Adjustable Robust Two-Stage Polynomial Optimization with Application to AC Optimal Power Flow

Olga Kuryatnikova* Bissan Ghaddar† Daniel K. Molzahn‡

Abstract

In this work, we consider two-stage polynomial optimization problems under uncertainty. In the first stage, one needs to decide upon the values of a subset of optimization variables (control variables). In the second stage, the uncertainty is revealed and the rest of optimization variables (state variables) are set up as a solution to a known system of possibly non-linear equations. This type of problem occurs, for instance, in optimization for dynamical systems, such as electric power systems. We combine tools from polynomial and robust optimization to provide a framework for general adjustable robust polynomial optimization problems. In particular, we propose an iterative algorithm to build a sequence of (approximately) robustly feasible solutions with an improving objective value and verify robust feasibility or infeasibility of the resulting solution under a semialgebraic uncertainty set. At each iteration, the algorithm optimizes over a subset of the feasible set and uses affine approximations of the second-stage equations while preserving the non-linearity of other constraints. The algorithm allows for additional simplifications in case of possibly non-convex quadratic problems under ellipsoidal uncertainty. We implement our approach for AC Optimal Power Flow and demonstrate the performance of our proposed method on Matpower instances.

Keywords: polynomial optimization, adjustable robust optimization, non-convex quadratic optimization, AC optimal power flow, uncertainty in energy systems.

1 Introduction

In many real world applications, data is not completely known in advance and making decisions requires one to consider the uncertainty in the data. There are two broadly used approaches to deal with data uncertainty in optimization, namely stochastic optimization and robust optimization. Methods for stochastic optimization assume that the probability distribution of the uncertain parameters is known or can be estimated in advance. Conversely, robust optimization does not require any knowledge about the distribution of the uncertain data but instead assumes that the uncertain data lies in a predefined set of scenarios and that constraints have to be satisfied for any realization of the uncertain data in that set. As an extension to robust optimization, the methodology of *adjustable* robust optimization (ARO) was introduced in Ben-Tal et al. 2004. In particular, there are cases of optimization models when there are multi-stage decisions where some variables represent decisions that must be made before the actual realization of the uncertain data becomes known while other variables can adjust themselves after the realization of the uncertain data. Two-stage ARO is a suitable way to model such a decision-making problem as it consists of two types

*Econometric Institute, Erasmus University Rotterdam, the Netherlands, kuryatnikova@ese.eur.nl
†Ivey Business School, Western University, bghaddar@ivey.ca
‡School of Electrical and Computer Engineering, Georgia Institute of Technology, molzahn@gatech.edu

of variables: the first-stage variables that are non-adjustable and the second-stage variables that are adjustable based on the first stage. We refer to these variables as control and state variables, respectively, since this setting is typical for optimal control problems. Furthermore, an Adjustable Robust Counterpart (ARC) of a problem generally yields a better objective function value than that of the robust counterpart because it gives rise to more flexible decisions and thus can be less conservative than robust optimization. This approach spans a wide range of applications, including power systems (Lorca and Sun 2016, Lee et al. 2019), gas networks (Aßmann et al. 2018), and inventory planning (Yanıkoğlu et al. 2019, Ardestani-Jaafari and Delage 2016).

Several approaches have been developed to solve ARO problems. The proposed solution approaches for ARO can be classified into two main categories: approximations and exact reformulations. The hardest part of the problem is defining the relation between the first- and second-stage decision variables, called *decision rules*, which are sometimes not specified and have to be optimized. For instance, Ben-Tal et al. 2004 proposes an approximation approach where the second-stage decision variables can be written as affine functions of the uncertain parameters. Exact reformulations of a robust counterpart can be obtained using approaches such as vertex enumeration algorithms (Bienstock and Özbay 2008), cutting plane algorithms (Bertsimas et al. 2012), or reformulation of each constraint (Zeng and Zhao 2013). There are several special cases as well that allow for an exact reformulation. Two-stage adjustable robust optimization problems are generally computationally intractable unless they are limited to some specific decision rules. Recent research in this area focused on studying two-stage adjustable robust linear programs (LP) with affine or piecewise affine decision rules, see Georghiou et al. 2020 for the latest developments. Such programs can be reformulated as single-stage problems that are computationally tractable. On the other hand, transforming two-stage adjustable linear programs with quadratic decision rules into single-stage robust problems yields numerically intractable non-convex quadratic optimization problems unless the decision rules or uncertainty sets are restricted to a certain class of rules or sets, respectively.

It is a frequent problem in adjustable robust optimization that the second-stage rules are not known and one therefore needs to determine the rules which correspond to the best objective. However, in certain problems occurring in practice, the second-stage variables need to be decided according to a set of predefined decision rules modeled by a set of equalities. In that case, it is not possible to choose any rule and stay robustly feasible. Some alternatives are to use solvability theorems (e.g., Farkas' Lemma) or to substitute the variables and eliminate these equalities, which reduces the problem to the classical robust optimization problem. These options are not possible in cases with non-linear equalities without known analytical solutions. This paper focuses on such cases by replacing the non-linear functions defining the equalities by their piece-wise affine approximations that would reflect the original functions as closely as needed. Recently, some progress has been made in that direction such as the work presented in Ardestani-Jaafari and Delage 2016, Aßmann et al. 2018, Lee et al. 2019, and Isenberg et al. 2021.

With rapidly increasing uncertainties in both the demand and supply of energy, networked infrastructure problems are important applications of ARO. Examples of relevant networked infrastructures include electricity (see Bienstock et al. 2014), natural gas (see Misra et al. 2020), and water (see Stuhlmacher and Mathieu 2020). Operators must ensure that these systems remain in acceptable states despite significant uncertainties while simultaneously considering performance criteria such as operating costs. ARO provides a natural approach that is often applied to balance these potentially competing concerns in networked infrastructure problems, as discussed in Yanıkoğlu et al. 2019 and Misra et al. 2020.

As an illustrative example of ARO, this paper focuses on so-called *robust AC optimal power flow* (ACOPF) problems which provide minimum cost operating points for electric power systems. ACOPF problems can be formulated as polynomial optimization (PO) problems, as discussed in Ghaddar et al. 2016. Even in the absence of uncertainties, ACOPF problems are non-convex and NP-Hard (see Bienstock and Verma 2019). To address these challenges, our approach builds on prior successes in applying polynomial optimization theory in power system optimization, e.g., the work described in Kuang et al. 2017, Ghaddar et al. 2016, and Josz and Molzahn 2018, which show that low-order relaxations in moment/sum-of-squares hierarchies are capable of computing global optima for a variety of difficult ACOPF problems. Extending this prior work, the approach proposed in this paper computes operating points that are certified to be robustly feasible with respect to the engineering inequality constraints in ACOPF problems. This allows us to leverage the substantial and ongoing advances in polynomial optimization techniques to improve robust ACOPF algorithms. We also note that, to the best of our knowledge, none of the existing literature investigates the question of infeasibility of robust ACOPF problems. We address this question for certain robust ACOPF problems using a polynomial optimization technique.

We next describe the contribution of our work with respect to what has been done recently in the literature.

1. We present a general framework for solving two-stage polynomial ARO problems where the second-stage variables are linked to the first-stage variables through a set of equality constraints. Once the first-stage variables are fixed and the uncertainty is realized, finding the second-stage variables amounts to solving a system of polynomial equalities. The framework approximates such an ARO problem by a sequence of PO problems in the control variables.

2. Verifying robust feasibility in the presence of general polynomial equalities is a challenging problem. Therefore, we address the robust feasibility of the equality and inequality constraints separately. We suggest using piece-wise affine approximations of the non-linear equalities with respect to the state variables to allow for closed-form decision rules. We partition the domain of the state variables into subsets small enough to represent each polynomial equation well by its first-order Taylor approximation. Thus, the non-linear equations are replaced by linear equations for every subset in the partition. Next we eliminate these equations and obtain a classical robust PO problem with control and uncertainty variables instead of an ARO. The obtained problem is then reformulated (or conservatively approximated) as a PO problem with tractable conic constraints over the control variables only. To speed-up the solution process, we design a dynamic scheme to generate the piece-wise affine approximations. To analyze the feasibility or infeasibility of the approximate solution with respect to the original problem, we generalize the techniques from Aßmann et al. 2018.

3. We propose an algorithm to find locally optimal solutions for semidefinite programs (SDP) with non-convex equalities. Such problems occur as a result of applying our approximation technique to ARO problems with *general equality* constraints, *quadratic inequality* constraints, and *ellipsoidal uncertainty*. We apply the proposed framework to the AC optimal power flow (ACOPF) problem, which is a non-convex quadratically constrained quadratic problem. The results demonstrate the effectiveness of our approach on small to moderate size instances from the literature ranging from 5 to 118 buses, with the potential to consider larger instances.

Next we outline some differences between our approach and existing methods. First, we do not consider any assumptions on convexity and concavity (cf. Georghiou et al. 2020) and hence target general polynomial ARO problems. Additionally, we do not use convex approximations of inequality constraints, such as in Lasserre 2015, Lee et al. 2020, and Lorca and Sun 2018. Our approach is also different from other robust optimization methods where non-linear constraints are linearized, such as Louca and Bitar 2019. First, we linearize the equality constraints while keeping the original non-linear inequalities. Second, we linearize locally, and our approximations are closely related to the original constraints via Taylor series. The results in this paper are distinct from Molzahn and Roald 2018 since in the latter work a robust solution is obtained by iteratively tightening the inequality constraints. In Isenberg et al. 2021, the authors also tackle general non-linear optimization problems but use an alternative formulation of ARO problems and thus a distinct solution strategy. Our approach is close in spirit to the approach in Roald and Andersson 2018. However, the authors of the latter paper consider chance constraints in a particular problem type using a linearization approach that iteratively adjusts the inequality constraints. Finally, our approach is complementary to approaches where the uncertainty set is partitioned into subsets, such as Postek and Hertog 2016. Since we approximate the second-stage decision rules not with respect to the uncertainty but rather with respect to the state variables, the approaches could be combined for large uncertainty sets; see Remark 1 in Section 2.

This paper is organized as follows. In Section 2, we present the formulation of the problem we are interested in and motivate our solution approach. In Section 3, we describe the proposed dynamic algorithm in detail. In Section 4, we certify feasibility and infeasibility of the solution. In Section 5, we evaluate the proposed approaches on ACOPF instances. Finally, we give conclusions in Section 6 and discuss how the approach presented could be further improved via future work.

2 Problem formulation and general framework

For a matrix A, we denote its range by $\mathcal{R}(A)$. We denote the space of $n\times n$ symmetric matrices by $\mathbb{S}^n$. For $A, B \in \mathbb{S}^n$, the trace inner product of A and B is denoted by $\langle A, B\rangle := \mathrm{trace}(AB)$. We use the notation $[n]$ for the set $\{1, \dots, n\}$. We denote by $\mathbb{R}[x]$ the set of polynomials in variables x with real coefficients, and by $\mathbb{R}_d[x]$ (respectively $\mathbb{R}_{=d}[x]$) the subset of these polynomials of degree not larger than (resp. equal to) d. For $S \subseteq \mathbb{R}^n$, let $\mathcal{P}(S) = \{p \in \mathbb{R}[x] : p(x) \geq 0 \text{ for all } x \in S\}$ be the set of polynomials non-negative on S, and $\mathcal{P}^+(S) = \{p \in \mathbb{R}[x] : p(x) > 0 \text{ for all } x \in S\}$ be the set of polynomials positive on S. Throughout the paper, for a vector V of length n, we denote the i^{th} entry of V by V_i. We denote by $\mathrm{Diag}(V)$ the operator that creates a diagonal matrix with V as the diagonal. We say that a sequence $(a_k)_{k=1}^{\infty}$ converges *linearly* to a if $\lim_{k\to\infty} \frac{\|a_{k+1}-a\|}{\|a_k-a\|} < 1$. Finally, in this work, we sometimes use the Lebesgue measure as the standard measure of volume of sets. For the detailed definition and properties of the Lebesgue measure we refer the readers to Chapter 1 of Folland 1999.

We let n_ζ, n_y, n_x, m_{eq}, and m_{in} be natural numbers and consider $\Omega \in \mathbb{R}^{n_\zeta}$, $S_y \in \mathbb{R}^{n_y}$, and $S_x \in \mathbb{R}^{n_x}$. Consider the *polynomial* mappings $f : \mathbb{R}^{n_y} \to \mathbb{R}$, $G : \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{m_{in}}$, and $L : \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{m_{eq}}$. We aim to solve the following problem:

Problem 1 (Two-stage adjustable robust polynomial optimization problem)**.**

$$
\begin{aligned}
\alpha = \inf_{y} \quad & f(y) \\
\text{s.t.} \quad & y \in S_y \\
& \textit{and for any } \zeta \in \Omega \textit{ there exists } x \textit{ such that the following holds:} \\
& L_i(y, \zeta, x) = 0 \textit{ for all } i \in [m_{eq}] && \text{(1a)} \\
& G_i(y, \zeta, x) \geq 0 \textit{ for all } i \in [m_{in}] && \text{(1b)} \\
& x \in S_x. && \text{(1c)}
\end{aligned}
$$

Problem (1) is an adjustable robust optimization problem with two stages and the *uncertain parameter* ζ. The first stage happens before the uncertainty realization. At this stage, one assigns values to the *control* variables $y \in S_y$. The second stage happens after the uncertainty realization. At this stage, one has to choose the best feasible value of the *state* variables $x \in S_x$ for the given uncertainty realization. The goal is to select a value for $y \in S_y$ such that there would be a feasible solution $x \in S_x$ in the second stage for any uncertainty realization $\zeta \in \Omega$. The set of control variables y can include the auxiliary variables coming from the worst-case reformulation of the part of the objective function that depends on x; see Section 5 for an example. Any solution to problem (1) is feasible for the underlying *nominal problem* and robust against potential uncertainty. We obtain the nominal problem by setting $\zeta = 0$ in problem (1). For the presented problem, we work under the following rather general assumptions.

