 0$. We will show that $z^{\{i\},\emptyset}(y)$ and $z^{\emptyset,\{i\}}(y)$ belong to $\mathrm{M}^{k-1}(P)$. To this end, we show that $z^{\{i\},\emptyset}(y)$ and $z^{\emptyset,\{i\}}(y)$ satisfy the constraints SA-0, SA-1, and SA-2 (see Section A.8.1). Notice that $z_\emptyset^{\{i\},\emptyset}(y) = \frac{y_\emptyset^{\{i\},\emptyset}}{y_\emptyset^{\{i\},\emptyset}} = 1$ and $z_\emptyset^{\emptyset,\{i\}}(y) = \frac{y_\emptyset^{\emptyset,\{i\}}}{y_\emptyset^{\emptyset,\{i\}}} = 1$, hence we are done verifying SA-0.

<u>SA-1 Constraints</u> Fix some $T \subseteq [n]$ of cardinality k and consider an arbitrary 2-partition $T_0 \sqcup T_1 = T$. Recall that every 2-partition induces a k-junta. We note that it suffices to consider k-juntas since they imply j-juntas with $j \leq k$ by Lemma A.7. We show that $z^{\{i\},\emptyset}(y)$ satisfies SA-1 inequalities.

Indeed, consider the sum

$$\begin{aligned}
\sum_{H \subseteq T_0} (-1)^{|H|} z^{\{i\},\emptyset}_{T_1 \cup H}(y) &= \frac{1}{y^{\{i\},\emptyset}_{\emptyset}} \left(\sum_{H \subseteq T_0} (-1)^{|H|} (y_{T_1 \cup H} - y_{T_1 \cup H \cup \{i\}}) \right) \\
&= \frac{1}{y^{\{i\},\emptyset}_{\emptyset}} \left(\sum_{H \subseteq T_0 \cup \{i\} \,:\, H \not\ni i} (-1)^{|H|} y_{T_1 \cup H} + \sum_{H \subseteq T_0 \,:\, H \not\ni i} (-1)^{|H|+1} y_{T_1 \cup H \cup \{i\}} \right) \\
&= \frac{1}{y^{\{i\},\emptyset}_{\emptyset}} \left(\sum_{H \subseteq T_0 \cup \{i\} \,:\, H \not\ni i} (-1)^{|H|} y_{T_1 \cup H} + \sum_{H' \subseteq T_0 \cup \{i\} \,:\, H' \ni i} (-1)^{|H'|} y_{T_1 \cup H'} \right) \\
&= \frac{1}{y^{\{i\},\emptyset}_{\emptyset}} \left(\sum_{H \subseteq T_0 \cup \{i\}} (-1)^{|H|} y_{T_1 \cup H} \right). \qquad \text{(A.17)}
\end{aligned}$$

We claim that the RHS of (A.17) is nonnegative. Indeed, if $i \notin T_1$, then $(T_0 \cup \{i\}) \sqcup T_1$ is a 2-partition such that $|(T_0 \cup \{i\}) \sqcup T_1| \leq k+1$; hence, the RHS is nonnegative since $y \in \mathrm{M}^k(P)$. If $i \in T_1$ then by Lemma A.7 the RHS is 0.

Next, we show that $z^{\emptyset,\{i\}}(y)$ satisfies SA-1 inequalities. We can write

$$\sum_{H \subseteq T_0} (-1)^{|H|} z^{\emptyset,\{i\}}_{T_1 \cup H}(y) = \frac{1}{y^{\emptyset,\{i\}}_{\emptyset}} \left(\sum_{H \subseteq T_0} (-1)^{|H|} y_{T_1 \cup H \cup \{i\}} \right). \qquad \text{(A.18)}$$

We claim that the RHS of (A.18) is nonnegative. Indeed, if $i \notin T_0$, then $T_0 \sqcup (T_1 \cup \{i\})$ is a 2-partition such that $|T_0 \sqcup (T_1 \cup \{i\})| \leq k+1$; hence, the RHS is nonnegative since $y \in \mathrm{M}^k(P)$. If $i \in T_0$ then again by Lemma A.7 the RHS is 0.

