Two-Stage Adjustable Robust Optimization Applied to Optimal Power Flow Problems

Olga Kuryatnikova

Econometric Institute, Erasmus University Rotterdam, the Netherlands, kuryatnikova@ese.eur.nl

Bissan Ghaddar

Ivey Business School, Western University, bghaddar@ivey.ca

Daniel K. Molzahn

School of Electrical and Computer Engineering, Georgia Institute of Technology, molzahn@gatech.edu

In this work, we consider two-stage optimization problems with polynomial inequality and equality constraints 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, gas or water networks. We propose an iterative algorithm to build a sequence of approximately robustly feasible solutions with an improving objective value 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. We implement our approach for AC Optimal Power Flow and demonstrate the performance of our proposed method on MATPOWER instances. This paper focuses on problems with polynomial equalities, quadratic inequalities and ellipsoidal uncertainty, but the approach is suitable for more general setups.

Key words: 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 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, ARO gives rise to more flexible decisions and thus can be less conservative than robust optimization.

With rapidly increasing uncertainties in both the demand and supply of resources, networked infrastructure problems are important applications of ARO. Examples of relevant networked infrastructures include electricity (see Bienstock et al. (2014)), natural gas (see Aßmann et al. (2018), 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)). Therefore, it is still a challenge to solve such problems under uncertainty for problems of realistic size. The approach proposed in this paper computes operating points that are approximately robustly feasible with respect to the inequality constraints in ACOPF problems. Even though we sacrifice some precision by using approximations, we gain efficiency as a result, so our approach runs in short time for middle-sized ACOPF instances, and we propose an extension to apply the approach to large instances.

1.1. Existing ARO solution approaches

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 optimization problem can be obtained using approaches such as vertex enumeration

algorithms (Bienstock and Özbay 2008), cutting plane algorithms (Bertsimas et al. 2012), or refor-mulation 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. However, 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). This paper focuses on such cases by replacing the non-linear functions defining the equalities by their piecewise affine approximations that would reflect the original functions as closely as needed.

1.2. Contribution

We begin this section by describing our main contributions.

1. This paper presents a general framework for approximately solving two-stage polynomial ARO problems where the second-stage variables are linked to the first-stage variables through a set of equality and inequality constraints, see Section 2. The presence of equalities increases the difficulty of ARO. Therefore, we address equality and inequality constraints separately. We propose exploiting piecewise affine approximations of the non-linear equalities with respect to the state variables to allow for closed-form decision rules. Using a dynamic scheme, we partition the domain of the state variables into subsets small enough to represent each equation well by its first-order Taylor approximation in state variables. Thus, the non-linear equations are replaced by linear (in state variables) 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 as a problem with polynomial equalities and tractable conic constraints over the control variables only. Finally, we design an alternating projections algorithm to find locally optimal solutions for such conic problems with polynomial equalities.

2. We apply the proposed framework to the AC optimal power flow (ACOPF) problem with uncertainty in power supply and demand. The ACOPF is formulated as a non-convex quadratically constrained quadratic problem with quadratic equality and inequality constraints, and an ellipsoidal uncertainty set is considered. As shown in Section 5, Our numerical results for ACOPF 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 we target ARO problems with general equalities and polynomial inequalities. Additionally, we do not use convex approximations of inequality constraints, such as in Lasserre (2015), Lee et al. (2021), 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 and use 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 the work of 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 2 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 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

We start this section by defining the notation. We denote the range of matrix A by $\mathcal{R}(A)$. We denote the space of $n \times n$ symmetric matrices by $\mathbb{S}^n$ and for $A, B \in \mathbb{S}^n$, the trace inner product of A and B is denoted by $\langle A, B \rangle := \text{trace}(AB)$. We use the notation $[n]$ for the set $\{1, \dots, n\}$. For a vector V of length n, we denote the ith entry of V by V_i and $\text{Diag}(V)$ the operator that creates a diagonal matrix with V as the diagonal. 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 $\sup_{x \in A} \|f(x) - g(x)\| \le \varepsilon$ for some given norm $\|\cdot\|$. We say that two vectors a, b are ε-close to each other if $\|a - b\| \le \varepsilon$.

We let n_ζ, n_y, n_x, m_{eq}, and m_{in} be natural numbers and consider semialgebraic sets $S_y \in \mathbb{R}^{n_y}$, $S_x \in \mathbb{R}^{n_x}$. Let m_x be the number of inequalities defining S_x. Consider *continuous* mappings $f : \mathbb{R}^{n_y} \to \mathbb{R}$, $G : \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{m_{in}}$, $L_x : \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_x} \to \mathbb{R}^{m_{eq}}$ and $L_y : \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_y} \to \mathbb{R}^{m_{eq}}$, where G and f are polynomial maps, and L is polynomial in y.

We begin with the following *two-stage nominal* problem, in other words, the two-stage problem without uncertainty.

PROBLEM 1 (NOMINAL TWO-STAGE POLYNOMIAL OPTIMIZATION PROBLEM).

$$
\begin{aligned}
\inf_{y,x} \quad & f(y) \\
\text{s.t.} \quad & y \in S_y \\
& L_i(y, x) = 0 \text{ for all } i \in [m_{eq}] \\
& G_i(y, x) \ge 0 \text{ for all } i \in [m_{in}] \\
& x \in S_x.
\end{aligned}
$$

Problem (1) occurs frequently in dynamic systems optimization, which inspired this paper. In such problems, an operator sets up the values of control variables y, and the state variables x are determined afterwards according to a system of equations defining the equilibrium. Some examples of dynamic systems optimization are energy problems (ACOPF as described in Section 5, optimal power dispatch presented by Bingane et al. (2019)), water problems (the valve placement problem by Ghaddar et al. (2017)) or gas problems (passive gas network feasibility problem studied by Aßmann et al. (2018)). Moreover, problem (1) describes the more general class of *bilevel optimization problems* where after setting the values of y in the first stage, the second stage variables are chosen from the set of optimal solutions of the second-stage optimization problem. The KKT optimality conditions of the second-stage problem can be written as a system of equalities, where the final set of second-stage variables x consists of the original second-stage variables and Lagrange multipliers of the second-stage problem. As a result one obtains the so-called *complementarity formulation* of the initial bilevel optimization problem, which has the form of problem (1). A typical example of a bilevel optimization problem is the Stackelberg competition in economics. Some examples of bilevel problems in engineering can be found in Raghunathan and Biegler (2003), Baumrucker and Biegler (2010).

REMARK 1. Problem (1) allows for binary variables and absolute values. One can write binary constraints on a variable a as quadratic equality constraints $a = a^2$. One can write the constraint $a = |b|$ as $a^2 = b^2, a \geq 0$.