Assumption 1 (Assumptions about problem structure)**.**

(a) S_x*,* S_y*, and* Ω *are compact semialgebraic sets with* S_x *defined by inequalities.*

If the definition of S_x *contains equalities, one can move them to* (1a) *as a preprocessing step.*

(b) Equality constraints (1a) *are separable in* (x, ζ) *and* (y, ζ)*, i.e.,*

$$L(y, \zeta, x) = L_1(\zeta, x) + L_2(y, \zeta), \tag{2}$$

and one of the following holds: either $L_1(\zeta, x)$ *is independent of* ζ*, or* Ω *is small in the sense that for any* (ζ_0, x_0)*, the first-order Taylor approximation of* $L_1(\zeta, x)$ *around* (ζ_0, x_0) *is sufficiently close to* $L_1(\zeta, x)$ *for all* $\zeta \in \Omega$ *and some open neighbourhood in* S_x *around* x_0*. The assumption is used for simplicity, as our approach is straightforward to generalize; see Remark 1. Notice that there is no separability requirement for the inequality constraints* (1b)*.*

(c) $n_x = n_{eq}$ *and there is no redundant equality constraints in* (1a)*.*

If $n_x \leq n_{eq}$ *and there are redundant constraints, they should be eliminated as a preprocessing step. If* $n_x > n_{eq}$*, then some state variables are free. One could add these variables to the pool of control variables and redefine the set* S_y *as a preprocessing step.*

Remark 1. *Assumption 1 (b) is used to restrict our approach to approximations over* S_x *and work with the functions of the form* $Ax + L_2(y, \zeta)$ *so that we can eliminate* x*. The approach is straightforward to generalize to cases with products of* x *with* y *or larger uncertainty sets* Ω *by constructing approximations over the set* $S_x \times S_y \times S_\zeta$*. In this case, one can, for instance, use the results from Postek and Hertog 2016 and related research. Our assumption implies that splitting the set* S_x *as done later in Algorithms 1–3 is enough if either there are no products of the uncertainty with state variables or the values of* ζ *appearing in products with* x *are small. This assumption reflects the specifics of the uncertainty in the ACOPF problem, which inspired this paper and for which we test our approach in Section 5.2. In ACOPF, the uncertainty usually comes from either renewable power sources, deviations in power demand, or imprecise values in admittance matrix. The first two types of uncertainty are not multiplied with* x *and can lead to large fluctuations. The last type of uncertainty results in products of* ζ *and* x*, but the uncertainty size is expected to be small, i.e., network parameter values are typically well characterized.*

We next provide a general description of our solution approach for problem (1). We begin with the intuition behind the approach. If we could analytically solve equalities (1a) for the second-stage variables x, we would express x as a function of y and ζ. We call such a function *the second-stage decision rule*. Assuming that the equalities (1a) in problem (1) are such that there exists a known second-stage decision rule $h : \mathbb{R}^{n_y \times n_\zeta} \to \mathbb{R}^{n_x}$, then (1) is equivalent to the following problem:

Problem 3.

$$\begin{aligned}
\alpha = \inf_{y} \quad & f(y) \\
\text{s.t.} \quad & y \in S_y \\
& G_i(y, \zeta, h(y, \zeta)) \geq 0 \textit{ for all } \zeta \in \Omega,\ i \in [m_{in}] \\
& h(y, \zeta) \in S_x \textit{ for all } \zeta \in \Omega.
\end{aligned}$$

That is, if a known second-stage decision rule exists, then we could substitute it in the problem and eliminate the equalities and state variables to obtain a classical (not adjustable) robust optimization problem. Clearly, if the second-stage variables are determined from a system of non-linear equalities, there might be no general rule in the problem. However, such rules can exist on small subsets of $S_x \times S_y \times S_\zeta$ according to the implicit function theorem (see, e.g., Spivak 1995). We work on such subsets and replace all terms in L which include monomials in x of degrees larger than one by their linear approximations. Recall that the functions L have the form described in (2). As a result, on each subset we obtain $L \approx Ax + L_2(y, \zeta)$, for some matrix A. With such approximations, explicit rules are either easy to derive or clearly do not exist.

Our goal is to obtain the adjustable robust counterpart (ARC) that has a form of *a classical PO problem in control variables y*. That is, we eliminate all variables from the problem except for the control variables y, since the latter are in fact the only variables we decide upon. This procedure brings two advantages: first, we obtain a classical PO problem instead of the original PO problem with infinitely many constraints, one for each realization of ζ. Second, even though the new problem might have a higher degree than the original one (at most the product of the degrees of the original problem in y and x), the number of variables might be substantially reduced since the number of control variables in the applications we consider is substantially smaller than the number of state variables. Moreover, by increasing the number of variables in ARC, one can make the degree of ARC equal to the degree of the original problem. If there exists a solution algorithm for the original non-robust PO problem (which is usually the case if one decides to solve a robust version instead of the nominal one), similar algorithms could be applicable to ARC.

Next, we present an approach to obtain decision rules that are linear in x which we substitute in problem (1) to eliminate the equality constraints and solve a problem of the form (3). We first approximate the non-linear equality constraints (1a) by piecewise affine functions and then build conservative approximations of the inequality constraints (1b) and (1c) using PO.

3 Solution approach

3.1 Piecewise affine approximations of equality constraints

In the first step of our approach, we replace the non-linear terms in L that involve x by their affine approximations. We say that two continuous maps f and g on a compact set $A \subset \mathbb{R}^n$ are *ε-close to each other* if $\sup_{x \in A} \|f(x) - g(x)\| \leq \varepsilon$ for some given norm $\|\cdot\|$. Our piecewise approximation decision is based on the following general result.

Proposition 1. *Let $A \subset \mathbb{R}^n$ be compact. For any polynomial map $F : A \to \mathbb{R}^m$ and $\varepsilon > 0$, there exists a piecewise affine map that is ε-close to F.*

Proof. Since $A \subset \mathbb{R}^n$ is compact, by the Stone-Weierstrass theorem, the set of piecewise constant functions on A is dense in the set of continuous functions on A, see, e.g., Theorem 4.45 in Folland 1999. Hence the set of piecewise affine functions is dense as well. Therefore, for any $\varepsilon > 0$ and each polynomial $F_i,\ i \in [m]$ in F, there exists a piecewise affine function that is $\frac{\varepsilon}{m}$-close to F_i. Combining piecewise affine approximations for all $i \in [m]$ would provide a piecewise affine map that is ε-close to F. □

In this work, we illustrate the idea of piecewise affine equality constraint approximations using first-order Taylor approximations. In addition to its simplicity, the main advantage of the Taylor approximation is its good fit for the original function around the approximation point. Under Assumption 1 (b), we only need to approximate the part of L that is non-linear in (ζ, x) and we can keep the rest of the non-linearities in the equality constraints. If the Jacobian of the linear approximation with respect to (ζ, x) is invertible at the point of approximation (ζ_0, x_0), then one can replace the non-linear equality constraints by their Taylor approximations within each subset and solve the resulting linear system to obtain decision rules (an approximate solution $x = h(y, \zeta)$) that are close to the true decision rules. Constructing a high-precision approximation with other desirable properties, e.g., conservativeness with respect to the inequality constraints after variable substitution, is a topic for a separate work, and we keep this topic as a question for future research.

Since either Ω is small or ζ is independent of x, it is enough to partition S_x into J small compact subsets $S_x^j,\ j \in [J]$, over which we approximate the original functions in L using the Taylor expansion. Let $S_x = \bigcup_{j=1}^{J} S_x^j$ for some $J > 0$. Recall that $L(y, \zeta, z) = L_1(\zeta, x) + L_2(y, \zeta)$. Denote by $T(x_0, \zeta_0, L)(\zeta, x)$ the Taylor approximation in (x_0, ζ_0) of $L_1(\zeta, x)$. Now, we approximate the equality constraints (1a) on $S_x^j \times \Omega$ by the following J systems of linear equalities in x:

$$L(y, \zeta, x) \approx T^j(x_0, \zeta_0, L)(\zeta, x) + L_2^j(y, \zeta) = A^j x + \hat{L}_2^j(y, \zeta) = 0 \text{ for all } j \in [J], \tag{4}$$

where $\hat{L}_2^j(y, \zeta)$ is $L_2^j(y, \zeta)$ with additional terms from $T^j(x_0, \zeta_0, L)(\zeta, x)$. If A^j is not invertible, there might be no second-stage decision rule according to the implicit function theorem (see, e.g., Spivak 1995, Theorem 2-12). In this case, the optimization subproblem over S_x^j might be infeasible, and thus we move from the set S^j to the next set S^{j+1}. If the matrix A^j from (4) is invertible, defining its inverse as A_{inv}^j, we can eliminate all state variables x from the problem using (4) and Assumption 1 (c). This process is described in the following algorithm.

Algorithm 1: Piecewise affine approximation of problem (1)

---

1 Partition S_x into subsets $S_x^1, \dots, S_x^J$;

2 Split problem (1) into J separate optimization subproblems over $x \in S_x^j,\ j \in [J]$;

3 Replace the non-linear equality constraints $L(y, \zeta, x)$ in each subproblem by (4);

4 Eliminate x from the subproblem, solve for all $j \in [J]$;

Problem 5.

$$\begin{aligned} \alpha_j = \inf_{y} \quad & f(y) + t && \text{(5a)} \\ \text{s. t.} \quad & y \in S_y, \\ & G_i\left(y, \zeta, A_{inv}^j \hat{L}_2^j(y, \zeta)\right) \geq 0 \textit{ for all } \zeta \in \Omega,\ i \in [m_{in}] && \text{(5b)} \\ & A_{inv}^j \hat{L}_2^j(y, \zeta) \in S_x^j \textit{ for all } \zeta \in \Omega, && \text{(5c)} \end{aligned}$$

and α_j is set to ∞ if the corresponding problem is infeasible;

5 The minimal optimal value among all subproblems and the corresponding solution are the optimal value and solution to the piecewise affine approximation, and $\alpha \approx \min_{j \in [J]} \alpha_j$;

If the original problem (1) is linear in x and contains no products between x and ζ or y, then no splitting is needed and problem (5) can be solved only once over the whole S_x. We also note that using problem (5) requires a trade-off since each such problem considers one affine piece at a time. On the one hand, we are interested in finer approximations to reflect the equality constraints better. On the other hand, if the approximations are too fine, problem (5) will be infeasible for realistic uncertainty sets. We have to balance out these two factors, and more precise analysis of their influence is needed.

If Ω is not a finite set, problem (5) is a PO problem with infinitely many constraints. Next, we reformulate this problem into a PO problem with finitely many constraints by including additional convex constraints which are tractable using state-of-the-art optimization solvers.

3.2 Conservative approximations of inequality constraints

To solve the piecewise affine approximation, we need to solve subproblems (5) for all $j \in [J]$. Each subproblem has a number of inequality constraints of the following general form:

$$h(y, \zeta) \geq 0 \text{ for all } \zeta \in \Omega. \tag{6}$$

By Assumption 1 (a), we can write

$$\Omega = \{\zeta \in \mathbb{R}^{n_\zeta} : g_j(\zeta) \geq 0,\ j \in [m_\zeta]\}.$$

Our goal in this section is to eliminate the variables ζ from (6) while keeping the shape of the polynomial h in y unchanged. In this way, we obtain a problem amenable to existing optimization approaches. Morever, this problem may have significantly fewer variables than the original problem. Since (6) concerns the non-negativity of a polynomial on the set Ω, whose description does not involve y, in the first part of this section we treat y as a parameter and answer the question of when a polynomial in ζ is non-negative on the set Ω. The answer to this question leads to elimination of ζ from problem (5). After that, we stop treating y as a parameter, and y becomes the only variable left in the modified optimization problem. As discussed in Proposition 2, it is straightforward to reformulate (6) if the degree of ζ in h or $g_1, \ldots, g_{n_\zeta}$ is equal to one.

Proposition 2 (Constraints linear in uncertainty and polyhedral uncertainty sets)**.** *Let $h(y, \zeta) := h_1(y) + \zeta^\top h_2(y)$ and $\Omega := \{\zeta \in \mathbb{R}^{n_\zeta} : A\zeta \leq b\}$. Then (6) holds if and only if there exist $z \geq 0$ such that $b^\top z - h_1(y) \leq 0,\ A^\top z + h_2(y) = 0$.*

Proof. Since $h(y, \zeta) \geq 0$ for all $\zeta \in \Omega$ if and only if $\inf_{\zeta \in \Omega} h(y, \zeta) \geq 0$, the result follows by LP duality. □

When the degree of ζ in h or $g_1, \ldots, g_{n_\zeta}$ is greater or equal than two, (6) is a hard but classical polynomial constraint, for which the following result is well-known (see Lasserre 2007).