SA-2 Constraints Fix $T \subseteq [n]$ with $|T| = k-1$ and consider a 2-partition $T_0 \sqcup T_1 = T$. Let $a^\top x \geq b$ be an arbitrary inequality of P. We show that $z^{\{i\},\emptyset}(y)$ satisfies the associated SA-2 type constraint. We can write

$$
\begin{aligned}
&\sum_{H\subseteq T_0}(-1)^{|H|}\left(\sum_{j=1}^{n}a_j z^{\{i\},\emptyset}_{T_1\cup\{j\}\cup H}(y)-bz^{\{i\},\emptyset}_{T_1\cup H}(y)\right)\\
=&\frac{1}{y^{\{i\},\emptyset}_{\emptyset}}\sum_{H\subseteq T_0}(-1)^{|H|}\left(\sum_{j=1}^{n}(a_jy_{T_1\cup\{j\}\cup H}-a_jy_{T_1\cup\{j\}\cup H\cup\{i\}})-(by_{T_1\cup H}-by_{T_1\cup H\cup\{i\}})\right)\\
=&\frac{1}{y^{\{i\},\emptyset}_{\emptyset}}\sum_{H\subseteq T_0\cup\{i\}\,:\,H\not\ni i}(-1)^{|H|}\left(\sum_{j=1}^{n}a_jy_{T_1\cup\{j\}\cup H}-by_{T_1\cup H}\right)\\
+&\frac{1}{y^{\{i\},\emptyset}_{\emptyset}}\sum_{H\subseteq T_0\cup\{i\}\,:\,H\ni i}(-1)^{|H|}\left(\sum_{j=1}^{n}a_jy_{T_1\cup\{j\}\cup H}-by_{T_1\cup H}\right)\\
=&\frac{1}{y^{\{i\},\emptyset}_{\emptyset}}\sum_{H\subseteq T_0\cup\{i\}}(-1)^{|H|}\left(\sum_{j=1}^{n}a_jy_{T_1\cup\{j\}\cup H}-by_{T_1\cup H}\right).
\end{aligned}
\tag{A.19}
$$

We claim that the RHS of (A.19) is nonnegative. Indeed, if $i\notin T_1$, then $(T_0\cup\{i\})\sqcup T_1$ is a 2-partition such that $|(T_0\cup\{i\})\sqcup T_1|\leq k$; hence, the RHS is nonnegative since $y\in \mathrm{M}^k(P)$. If $i\in T_1$ then by Lemma A.7 the RHS is 0.

Next, we show that $z^{\emptyset,\{i\}}(y)$ satisfies the associated SA-2 type constraint. Indeed,

$$
\begin{aligned}
&\sum_{H\subseteq T_0}(-1)^{|H|}\left(\sum_{j=1}^{n}a_j z^{\emptyset,\{i\}}_{T_1\cup\{j\}\cup H}(y)-bz^{\emptyset,\{i\}}_{T_1\cup H}(y)\right)\\
=&\frac{1}{y^{\emptyset,\{i\}}_{\emptyset}}\sum_{H\subseteq T_0}(-1)^{|H|}\left(\sum_{j=1}^{n}a_jy_{T_1\cup\{j\}\cup H\cup\{i\}}-by_{T_1\cup H\cup\{i\}})\right)
\end{aligned}
\tag{A.20}
$$

We claim that the RHS of (A.20) is nonnegative. Indeed, if $i\notin T_0$, then $T_0\sqcup(T_1\cup\{i\})$ is a 2-partition such that $|T_0\sqcup(T_1\cup\{i\})|\leq k$; hence, the RHS is nonnegative since $y\in \mathrm{M}^k(P)$. If $i\in T_0$ then again by Lemma A.7 the RHS is 0.

In consequence, 1) follows if $y^{\{i\},\emptyset}_{\emptyset}, y^{\{i\},\emptyset}_{\emptyset}>0$. Finally, notice 2) is satisfied by construction since

$$
y^{\{i\},\emptyset}_{\emptyset}z^{\{i\},\emptyset}_{I}(y)+y^{\emptyset,\{i\}}_{\emptyset}z^{\emptyset,\{i\}}_{I}(y)=y^{\{i\},\emptyset}_{I}+y^{\emptyset,\{i\}}_{I}=(y_I-y_{I\cup\{i\}})+y_{I\cup\{i\}}=y_I. \tag{A.21}
$$

for $I \subseteq [n]$ with $|I| \leq k$, where $y_\emptyset^{\{i\},\emptyset} + y_\emptyset^{\emptyset,\{i\}} = 1$ since $y_\emptyset = 1$.