Next we show some concrete examples of problem (1). The examples have a special form: the equality constraints are separable in x and y. While such a requirement is not necessary for our approach, the approach works most efficiently on such problems as seen in Section 3. We begin with the ACOPF problem, which is the main use case considered in this paper. It can be written as follows using the general notation above.

$$
\begin{aligned}
\min_{y,x} \quad & y^\top P y + p^\top y + p_0 \\
\text{s.t.} \quad & Ay \leq b \\
& x^\top Q_i x + r_i = y_i \quad \text{for all } i \in [m_{eq}] \\
& x^\top Q_j x + r_j \geq 0 \quad \text{for all } j \in [m_{in}],
\end{aligned}
$$

where y are active powers and voltage magnitudes on PV buses, and x are voltages in rectangular form, see Section 5 for the full problem formulation. Another example is the valve setting problem in water distribution networks from Ghaddar et al. (2017) with the following formulation:

$$
\begin{aligned}
\min_{y,x} \quad & d^\top y \\
\text{s.t.} \quad & A_1 y \leq b_1, A_2 x \leq b_2 \\
& x^\top Q_i x + q_i^\top x + c_i^\top y + r_i = 0 \quad \text{for all } i \in [m_{eq}] \\
& x^\top Q_j x + q_j^\top x + y^\top C_j y + c_j^\top y + x^\top R_j y + r_j \leq 0 \quad \text{for all } j \in [m_{in}],
\end{aligned}
$$

where y are pressure heads and valve placement indicators, and x are flow rates and absolute values of flow rates.

In practice given the uncertain demand and supply that is encountered in such applications, one often aims at solving problem (1) under uncertainty, which results in the next formulation.

PROBLEM 2 (TWO-STAGE POLYNOMIAL ARO PROBLEM).

$$
\begin{aligned}
z = \inf_{y} \quad & f(y) \\
\text{s.t.} \quad & y \in S_y \\
& \text{and for any } \zeta \in \Omega \text{ there exists } x \text{ such that the following holds:} \\
& L_i(y, \zeta, x) = 0 \text{ for all } i \in [m_{eq}] && \text{(2a)} \\
& G_i(y, \zeta, x) \geq 0 \text{ for all } i \in [m_{in}] && \text{(2b)} \\
& x \in S_x. && \text{(2c)}
\end{aligned}
$$

Problem (2) 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 $\zeta \in \Omega$. Any solution to problem (2) is feasible for the underlying nominal problem and robust against potential uncertainty. We obtain the nominal problem by setting $\zeta = 0$ in problem (2). Now we summarize main assumptions used in this paper.

ASSUMPTION 1 **(The characteristics of the objective and the constraints)**.

(a) Constraints (2a), (2b) are polynomial in y. Constraints (2b) and (2c) are polynomial in x. Sets S_y, S_x, Ω are compact and basic semialgebraic.

(b) Constraints (2a), (2b) are polynomial of degree at most two in ζ.

(c) Ω is an ellipsoidal uncertainty set of the form:

$$\Omega = \left\{ \zeta \in \mathbb{R}^{n_\zeta} : \zeta^\top \Sigma \zeta + \sigma^\top \zeta + r \geq 0, \ j \in [m_\zeta] \right\}, \tag{3}$$

where Σ is positive semidefinite.

We need assumptions on the degree of ζ and the shape of Ω to efficiently eliminate ζ from the problem, see Section 4.1. We have chosen the ellipsoidal uncertainty as the base case since it is convenient for our approach, frequently appears in the literature being less conservative than, for instance, box uncertainty, and has interpretations from both robust and chance constrained perspectives, see, e.g, (Golestaneh et al. 2018, Chen et al. 2010). In Section EC.1 we relax Assumptions 1 (b) and (c), and we also consider more shapes of Ω in Section 4.1.2 for problems that are linear in the uncertainty.

ASSUMPTION 2 **(Assumptions without loss of generality)**.

(a) S_x is defined by inequalities.

   The assumption is w.l.o.g. since if the definition of S_x contains equalities, they could be moved to (2a) as a preprocessing step.

(b) $n_x = n_{eq}$ and there are no redundant equality constraints in (2a).

   The assumption is w.l.o.g. If $n_x \leq n_{eq}$ and there are redundant constraints, they could be detected and eliminated as a preprocessing step. If $n_x > n_{eq}$, then some state variables are free and could be added to the pool of control variables.

(c) Inequality constraints (2b), (2c) have degree at most two in x.

   The assumption is w.l.o.g. since there is no restriction on the degree of x in the equality constraints (2a). If some monomials have higher degree in x, they can be eliminated using variable substitution and increasing the number of state variables and equality constraints. We emphasize that this procedure does not influence the final size of the problems we solve

since this size depends on the number of control and uncertainty variables only. We relax the assumption in Section EC.1.

(d) *Objective f is amenable for optimization (e.g., convex quadratic or linear).*

The assumption is w.l.o.g. since if f is not convex, we introduce the epigraph control variable and add the epigraph constraint to S_y.

Our goal is to approximate the original problem by *a PO problem in control variables y*, since they are in fact the only variables we need to decide upon. This procedure brings two advantages: first, we obtain a PO problem instead of the original 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 optimal control applications is smaller than the number of state variables.

We conclude this section with a general description of our approach and the intuition behind it. If we could analytically solve equalities (2a) 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*. If we assume that equalities (2a) are such that there exists a known second-stage decision rule $\delta : \mathbb{R}^{n_y \times n_\zeta} \to \mathbb{R}^{n_x}$, then problem (2) is equivalent to the following problem:

PROBLEM 4.

$$z = \inf_y \quad f(y)$$
$$\text{s.t.} \quad y \in S_y$$
$$\qquad G_i(y, \zeta, \delta(y, \zeta)) \geq 0 \text{ for all } \zeta \in \Omega, \ i \in [m_{in}]$$
$$\qquad \delta(y, \zeta) \in S_x \text{ for all } \zeta \in \Omega.$$

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 unique analytical expression for the decision rule in the problem. However, such expressions exist on small subsets of $S_x \times S_y \times S_\zeta$ under known conditions according to the implicit function theorem (see, e.g., Spivak (1995)). We work on such subsets and replace all terms in the equalities which include monomials in x of degrees larger than one by their linear approximations. As a result, on each subset we approximate (2a) as $L_1(y, \zeta) + Ax$, for some matrix A and explicitly derive $x = \delta(y, \zeta)$ if A is invertible. We substitute δ in problem (2), eliminate the equality constraints and solve a problem of the form (4). We present the approach in detail in the next sections.

3. 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. Our piecewise approximation decision is based on the following general result.