Lemma 1. *For $r = 0, 1, 2, \ldots,$ let $\mathcal{K}^r(\Omega) \subset \mathcal{P}(\Omega)$, and let $\bigcup_{r=0}^{\infty} \mathcal{K}^r(\Omega) \supseteq \mathcal{P}(\Omega)$. For $y \in S_y$, inequality (6) holds if and only if there exists $r \geq 0$ such that*

$$h(y, \zeta) \in \mathcal{K}^r(\Omega). \tag{7}$$

The main idea for solving PO problems is to relax $\mathcal{P}(\Omega)$ to $\mathcal{K}^r(\Omega)$ to be able to solve an easier problem. For a given r, a set $\mathcal{K}^r(\Omega) \subset \mathcal{P}(\Omega)$ is called *a certificate of non-negativity of polynomials on a semialgebraic set* Ω. We define

$$\mathcal{K}^r(\Omega) := \left\{p \in \mathbb{R}[x] : p = F(g_1, \ldots, g_{m_\zeta}, r),\ \text{where } F \text{ is non-negative on } \Omega\right\}.$$

Then (7) is equivalent to

$$h(y,\zeta) = F(g_1(\zeta), \ldots, g_{m_\zeta}(\zeta), r) \text{ for all } \zeta. \tag{8}$$

Finding a certificate of non-negativity that provides a good approximation of $\mathcal{P}(\Omega)$ is a key factor in obtaining a good bound on the problem. At the same time, we need a tractable approximation that can be solved efficiently using state-of-the-art methods. As an example, consider $F(g_1, \ldots, g_{m_\zeta}, r) = \sum_{i=1}^{m_\zeta} c_i g_i$ where $c_i \in \mathbb{R}_+$. Lemma 1 implies that replacing (6) by (7) results in a *strengthening* of the initial constraint (6). Moreover, if $\bigcup_{r=0}^{\infty} \mathcal{K}^r(\Omega) \supseteq \mathcal{P}(\Omega)$, then, for sufficiently large r, replacing (6) by (7) results in a *reformulation* of the initial constraint (6).

Based on (8), using a certificate is equivalent to saying that the coefficients in ζ of h and F are equal to each other. Writing down the coefficient-wise equalities, we can eliminate ζ from the problem. As a result, our approach transforms each subproblem (5) into an optimization problem that is polynomial in the control variables y. This transformation makes sense if the certificate F has a representation amenable for optimization purposes. For instance, $F(g_1, \ldots, g_{m_\zeta}, r) = \sum_{i=1}^{m_\zeta} c_i g_i,\ c_i \in \mathbb{R}_+$ only involves unknown constant terms and would not create any difficulties for solving such optimization problems.

Proposition 3. *Independent of the choice of the certificate, the maximum degree of y after using this certificate is the product of the degrees of x and y in the initial problem* (1)*.*

Proposition 3 implies that if the initial problem is linear in either x or y, the degree of the strengthening is not larger than the degree of the original problem. In particular, if problem (1) is linear in (x, y), we obtain *the classical robust LP* (see, e.g., Ben-Tal et al. 2004), with an additional check for the rank of the matrix A^j in (4).

Some certificates of non-negativity exist for all *non-negative* polynomials. However, such certificates ask for additional structure from (6). For instance, there might be requirements on the geometry of the set (Schweighofer 2005), derivatives of the defining polynomials (Marshall 2006), types of the defining polynomials (Pólik and Terlaky 2007), or zeros of the defining polynomials (Scheiderer 2005). On the other hand, there is a large number of certificates for which

$$\bigcup_{r=0}^{\infty} \mathcal{K}^r(\Omega) \supseteq \mathcal{P}^+(\Omega). \tag{9}$$

With such certificates, any y for which (7) holds is robustly feasible. The solution may be conservative if, for the optimal y^*, it happens to be the case that $h(y^*, \zeta) = 0$ for some $\zeta \in \Omega$. In general, existence of such situations can be tested by making the uncertainty set a little larger or the inequality constraints slightly tighter and checking whether the optimal y changes. Such solutions are robustly feasible, but any small perturbation of the uncertainty set can make them infeasible. Hence, it may be reasonable to eliminate such solutions.

In this work, we illustrate our approach using the most common certificate of non-negativity for which (9) holds under some non-restrictive conditions. This is the certificate by Putinar 1993, which is the dual equivalent of the well-known Lasserre's hierarchy proposed in Lasserre 2001. From here on we call constraints of the form $x_i - l_i \geq 0,\ u_i - x_i \leq 0,\ i \in [n]$ for some $u, l \in \mathbb{R}^n$ *box constraints* and constraints of the form $u - \|x\|^2 \geq 0$ *ball constraints.*

Theorem 1 (Putinar's Positivstellenzatz presented in Putinar 1993)**.** *Let $g_1, \ldots, g_{m_\zeta} \in \mathbb{R}[x]$ be such that $\Omega = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_{m_\zeta}(x) \geq 0\}$ is a non-empty compact set. Assume that ball or box constraints are included in the description of the set. If $h \in \mathcal{P}^+(\Omega)$, then there exists $r \geq 0$*

such that $h \in \mathcal{K}^r(\Omega)$, where

$$\mathcal{K}^r(\Omega) = \left\{ g \in \mathbb{R}[\zeta] : g = \sigma_0 + \sum_{j=1}^{m_\zeta} \sigma_j g_j,\ \sigma_0, \ldots, \sigma_{m_\zeta} \in \mathbb{R}_{2r}[\zeta] \text{ are sums-of-squares} \right\}$$
$$= \sum_{j=0}^{m_\zeta} m_r(\zeta)^\top S_k m_r(\zeta),\ S_k \succeq 0 \text{ for all } k \in \{0, \ldots, m_\zeta\}, \tag{10}$$

$m_r(\zeta)$ is the vector of all monomials in ζ of degree up to r.

A given polynomial $p \in \mathbb{R}_{2r}[x]$ is a *sum-of-squares polynomial (SOS)* if $p(x) = \sum_{i \leq l} q_i(x)^2$ for some $q_1, \ldots, q_l \in \mathbb{R}_{2r}[x]$, $l \in \mathbb{N}$. SOS certificates are actively used in optimization since they can be written using semidefinite programming (SDP). It is easy to show that a polynomial $p \in \mathbb{R}_{2r}[x]$ is an SOS if and only if $p(x) = m_r(x)^\top S\, m_r(x)$, where $S \succeq 0$ and $m_r(x)$ is the vector of all monomials in x of degree up to r. In turn, the optimal values and solutions of SDP problems can be approximated to any chosen precision using interior point methods. Applying (10) to (6), we obtain

$$h(y, \zeta) = \sum_{j=0}^{m_\zeta} m_r(\zeta)^\top S_k\, m_r(\zeta),\ S_k \succeq 0 \text{ for all } k \in \{0, \ldots, m_\zeta\}, \tag{11}$$

where $m_r(\zeta)$ is the vector of all monomials in ζ of degree up to r. The size of the SDP in SOS certificates is defined by the number of monomials in the given set of variables, and thus grows exponentially in the number of variables and the degree. Hence, SOS certificates can become computationally burdensome when the degree of the problem at hand is high or the problem has many variables. Many other certificates have been proposed that use alternatives to SOS, such as non-negative constants as inPowers and Reznick 2001 and Dickinson and Povh 2018, sums of non-negative circuit polynomials (SONC) as in Dressler et al. 2017, sums of arithmetic-geometric-mean-exponential (SAGE) polynomials such as the work of Chandrasekaran and Shah 2016, and copositive polynomials as done by Kuryatnikova et al. 2019. The alternatives to SOS can be especially suitable if the underlying problem has a higher degree or some inherent sparsity structure.

To prove the next result, we need the celebrated *S-lemma* by Yakubovich 1977.

Theorem 2 (Yakubovich 1977)**.** *Let $p, q \in \mathbb{R}_2[x]$ and suppose there is $x \in \mathbb{R}^n$ such that $g(x) < 0$. Then the following two statements are equivalent:*

1. *$p(x) \geq 0$ for all $x \in \mathbb{R}^n$ such that $q(x) \geq 0$.*
2. *There is $\lambda \in \mathbb{R}_+$ such that $p(x) - \lambda q(x) \geq 0$ for all $x \in \mathbb{R}^n$.*

Proposition 4 (Non-linear constraints in uncertainty and general uncertainty sets)**.** *Let $y \in S_y$, $j \in [J]$ and $r \geq 0$. First,* (11) *implies* (6)*. Let the description of Ω include ball or box constraints (imposed artificially if needed). If one of the following holds, then there exists $r > 0$ such that* (6) *implies* (11)*.*

(a) $h(y, \zeta) > 0$ for all $\zeta \in \Omega$;

(b) $h(y, \zeta)$ is quadratic in ζ, Ω is defined by one quadratic constraint $g_1(\zeta) \geq 0$ and there is $x \in \mathbb{R}^n$ such that $g_1(x) < 0$; in this case $r = 1$;

(c) For all combinations of $h, g_1, \ldots, g_{m_\zeta}$, except for a subset with Lebesgue measure zero.

Proof. First, (11) implies (6) by construction. Next, statement (a) holds by Lemma 1, and statement (b) follows from Theorem 2 since a quadratic function is always non-negative if and only it is an SOS, see, e.g., Lasserre 2001. Using the latter argument, Theorem 2 implies the existence of the Putinar's certificate (11) for p on $q(x) \geq 0$ for $r = 1$. Finally, item (c) is proven in Nie 2014. ☐

We use the S-lemma as an illustration as it requires minimal assumptions, but there exists many generalizations of the S-lemma, such as the work of Pólik and Terlaky 2007. We conclude this subsection by showing how to use alternatives to SOS and presenting the general form of the resulting PO problem in the control variables.

Theorem 3 (Corollary 5 by Kuryatnikova et al. 2019)**.** *Let $\mathcal{K}$ where $\mathbb{R}_+ \subseteq \mathcal{K} \subset \mathcal{P}(\mathbb{R}^n_+)$ be given. Let $g_1, \ldots, g_m \in \mathbb{R}[x]$ be such that $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$ is a non-empty compact set. Let $N \in \mathbb{R}^n$ and $M \in \mathbb{R}$ be such that $S \subseteq \{x \in \mathbb{R}^n : N \leq x,\, e^\top x \leq M\}$. For any $h \in \mathcal{P}^+(S)$ there exist $r \geq 0$ and $c_{\alpha,\beta,\gamma} \in \mathcal{K}$ for $(\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}_r$ such that*

$$p(x) = \sum_{(\alpha,\beta,\gamma)\in\mathbb{N}^{n+m+1}_r} c_{\alpha,\beta,\gamma}(x)(x - N)^\alpha g(x)^\beta (M - e^\top x)^\gamma, \tag{12}$$

where $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

In the theorem above, $\mathcal{K}$ can be, for example, $\mathbb{R}_+$, SOS, SONC, or SAGE polynomials. Using general certificates of non-negativity, such as Theorem 3, we obtain Algorithm 2 that represents the adjustable counterpart of the linear approximation described in Algorithm 1.

Algorithm 2: Adjustable robust counterpart (ARC) of Algorithm 1

---

1 Partition S_x into subsets $S_x^1, \ldots, S_x^J$;

2 Split problem (1) into J separate optimization subproblems over $x \in S_x^j,\ j \in [J]$;

3 Fix r and solve the following optimization subproblem for each $j \in [J]$;

Problem 13.

$$\begin{aligned} \alpha_j^r = \inf_{y,S} \quad & f(y) && (13a)\\ \text{s.t.} \quad & y \in S_y, && (13b)\\ & H(y) = F(S) && (13c)\\ & S \in \mathcal{C}_r, && (13d) \end{aligned}$$

where H is a polynomial mapping, $F : \mathcal{C}_r \to \mathbb{R}^l$ is a linear mapping, $\mathcal{C}_r$ is a convex cone, and l is the total number of equalities obtained as a result of substituting a certificate, such as (12)*, into all inequalities* (5b), (5c)*;*

4 $\alpha \approx \min_{j\in[J]} \alpha_j \leq \min_{j\in[J]} \alpha_j^r$

The type of the cone $\mathcal{C}_r$ depends on the chosen certificate of non-negativity. For instance, Proposition 2 results in a polyhedral cone, and in Proposition 4 we obtain an SDP cone. We formalize our results in the next lemma.