Next, we prove the cases where either $y_\emptyset^{\{i\},\emptyset}$ or $y_\emptyset^{\{i\},\emptyset}$ is equal to 0 (notice that both cannot be 0 by Remark A.6 and the fact that $y_\emptyset = 1$). First, consider the case $y_\emptyset^{\{i\},\emptyset} = 0$, i.e., $y_i = 1$. This implies $z^{\emptyset,\{i\}}(y) = y^{\emptyset,\{i\}}$. Since $y \in \mathrm{M}^k(P)$, the equalities (A.18) and (A.20) imply that $z^{\emptyset,\{i\}} \in \mathrm{M}^{k-1}(P)$. Moreover, 2) follows since, for $I \subseteq [n]$ with $|I| \leq k$,

$$y_\emptyset^{\{i\},\emptyset} z_I^{\{i\},\emptyset}(y) + y_\emptyset^{\emptyset,\{i\}} z_I^{\emptyset,\{i\}}(y) = y_I^{\emptyset,\{i\}} = y_{I\cup\{i\}} = y_I. \tag{A.22}$$

where the last equality is implied by Lemma A.9.

An analogous argument can be used for the case $y_\emptyset^{\emptyset,\{i\}} = 0$, i.e., $y_i = 0$. We have that $z^{\{i\},\emptyset}(y) = y^{\{i\},\emptyset}$, hence $z^{\{i\},\emptyset}(y) \in \mathrm{M}^k(P)$ since $y \in \mathrm{M}^k(P)$ and equations (A.17) and (A.19) hold. Condition 2) is satisfied since

$$y_\emptyset^{\{i\},\emptyset} z_I^{\{i\},\emptyset}(y) + y_\emptyset^{\emptyset,\{i\}} z_I^{\emptyset,\{i\}}(y) = y_I^{\{i\},\emptyset} = y_I - y_{I\cup\{i\}} = y_I.$$

for $I \subseteq [n]$ with $|I| \leq k$. The last equality above also follows by Lemma A.9.

1) and 2) imply $y \in \mathrm{M}^k(P)$. We show that $y \in \mathrm{M}^k(P)$. First, we check that $y_\emptyset = 1$. Indeed, we have that for arbitrary $i \in [n]$,

$$y_\emptyset = y_\emptyset^{\{i\},\emptyset} + y_\emptyset^{\emptyset,\{i\}} = 1 \tag{A.23}$$

where the first equality follows from Remark A.6 with $S = \{i\}$, and the second follows from the convex combination assumption 2).

We check that y satisfies constraints SA-1. Consider an arbitrary $T \subseteq [n]$ with $|T| = k+1$ and a 2-partition $T_0 \sqcup T_1 = T$. First, suppose that $T_0 \neq \emptyset$. Let $i \in T_0$ be arbitrary. Then, we may write $(T_0' \sqcup \{i\}) \sqcup T_1 = T$. If $y_\emptyset^{\{i\},\emptyset} > 0$, then equation (A.17) implies that y satisfies the corresponding SA-1 constraint for the 2-partition $(T_0' \sqcup \{i\}) \sqcup T_1 = T$, since $z^{\{i\},\emptyset}(y)$ satisfies the corresponding constraint for the 2-partition $T_0' \sqcup T_1 = T \setminus \{i\}$. On the other hand, if $y_\emptyset^{\{i\},\emptyset} = 0$, then $y_\emptyset^{\emptyset,\{i\}} = 1$ by equation (A.23), and $y_I = y_I^{\emptyset,\{i\}} = y_{I\cup\{i\}}$ for all $I \subseteq [n]$ such that $|I| \leq k$, follows by (A.16). Moreover, by assumption $|T_0 \sqcup T_1| \leq k+1$, which implies that $|T_0' \sqcup T_1| \leq k$. Hence, by Lemma A.9, we have for all $H \subseteq T_0'$,