PROPOSITION 1. *Let $A \subset \mathbb{R}^n$ be compact, and let the map $F : A \to \mathbb{R}^m$ be such that F_i, $i \in [m]$ are continuous functions on A. For any $\varepsilon > 0$ 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., Corollary 4.50 in (Folland 1999). Hence the set of piecewise affine functions is dense as well. Therefore, for any $\varepsilon > 0$ and each continuous function 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. $\square$

In this work, we illustrate the idea of piecewise affine equality constraint approximations using first-order Taylor approximations on partitions of S_x. In addition to its simplicity, the main advantage of the Taylor approximation is its good fit for the original function around the approximation point. For our approach, we only need to approximate the part of L (2a) that is non-linear in x, and we keep the rest of the non-linearities in the problem. We split 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. If the Jacobian of the linear approximation on each subset of a given partition of S_x is invertible at the point of approximation (y_0, ζ_0, x_0), then one can replace the non-linear equality constraints $L(y, \zeta, x)$ by their Taylor approximations $Ax + f(y, \zeta)$ within each subset, solve the resulting linear system, and obtain decision rules (an approximate solution $x = h(y, \zeta)$) that are close to the true decision rules.

If the matrix A^j from (5) is invertible, we can eliminate all state variables x from the problem using (5) and Assumption 1 (b) and solve J subproblems of the form (4) obtaining the optimal values z_j. The minimal optimal value z_j among all subproblems and the corresponding solution are the optimal value and solution to the piecewise affine approximation.

To implement the approach, we need to partition S_x into subsets, which can be done in various ways and is not a trivial step. For instance, the partition could be too large or suboptimal and the optimization over all subsets in such partitions can be time consuming and imprecise. To overcome these difficulties, instead of working with a given partition, we generate it dynamically. Namely, since the nominal problem (1) for which we want to find a robust solution is usually feasible, we start from some *nominal feasible solution* at hand, that is, the control solution $\hat{y}_1$ feasible for $\hat{\zeta} = 0$. In particular, we suggest choosing a solution that is close to the original optimal solution and we obtain the corresponding $\hat{x}_1$ such that $L(\hat{x}_1, \hat{y}_1, 0) = 0$. If the Jacobian of L with respect to

---

Algorithm 1: Piecewise affine approximations of L in (2a) using Taylor series

---

Preprocessing: If $L(y, \zeta, x)$ contains products of ζ or y with x, replace $L(x, y, \zeta)$ with $G(a, x) + F(y, \zeta)$, where $F(y, \zeta)$ is the part of L that depends on (y, ζ) only, and $a = a(y, \zeta)$ is the variable that represents the monomials in (y, ζ) (see Assumptions 1 and(a), (b)) that occur in products with x.

Input   : $G(a, x)$, $a(y, \zeta)$, $S_x = \bigcup_{j=1}^{J} S_x^j$ for some $J > 0$. Vectors $\hat{\zeta} \in \Omega$, $\hat{x}_1 \in S_x^1, \ldots, \hat{x}_J \in S_x^J$, $\hat{y}_1, \ldots, \hat{y}_J \in S_y$ such that $L(\hat{x}_j, \hat{y}_j, \zeta) = 0$ for $j = 1, \ldots, J$.

1 Calculate $\hat{a}_j := a(\hat{y}_j, \hat{\zeta})$ for $j = 1, \ldots, J$;
2 for $j = 1, \ldots, J$ **do**
3   Define the Taylor approximation of $G(a, x)$ in $(\hat{a}_j, \hat{x}_j)$ by
4 $$T(\hat{a}_j, \hat{x}_j, G)(a, x) := A^j x + B^j a + c_j ;$$
5   Approximate non-linear equalities L in (2a) on $S_x^j \times S_y \times \Omega$ by the following system of equalities linear in x:

$$L(x, y, \zeta) \approx A^j x + \hat{F}^j(y, \zeta) := T(\hat{a}_j, \hat{x}_j, G)(a(y, \zeta), x) + F(y, \zeta). \qquad (5)$$

---

x at $\hat{x}_1$ is invertible, we optimize over a small set around $\hat{x}_1$ defined by $S_x^1 = \{x \in S_x : \|\hat{x}_1 - x\| \leq \varepsilon\}$ for a chosen $\varepsilon > 0$ and norm type. If we find a robustly feasible solution $\hat{y}_2 \neq \hat{y}_1$, we set $\hat{x}_2$ such that $L(\hat{y}_2, \hat{x}_2, 0) = 0$ and repeat the procedure. The process is described in the following algorithm.

THEOREM 1. *Algorithm 2 stops after a finite number of iterations. The obtained control solution is robustly feasible for the piecewise affine approximation and ensures nominal feasibility of equations* (2a).

Proof.   The algorithm stops after finitely many iterations since *tol* > 0 and S_x is bounded. The other statements follow directly from the steps of the algorithm. $\square$

Algorithm 2 ensures nominal feasibility of the equality constraints, as well as robust feasibility of the affine approximations to those constraints. The existing algorithms for problem (1) under uncertainty ensure nominal feasibility of equations (2a) only see, e.g., (Molzahn and Roald 2018, Weisser et al. 2018). In this sense, our approximations are at least as good as the existing results. If the affine approximations are close enough to the original constraints, full-rank Jacobian of the approximation is likely to guarantee robust feasibility of the original equality constraints.

We finish this section by mentioning some limitations of Algorithm 2. First, it 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. One can restart the algorithm and consider heuristics that help to move away from local optima. We also note that the algorithm requires a trade-off since each subproblem (6) considers one affine piece at a time. On one hand, we are interested in finer approximations to

---

Algorithm 2: Dynamic piecewise affine approximation of problem (2)

Input : $tol > 0$, f_0, norm type $\|\cdot\|$, a rule for ε_j, and the initial nominal feasible solution $(\hat{y}_1, \hat{x}_1)$ such that the Jacobian in (5) has a full rank at $(\hat{x}_1, \hat{y}_1, 0)$

Output: y^*

1 Set $j := 1$, $y^* = \hat{y}_1$, $f_1 = f(\hat{y}_1)$;

2 while $f_j - f_{j-1} \leq tol$ **do**

3   Replace the non-linear equality constraints $L(y, \zeta, x)$ in each subproblem using Algorithm 1. If needed, use Remark 2 to make approximations more precise ;

4   **if** *Jacobian of the approximation is not full-rank* **then**

5   Adjust S_x^{j-1} and reoptimize to obtain x_j with full Jacobian rank ;

6   **else**

7   Update ε_j according to the rule in the Input ;

8   Define $S_x^j := \{x \in S_x : \|x_j - x\| \leq \varepsilon_j\}$;

9   Solve:

   PROBLEM 6.

$$z_j = \inf_y \quad f(y) \tag{6a}$$

$$\text{s.t.} \quad y \in S_y,$$

$$G_i\left(y, \zeta, (A^j)^{-1} \hat{F}^j(y, \zeta)\right) \geq 0 \text{ for all } \zeta \in \Omega, \ i \in [m_{in}] \tag{6b}$$

$$(A^j)^{-1} \hat{F}^j(y, \zeta) \in S_x^j \text{ for all } \zeta \in \Omega, \tag{6c}$$