Lemma 2. *Let $j \in [J]$ and $r \geq 0$, and let and the cone $\mathcal{C}_r$ correspond to some certificate of non-negativity. Any feasible solution to problem* (13) *is feasible for problem* (5) *and approximately feasible for problem* (1)*. Consequently, $\alpha \approx \min_{j\in[J]} \alpha_j \leq \min_{j\in[J]} \alpha_j^r$. Moreover, under the conditions in Proposition 2, Proposition 4, or Theorem 3, there exists $r \geq 0$ such that $\alpha_j^r = \alpha_j$.*

As an example, using Putinar's certificate (11) in problem (13), we obtain

$$\begin{aligned} \alpha_j^r = \inf_{y,\{S_k\}} \quad & f(y) + t \\ \text{s.t.} \quad & y \in S_y, \\ & H_i(y) = \sum_{k=0}^{m_\zeta} \langle A_{i,k}, S_k \rangle \text{ for all } i \in \left[\binom{2r}{n_\zeta} \right], \\ & S_k \succeq 0 \text{ for all } k \in \{0, \ldots, m_\zeta\}, \end{aligned}$$

where the polynomial mapping $H : \mathbb{R}^{n_y} \to \left[\binom{2r}{n_\zeta}\right]$ is obtained from (5b), (5c).

3.3 Dynamic algorithm

This section describes the algorithm we use to solve problem (1). To begin with, we need to partition S_x into subsets, which can be done in various ways. As a result, the partition could be too large or suboptimal. Optimization over all subsets in such partition can be time consuming and imprecise. To overcome these difficulties, instead of working with a given partition, we generate it dynamically. Namely, since robust optimization problems are usually feasible for $\zeta = 0$, we start from some *nominal feasible solution* at hand, that is, the control solution y_0 feasible for $\zeta = 0$. In particular, we suggest choosing a solution that is close to the original optimal solution since the robust solution could be close to the nominal solution if the uncertainty set is not too large. We obtain the corresponding x_0 such that $L(y_0, 0, x_0) = 0$. If the Jacobian of L with respect to x at x_0 is invertible, we optimize over a small set around x_0 defined by $S_x^0 = \{x \in S_x : \|x_0 - x\| \le \varepsilon\}$ for a chosen $\varepsilon > 0$ and norm type. If we are able to find a robustly feasible solution $y_1 \neq y_0$, we set x_1 such that $L(y_1, 0, x_1) = 0$ and repeat the procedure. A detailed description of our algorithm follows below.

Algorithm 3: Dynamic algorithm to solve ARC

Input: $tol > 0$, f_0, norm type $\|\cdot\|$, a rule for ε_j, and the initial nominal feasible solution (y_0, x_0) such that the Jacobian in (4) has a full rank at $(0, x_0)$

```
Set j := 0;
while f_j − f_{j−1} ≤ tol or j = 0 do
    Compute Taylor approximation (4);
    if Jacobian of the approximation is not full-rank then
        Adjust S_x^j and reoptimize until a new x_j with full Jacobian rank is obtained
    else
        Set j := j + 1
        Update ε_j according to the rule in the Input
        Define S_x^j := {x ∈ S_x : ‖x_{j−1} − x‖ ≤ ε_j}
    Solve problem (13);
    if Problem (13) is feasible then
        Save the optimal solution y_j
    else
        Set y_j := ∞, f_j := ∞
```

Theorem 4. *Algorithm 3 stops after a finite number of iterations. The obtained solution is robustly feasible for the piecewise affine relaxation from Algorithm 1.*

Proof. The algorithm stops after finitely many iterations since $tol > 0$ and S_x is bounded. The obtained solution is robustly feasible for the approximation in Algorithm 1 by Lemma 2. □

Algorithm 3 is a local search algorithm in the sense that it might not find the solution with the best objective if that solution is far from the starting point. Therefore one might have to restart the algorithm or consider some heuristic approaches that help to move away from local optima.

4 Certifying robust feasibility or infeasibility

For any choice of $S_x^j, j \in [J]$ that cover S_x, Algorithm 3 yields an approximate solution to problem (1). In this work, we do not derive universal approximation guarantees for our framework but suggest to do a posterior verification of the obtained solution to problem (13) to see if it is feasible for the original problem (1).

4.1 Certifying robust feasibility

First, consider certifying robust feasibility of the equality constraints, that is, (1a), for a given $y \in S_y$. If we have the simple situation where the constraints are linear in x and have no products between x and ζ, it is possible to check their feasibility in advance by looking at the range of the Jacobian of (1a) with respect to x and possible realizations of $\zeta \in \Omega$. Now, let (1a) be non-linear. In this paper we assume that the sets S_x^j, $j \in [J]$, are small enough so that the piecewise affine approximations reflect the real system of equalities. Thus if our piecewise affine approximation has a solution, then the original system does as well. We use this assumption since in general verifying solvability of a given system of linear equalities for all realizations of some parameter is hard. However, there exist instruments for special cases. For instance, Dvijotham et al. 2018 discuss the case where (1a) is quadratic in x and linear in ζ, and Goldsztejn and Jaulin 2006 consider the case where Ω is a box or one knows inscribed or circumscribed boxes for the uncertainty set.

Now, fix $y \in S_y$ and assume that (1a) holds for all $\zeta \in \Omega$. We need to verify feasibility of the inequalities (1b) and (1c). For this purpose we use strengthenings of the inequality constraints based on the approach in Aßmann et al. 2018, where PO is used for gas networks optimization problems. For $y \in S_y$ define

$$\mathcal{E}_y = \Big\{(\zeta, x) : \zeta \in \Omega,\ x \in \tilde{S}_x,\ L_i(y, \zeta, x) = 0 \text{ for all } i \in [m_{eq}]\Big\}, \tag{15}$$

where $\tilde{S}_x \supset S_x$ and the constraints $x \in \tilde{S}_x$ ensure that the feasibility set is compact and prevent physically unrealistic solutions. Such constraints are generally possible to find for real-life problems. They could either come from the widest realistic bounds on x or could be tailored using an expert opinion. For instance, in the robust ACOPF application considered later in this paper, these constraints describe limits on the voltage magnitudes and phase angle differences corresponding to realistic solutions to the power flow equations.

Assume that the set $\mathcal{E}_y$ is non-empty and (1a) holds for any given $\zeta \in \Omega$. For each constraint (1b), consider the following subproblem (analogous subproblems can be written for (1c)):

Problem 16.

$$\begin{aligned} &\inf_t\ t \\ &\text{s.t. } G_i(y, \zeta, x) + t \geq 0 \textit{ for all } (\zeta, x) \in \mathcal{E}_y. \end{aligned}$$

Proposition 5. *If subproblems* (16) *for all constraints* (1b), (1c) *have non-positive optimal values, then y is feasible for problem* (1), *given that* (1a) *is feasible for all $\zeta \in \Omega$. If the subproblem* (16) *corresponding to some constraint is unbounded, then the set $\mathcal{E}_y$ is empty.*

Proof. Every feasible solution to (16) is feasible for the corresponding constraint in problem (1). Hence the first part of the proposition follows. The second result follows from the fact that $\mathcal{E}_y$ is compact, hence an optimization problem over $\mathcal{E}_y$ can be unbounded only if the set is empty. □

Proposition 5 can be used to confirm robust feasibility or detect if the obtained solution is infeasible even for the nominal problem, which could happen if coarse affine approximations were used. One can solve (16) by replacing the non-negativity condition by a certificate of non-negativity as in Section 3.2. A certificate will provide an upper bound on the original subproblem, and thus we conclude the following:

Corollary 1. *Let us use a certificate of non-negativity to solve* (16). *If the certificate yields non-positive optimal values for all* (1b), (1c), *we conclude that y is feasible for problem* (1), *given that* (1a) *is also feasible for all $\zeta \in \Omega$. If the optimal values of some subproblems are positive, the feasibility check is inconclusive.*

4.2 Certifying robust infeasibility

If we are not able to find a robustly feasible solution, it makes sense to verify if problem (1) is infeasible. Next, we consider a general approach to verify in advance whether the problem is infeasible for a given uncertainty set. We note that this approach generalizes the approach presented in Aßmann et al. 2018. Assume we know the Lebesgue measure λ on Ω. This assumption is realistic if Ω corresponds to box or ellipsoidal uncertainty, otherwise we could approximate the volume of the set using, for instance, the results of Henrion et al. 2009. For a given $y \in S_y$, denote by U_y the feasible set of the nominal problem for problem (1):

$$U_y := \big\{(\zeta, x) : \zeta \in \Omega,\ x \in S_x,\ L(y, \zeta, x) = 0,\ G(y, \zeta, x) \leq 0\big\}.$$

Notice that U_y is non-empty as long as the nominal problem is feasible. Consider the following problem:

Problem 16.

$$\begin{aligned} \inf_{p \in \mathbb{R}_d[\zeta, x]} \quad & \sum_\alpha p_{\alpha,\beta} \int_{\zeta \in \Omega} \zeta^\alpha x^\beta d\lambda \\ \text{s.t.} \quad & p_{\alpha,\beta} = 0 \ \ \textit{when } \beta \neq 0,\ ||p|| \leq 1 \\ & p(\zeta, x) \geq 0 \textit{ for all } (\zeta, x) \in U_y. \end{aligned} \tag{16a}$$

The norm constraint ensures that the problem is bounded. Define

$$\mathrm{Proj}_\zeta(U_y) = \{\zeta \in \mathbb{R}^{n_\zeta} : (\zeta, x) \in U_y \text{ for some } x \in \mathbb{R}^{n_x}\}. \tag{17}$$

If the optimal value of this problem is negative, then there exists a polynomial that depends on ζ only and that is non-negative on U_y but negative on some subset of Ω. The latter implies that $\Omega \setminus \mathrm{Proj}_\zeta(U_y) \neq \emptyset$, and therefore problem (1) is infeasible.

Theorem 5 (Proposition 3.4 in Aßmann et al. 2018)**.** *Let Ω, U_y be compact sets such that $\Omega \setminus Proj_\zeta(U_y) \neq \emptyset$. Then there exists d such that the optimal value of* (16) *is negative. Moreover, assume that the set U_x has explicit ball or box constraints (possibly added as redundant constraints). Then the optimal value of the problem where the constraint* (16a) *is replaced by Putinar's certificate* (11) *is negative for some finite level r. If the optimal value is non-negative for the given r, the result of the certificate is inconclusive.*

By Assumption 1 (a), we can use Theorem (5) to verify infeasibility of problem (1). Now we can generalize the result above to see if the uncertainty set is large enough to induce infeasibility of the original robust optimization problem. Denote the full feasible set of the nominal problem for problem (1) by U:

$$U := \big\{(y, \zeta, x) : y \in S_y,\ \zeta \in \Omega,\ x \in S_x,\ L(y, \zeta, x) = 0,\ G(y, \zeta, x) \leq 0\big\}.$$

Next we consider a generalization of problem (16) that follows from Theorem 5.

Problem 18.

$$\inf_{p \in \mathbb{R}_r[y, \zeta, x]} \quad \sum_\alpha p_{\alpha,\beta,\gamma} \int_{\zeta \in \Omega} \zeta^\alpha x^\beta y^\gamma d\lambda \tag{18a}$$

$$\text{s.t.} \quad p_{\alpha,\beta,\gamma} = 0 \ \ \textit{when } \beta + \gamma \neq 0,\ ||p|| \leq 1 \tag{18b}$$

$$p(y, \zeta, x) \geq 0 \textit{ for all } (y, \zeta, x) \in U. \tag{18c}$$

Corollary 2. *Let Ω, U be compact sets such that $\Omega \setminus Proj_\zeta(U) \neq \emptyset$. Then there exists d such that the optimal value of problem* (16) *is negative. Moreover, assume that the set U has explicit ball or box constraints (possibly added as redundant constraints). Then the optimal value of the problem where the constraint* (18c) *is replaced by Putinar's certificate* (11) *is negative for some finite r.*

If the optimal value (or upper bound) of problem (18) is negative, then the original problem (1) is infeasible for some $\zeta \in \Omega$ and any $y \in S_y$, $x \in S_x$. We note that the certificate (18) will not necessarily detect infeasibility of problems where different values of $\zeta \in \Omega$ make (1) infeasible for various values of y. This result is interesting for detecting the largest possible levels of uncertainty admissible for the original problem. The choice of the uncertainty set is another hard and open question in the literature, and our generalization is convenient for obtaining limits on the maximum size of the uncertainty set for problem (1); see, e.g., Lee et al. 2019.

5 Quadratic inequality constraints with ellipsoidal uncertainty in application to ACOPF

In this section, we derive the solution method for problem (5) when the inequality constraints in this problem are quadratic and Ω is an ellipsoid. In this case, our approach to the inequality constraints yields an equivalent reformulation (13) of problem (5) by Proposition 4. A quadratic problem (5) occurs, for example, in energy optimization problems such as ACOPF. We work with ellipsoidal uncertainty since it is less conservative than polyhedral, has interpretations from both robust and chance constrained perspectives (see, e.g, Golestaneh et al. 2018), and is simple to use via Proposition 4 (b). We first derive theoretical results for the described setting and then show how our theory applies to the robust ACOPF problem.

5.1 Quadratic inequality constraints with ellipsoidal uncertainty model

Let the constraints in (1b), (1c) be quadratic and consider

$$\Omega = \{\zeta \in \mathbb{R}^{n_\zeta} : (\zeta - \zeta^*)^\top \Sigma(\zeta - \zeta^*) \le r\}. \tag{19}$$

That is, Ω is an ellipsoid centered at ζ^*. Let m_x be the number of inequalities defining S_x. Since G_i in (1b) are quadratic in x and ζ, after substituting x from the first-order Taylor approximation in each inequality constraint (5b), (5c), we obtain a constraint that is quadratic in ζ and, for all $i \in [m_{in} + m_x]$, has the form

$$\hat{G}_i(y, \zeta) := \zeta^\top A_i \zeta + (y^\top B_i + b_i^\top)\zeta + (y^\top C_i + c_i^\top)y + d_i \ge 0, \tag{20}$$

for some given parameters $A_i,\ B_i,\ C_i,\ b_i,\ c_i,\ d_i$. Now we can use Proposition 4 statement (b) to obtain the following form of problem (13).