$$y_{T_1 \cup H \cup \{i\}} = y_{T_1 \cup H}. \tag{A.24}$$

Therefore,

$$\begin{aligned}
\sum_{H \subseteq T_0' \cup \{i\}} (-1)^{|H|} y_{T_1 \cup H} &= \sum_{H \subseteq T_0' \cup \{i\} \,:\, H \ni i} (-1)^{|H|} y_{T_1 \cup H} + \sum_{H \subseteq T_0' \cup \{i\} \,:\, H \not\ni i} (-1)^{|H|} y_{T_1 \cup H} \\
&= \sum_{H \subseteq T_0'} (-1)^{|H|+1} y_{T_1 \cup H \cup \{i\}} + \sum_{H \subseteq T_0'} (-1)^{|H|} y_{T_1 \cup H} \\
&= -\sum_{H \subseteq T_0'} (-1)^{|H|} y_{T_1 \cup H} + \sum_{H \subseteq T_0'} (-1)^{|H|} y_{T_1 \cup H} \quad \text{by (A.24)} \\
&= 0.
\end{aligned}$$

Next, assume that $T_1 \neq \emptyset$. Let $i \in T_1$ be arbitrary. Then, we may write $T_0 \sqcup (T_1' \sqcup \{i\}) = T$. If $y_\emptyset^{\emptyset,\{i\}} > 0$, then $z^{\emptyset,\{i\}}(y) \in M^{k-1}(P)$ and equation (A.18) imply that y satisfies SA-1 constraints. Suppose instead that $y_\emptyset^{\emptyset,\{i\}} = 0$, then equation (A.16) implies that $y_I = y_I^{\{i\},\emptyset}$ for all $I \subseteq [n]$ with $|I| \leq k$. Since $y_I^{\{i\},\emptyset} = y_I - y_{I \cup \{i\}}$, it follows that $y_{I \cup \{i\}} = 0$. Therefore, y satisfies the SA-1 constraint associated to the 2-partition $T_0 \sqcup (T_1' \sqcup \{i\})$. The above shows that y satisfies all the SA-1 constraints.

To check that y satisfies SA-2 constraints the exact same arguments can be used by considering any subset $T \subseteq [n]$ with $|T| = k$ and an arbitrary 2-partition $T_0 \sqcup T_1 = T$. $\square$

Finally, we restate and prove the following general version of the Decomposition Lemma, which appeared as Theorem 3.9 in Section. This result allows us to induce on subsets of variables of cardinality $1 \leq d \leq k$ and thereby express the restriction of y to its first $\binom{n}{\leq k-d+1}$ coordinates as a convex combination of points in lower levels of the Sherali-Adams hierarchy.

Theorem A.11 (Decomposition Lemma)**.** *Let $P \subseteq [0,1]^n$, $d, k \in \mathbb{Z}_{\geq 0}$ such that $d \leq k$ and*

$y \in [0,1]^{\binom{n}{\leq k+1}}$. *Then* $y \in \mathrm{M}^k(P)$ *if and only if for every* $S \subseteq [n]$ *of cardinality* d*, the following hold:*

1. *For every* 2*-partition* $J_0 \sqcup J_1 = S$*, we have* $z^{J_0,J_1}(y) \in \mathrm{M}^{k-d}(P)$ *whenever* $y_\emptyset^{J_0,J_1} > 0$*; and*

2. *The restriction of* y *to its first* $\binom{n}{\leq k-d+1}$ *coordinates can be expressed as the convex combination of points* $z^{J_0,J_1}(y)$*, namely,*

$$y_I = \sum_{J_0,J_1 \subseteq [n]\,:\, J_0 \sqcup J_1 = S} y_\emptyset^{J_0,J_1} z_I^{J_0,J_1}(y), \tag{A.25}$$

for every $I \subseteq [n]$ *with* $|I| \leq k-d+1$.

Proof. The sufficiency follows immediately, since the assumptions hold in particular for the case $d=1$, and Lemma A.10 then implies that $y \in \mathrm{M}^k(P)$. We now prove necessity. We proceed by induction on d, for $1 \leq d \leq k$, with k and n fixed. The base case follows from Lemma A.10. Assume that the statement holds for $d-1$. We prove it holds for d.