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

   if *Problem* (6) *is feasible* **then**

     Substitute the optimal control solution $\hat{y}_j$ and $\zeta = 0$ in problem (2) and solve equations (2a) for state variables $\hat{x}_j$;

     if *Solution* $(\hat{x}_j, \hat{y}_j, 0)$ *is feasible for problem* (2) **then**

       Save $(\hat{x}_j, \hat{y}_j)$, set $y^* := \hat{y}_j$, $f_j = f(\hat{y}_j)$

     else

       Set $f_j := \infty$

   else

     Set $f_j := \infty$

   Set $j := j + 1$;

---

reflect the equality constraints better and on the other hand, if the approximations are too fine, problem (6) will be infeasible for realistic uncertainty sets. Therefore, we have to balance out these two factors, and additional analysis of their influence is needed. Finally, in Algorithm 1 we do not partition S_y and Ω into subsets to obtain piecewise affine approximations, we partition S_x only. As a result, our approach is *most efficient* when the equality constraints (2a) are *separable* in x and (y, ζ). Otherwise the approach might be less precise but still applicable.

REMARK 2. Our approach can be made more precise for problems where equalities (2a) contain products of x with ζ, y or to work with larger uncertainty sets by constructing piecewise affine 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) to split the uncertainty set.

4. Quadratic inequality constraints under uncertainty

Solving the piecewise affine approximation of problem (2) amounts to solving J problems of the form (6), which are quadratic problems under uncertainty by Assumptions 1 (a), (b). The problem has complicating constraints (6b) and (6c). Under our assumptions, Ω is not a finite set, so each constraint in (6b) and (6c) results in infinitely many quadratic inequality constraints. Our first step is to eliminate ζ from (6b), (6c). The constraints in question have the following general form:

$$h(y,\zeta) \geq 0 \text{ for all } \zeta \in \Omega. \tag{7}$$

For example, for constraints (6b) we have $h_i(y,\zeta) := G_i\left(y,\zeta,(A^j)^{-1}\hat{F}^j(y,\zeta)\right)$, $i \in [m_{in}]$. Since (7) concerns the non-negativity of a polynomial on the set Ω, whose description does not involve y, in the next subsection 4.1 we treat y as a parameter and answer the question of when a polynomial of degree at most two in ζ is non-negative on Ω. The answer leads to a reformulation of (7) that contains no ζ. In this way we eliminate the uncertainty ζ from problem (6). After that, we stop treating y as a parameter, and y becomes the only variable left in the reformulated problem (6), which we solve in subsection 4.2.

4.1. Eliminating the uncertainty from the problem

Under Assumptions 1 and 2, all inequalities in problem (2) are of degree at most two in x and ζ. We begin uncertainty elimination with the most general case when all inequalities in problem (2) are quadratic in x and ζ, and therefore h in (7) is quadratic in ζ. After that we consider the special situation where the inequalities in (2) are linear in x and ζ, and thus h in (7) is linear in ζ. The latter case allows us to exploit not only ellipsoidal but also other types of uncertainty sets.

4.1.1. Quadratic inequality constraints and quadratic uncertainty

When all inequalities in problem (2) are quadratic in x and ζ, we obtain h in (7) that is quadratic in ζ.

$$h(y,\zeta) := \zeta^\top A \zeta + (y^\top B + b^\top)\zeta + c^\top y + d + g(y) \geq 0, \tag{8}$$

for some given parameters A, B, b, c, d, and a function g that contains all monomials non-linear in y. Recall that Ω has the form (3). Such a combination of h and Ω allows reformulating (7) and eliminating the uncertainty from it due to the *S-lemma*.

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

1. *$h(x) \geq 0$ for all $x \in \mathbb{R}^n$ such that $g(x) \geq 0$.*

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

S-lemma is well known in robust optimization (see Ben-Tal et al. (2009)) but is usually applied to convex problems in y, ζ while we use it for a general quadratic constraint (7).

PROPOSITION 3. *Constraint* (7) *with h as in* (8) *and the uncertainty set* (3) *holds if and only if there exist* λ, γ *such that*

$$\begin{bmatrix} \gamma + c^\top y + d + \lambda r & \frac{1}{2}(y^\top B + b^\top + \lambda \sigma^\top) \\ \frac{1}{2}(B^\top y + b + \lambda \sigma) & \lambda \Sigma + A \end{bmatrix} \succeq 0$$
$$\lambda \geq 0,$$
$$g(y) = \gamma,$$

where Σ, σ *and r are the parameters from* (3), *and other parameters are defined in* (8).

Proof. We have

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

$$\overset{\text{Prop. 2}}{\iff} \zeta^\top A \zeta + (y^\top B + b^\top) \zeta + c^\top y + d + g(y) \geq 0 \text{ for all } \zeta \in \mathbb{R}^{n_\zeta} : \zeta^\top \Sigma \zeta + \sigma^\top \zeta + r \leq 0$$

$$\iff \zeta^\top A \zeta + (y^\top B + b^\top) \zeta + c^\top y + d + g(y) - \lambda(-r - \zeta^\top \Sigma \zeta - \sigma^\top \zeta) = \begin{bmatrix} 1 \\ \zeta \end{bmatrix}^\top S \begin{bmatrix} 1 \\ \zeta \end{bmatrix},$$

$$S \succeq 0, \ \lambda \geq 0$$

$$\iff \begin{bmatrix} \gamma + c^\top y + d + \lambda r & \frac{1}{2}(y^\top B + b^\top + \lambda \sigma^\top) \\ \frac{1}{2}(B^\top y + b + \lambda \sigma) & \lambda \Sigma + A \end{bmatrix} \succeq 0, \ g(y) = \gamma, \ \lambda \geq 0. \qquad \square$$

4.1.2. Linear inequality constraints and conic uncertainty

Now we consider a special case when h in (7) is linear in ζ

$$h(y, \zeta) := (y^\top B + b^\top) \zeta + c^\top y + d + g(y) \geq 0, \tag{10}$$

for some given parameters B, b, c, d, and a function g that contains all monomials non-linear in y. In this case we can relax Assumption 1 (c) to define Ω by a conic constraint:

$$\Omega = \{\zeta \in \mathbb{R}^{n_\zeta} : \Sigma \zeta + \sigma \in \mathcal{C}\}, \tag{11}$$

where Σ, σ are parameters and $\mathcal{C}$ is a *proper semialgebraic* cone. When Ω is an ellipsoid as in Assumption 1 (c), $\mathcal{C}$ is the second-order cone. The cone $\mathcal{C}$ can also be $\mathbb{R}_+^{n_\zeta}$, then Ω is a box or budget uncertainty set. These uncertainty sets are common in robust optimization, see, e.g., Ben-Tal et al. (2009). Positive semidefinite cone can be used as $\mathcal{C}$ too.