Problem 21.

$$\begin{aligned}
\beta_k = \inf_{y,\{\lambda_i\},\{\gamma_i\}} \quad & f(y) && \text{(21a)}\\
\text{s.t.} \quad & y \in S_y, && \text{(21b)}\\
& \begin{bmatrix} \gamma_i + c_i^\top y + d_i - \lambda_i(r - (\zeta^*)^\top P \zeta^*) & \frac{1}{2}(y^\top B_i + b_i^\top - 2\lambda_i(\zeta^*)^\top P) \\ \frac{1}{2}(B_i^\top y + b_i - 2\lambda_i P\zeta^*) & \lambda_i \Sigma + A_i \end{bmatrix} \succeq 0 && \text{(21c)}\\
& \lambda_i \ge 0, \quad \textit{for all } i \in [m_{in} + m_x] && \text{(21d)}\\
& y^\top C_i y = \gamma_i \quad \textit{for all } i \in [m_{in} + m_x], && \text{(21e)}
\end{aligned}$$

where P and r are the ellipsoid parameters from (19), and other parameters are defined in (20).

Proposition 6. *Let the constraints* (5b), (5c) *in problem* (5) *be of the form* (20). *Then problem* (5) *is equivalent to problem* (13) *and to problem* (21).

Proof. By Proposition 4 statement (b), for each $i \in [m_{in} + m_x]$ we have

$$\begin{aligned}
& \hat{G}_i(y, \zeta) \ge 0 \text{ for all } \zeta \in \Omega \\
\iff & \zeta^\top A_i \zeta + (y^\top B_i + b_i^\top)\zeta + (y^\top C_i + c_i^\top)y + d_i \ge 0 \text{ for all } \zeta \in \mathbb{R}^{n_\zeta} : (\zeta - \zeta^*)^\top \Sigma(\zeta - \zeta^*) \le r \\
\overset{\text{Prop. 4}}{\iff} & \zeta^\top A_i \zeta + (y^\top B_i + b_i^\top)\zeta + (y^\top C_i + c_i)y + d_i - \lambda_i(r - (\zeta - \zeta^*)^\top \Sigma(\zeta - \zeta^*)) = \begin{bmatrix} 1 \\ \zeta \end{bmatrix}^\top S_i \begin{bmatrix} 1 \\ \zeta \end{bmatrix}, \\
& S_i \succeq 0,\ \lambda_i \ge 0 \\
\iff & \begin{bmatrix} \gamma_i + c_i^\top y + d_i - \lambda_i(r - (\zeta^*)^\top P \zeta^*) & \frac{1}{2}(y^\top B_i + b_i^\top - 2\lambda_i(\zeta^*)^\top P) \\ \frac{1}{2}(B_i^\top y + b_i - 2\lambda_i P\zeta^*) & \lambda_i \Sigma + A_i \end{bmatrix} \succeq 0,\ y^\top C_i y = \gamma_i,\ \lambda_i \ge 0.
\end{aligned}$$

The final formulation follows from substituting the constraints above in problem (13) and the equivalence of problems (5) and (13) by Proposition 4. $\square$

Problem (21) is an SDP problem with potentially complicating non-convex constraints (21e). We do not specify the form of the objective f since it does not cause additional complications; see Remark 3. The number of equality constraints (21e) is finite and has a special structure. To begin with, we can use any known relaxation of the non-convex quadratic equality constraints, such as SDP relaxations or McCormick envelopes. Below is the SDP relaxation we use later in the computational experiments:

Problem 22.

$$
\begin{aligned}
\beta_k = \inf_{\substack{y,Y,\\ \{\lambda_i\},\{\gamma_i\}}} \quad & f(y) \\
\text{s.t.} \quad & (21b) - (21d) \\
& \langle Y, C_i \rangle = \gamma_i \quad \textit{for all } i \in [m_{in} + m_x] \\
& Y \succeq 0, \; Y(:,1) = \left[\begin{smallmatrix} 1 \\ y \end{smallmatrix}\right],
\end{aligned} \tag{22a}
$$

where $Y(:,1)$ is the first column of Y. The variable Y represents $\left[\begin{smallmatrix} 1 & y \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 \\ y \end{smallmatrix}\right]$. If we add the constraint $Y = \left[\begin{smallmatrix} 1 & y \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 \\ y \end{smallmatrix}\right]$ to (22), we obtain a reformulation of (21). However, since that constraint is non-convex, we omit it and obtain a lower bound relaxation; see Wolkowicz et al. 2000 for details.

Any lower-bound relaxation might provide a solution that is infeasible for (1). To obtain a feasible solution from the solution to the relaxation, we use the alternating projection method as presented in Algorithm 4. Potentially, the method finds a local minimum of problem (21). Define

$$
\mathcal{A} := \{(y, \gamma, \lambda) : (21a) - (21d) \text{ holds}\}, \tag{23}
$$

$$
\mathcal{B} := \{(y, \gamma, \lambda) : (21e) \text{ holds}\} \tag{24}
$$

Let $y := (y^{nc}, y^c)$, where y^{nc}, y^c are the subsets of variables y that are involved (resp. not involved) in the non-convex constraints (21e).

Algorithm 4: Alternating projections with line search for problem (21)

Input: $tol > 0$, $f_0 \geq \beta_k$, $(\nu_i)_{i=1}^N \in (0, 1]$, $N \geq 1$

Solve any lower bound relaxation of problem (21) and denote its solution by $(y_0, \gamma_0, \lambda_0)$ and its objective value by β_k^l;
if *the lower bound relaxation is infeasible,* **then**
  Problem (21) is infeasible, return "Problem infeasible"
else
  Go to the next step
Set $i = 0$ and choose $(y_1, \gamma_1, \lambda_1)$ such that $\|(y_0^{nc}, \gamma_0) - (y_1^{nc}, \gamma_1)\| > tol$;
while $\|(y_0^{nc}, \gamma_0) - (y_1^{nc}, \gamma_1)\| > tol$ *and* $i \leq N$ **do**
  Set $i := i + 1$;
  Project on $\mathcal{B}$: $(y_1, \gamma_1, \lambda_1) := (y_0, y_0^\top C_1 y_0, \ldots, y_0^\top C_{[m_{in}+m_x]} y_0, \lambda_0)$;
  Project on $\mathcal{A}$: $(y_0, \gamma_0, \lambda_0) := \arg\min_{(y,\gamma,\lambda)\in\mathcal{A},\ f(y)\leq f_0} \|(y^{nc}, \gamma) - (y_1^{nc}, \gamma_1)\|$;
  if $\|(y_0^{nc}, \gamma_0) - (y_1^{nc}, \gamma_1)\| < tol$, **then**
    Find the best y^c given (y_0^{nc}, γ_0) by solving $(y_0^c, \lambda_0) = \arg\min_{(y_0^{nc}, y^c, \gamma_0, \lambda)\in\mathcal{A}} f(y_0^{nc}, y^c)$;
    Save $(y^*, \gamma^*, \lambda^*) := (y_0^{nc}, y_0^c, \gamma_0, \lambda_0)$ as the current best feasible solution;
    Try to find a feasible solution with a better objective by setting $f_0 = \nu f(y_0) + (1 - \nu_i)\beta_k^l$;
    Set $(y_1, \gamma_1, \lambda_1) := \frac{1}{\nu_i}(y_0, \gamma_0, \lambda_0)$
if *No feasible solution is obtained,* **then**
  Problem (21) could be infeasible, return "Inconclusive, out of iterations"
else
  Return "The best obtained solution is $(y^*, \gamma^*, \lambda^*)$"

Theorem 6. *Let $\mathcal{A} \cap \mathcal{B} \neq \emptyset$, and let N be large enough and tol be small enough. Then the following holds:*

(a) Algorithm 4 stops after finitely many iterations. The algorithm can either find a feasible solution, or report an infeasible problem, or not be able to find a solution in the given number of iterations.

(b) If Algorithm 4 starts at a point $(y_0, \gamma_0, \lambda_0)$ that is sufficiently close to $\mathcal{A} \cap \mathcal{B}$, then the algorithm converges to a point $(y^*, \gamma^*, \lambda^*) \in \mathcal{A} \cap \mathcal{B}$, up to the given tolerance tol.

(c) If $\mathcal{A}$ has a non-empty interior and $(y_0, \gamma_0, \lambda_0)$ is sufficiently close to $\mathcal{A} \cap \mathcal{B}$, Algorithm 4 converges to a point in $\mathcal{A} \cap \mathcal{B}$ linearly.

(d) Let $(\hat{y}, \hat{\gamma}, \hat{\lambda})$ be a local optimum of problem (21). There is a sequence $(\nu_i)_{i=1}^N$ such that Algorithm 4 converges to the point $(y^*, \gamma^*, \lambda^*)$ with $\|(\hat{y}, \hat{\gamma}, \hat{\lambda}) - (y^*, \gamma^*, \lambda^*)\| \leq tol$.

Proof. Item (a) follows by construction of the algorithm since we limit the number of iterations. Item (b) follows from Theorem 7.3 in Drusvyatskiy et al. 2015 as the sets $\mathcal{A}$ and $\mathcal{B}$ are closed and semialgebraic. Item (c) follows from the fact that $\mathcal{A}$ is convex. Hence the normal cone at any point in the interior of A equals $\{0\}$, and thus $\mathcal{A}$ and $\mathcal{B}$ satisfy the conditions of Theorem 2.1 in Drusvyatskiy et al. 2015, which implies linear convergence. Next we prove item (d). For a given ν, let $S := \mathcal{A} \cap \mathcal{B} \cap \{(y, \gamma, \lambda) : f(y) \leq \nu f(y_0) + (1-\nu)\beta_k^l\}$. If $(y^*, \gamma^*, \lambda^*)$ is a local optimum, then for a $(\nu_i)_{i=1}^N$ small enough the algorithm cannot find a better feasible solution, stops after N iterations, and returns $(y^*, \gamma^*, \lambda^*)$. Now, assume that $(y^*, \gamma^*, \lambda^*)$ is not a local optimum. Then for any $\delta > 0$ there exists $(y, \gamma, \lambda) \in \mathcal{A} \cap \mathcal{B} \cap \{y, \gamma, \lambda : \|(y^*, \gamma^*, \lambda^*) - (y, \gamma, \lambda)\| \leq \delta\}$ and $f(y) < f(y^*)$. Hence there is a sufficiently small ν_i such that $S \neq \emptyset$ and the distance between $(y^*, \gamma^*, \lambda^*)$ and S satisfies the conditions of item (b). Therefore, using $(y^*, \gamma^*, \lambda^*)$ as the new starting point, the algorithm converges to a new feasible point in S with an improved objective value. Since the tolerance $tol > 0$ is fixed and N is large, the process can be repeated until no improvement within tol is obtained. □

Remark 2. *In Algorithm 4, we write $y := (y^{nc}, y^c)$ and use two subsets of the variables to speed up the algorithm by skipping the iterations in which only y^c changes.*

Remark 3. *If the objective f is not convex, we can introduce the epigraph variable and add another constraint to the set $\mathcal{B}$ (24). Even though the latter constraint may be not quadratic, Algorithm 4 will handle it since the degree of the equality constraints is not important for the alternating projections method. In fact, in Algorithm 4, the non-linear equality constraints can be of any degree as long as it is possible to obtain some lower bound for a good starting point.*

In the next subsection, we demonstrate the performance of Algorithm 4 on the ACOPF problem using the SDP lower bound for Step 2 of the algorithm. The experiments show that in the majority of cases Algorithm 4 finds a feasible solution with an objective value close to the lower bound in few iterations.

5.2 Adjustable ACOPF with uncertain renewable generation and load demands

Optimal power flow (OPF) is one of the key optimization problems relevant to the operation of electric power systems. OPF solutions provide minimum cost operating points that satisfy both equality constraints termed the "power flow equations" which model the power system network and inequality constraints that impose limits on line flows, generator outputs, voltage magnitudes, etc.

Accurately modeling the steady-state behavior of power systems requires the non-linear *AC power flow equations*, which can be formulated as a system of quadratic polynomial equality constraints. To bypass the analytical and practical challenges associated with the AC power flow equations, many OPF formulations employ the so-called *DC power flow approximation*, which linearizes the AC power flow equations using assumptions regarding typical operating characteristics for electric transmission systems (see Stott et al. 2009). The resulting DC OPF problems are formulated as convex quadratic programs that can be efficiently solved. This improvement in tractability comes at the cost of solution accuracy and modeling flexibility, with DC OPF problems being inapplicable to certain settings such as electric distribution systems. Researchers have proposed a wide variety of other approximations and convex relaxations of the AC power flow equations. See Molzahn and Hiskens 2019 for a recent survey. While these alternative representations are useful in various contexts, there are many situations which require directly addressing the non-linear AC power flow equations.