By our induction hypothesis we have that $y \in \mathrm{M}^k(P)$ if and only if for any $S \subseteq [n]$ of cardinality $d-1$ the statements 1) and 2) hold. Consider an arbitrary $S \subseteq [n]$ of cardinality d and fix any $i \in S$. We denote $S' \coloneqq S \setminus \{i\}$. Therefore, we have that, for every $I' \subseteq [n]$ with $|I'| \leq k-(d-1)+1$,

$$y_{I'} = \sum_{J_0,J_1 \subseteq [n]\,:\, J_0 \sqcup J_1 = S'} y_\emptyset^{J_0,J_1} z_{I'}^{J_0,J_1}(y). \tag{A.26}$$

For $J_0 \sqcup J_1 = S'$, set $w(J_0,J_1) \coloneqq z^{J_0,J_1}(y) \in \mathrm{M}^{k-(d-1)}(P)$ if $y_\emptyset^{J_0,J_1} > 0$. Since $i \in [n] \setminus S'$ we can assume that $0 < w_\emptyset(J_0,J_1)^{\emptyset,\{i\}} < 1$ is fractional (we later elaborate on the case $w_\emptyset(J_0,J_1)^{\emptyset,\{i\}} \in \{0,1\}$). Hence, for each 2-partition $J_0 \sqcup J_1 = S'$, and corresponding $w(J_0,J_1)$, we can apply Lemma A.10 and obtain: for $I \subseteq [n]$ with $|I| \leq k-d+1$,

$$w_I(J_0,J_1) = w_\emptyset(J_0,J_1)^{\{i\},\emptyset} z_I^{\{i\},\emptyset}(w(J_0,J_1)) + w_\emptyset(J_0,J_1)^{\emptyset,\{i\}} z_I^{\emptyset,\{i\}}(w(J_0,J_1))$$

where $z^{\{i\},\emptyset}(w(J_0, J_1))$ and $z^{\emptyset,\{i\}}(w(J_0, J_1))$ belong to $\mathrm{M}^{k-d}(P)$. Next, note that we can write

$$
\begin{aligned}
z_I^{\{i\},\emptyset}(w(J_0, J_1)) &= z_I^{\{i\},\emptyset}\left(\frac{y^{J_0,J_1}}{y_\emptyset^{J_0,J_1}}\right) \\
&= \frac{1}{y_\emptyset^{J_0,J_1}} \frac{1}{w_\emptyset(J_0, J_1)^{\{i\},\emptyset}} \left(y^{J_0,J_1}\right)_I^{\{i\},\emptyset} \\
&= \frac{1}{y_\emptyset^{J_0,J_1}} \frac{1}{w_\emptyset(J_0, J_1)^{\{i\},\emptyset}} (y_I^{J_0,J_1} - y_{I\cup\{i\}}^{J_0,J_1}) \\
&= \frac{1}{y_\emptyset^{J_0,J_1}} \frac{1}{w_\emptyset(J_0, J_1)^{\{i\},\emptyset}} y_I^{J_0\cup\{i\},J_1}, \qquad \text{by Lemma A.7} \qquad (A.27)
\end{aligned}
$$

and

$$
\begin{aligned}
z_I^{\emptyset,\{i\}}(w(J_0, J_1)) &= z_I^{\emptyset,\{i\}}\left(\frac{y^{J_0,J_1}}{y_\emptyset^{J_0,J_1}}\right) \\
&= \frac{1}{y_\emptyset^{J_0,J_1}} \frac{1}{w_\emptyset(J_0, J_1)^{\emptyset,\{i\}}} \left(y^{J_0,J_1}\right)_I^{\emptyset,\{i\}} \\
&= \frac{1}{y_\emptyset^{J_0,J_1}} \frac{1}{w_\emptyset(J_0, J_1)^{\emptyset,\{i\}}} y_{I\cup\{i\}}^{J_0,J_1} \\
&= \frac{1}{y_\emptyset^{J_0,J_1}} \frac{1}{w_\emptyset(J_0, J_1)^{\emptyset,\{i\}}} y_I^{J_0,J_1\cup\{i\}}. \qquad \text{by Lemma A.7} \qquad (A.28)
\end{aligned}
$$