PROPOSITION 4 (**Constraints linear in uncertainty and conic uncertainty sets**).
Constraint (7) *with h as in* (10) *and the uncertainty set* (11) *holds if and only if there exist* z, γ *such that*

$$-\sigma^\top z + c^\top y + d + \gamma \geq 0$$
$$\Sigma^\top z = B^\top y + b$$
$$z \in \mathcal{C}^*$$
$$g(y) = \gamma,$$

where Σ, σ are the parameters from (11)*, $\mathcal{C}^*$ is the dual cone of $\mathcal{C}$, and other parameters are defined in* (10)*.*

Proof.   The result follows from conic duality.

$$(y^\top B + b^\top)\zeta + c^\top y + d + g(y) \geq 0 \text{ for all } \zeta \in \Omega = \{\zeta \in \mathbb{R}^{n_\zeta} : \Sigma\zeta + \sigma \in \mathcal{C}\}$$
$$\Longleftrightarrow c^\top y + d + g(y) + \min_{\zeta \in \mathbb{R}^{n_\zeta}} \{(y^\top B + b^\top)\zeta : \Sigma\zeta + \sigma \in \mathcal{C}\} \geq 0$$
$$\Longleftrightarrow c^\top y + d + g(y) + \max_z \{-\sigma^\top z : \Sigma^\top z = (y^\top B + b^\top), z \in \mathcal{C}^*\} \geq 0$$
$$\Longleftrightarrow c^\top y + d + g(y) - \sigma^\top z \geq 0$$
$$\Sigma^\top z = (y^\top B + b^\top), \ z \in \mathcal{C}^*. \qquad\qquad \square$$

As we say in Assumption 2(c), one can eliminate all non-linear terms in x from the inequality constraints (2b) by increasing the number of state variables and equality constraints. Hence it is always possible to make $h(y, \zeta)$ linear in ζ and use the result from this subsection if the original problem (2) is linear in ζ. This procedure could lead to a loss of precision in comparison to the quadratic formulation from the previous section due to the larger number of approximated state variables in the problem. On the other hand, the reformulation for the ellipsoidal uncertainty becomes SOCP instead of SDP (c.f. Propositions 3 and 4), and thus could potentially be solved faster.

For problems (2) linear in ζ, the choice between the quadratic (Proposition 3) and linear (Proposition 4) formulations depends on the shape of the uncertainty set and problem size. Our application in Section 5 is quadratic in x and linear in ζ, we use the quadratic approach as it is more straightforward, and it performs well for middle-sized instances. The linear approach as in Proposition 4 could be considered in the future research to work with large-scale instances.

4.2. Alternating projections algorithm

In Section 4.1 we eliminate ζ from constraints (6b) and (6c) in problem (6) and add auxiliary control variables (e.g., λ, γ in Proposition 3). Next we *reformulate* problem (6) as

PROBLEM 13.

$$z_j = \inf_y \quad f(y) \tag{13a}$$
$$\text{s.t.} \quad H(y) = \gamma \tag{13b}$$
$$F \begin{bmatrix} y \\ \gamma \end{bmatrix} \in \mathcal{C}, \tag{13c}$$

where γ are auxiliary control variables as in Propositions 4 and 3, and y includes other original and auxiliary control variables; H is a polynomial mapping, F is a linear mapping, and $\mathcal{C}$ is a proper semialgebraic cone amenable for optimization (e.g., positive semidefinite cone, second-order cone or the non-negative orthant).

By Assumptions 1 (a) and 2 (d), the problem is polynomial, has finitely many equality constraints and a convex objective function. We fist obtain a solution feasible for (13c) and then iteratively transform it into a solution feasible for the whole problem (13). Notice that problem (13) cannot be unbounded as the feasible set of its nominal problem is compact by Assumption 1 (a). We begin with a lower bound for the problem, which can be obtained using any relaxation of the polynomial equality constraints. One can use lifting techniques, e.g., the reformulation-linearization technique as in Sherali and Tuncbilek (1992). The lower-bound relaxation might provide a solution that is infeasible for (13b). To obtain a feasible solution from the solution to the relaxation, we use the alternating projection method as presented in Algorithm 3. Let z^U be an upper bound on $f(y)$ in problem (13). Define two sets

$$\mathcal{A} := \{(y, \gamma) : (13c) \text{ holds}, f(y) \leq z^U\}, \tag{14}$$

$$\mathcal{B} := \{(y, \gamma) : (13b) \text{ holds}\} \tag{15}$$

We denote $y := (y^{nc}, y^c)$, where y^{nc}, y^c are the subsets of variables y that are involved (resp. not involved) in the non-convex constraints (13b).

Algorithm 3: Alternating projections algorithm for problem (13)

Input: small $tol > 0$, upper bound $z^U \geq z_j$, $(\nu_i)_{i=1}^N \in (0,1]$, $N \geq 1$

1 Solve a lower bound relaxation of problem (13) and denote its solution by (y_0, γ_0) and its objective value by z^L;

2 **if** *the lower bound relaxation is infeasible,* **then**

3 $\quad$ Lower-bound relaxation is infeasible, stop, return "Problem (13) infeasible"

4 Set $i = 0$ and choose (y_1, γ_1) such that $\|(y_0^{nc}, \gamma_0) - (y_1^{nc}, \gamma_1)\| > tol$;

5 **while** $\|(y_0^{nc}, \gamma_0) - (y_1^{nc}, \gamma_1)\| > tol$ *and* $i \leq N$ **do**

6 $\quad$ Set $i := i + 1$;

7 $\quad$ Project on $\mathcal{B}$: $(y_1, \gamma_1) := (y_0, H(y_0))$;

8 $\quad$ Project on $\mathcal{A}$: $(y_0, \gamma_0) := \arg\min_{(y,\gamma) \in \mathcal{A}} \|(y^{nc}, \gamma) - (y_1^{nc}, \gamma_1)\|$;

9 $\quad$ **if** $\|(y_0^{nc}, \gamma_0) - (y_1^{nc}, \gamma_1)\| < tol$, **then**

10 $\quad\quad$ Find the best y^c given (y_0^{nc}, γ_0) by solving $y_0^c = \arg\min_{(y_0^{nc}, y^c, \gamma_0) \in \mathcal{A}} f(y_0^{nc}, y^c)$;

11 $\quad\quad$ Save $(y^*, \gamma^*) := (y_0^{nc}, y_0^c, \gamma_0)$ as the current best feasible solution;

12 $\quad\quad$ Try to find a feasible solution with a better objective value:

13 $\quad\quad$ Update $z^U := \nu_i f(y^*) + (1 - \nu_i) z^L$ $\quad$ (to decrease the upper bound) ;