Compounding the difficulties posed by the power flow non-linearities, rapidly increasing quantities of wind and solar generation are introducing significant amounts of power injection uncertainties into electric grids. To address these uncertainties, researchers have studied a wide range of stochastic and robust OPF problems, many of which use the DC power flow approximation. (See, e.g., Vrakopoulou et al. 2013 and Bienstock et al. 2014 for several relevant examples). This linear power flow representation permits the application of stochastic and robust optimization techniques developed for linear programs. Alternative approaches replace the AC power flow equations with other more sophisticated approximations, such as the work in Mühlpfordt et al. 2019 and Roald and Andersson 2018, or convex relaxations, such as the work in Venzke et al. 2018. Such approaches can provide useful solutions in many contexts, particularly when the approximations are iteratively updated or adaptively adjusted. However, the quality guarantees from these approaches are provided with respect to the approximation or convex relaxation as opposed to the original non-convex ACOPF problem. Thus, the resulting solutions may lead to unacceptable constraint violations during operation in the physical system.

The power systems literature also includes approaches that directly address the non-linear AC power flow equations. These approaches can provide high-quality solutions in certain instances but may be limited to special classes of problems, such as systems that satisfy restrictive requirements on the power injections at each bus as in Louca and Bitar 2019. Other approaches use scenario-based techniques that enforce feasibility for selected uncertainty realizations, possibly obtained via subproblems that compute worst-case uncertainty realizations with local solvers as in Capitanescu et al. 2012 or convex relaxations as in Lorca and Sun 2018. Certifying robustness with such approaches is challenging due to the possibilities of local solutions and inexact relaxations. Rather than seeking the worst-case uncertainty realizations, the approach in Molzahn and Roald 2018 instead bounds the worst-case impacts of the uncertainties with respect to each constrained quantity. While this approach provides guarantees regarding the satisfaction on the engineering inequality constraints, each iteration requires the solution of many computationally expensive subproblems. We also note recent work in Lee et al. 2020 that uses so-called "convex restriction" techniques (see Lee et al. 2019) to compute robustly feasible ACOPF solutions. While promising, this approach is undergoing continuing development and requires specialization to the particular non-linearities in each class of problems.

Exploiting the polynominal representation of the AC power flow equations, we next apply our PO based approach described in this paper to the robust ACOPF problem, beginning with our notation and the problem formulation. Consider a power network $P = (N, E)$ with the set of buses $N = \{1, \dots, n\}$ and the set of lines connecting these buses E. Denote the set of buses with generators by G. Denote the active and reactive power demand (load) at each bus $k \in N$ by P_k^d and Q_k^d, respectively. Denote the index of the reference bus by s. To implement thermal restrictions on the transmission lines, we impose line current limits, see Zimmerman et al. 2010. Our objective is to minimize the cost of power generation, which is one of classical objectives in OPF problems. Denote the active and reactive power injections due to load or generation fluctuations by P_k^r and Q_k^r, respectively, for all $k \in N$. In the nominal ACOPF without uncertainty P_k^r and Q_k^r are known and fixed. Next we define the ACOPF problem as a quadratic optimization problem.

Problem 25.

$$
\begin{aligned}
\alpha^{nom} = \inf_{x, P^g, Q^g} \quad & \sum_{k \in G} c_k^2 (P_k^g)^2 + c_k^1 P_k^g + c_k^0 \\
\text{s.t.} \quad & P_k^{\min} \leq P_k^g \leq P_k^{\max} && \textit{for all } k \in N \\
& Q_k^{\min} \leq Q_k^g \leq Q_k^{\max} && \textit{for all } k \in N \\
& (V_k^{\min})^2 \leq x_k^2 + x_{k+n}^2 \leq (V_k^{\max})^2 && \textit{for all } k \in N \\
& tr(Y_{lm} x x^\top) \leq S_{lm}^{\max} && \textit{for all } \{lm\} \in E \\
& P_k^g + P_k^r = P_k^d + tr(Y_k x x^\top) && \textit{for all } k \in N \\
& Q_k^g + Q_k^r = Q_k^d + tr(\bar{Y}_k x x^\top) && \textit{for all } k \in N \\
& x(s) = 0,
\end{aligned}
$$

where the last constraint sets the phase angle of the reference bus to zero. Now, let the active power fluctuations for each $k \in N$ be

$$P_k^r = \bar{P}_k^r + \zeta_k,$$

where ζ is uncertain. We assume that the power injection uncertainties from the load and generation at each bus $k \in N$ are modeled via a constant power factor $\cos \phi_k$ so that the reactive power fluctuations are

$$Q_k^r = \bar{Q}_k^r + \gamma_k \zeta_k, \ \gamma_k := \begin{cases} 0 & \text{if } P_k^d = 0, \\ \frac{\sqrt{1-\cos^2 \phi_k}}{\cos \phi_k} & \text{otherwise.} \end{cases}$$

Without loss of generality, we let $\bar{P}_k^r = \bar{Q}_k^r = 0$, otherwise one can adjust the loads P_k^d and Q_k^d. We denote by δ the total change in the active power losses due to the redistribution of power flows from the uncertain power injection fluctuations relative to the losses from the nominal power flows. Note that δ is typically near zero, as the losses themselves are usually small and the changes in losses are even smaller. For an operating point to be robustly feasible, the generators must account for the total change in the power injections, $\sum_{i=1}^n \zeta_i - \delta$, associated with each uncertainty realization without leading to constraint violations. We adopt a "participation factor" model where this change in power injections is distributed among all generators according to a linear recourse policy with specified participation factors α_k for each generator k. Thus, for each $k \in N$, the actual active power generation consists of the nominal power P_k^g and an adjustment in generation due to uncertainty:

$$P_k^g - \alpha_k \left(\sum_{i=1}^n \zeta_i - \delta \right), \ \alpha_i \geq 0, \ \sum_{i=1}^n \alpha_i = 1, \tag{26}$$

Thus, when introducing uncertainty to problem (25), we replace P_k^g in this problem with (26). We note that this model represents the steady-state behavior of widely used automatic generation control (AGC) (see Jaleeli et al. 1992) and is adopted in many robust and stochastic OPF formulations, e.g, those used in Venzke et al. 2018, Roald and Andersson 2018, and Molzahn and Roald 2018. To model the uncertainty, we let the uncertain parameters ζ belong to the region

$$\Omega = \{\zeta \in \mathbb{R}^{n_\zeta} : \zeta^\top \Sigma \zeta \leq 1\},$$

where P is a covariance matrix. That is, our uncertainty region is an ellipsoid centered on the point with no fluctuations.

For $k \in N$, we denote by $V_k^g := x_k^2 + x_{k+n}^2$ the squared voltage magnitude at bus k. Following traditional power system modeling practices, we consider three types of buses: PV, PQ and the reference bus. If k is a PV bus, the active power P_k^g and squared voltage magnitudes V_k^g are set by the operator while the reactive power Q_k^g can change. If k is a PQ bus, then the active power and reactive power are fixed to constant values while the voltage magnitude can change. Without loss of generality, we assume that active and reactive power generation at PQ buses is zero, otherwise the loads can be adjusted. Finally, the operator selects the voltage magnitude at the reference bus while the active and reactive powers are free to vary. We also introduce a variable t that denotes the worst-case upper bound on the active power on the reference bus. We use this bound to estimate the worst-case objective value over the uncertain power injection fluctuations, as is typical in robust optimization problems. As a result, the control variables y in the problem include t, P_k^g, where k belongs to the set of PV buses, and V_k^g, where k belongs to the union of PV and reference buses.

For ease of notation and implementation, we make several assumptions, which are straightforward to relax. First, we assume that all buses with generators are either reference or PV buses and all buses without generators are PQ buses. That is, the set of PQ buses is $N \setminus G$. This assumption holds for the cases we use to test our approach. We use the formulation in MATPOWER (see Zimmerman et al. 2010) with one reference bus that accounts for the change in network losses, δ, as opposed to an alternative "distributed slack" approach. Problem (1) for the ACOPF becomes:

Problem 27.

$$
\begin{aligned}
\alpha = \inf_{P^g, V^g, t} \quad & \sum_{k\in G\setminus\{s\}} c_k^2 (P_k^g)^2 + c_k^1 P_k^g + c_k^0 + c_s^2 t^2 + c_s^1 t + c_s^0 \\
\text{s. t.} \quad & P_k^{\min} \le P_k^g \le P_k^{\max} && \textit{for all } k \in G\setminus\{s\} \\
& (V_k^{\min})^2 \le V_k^g \le (V_k^{\max})^2 && \textit{for all } k \in G \\
& P_s^{\min} \le t \le P_s^{\max} \\
& \textit{For any } \zeta \in \Omega \textit{ there exists } x \textit{ such that} \\
& P_s^{\min} \le P_s^d + tr(Y_s x x^\top) - \zeta_s + \alpha_k \sum_{i=1}^n \zeta_i \le t \\
& Q_k^{\min} \le Q_k^d + tr(\bar{Y}_k x x^\top) - \gamma_k \zeta_k \le Q_k^{\max} && \textit{for all } k \in G \\
& (V_k^{\min})^2 \le x_k^2 + x_{k+n}^2 \le (V_k^{\max})^2 && \textit{for all } k \in N\setminus G, \\
& tr(Y_{lm} x x^\top) \le S_{lm}^{\max} && \textit{for all } \{lm\} \in E \\
& P_k^g = P_k^d + tr(Y_k x x^\top) - \zeta_k + \alpha_k \sum_{i=1}^n \zeta_i && \textit{for all } k \in G\setminus\{s\} \\
& 0 = P_k^d + tr(Y_k x x^\top) - \zeta_k && \textit{for all } k \in N\setminus G \\
& 0 = Q_k^d + tr(\bar{Y}_k x x^\top) - \gamma_k \zeta_k && \textit{for all } k \in N\setminus G \\
& V_k^g = x_k^2 + x_{k+n}^2 && \textit{for all } k \in G \\
& x(s) = 0.
\end{aligned}
$$

The control variables are $y := (t, P^g, V^g)$. The state variables are x, the real and imaginary parts of voltage phasors. There are $2n$ of state variables and equality constraints in the problem.

Proposition 7. *Any feasible solution to problem* (27) *is feasible for problem* (25)*, and the problems are equivalent for* $\Omega = \{0\}$.

Now we define the problem in the same form as (1) to more easily use the results from the earlier sections. This yields the following:

$$
f(t, P^g, V^g) = \sum_{k\in G\setminus\{s\}} c_k^2 (P_k^g)^2 + c_k^1 P_k^g + c_k^0 + c_s^2 t^2 + c_s^1 t + c_s^0, \tag{28}
$$

$$
\begin{aligned}
S_y = & \big\{ (P^g, V^g) : P_k^{\min} \le P_k^g \le P_k^{\max} \ \text{ for all } k \in G\setminus\{s\}, \\
& (V_k^{\min})^2 \le V_k^g \le (V_k^{\max})^2 \ \text{ for all } k \in G,\ P_s^{\min} \le t \le P_s^{\max} \big\},
\end{aligned} \tag{29}
$$

$$
L(P^g, V^g, \zeta, x) = \begin{bmatrix} P_k^d + \operatorname{tr}(Y_k x x^\top) - \zeta_k - P_k^g + \alpha_k \sum_{i=1}^n \zeta_i & \text{for all } k \in G\setminus\{s\} \\ P_k^d + \operatorname{tr}(Y_k x x^\top) - \zeta_k + \alpha_k \sum_{i=1}^n \zeta_i & \text{for all } k \in N\setminus G \\ Q_k^d + \operatorname{tr}(\bar{Y}_k x x^\top) - \gamma_k \zeta_k & \text{for all } k \in N\setminus G \\ x_k^2 + x_{k+n}^2 - V_k^g & \text{for all } k \in G \\ x(s) & \end{bmatrix}, \tag{30}
$$

$$
G(t, P^g, V^g, \zeta, x) = \begin{bmatrix} -P_s^{\min} + P_s^d + \operatorname{tr}(Y_s x x^\top) - \zeta_s + \alpha_k \sum_{i=1}^n \zeta_i & \\ t - P_s^d - \operatorname{tr}(Y_s x x^\top) + \zeta_s - \alpha_k \sum_{i=1}^n \zeta_i & \\ -Q_k^{\min} + Q_k^d + \operatorname{tr}(\bar{Y}_k x x^\top) - \gamma_k \zeta_k & \text{for all } k \in G \\ Q_k^{\max} - Q_k^d - \operatorname{tr}(\bar{Y}_k x x^\top) + \gamma_k \zeta_k & \text{for all } k \in G \\ (V_k^{\max})^2 - x_k^2 + x_{k+n}^2 & \text{for all } k \in N\setminus G \\ x_k^2 + x_{k+n}^2 - (V_k^{\min})^2 & \text{for all } k \in N\setminus G \end{bmatrix}, \tag{31}
$$

$$
S_x = \big\{ x \in \mathbb{R}^{2n} : (V_k^{\min})^2 \le x_k^2 + x_{k+n}^2 \le (V_k^{\max})^2 \text{ for all } k \in N\setminus G \big\} \tag{32}
$$

$$
\tilde{S}_x = \big\{ x \in \mathbb{R}^{2n} : \tfrac{1}{2}(V_k^{\min})^2 \le x_k^2 + x_{k+n}^2 \le \tfrac{3}{2}(V_k^{\max})^2 \text{ for all } k \in N \big\}. \tag{33}
$$

In the next subsection, we run numerical experiments solving problem (1) with the inputs defined above and the instances from Matpower.