Moreover, since $z^{\{i\},\emptyset}(w(J_0, J_1))$ and $z^{\emptyset,\{i\}}(w(J_0, J_1))$ belong to $M^{k-d}(P)$, we have that

$$z_\emptyset^{\{i\},\emptyset}(w(J_0, J_1)) = z_\emptyset^{\emptyset,\{i\}}(w(J_0, J_1)) = 1.$$

From equations (A.27)-(A.28) we deduce that

$$y_\emptyset^{J_0,J_1} w_\emptyset(J_0, J_1)^{\{i\},\emptyset} = y_\emptyset^{J_0\cup\{i\},J_1}, \quad y_\emptyset^{J_0,J_1} w_\emptyset(J_0, J_1)^{\emptyset,\{i\}} = y_\emptyset^{J_0,J_1\cup\{i\}}.$$

Hence, $z_I^{\{i\},\emptyset}(w(J_0, J_1)) = z_I^{J_0\cup\{i\},J_1}(y)$ and $z_I^{\emptyset,\{i\}}(w(J_0, J_1)) = z_I^{J_0,J_1\cup\{i\}}(y)$.

Finally, invoking the induction hypothesis (A.26) together with the identities established above, we obtain that for any $I \subseteq [n]$ with $|I| \leq k - d + 1$,

$$
\begin{aligned}
y_I &= \sum_{J_0, J_1 \subseteq [n] \,:\, J_0 \sqcup J_1 = S'} y_\emptyset^{J_0, J_1} w_I(J_0, J_1) \\
&= \sum_{J_0, J_1 \subseteq [n] \,:\, J_0 \sqcup J_1 = S'} y_\emptyset^{J_0, J_1} \Big(w_\emptyset(J_0, J_1)^{\{i\}, \emptyset} z_I^{\{i\}, \emptyset}(w(J_0, J_1)) + w_\emptyset(J_0, J_1)^{\emptyset, \{i\}} z_I^{\emptyset, \{i\}}(w(J_0, J_1)) \Big) \\
&= \sum_{J_0, J_1 \subseteq [n] \,:\, J_0 \sqcup J_1 = S'} y_\emptyset^{J_0 \cup \{i\}, J_1} z_I^{J_0 \cup \{i\}, J_1}(y) + y_\emptyset^{J_0, J_1 \cup \{i\}} z_I^{J_0 \cup \{i\}, J_1}(y) \qquad \text{(A.29)} \\
&= \sum_{T_0, T_1 \subseteq [n] \,:\, T_0 \sqcup T_1 = S} y_\emptyset^{T_0, T_1} z_I^{T_0, T_1}(y).
\end{aligned}
$$

Finally, we verify the case where $w_\emptyset(J_0, J_1)^{\emptyset, \{i\}} = 1$ (the case $w_\emptyset(J_0, J_1)^{\emptyset, \{i\}} = 0$ is analogous). By definition we have that

$$
z_I^{\emptyset, \{i\}}(w(J_0, J_1)) = \frac{1}{y_\emptyset^{J_0, J_1}} y_I^{J_0, J_1 \cup \{i\}}.
$$

By Lemma A.10 $z_I^{\emptyset, \{i\}}(w(J_0, J_1)) \in \mathrm{M}^{k-d}(P)$, hence we deduce that $y_\emptyset^{J_0, J_1 \cup \{i\}} = y_\emptyset^{J_0, J_1} > 0$. Moreover, $y_\emptyset^{J_0, J_1 \cup \{i\}} = y_i^{J_0, J_1}$ and $y_\emptyset^{J_0 \cup \{i\}, J_1} = y_\emptyset^{J_0, J_1} - y_i^{J_0, J_1}$ follow by Lemma A.7. Hence, $y_\emptyset^{J_0 \cup \{i\}, J_1} = 0$. The above also implies $z_I^{\emptyset, \{i\}}(w(J_0, J_1)) = z_I^{J_0, J_1 \cup \{i\}}(y)$. Therefore, the summand in (A.29) can be written as

$$
y_\emptyset^{J_0 \cup \{i\}, J_1} z_I^{J_0 \cup \{i\}, J_1}(y) + y_\emptyset^{J_0, J_1 \cup \{i\}} z_I^{J_0 \cup \{i\}, J_1}(y) = y_\emptyset^{J_0, J_1 \cup \{i\}} z_I^{J_0 \cup \{i\}, J_1}(y).
$$

This concludes the proof. $\square$