14 $\quad\quad$ Update $\mathcal{A} := \mathcal{A} \cap \{(y, \gamma) : f(y) \leq z^U\}$ $\quad$ (to work with lower objective values) ;

15 $\quad\quad$ Adjust $(y_1, \gamma_1) := \frac{1}{tol}(y_0, \gamma_0)$ $\quad$ (to proceed with the while loop)

16 **if** *No feasible solution is obtained,* **then**

17 $\quad$ Problem (13) could be infeasible, return "Inconclusive, out of iterations"

18 **else**

19 $\quad$ Return "The best obtained solution is (y^*, γ^*)"

THEOREM 2. *Let $\mathcal{A} \cap \mathcal{B} \neq \emptyset$. Then the following holds:*

(a) *Algorithm 3 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 3 starts at a point (y_0, γ_0) that is sufficiently close to $\mathcal{A} \cap \mathcal{B}$, then it stops in a point (y^*, γ^*) that is tol-close to $\mathcal{A} \cap \mathcal{B}$ for N large enough.*

(c) *If $\mathcal{A}$ has a non-empty interior and (y_0, γ_0) is sufficiently close to $\mathcal{A} \cap \mathcal{B}$, Algorithm 3 converges to a point in $\mathcal{A} \cap \mathcal{B}$ linearly. We say that a sequence $(a_k)_{k=1}^{\infty}$ converges linearly to a if $\lim_{k \to \infty} \frac{\|a_{k+1} - a\|}{\|a_k - a\|} < 1$.*

(d) *For any $\delta > 0$ there exists N large enough and tol, $(\nu_i)_{i=1}^N$ small enough such that the solution obtained by Algorithm (3) is δ-close to a local minimum.*

Proof. Item (a) follows by construction of the algorithm since we limit the number of iterations. Item (b) follows from Assumption 1 (a) and 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 $S_\nu := \mathcal{A} \cap \mathcal{B} \cap \{(y, \gamma) : f(y) \leq \nu f(y^*) + (1 - \nu)z^L\}$. Let (y^i, γ^i) be the best solution of the algorithm at iteration i, and assume that it is not the closest local optimum denoted by $(\hat{y}, \hat{\gamma})$. Then $f(y^i) < f(\hat{y})$ and there exist $\nu_i > 0$ such that $S_{\nu_i} \neq \emptyset$ and (y^i, γ^i) is close enough to S_{ν_i} to satisfy the conditions of item (b). Notice that the latter argument might fail if we require $(\hat{y}, \hat{\gamma})$ to be a global optimum. Now, using ν_i, the algorithm finds the next point (y^{k+1}, γ^{k+1}) that is *tol*-close to S_{ν_i} and such that $f(y^i) < f(y^{i+1}) \leq f(\hat{y})$. Since the interval $(f(y^i), f(\hat{y}))$ is bounded, for any $\delta > 0$ there are N large enough, *tol* small enough and a sequence $(\nu_i)_{i=1}^N$ such that $f(y^N)$ is δ-close to $f(\hat{y})$. $\square$

REMARK 3. In Algorithm 3, 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.

In the next subsection, we demonstrate the performance of the combination of Algorithm 2 with Algorithm 3 to solve problem (6) on the ACOPF. The experiments show that in the majority of cases Algorithm 3 finds a feasible solution with an objective value close to the lower bound in few iterations, that is, for small N.

5. 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.

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, 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. 2021) 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.

5.1. Robust ACOPF formulation

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 $N = \{1, \ldots, n\}$ and the set of lines connecting these buses E. We denote the set of buses with generators by G and the active and reactive power demand (load) at each bus $k \in N$ by P_k^d and Q_k^d, respectively and 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 fluctuation 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 16.

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

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 represents the uncertainty. 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

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

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 active 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 the uncertainty:

$$P_k^g - \alpha_k \left(\sum_{i=1}^n \zeta_i - \delta \right), \ \alpha_i \geq 0, \ \sum_{i=1}^n \alpha_i = 1, \tag{17}$$

Thus, when introducing uncertainty to problem (16), we replace P_k^g in this problem with (17). 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 by 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 Σ 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.

Now we define the problem in the same form as (2) to more easily use the results from the earlier sections. This yields the following:

$$f(t, P^g, V^g) = \sum_{k \in G \setminus \{s\}} c_k^2 (P_k^g)^2 + c_k^1 P_k^g + c_k^0 + c_s^2 t^2 + c_s^1 t + c_s^0, \tag{18}$$

$$S_y = \big\{ (P^g, V^g) : P_k^{\min} \leq P_k^g \leq P_k^{\max} \ \text{ for all } k \in G \setminus \{s\},$$

$$(V_k^{\min})^2 \leq V_k^g \leq (V_k^{\max})^2 \ \text{ for all } k \in G, \ P_s^{\min} \leq t \leq P_s^{\max} \big\}, \tag{19}$$

$$L(P^g, V^g, \zeta, x) = \begin{bmatrix} P_k^d + \mathrm{tr}\,(Y_k x x^\top) - \zeta_k - P_k^g + \alpha_k \sum_{i=1}^n \zeta_i \text{ for all } k \in G \setminus \{s\} \\ P_k^d + \mathrm{tr}\,(Y_k x x^\top) - \zeta_k + \alpha_k \sum_{i=1}^n \zeta_i \quad \text{ for all } k \in N \setminus G \\ Q_k^d + \mathrm{tr}\,(\bar{Y}_k x x^\top) - \gamma_k \zeta_k \quad \text{ for all } k \in N \setminus G \\ x_k^2 + x_{k+n}^2 - V_k^g \quad \text{ for all } k \in G \\ x(s) \end{bmatrix}, \tag{20}$$

$$G(t, P^g, V^g, \zeta, x) = \begin{bmatrix} -P_s^{\min} + P_s^d + \operatorname{tr}(Y_s x x^\top) - \zeta_s + \alpha_k \sum_{i=1}^n \zeta_i \\ t - P_s^d - \operatorname{tr}(Y_s x x^\top) + \zeta_s - \alpha_k \sum_{i=1}^n \zeta_i \\ -Q_k^{\min} + Q_k^d + \operatorname{tr}(\bar{Y}_k x x^\top) - \gamma_k \zeta_k \quad \text{for all } k \in G \\ Q_k^{\max} - Q_k^d - \operatorname{tr}(\bar{Y}_k x x^\top) + \gamma_k \zeta_k \quad \text{for all } k \in G \\ \operatorname{tr}(Y_{lm} x x^\top) \le S_{lm}^{\max} \quad \text{for all } \{lm\} \in E \\ (V_k^{\max})^2 - x_k^2 + x_{k+n}^2 \quad \text{for all } k \in N \setminus G \\ x_k^2 + x_{k+n}^2 - (V_k^{\min})^2 \quad \text{for all } k \in N \setminus G \end{bmatrix}, \qquad (21)$$

$$S_x = \left\{ x \in \mathbb{R}^{2n} : (V_k^{\min})^2 \le x_k^2 + x_{k+n}^2 \le (V_k^{\max})^2 \text{ for all } k \in N \setminus G \right\} \qquad (22)$$

In the next subsection, we run numerical experiments solving problem (2) with the inputs defined above and the instances from MATPOWER.