5.3 Numerical results

In this section, we implement the dynamic Algorithm 3, where we solve problem (21) at Step 10 (see Proposition 6). We also implement the PO feasibility and infeasibility checks from Section 4 using Putinar's certificate of non-negativity from Theorem 1. We use MATPOWER instances of the size up to 118 buses. All computations are done using MATLAB R2020a and Yalmip (see Löfberg 2004) on a computer with the processor Intel® Core® i7-6820HQ CPU @ 2.7GHz and 24 GB of RAM. Semidefinite programs are solved with MOSEK, Version 9.2.28 ApS 2019.

In Algorithm 3 we use the Euclidean norm, set $tol = 1e^{-5}$, the update rule $\varepsilon_j = \|x_{j-1}\|/10$ for $|N| < 30$ and $\varepsilon_j = \|x_{j-1}\|/30$ for $|N| \geq 30$ since the solution norm grows with the instance size. We exploit a warm start by setting the initial feasible nominal solution to the solution of the nominal problem (25) where the absolute values of all bounds are reduced by 0.5%. In other words, we "squeeze" all bounds to obtain a potentially robust solution. As a result, the first iteration of the algorithm often provides good feasible solutions, which could not be substantially improved at the second and next iterations. Having observed this pattern, we decided to omit Step 2 of Algorithm 3 in the experiments and run only one iteration of the algorithm. Of course, in realistic settings one could try to improve the solution by, for instance, accepting a solution with a worse objective value with a certain probability (as in the simulated annealing heuristic) or repeating the algorithm from different starting points.

To solve problem (21), we use Algorithm 4 with $f_0 = 1e^5$, $tol = 1e^{-5}$, $N = 100$, $\nu_i = 1$ for all $i \in [N]$. To compute the lower bound in Algorithm 4, we use relaxation (22). Setting $\nu_i = 1$ implies that we do not try to improve the objective value after finding the first feasible solution. We considered that option and obtained the following results. For the tested cases, the initially obtained objective value was close to the lower bound. Attempts to improve the objective values resulted in using substantially more or all N iterations with negligible or no improvement in the objective value. This result can be explained by the high quality of the lower bound from the SDP relaxation (22), which frequently provides a feasible (see the cases with "Num iter"= 1 in the tables with results) or close to feasible solution to problem (21). Therefore, we decided to stop at the first obtained feasible solution by setting $\nu_i = 1$ for all $i \in [N]$. Finally, in ACOPF problem, some matrices C_i in (20) are positive semi-definite. We do not have to project on the corresponding equality constraints in (21e). Instead, we replace the equality sign by "$\leq$" and add the resulting constraints to the definition in (23).

In the feasibility and infeasibility checks, we use Putinar's certificate (10) of degree four for $|N| \leq 9$ and of degree two for $|N| = 14$. In all cases we restrict the degree of the largest SOS σ_0 to two. For each subproblem (16) we add the redundant constraint generated by the previous subproblem in order to obtain stronger certificates, otherwise all feasibility checks were inconclusive. We construct the objective of the infeasibility check problem (16) using the standard change of variables from the ellipsoid to the unit hyper-sphere and then computing the integral over the hyper-sphere. We do not apply the feasibility and infeasibility checks for larger instances since our Matlab implementation of PO problems cannot be constructed if the number of variables ($n_x + n_\zeta$ in case of problem (16)) is larger than 40. In general, the checks with Putinar's certificate of degree two could be implemented for larger instances since even the interior point method can handle SDP problems of up to approximately 400 variables. Even more variables can be handled in presence of sparsity, which is a characteristic of ACOPF problems.

In the experiments, we need to choose the vector α and the parameters defining the ellipsoid Ω. We set α_k for all $k \in N$ equal to the ratio of the difference between the maximum and minimum active power generation at bus k to the sum of such differences over all buses in the system. Therefore, $\alpha_k = 0$ for all $k \in N \setminus G$ since active power generation at those buses is fixed. For Ω, we consider the case where the uncertainty only occurs at the buses with positive active power loads. This is straightforward to change. For instance, we could also consider uncertainty in generation by allowing additional sources of uncertainty at generator buses. We use two options for Σ in the definition of Ω. For the first option, we set

$$\Sigma = \text{Diag } \sigma, \ \sigma_k = \frac{1}{(wP_k^d)^2},$$

and w varies from 0.01 (1% of the load at the corresponding bus) to 0.5 (50% of the load at the corresponding bus). In the second option, we allow correlations among the courses of uncertainty equal to $1/|N|$, that is,

$$\Sigma = \mathrm{Diag}\sqrt{\sigma} \begin{bmatrix} 1 & 1/|N| & \dots & 1/|N| \\ 1/|N| & 1 & \dots & 1/|N| \\ \dots & \dots & \dots & \dots \\ 1/|N| & 1/|N| & \dots & 1 \end{bmatrix} \mathrm{Diag}\sqrt{\sigma}, \ \sigma_k = \frac{1}{(wP_k^d)^2},$$

Our set-up implies that the more sources of uncertainty, the less correlation between each two of them. We use this correlation pattern and not random correlation matrices to avoid numerical instability. In real-life applications, one can choose a correlation matrix that is suitable for the given application; for instance, one can assume that correlations are proportional to distances between buses.

The tables with results contain the following abbreviations:

- LNF – Lower bound subproblem (22) is not feasible. This implies that problem (5) is infeasible by Proposition 6. Hence, the linear approximation over the given small subset of S_x is infeasible.
- NC – Algorithm 4 did not converge in 100 iterations of alternating projections. Having NC and the number of iterations smaller than 100 means that the convergence was so slow that the algorithm terminated prematurely. We added this stopping criterion to speed up the implementation.
- F – abbreviation for “feasible”.
- NF – abbreviation for “not feasible”.
- IC – indicates that the corresponding feasibility or infeasibility check was inconclusive; see Corollary 1 and Theorem 5.
- NP – Mosek experienced numerical problems.

The computational results are presented in Tables 1–3. Tables 1 and 2 cover power networks with up to 14 buses, for which we run one iteration of Algorithm 3 as described above and apply the feasibility and infeasibility checks from Section 4. Table 3 shows the results for larger power networks after one iteration of Algorithm 3. In all tables, the first column denotes the values of the uncertainty as a fraction of load, mentioned earlier as w, in percent. The notation “Nom lower bound” is used for the SDP lower bound of the nominal problem; “Num iter” denotes the number of iterations of the alternating projections method in Algorithm 4 until convergence. The fifth and sixth columns denote the running time per iteration and the total running time. Notice that the former is not equal to the latter divided by the number of iterations since the total time includes the time spent computing the lower bound (22). Finally, the columns “Feas check” and “Infeas check” in Tables 1 and 2 show the results of the feasibility and infeasibility checks.

The numerical results show that our approximation approach in Section 3 provides a potentially robust solution in reasonable time even for the instance with 118 buses. At the same time, we could not obtain certificates of feasibility or infeasibility for the majority of cases. The results for the certificate of infeasibility are expected: if the obtained solutions are feasible, the certificate of infeasibility should be inconclusive while the certificate of feasibility could potentially exist. However, the certificate of feasibility turns out very time consuming and still inconclusive. Such long running times are the result of solving many subproblems (16) sequentially and adding a new constraint to the feasible set at every iteration. To speed up the implementation, one could solve the problems in parallel, then add redundant constraints to the subproblems that were infeasible and resolve only those subproblems.

For all but one of the cases, we obtained a potentially robust feasible solution after running one iteration of Algorithm 3. We could certify feasibility of the obtained solutions for one case (“case9”) with various levels of uncertainty. In particular, we obtained a certified feasible solution even when the uncertainty on some bus could amount to 50% of the load. This case has relatively lose bounds, so such a result is expected. We could obtain a certificate of infeasibility of one instance, for “case14” with 5% uncertainty.

For smaller uncertainty sets, the bounds are closer to the nominal values, and the larger the case, the larger the deviation from the nominal value even for small amounts of uncertainty. The only case for which one iteration of Algorithm 3 did not provide a solution is “WB5”. In fact, we could obtain feasible solutions when the uncertainty was 1% of the loads, but this was only achieved after about 150 iterations. Also, this is one of the two cases where the alternating projections stopped prematurely. This test case is known to be especially challenging in various contexts (see Molzahn 2017), so it is not surprising that it requires more iterations of Algorithm 3. We notice that for both large and small cases, we are either able to obtain a solution in several iterations or cannot obtain a solution at all. Therefore, it seems reasonable to terminate Algorithm 4 after about 50 alternating projections iterations instead of 100.

Finally, we see that adding correlations to the uncertainty ellipsoid does not substantially change the bounds or feasibility results in small cases but increases the running times and hampers numerical stability. This is related to the fact that the projection subproblem at Step 10 of Algorithm 4 becomes less numerically stable. When the instance size grows, both feasibility and numerical stability are influenced. In particular, for “case30”, “case57” and “case118” we could not obtain feasible bounds using the matrix Σ with correlations. The results for ‘case39” did not change substantially with or without correlations, except for the running times. Therefore, we do not provide the table with results with correlations for larger cases.

Table 1: Results for instances with up to 14 buses, *without correlation*. All objective values are divided by 100 in comparison to the original data.

Uncert, % of load	Nom. lower bound	Upper bound	Num iter	Average time per iter, sec	Full time, sec	Feas check	Time feas check, sec	Infeas check	Time infeas check, sec
LMBM3, 3 buses									
1%	56.95	57.01	1	1.4	2.0	IC	44.9	IC	7.6
5%	57.95	57.04	1	0.7	1.9	IC	45.1	IC	1.1
10%	58.95	57.08	1	0.6	0.9	IC	45.6	IC	1.0
20%	59.95	57.16	1	0.5	1.1	IC	46.0	IC	1.0
30%	59.95	57.26	1	0.7	1.1	IC	46.5	IC	1.7
40%	59.95	57.36	1	0.5	0.9	IC	46.9	IC	0.7
50%	59.95	57.49	1	0.6	1.0	IC	47.0	IC	0.7
WB5, 5 buses									
1-10%	10.70	NC	101	0.95	85.5	–	–	–	–
≥20%	10.70	LNF	–	–	1.0	–	–	–	–
case6ww, 6 buses									
1%	31.26	31.59	1	0.8	1.6	IC	667.0	IC	3.1
5%	31.26	31.79	1	0.7	1.2	IC	706.2	IC	3.1
≥10%	31.26	LNF	–	–	0.7	–	–	–	–
case9, 9 buses									
1%	52.97	53.13	1	0.9	1.6	**F**	1,437.2	–	–
5%	52.97	53.16	1	0.8	1.3	**F**	1,603.1	–	–
10%	52.97	53.18	1	0.7	1.3	**F**	1,762.7	–	–
20%	52.97	53.24	1	0.7	1.4	**F**	1,934.4	–	–
30%	52.97	53.31	1	0.7	1.3	**F**	2,102.4	–	–
40%	52.97	53.39	1	0.8	1.4	**F**	2,295.7	–	–
50%	52.97	53.47	1	0.7	1.4	**F**	2,487.7	–	–
case14, 14 buses									
1%	80.82	81.20	7	0.5	4.4	IC	86.1	IC	1.7
5%	80.82	81.31	8	0.5	4.8	IC	90.7	**NF**	1.5
10%	80.82	81.46	8	0.5	4.8	IC	96.5	IC	1.0
20%	80.82	81.77	5	0.5	3.3	IC	100.5	IC	0.9
30%	80.82	82.14	7	0.6	4.8	IC	107.3	IC	1.0
40%	80.82	82.53	14	0.5	8.0	IC	112.4	IC	0.9
50%	80.82	LNF	–	–	0.7	–	–	–	–

Table 2: Results for instances with up to 14 buses, *with correlation*. All objective values are divided by 100 in comparison to the original data.

Uncert, % of load	Nom. lower bound	Upper bound	Num iter	Average time per iter, sec	Full time, sec	Feas check	Time feas check, sec	Infeas check	Time infeas check, sec
LMBM3, 3 buses									
1%	56.95	57.01	1	1.5	3.6	IC	29.6	IC	10.9
5%	57.95	57.05	1	1.3	2.3	IC	29.5	IC	9.6
10%	58.95	57.09	1	1.3	2.1	IC	29.8	IC	9.7
20%	59.95	57.19	1	1.2	2.1	IC	29.8	IC	9.7
30%	59.95	57.30	1	1.3	2.2	IC	30.1	IC	9.8
40%	59.95	57.41	1	1.2	2.0	IC	30.0	IC	9.6
50%	59.95	57.73	1	1.2	2.0	IC	30.1	IC	9.6
WB5, 5 buses									
1-5%	10.70	NC	101	71.7	3.3	–	–	–	–
≥10%	10.70	LNF	101	0.9	1.6	–	–	–	–
case6ww, 6 buses									
1%	31.26	31.59	1	1.7	3.4	IC	612.0	IC	20.5
5%	31.26	31.80	1	1.8	3.3	IC	601.5	IC	20.5
≥10%	31.26	LNF	–	–	1.5	–	–	–	–
case9, 9 buses									
1%	52.97	53.14	1	1.8	3.3	**F**	1,265.5	–	–
5%	52.97	53.17	1	1.7	3.3	**F**	1,351.6	–	–
10%	52.97	53.20	1	1.8	3.3	**F**	1,452.1	–	–
20%	52.97	53.27	1	1.8	3.3	**F**	1,535.7	–	–
30%	52.97	53.34	1	1.8	3.4	**F**	1,624.9	–	–
40%	52.97	53.43	1	1.8	3.3	**F**	1,729.6	–	–
50%	52.97	53.51	1	1.8	3.5	**F**	1,826.6	–	–
case14, 14 buses									
1%	80.82	NP	–	–	1.7	–	–	–	–
5%	80.82	81.31	8	1.4	12.5	IC	678.9	**NF**	22.2
10%	80.82	81.45	8	1.4	12.3	IC	683.0	IC	22.0
20%	80.82	81.75	5	1.4	8.6	IC	688.2	IC	21.8
30%	80.82	82.10	7	1.4	11.5	IC	689.8	IC	21.9
40%	80.82	82.49	14	1.4	20.9	IC	694.2	IC	21.9
50%	80.82	LNF	–	–	1.7	–	–	–	–

6 Conclusions and directions for future research

In this paper, we proposed a framework to obtain approximately feasible solutions to general robust optimization problems. In particular, we approximate the original ARO problem by a sequence of classical PO problems with additional tractable conic constraints. When the original problem has (potentially nonconvex) inequality constraints, we design an algorithm that converges to a local optimum of our robust approximation. We implement the latter algorithm for ACOPF problems with uncertainty in loads and power generation in order to obtain potentially robustly feasible solutions for cases with up to 118 buses.

Table 3: Results for instances with more than 14 buses, *without correlation*. All objective values are divided by 100 in comparison to the original data.

Uncert, % of load	Nom. lower bound	Upper bound	Num iter	Average time per iter, sec	Full time, sec
case30, 30 buses					
1%	5.75	5.84	1	1.8	5.2
≥5%	5.75	LNF	–	–	2.7
case39, 39 buses					
1%	418.62	NP	–	–	3.9
5%	418.62	NP	28	2.1	63.5
10%	418.62	418.67	29	2.2	68.6
20%	418.62	420.91	63	2.4	152.8
Case57, 57 buses					
1%	417.38	423.67	3	3.5	18.0
5%	417.38	NC	101	3.5	354.4
Case118, 118 buses					
1%	1,296.55	1,299.98	2	50.1	216.8
5%	1,296.55	1,300.16	3	45.8	254.3
10%	1,296.55	1,300.16	4	45.6	304.8
20%	1,296.55	NC	40	47.2	2,005.4

One of the main open questions of this paper is approximation guarantees for our approach. It would be convenient to formulate the criteria under which our approach does not require additional feasibility checks since such checks can be time consuming, as our experiments show. Another alternative would be to improve the running times and efficiency of the feasibility check algorithms or suggest alternatives to them. Additionally, in Section 3.1, we note that our approach has to balance between the precision of the piecewise affine approximations and possible sizes of the uncertainty sets. As a result, the approximations cannot be as precise as desired. A natural remedy to increase precision would be to consider several linear pieces simultaneously in each subproblem (5).

It would be also important to explore the true scalability of our method for ACOPF problems by implementing the algorithms more efficiently and investigating possibilities to exploit the inherent sparsity structure of ACOPF problems. Finally, it would be interesting to look at the extensions of our method for multiperiod problems and to combine our approach with the approach proposed in Postek and Hertog 2016 to address larger uncertainty sets.

Acknowledgment

Olga Kuryatnikova and Bissan Ghaddar were supported by NSERC Discovery Grant 2017-04185. Daniel Molzahn was supported by the National Science Foundation under grant number 2023140.

References

ApS M (2019) *The MOSEK optimization toolbox for MATLAB manual. Version 8.0.0.81.* URL `https://www.mosek.com/documentation/`.

Ardestani-Jaafari A, Delage E (2016) Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations Research* 64(2):474–494.

Aßmann D, Liers F, Stingl M, Vera J (2018) Deciding robust feasibility and infeasibility using a set containment approach: An application to stationary passive gas network operations. *SIAM Journal on Optimization* 28(3):2489–2517.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* A(99):351–376.

Bertsimas D, Litvinov E, Sun XA, Zhao J, Zheng T (2012) Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems* 28(1):52–63.

Bienstock D, Chertkov M, Harnett S (2014) Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review* 56(3):461–495.

Bienstock D, Özbay N (2008) Computing robust basestock levels. *Discrete Optimization* 5(2):389–414.

Bienstock D, Verma A (2019) Strong NP-hardness of AC power flows feasibility. *Operations Research Letters* 47(6):494–501.

Capitanescu F, Fliscounakis S, Panciatici P, Wehenkel L (2012) Cautious operation planning under uncertainties. *IEEE Transactions on Power Systems* 27(4):1859–1869.

Chandrasekaran V, Shah P (2016) Relative entropy relaxations for signomial optimization. *SIAM Journal on Optimization* 26(2):1147–1173.

Dickinson PJ, Povh J (2018) A new approximation hierarchy for polynomial conic optimization. *Preprint* URL `http://www.optimization-online.org/DB_HTML/2013/06/3925.html`.

Dressler M, Iliman S, de Wolff T (2017) A Positivstellensatz for sums of nonnegative circuit polynomials. *SIAM Journal on Applied Algebra and Geometry* 1(1):536–555.

Drusvyatskiy D, Ioffe AD, Lewis AS (2015) Transversality and alternating projections for nonconvex sets. *Foundations of Computational Mathematics* 15(6):1637–1651.

Dvijotham K, Nguyen H, Turitsyn K (2018) Solvability regions of affinely parameterized quadratic equations. *IEEE Control Systems Letters* 2(1):25–30.

Folland GB (1999) *Real analysis: Modern techniques and their applications*, volume 40 (John Wiley & Sons).

Georghiou A, Tsoukalas A, Wiesemann W (2020) A primal–dual lifting scheme for two-stage robust optimization. *Operations Research* 68(2):572–590.

Ghaddar B, Marecek J, Mevissen M (2016) Optimal power flow as a polynomial optimization problem. *IEEE Transactions on Power Systems* 31(1):539–546.

Goldsztejn A, Jaulin L (2006) Inner and outer approximations of existentially quantified equality constraints. *International Conference on Principles and Practice of Constraint Programming*, 198–212 (Springer).

Golestaneh F, Pinson P, Azizipanah-Abarghooee R, Gooi HB (2018) Ellipsoidal prediction regions for multivariate uncertainty characterization. *IEEE Transactions on Power Systems* 33(4):4519–4530.

Henrion D, Lasserre JB, Savorgnan C (2009) Approximate volume and integration for basic semialgebraic sets. *SIAM Review* 51(4):722–743.

Isenberg NM, Akula P, Eslick JC, Bhattacharyya D, Miller DC, Gounaris CE (2021) A generalized cutting-set approach for nonlinear robust optimization in process systems engineering. *AIChE Journal* e17175.

Jaleeli N, VanSlyck LS, Ewart DN, Fink LH, Hoffmann AG (1992) Understanding automatic generation control. *IEEE Transactions on Power Systems* 7(3):1106–1122.

Josz C, Molzahn DK (2018) Lasserre hierarchy for large scale polynomial optimization in real and complex variables. *SIAM Journal on Optimization* 28(2):1017–1048.

Kuang X, Ghaddar B, Naoum-Sawaya J, Zuluaga LF (2017) Alternative LP and SOCP hierarchies for ACOPF problems. *IEEE Transactions on Power Systems* 32(4):2828–2836.

Kuryatnikova O, Vera JC, Zuluaga LF (2019) Copositive certificates of non-negativity for polynomials on unbounded sets. *Working paper, Tilburg University* .

Lasserre J (2001) Global optimization problems with polynomials and the problem of moments. *SIAM J. Optim.* 11(3):796–817.

Lasserre JB (2007) A sum of squares approximation of nonnegative polynomials. *SIAM Review* 49(4):651–669.

Lasserre JB (2015) Tractable approximations of sets defined with quantifiers. *Mathematical Programming* 151(2):507–527.

Lee D, Nguyen HD, Dvijotham K, Turitsyn K (2019) Convex restriction of power flow feasibility sets. *IEEE Transactions on Control of Network Systems* 6(3):1235–1245.

Lee D, Turitsyn K, Molzahn DK, Roald LA (2020) Robust AC optimal power flow with convex restriction. *arXiv preprint arXiv:2005.04835* .

Lee D, Turitsyn K, Slotine J (2019) Sequential convex restriction and its applications in robust optimization. *arXiv preprint arXiv:1909.01778* .

Löfberg J (2004) YALMIP : A toolbox for modeling and optimization in MATLAB. *Proceedings of the CACSD Conference* (Taipei, Taiwan).

Lorca A, Sun XA (2016) Multistage robust unit commitment with dynamic uncertainty sets and energy storage. *IEEE Transactions on Power Systems* 32(3):1678–1688.

Lorca Á, Sun XA (2018) The adaptive robust multi-period alternating current optimal power flow problem. *IEEE Transactions on Power Systems* 33(2):1993–2003.

Louca R, Bitar E (2019) Robust AC optimal power flow. *IEEE Transactions on Power Systems* 34(3):1669–1681.

Marshall M (2006) Representations of non-negative polynomials having finitely many zeros. *Annales de la Faculté des sciences de Toulouse: Mathématiques* 15(3):599–609.

Misra S, Vuffray M, Zlotnik A (2020) Monotonicity properties of physical network flows and application to robust optimal allocation. *Proceedings of the IEEE* 108(9):1558–1579.

Molzahn DK (2017) Computing the feasible spaces of optimal power flow problems. *IEEE Transactionson Power Systems* 32(6):4752–4763.

Molzahn DK, Hiskens IA (2019) A survey of relaxations and approximations of the power flow equations. *Foundations and Trends in Electric Energy Systems* 4(1-2):1–221.

Molzahn DK, Roald LA (2018) Towards an AC optimal power flow algorithm with robust feasibility guarantees. *20th Power Systems Computation Conference (PSCC)*, 1–7.

Mühlpfordt T, Roald L, Hagenmeyer V, Faulwasser T, Misra S (2019) Chance-constrained AC optimal power flow: A polynomial chaos approach. *IEEE Transactions on Power Systems* 34(6):4806–4816.

Nie J (2014) Optimality conditions and finite convergence of Lasserre's hierarchy. *Mathematical Programming* 146(1):97–121.

Pólik I, Terlaky T (2007) A survey of the S-lemma. *SIAM Review* 49(3):371–418.

Postek K, Hertog Dd (2016) Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing* 28(3):553–574.

Powers V, Reznick B (2001) A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *Journal of Pure and Applied Algebra* 164(1-2):221–229.

Putinar M (1993) Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal* 42(3):969–984.

Roald L, Andersson G (2018) Chance-constrained AC optimal power flow: Reformulations and efficient algorithms. *IEEE Transactions on Power Systems* 33(3):2906–2918.

Scheiderer C (2005) Distinguished representations of non-negative polynomials. *Journal of Algebra* 289(2):558–573.

Schweighofer M (2005) Certificates for nonnegativity of polynomials with zeros on compact semialgebraic sets. *manuscripta mathematica* 117(4):407–428.

Spivak M (1995) *Calculus on Manifolds. A Modern Approach to Classical Theorems of Advanced Calculus.* (Addison-Wesley Publishing Company).

Stott B, Jardim J, Alsaç O (2009) DC power flow revisited. *IEEE Transactions on Power Systems* 24(3):1290–1300.

Stuhlmacher A, Mathieu JL (2020) Water distribution networks as flexible loads: A chance-constrained programming approach. *21st Power Systems Computation Conference (PSCC)*, 1–8.

Venzke A, Halilbasic L, Markovic U, Hug G, Chatzivasileiadis S (2018) Convex relaxations of chance constrained AC optimal power flow. *IEEE Transactions on Power Systems* 33(3):2829–2841.

Vrakopoulou M, Margellos K, Lygeros J, Andersson G (2013) A probabilistic framework for reserve scheduling and $\mathrm{N}-1$ security assessment of systems with high wind power penetration. *IEEE Transactions on Power Systems* 28(4):3885–3896.

Wolkowicz H, Saigal R, Vandenberghe L, eds. (2000) *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications* (Kluwer Academic Publishers).

Yakubovich VA (1977) S-procedure in nonlinear control theory. *Vestnik Leningrad University* 4(1):73–93, english translation; original Russian publication in Vestnik Leningradskogo Universiteta, Seriya Matematika (1971), 62–77.

Yanıkoğlu İ, Gorissen BL, den Hertog D (2019) A survey of adjustable robust optimization. *European Journal of Operational Research* 277(3):799–813.

Zeng B, Zhao L (2013) Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters* 41(5):457 –461.

Zimmerman RD, Murillo-Sánchez CE, Thomas RJ (2010) MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems* 26(1):12–19.