5.2. Numerical results

In this section, we implement the dynamic Algorithm 2, where we reformulate problem (6) as problem (13) (see Proposition 3). 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 2 we use the Euclidean norm, set $tol = 1e^{-5}$, the update rule $\varepsilon_j = \sqrt{\|x_{j-1}\|/10}$ for $|N| < 30$ and $\varepsilon_j = \sqrt{\|x_{j-1}\|/30}$ for $|N| \ge 30$ 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 (16) 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 2 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 (13), we use Algorithm 3 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 3, we use the classical SDP relaxation of the quadratic problem (13) which amounts to solving the Lagrangian dual of this problem. 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, which frequently provides a feasible (see the cases with "Num iter"$= 1$ in the

tables with results) or close to feasible solution to problem (13). 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 A in (8) are negative semi-definite. We do not have to project on the corresponding equality constraints (13b). Instead, we replace the equality sign by "$\leq$" and add the resulting constraints to the definition of $\mathcal{A}$ in (14).

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 $\Sigma = \mathrm{Diag}\,\sigma$, $\sigma_k = \frac{1}{(wP_k^d)^2}$, 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 is not feasible. This implies that problem (6) is infeasible by Proposition 3. Hence, the linear approximation over the given small subset of S_x is infeasible.
- NC – Algorithm 3 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 ran out of the 2000 sec time limit we used in implementation.
- 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, and Table 3 shows the results for larger power networks. For all instances we run one iteration of Algorithm 2 as described above. 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

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.

Uncertainty, % of load	Lower bound	Upper bound	Num of altern proj iterations	Average time per iteration, sec	Full time, sec
LMBM3, 3 buses					
1%	56.95	57.01	1	1.4	2.0
5%	56.95	57.04	1	0.7	1.9
10%	56.95	57.08	1	0.6	0.9
20%	56.95	57.16	1	0.5	1.1
30%	56.95	57.26	1	0.7	1.1
40%	56.95	57.36	1	0.5	0.9
50%	56.95	57.49	1	0.6	1.0
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
5%	31.26	31.79	1	0.7	1.2
≥10%	31.26	LNF	–	–	0.7
case9, 9 buses					
1%	52.97	53.13	1	0.9	1.6
5%	52.97	53.16	1	0.8	1.3
10%	52.97	53.18	1	0.7	1.3
20%	52.97	53.24	1	0.7	1.4
30%	52.97	53.31	1	0.7	1.3
40%	52.97	53.39	1	0.8	1.4
50%	52.97	53.47	1	0.7	1.4
case14, 14 buses					
1%	80.82	81.20	7	0.5	4.4
5%	80.82	81.31	8	0.5	4.8
10%	80.82	81.46	8	0.5	4.8
20%	80.82	81.77	5	0.5	3.3
30%	80.82	82.14	7	0.6	4.8
40%	80.82	82.53	14	0.5	8.0
50%	80.82	LNF	–	–	0.7

"Lower bound" is used for the SDP lower bound of the nominal problem, which automatically provides a lower bound for the robust problem as well as our problem is minimization; "Num of altern proj iterations" denotes the number of iterations of the alternating projections method in Algorithm 3 until convergence. The fifth and sixth columns denote the running time per iteration of the alternating projections method 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 verifying the nominal feasibility of the equalities and computing the lower bound to obtain the initial solution for the alternating projections algorithm.

The numerical results show that our approximation approach provides a potentially robust solution in reasonable time even for the instance with 118 buses. For all but one of the cases, we obtained a potentially robust feasible solution after running one iteration of Algorithm 2. 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 2 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

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.

Uncertainty, % of load	Lower bound	Upper bound	Num of altern proj iterations	Average time per iteration, sec	Full time, sec
LMBM3, 3 buses					
1%	56.95	57.01	1	1.5	3.6
5%	56.95	57.05	1	1.3	2.3
10%	56.95	57.09	1	1.3	2.1
20%	56.95	57.19	1	1.2	2.1
30%	56.95	57.30	1	1.3	2.2
40%	56.95	57.41	1	1.2	2.0
50%	56.95	57.73	1	1.2	2.0
WB5, 5 buses					
1-5%	10.70	NC	101	71.7	3.3
$\geq$10%	10.70	LNF	101	0.9	1.6
case6ww, 6 buses					
1%	31.26	31.59	1	1.7	3.4
5%	31.26	31.80	1	1.8	3.3
$\geq$10%	31.26	LNF	–	–	1.5
case9, 9 buses					
1%	52.97	53.14	1	1.8	3.3
5%	52.97	53.17	1	1.7	3.3
10%	52.97	53.20	1	1.8	3.3
20%	52.97	53.27	1	1.8	3.3
30%	52.97	53.34	1	1.8	3.4
40%	52.97	53.43	1	1.8	3.3
50%	52.97	53.51	1	1.8	3.5
case14, 14 buses					
1%	80.82	NP	–	–	1.7
5%	80.82	81.31	8	1.4	12.5
10%	80.82	81.45	8	1.4	12.3
20%	80.82	81.75	5	1.4	8.6
30%	80.82	82.10	7	1.4	11.5
40%	80.82	82.49	14	1.4	20.9
50%	80.82	LNF	–	–	1.7

150 iterations. 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 2. 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 3 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 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 3 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 results with correlations for larger cases.

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.

Uncertainty, % of load	Lower bound	Upper bound	Num of altern proj iterations	Average time per iteration, 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

6. Conclusions and directions for future research

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

One of the main open questions of this work is approximation guarantees for our approach since we use Taylor approximations, which can be precise but do not have to be conservative. We address feasibility guarantees for equalities (2a) by taking piecewise affine approximations over small sets so that the approximations reflect well the real system of equalities. Thus if our piecewise affine approximation has a solution, then the original system (2a) is likely to have it as well. However, we cannot say much about guarantees for original *in*equality constraints (2b). Constructing high-precision approximations with other desirable properties, e.g., conservativeness with respect to the inequality constraints, is a question for future research.

To provide feasibility guarantees, one can ex-post verify feasibility of the obtained solution for the original problem (2). Alternatively, one could think about the criteria under which our approach does not require posterior feasibility checks. The latter way could be more promising since we did verify feasibility of the inequalities (2b) and (2c), for a given $y \in S_y$ in the numerical experiments and obtained demotivating results. We implemented the approach in (Aßmann et al. 2018), where PO is used for ARO in gas networks. While gas networks properties allowed for some simplifications, this was not the case for ACPOF. The approach turned out to be computationally intense and

inconclusive even for small ACOPF instances, therefore we do not present its results in the paper and remark that alternatives to posterior feasibility check could be of interest.

Additionally, in Section 3, 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 (6). Another option is to combine our approach with the one proposed in (Postek and Hertog 2016) to split larger uncertainty sets.

It would be also important to explore the true scalability of our method for ACOPF problems by implementing the algorithms more efficiently, investigating possibilities to exploit the inherent sparsity structure of ACOPF problems, and implementing the LP-based modification from section 4.1.2. Finally, it would be interesting to look at the extensions of our method for multiperiod problems.

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.

Baumrucker B, Biegler LT (2010) Mpec strategies for cost optimization of pipeline operations. *Computers & chemical engineering* 34(6):900–913.

Ben-Tal A, Ghaoui LE, Nemirovski A (2009) *Robust Optimization* (Princeton Series in Applied Mathematics. Princeton University Press).

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.

Bingane C, Anjos MF, Le Digabel S (2019) Tight-and-cheap conic relaxation for the optimal reactive power dispatch problem. *IEEE transactions on power systems* 34(6):4684–4693.

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.

Chen W, Sim M, Sun J, Teo CP (2010) From CVaR to uncertainty set: Implications in joint chance-constrained optimization. *Operations research* 58(2):470–485.

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.

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, Claeys M, Mevissen M, Eck BJ (2017) Polynomial optimization for water networks: Global solutions for the valve setting problem. *European Journal of Operational Research* 261(2):450 – 459.

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

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.

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.

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 (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 L (2021) Robust AC Optimal Power Flow with Convex Restriction. *IEEE Transactions on Power Systems* 36(6):4953–4966.

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 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.

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 Transactions on Power Systems* 32(6):4752–4763.

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.

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.

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

Raghunathan AU, Biegler LT (2003) Mathematical programs with equilibrium constraints (mpecs) in process engineering. *Computers & chemical engineering* 27(10):1381–1392.

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

Sherali HD, Tuncbilek CH (1992) A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique. *Journal of Global Optimization* 2(1):101–112.

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

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 $N-1$ security assessment of systems with high wind power penetration. *IEEE Transactions on Power Systems* 28(4):3885–3896.

Weisser T, Roald L, Misra S (2018) Chance-constrained optimization for non-linear network flow problems. *arXiv preprint 1803.02696* .

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.

Electronic Companion

EC.1. General semialgebraic problems and uncertainty sets

The assumptions of this paper require the initial problem (2) to be linear or quadratic in ζ and impose restrictions on the uncertainty set Ω. The assumptions are quite general but still do not cover some realistic cases, e.g., when the initial problem is quadratic in ζ, and Ω is a box uncertainty (defined by polyhedral constraints). In this section we show how to eliminate the uncertainty from such problems obtaining a strengthening of constraint (7) instead of reformulation. That is, for more general problems we obtain a more conservative constraint after eliminating the uncertainty.

We drop Assumtions 1 (b),(c) and 2 (c) and allow Ω to be a general semialgebraic set:

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

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. We use the following intuitive result.

LEMMA EC.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* (7) *holds if and only if there exists* $r \geq 0$ *such that*

$$h(y, \zeta) \in \mathcal{K}^r(\Omega). \tag{EC.1}$$

The main idea we use comes from polynomial optimization. We replace $\mathcal{P}(\Omega)$ with $\mathcal{K}^r(\Omega)$ and obtain an easier problem. A natural question is which sets $\mathcal{K}^r(\Omega)$ result in such easier problems. The answer is disappointing in the sense that the sets with the property $\bigcup_{r=0}^{\infty} \mathcal{K}^r(\Omega) \supseteq \mathcal{P}(\Omega)$ are rare, see, e.g., Kuryatnikova et al. (2019) for more details. There exist, however, many known and amenable for optimization sets $\mathcal{K}^r(\Omega)$ such that

$$\bigcup_{r=0}^{\infty} \mathcal{K}^r(\Omega) \supseteq \mathcal{P}^+(\Omega). \tag{EC.2}$$

Lemma EC.1 implies that replacing $\mathcal{P}(\Omega)$ with $\mathcal{K}^r(\Omega)$ of the type (EC.2) results in a *strengthening* of the initial constraint (7). The resulting solution y is robustly feasible but may be conservative if, for the optimal y^*, it turns out 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 smaller or the inequality constraints

slightly looser and checking whether the optimal y changes when using (EC.2). If a robustly feasible solution y^* exists that is excluded by using (EC.2), any small perturbation of the uncertainty set can make such solution infeasible. Therefore we do not lose much by excluding such solutions from consideration and using (EC.2).

Further we mention the most common set for which (EC.2) holds under non-restrictive conditions. This set was suggested by (Putinar 1993), and this set is the basis of the well-known Lasserre's hierarchy proposed in (Lasserre 2001). 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*.

PROPOSITION EC.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*

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

$$= \sum_{j=0}^{m_\zeta} m_r(\zeta)^\top S_k m_r(\zeta), \ S_k \succeq 0 \text{ for all } k \in \{0, \ldots, m_\zeta\}, \tag{EC.3}$$

$m_r(\zeta)$ 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 (EC.3) to (7), we obtain

$$h(y, \zeta) = \sum_{j=0}^{m_\zeta} m_r(\zeta)^\top S_k \, m_r(\zeta), \ S_k \succeq 0 \text{ for all } k \in \{0, \ldots, m_\zeta\}, \tag{EC.4}$$

where $m_r(\zeta)$ is the vector of all monomials in ζ of degree up to r. The equality in (EC.4) is equivalent to saying that the coefficients in ζ of h and the expression on the right-hand side are equal to each other. Writing down the coefficient-wise equalities, we can eliminate ζ and obtain the following equivalent expression.

$$H(y) = \sum_{k=0}^{m_\zeta} \langle A_k, S_k \rangle \text{ for all } i \in \left[\binom{2r}{n_\zeta} \right],$$
$$S_k \succeq 0 \text{ for all } k \in \{0, \ldots, m_\zeta\}, \tag{EC.5a}$$

where the polynomial mapping $H : \mathbb{R}^{n_y} \to \left[\binom{2r}{n_\zeta}\right]$ is obtained from the equality in (EC.4).

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. Therefore, recently some alternatives to SOS have appeared, such as (Dickinson and Povh 2018, Dressler et al. 2017, Chandrasekaran and Shah 2016, Kuryatnikova et al. 2019). We do not present details for these alternatives since they use the same reformulation strategy, only the cone in the constraint (EC.5a) changes, SDP constraint is then replaced by other conic constraints.

Acknowledgment