TWO-STAGE ROBUST QUADRATIC OPTIMIZATION WITH EQUALITIES AND ITS APPLICATION TO OPTIMAL POWER FLOW

OLGA KURYATNIKOVA*, BISSAN GHADDAR†, AND DANIEL K. MOLZAHN‡

Abstract. In this work, we consider two-stage quadratic optimization problems under ellipsoidal 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 the 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 as well as gas and water networks. We propose a convergent iterative algorithm to build a sequence of approximately robustly feasible solutions with an improving objective value. 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 and demonstrate its performance on MATPOWER instances of AC Optimal Power Flow. This paper focuses on quadratic problems, but the approach is suitable for more general setups.

Key words. non-convex quadratic optimization, two-stage robust optimization, AC optimal power flow, uncertainty in energy systems.

1. Introduction. In many optimization problems, data is not completely known in advance. One of the main approaches to deal with this lack of information is robust optimization (RO). It assumes that the data lies in a predefined set of scenarios and the constraints have to be satisfied for any realization of the data in that set. RO does not require any knowledge about the distribution of the uncertain data and is useful when the distribution is hard to estimate and feasibility for a certain set of parameters is important. In reality, some optimization variables represent decisions that must be made before the actual realization of the uncertain data while other variables can be adjusted after the uncertain data becomes known. To account for such situations in RO, *two-stage adjustable* robust optimization (ARO) was introduced in [5]. Two-stage ARO problem includes two types of variables: the first-stage variables that are fixed and the second-stage variables that may change depending on the uncertainty realization. We refer to these variables as control and state variables, respectively, since this setting is typical for optimal control problems. ARO gives rise to more flexible decisions than robust optimization and thus could be less conservative.

Several approaches have been developed to solve ARO problems, most of them are approximations since solving ARO to optimality can be NP-hard even if the original problem without uncertainty is a linear program [5]. The hardness comes from defining the relations between the first- and second-stage decision variables, called *decision rules*, which are usually not specified and must be optimized. In certain problems occurring in practice, the exact functional form of the second-stage rules is predetermined but unknown since it is defined *implicitly* via, for instance, a system of non-linear equalities that are challenging, or even impossible, to eliminate. Implicitly defined decision rules are rarely considered in the literature while they occur in many applications, some of which are described below. This paper aims to partly close this gap and studies general quadratic ARO with decision rules implicitly defined by a system of quadratic equalities.

With rapidly increasing uncertainties in both the demand and supply of resources, networked infrastructure problems are important applications of ARO with implicitly defined decision rules. Examples of relevant networked infrastructures include electricity (see [8]), natural gas (see [1, 30]),

*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).

and water (see [42]). Operators must ensure that these systems remain in acceptable states despite uncertainties while also considering performance criteria such as operating costs. ARO provides a way to balance these potentially competing concerns, as discussed in [48] and [30]. As an illustrative application of ARO, we have chosen the so-called *robust AC optimal power flow* (ACOPF) problem that provides minimum cost operating points for electric power systems. ACOPF can be formulated as a non-convex quadratic optimization problem, as discussed in [19]. Even in the absence of uncertainties, ACOPF is NP-Hard (see [10]), and solving such problems under uncertainty for instances of realistic size is still a challenge, which we explain in detail in Section 5.

1.1. Contribution. This paper presents two main contributions. First, we propose a convergent iterative solution approach for two-stage non-convex quadratic problems under uncertainty. The resulting solutions are feasible for the underlying problem without uncertainty (in other words, the *nominal problem*) and approximately feasible for the ARO version of the problem. Second, we implement the proposed approach for ACOPF, and it outperforms the benchmark methods.

Solution framework for quadratic ARO under ellipsoidal uncertainty with implicit decision rules defined by quadratic equality constraints. The equalities increase the difficulty of ARO. Therefore, we address equality and inequality constraints separately. We construct piecewise affine approximations of the implicit decision rules. In particular, we express the state variables as functions of the control and uncertainty variables using the first-order Taylor approximations of the original implicit rules on small subsets. To our knowledge, we present the first algorithm for ARO which uses implicit second-stage rules.
For each piece of the piecewise affine approximation, we eliminate the equalities and second-stage variables and obtain a standard non-linear quadratic problem in the *first-stage variables* under ellipsoidal uncertainty. This problem is reformulated into a semidefinite program (SDP) with (possibly) quadratic constraints. We suggest using an alternating projections algorithm to find locally optimal solutions for this SDP in the presence of quadratic constraints. Such a first-stage solution is feasible for the nominal problem, and possible constraint violations under uncertainty are limited by the parameters of our method.

Implementation for ACOPF and comparison with benchmarks. We apply the proposed framework to ACOPF with uncertainty in power supply and demand. The ACOPF is formulated as a non-convex quadratically constrained quadratic problem, and an ellipsoidal uncertainty set is considered. As shown in Section 5, our numerical results for ACOPF demonstrate the effectiveness of our approach in comparison to other approaches from the literature on small to moderate-size instances ranging from 6 to 118 buses, with the potential to consider larger instances.

1.2. Existing ARO solution approaches. One of the most popular ways to solve general ARO is an approximation where the second-stage decision variables can be written as affine functions of the uncertain parameters. This approach is proposed by [5], and its current state-of-the art is discussed in [17]. Besides affine decision rules, one could consider piecewise constant decision rules. Such rules are constructed by partitioning the uncertainty set into subsets and implementing a fixed second-stage decision for each subset. Naturally, the two types of decision rules could be combined into piecewise affine decision rules, which is suggested in [35]. Fixing the form of decision rules could restrict flexibility, so one could miss the optimal ARO solution. As an alternative, there exist convergent relaxations which gradually add constraint violating uncertain scenarios to the problem [9, 6, 49]. These approaches work best if the uncertainty set is polyhedral and the nominal problem can be solved efficiently, which is different from the setup in our paper.

Our approach has features of both described ARO frameworks. On the one hand, we use piecewise affine decision rules. However, those cannot be arbitrary good rules, as they must reflect an actual process represented via a system of equalities to prevent second-stage infeasibilities. Moreover, the equalities are not solvable for some first-stage decisions and uncertainty values. Thus, our method constructs piecewise rules that are close to the real ones when those exist and indicates a lack of second-stage solutions otherwise. Using approximations might still lead to second-stage infeasibilities, but one can control those by increasing the number of affine pieces. This behavior resembles the approaches ensuring feasibility on a subset of the uncertainty realizations.

Finally, let us discuss the literature addressing non-linear ARO with implicitly defined decision rules. Recently, some progress has been made in methods suitable for such ARO, and we outline the differences between our approach and these methods. First, we do not consider any assumptions on convexity and concavity, which makes our work different from all results in convex and linear ARO. Another advantage is that our approach, while using SDP, is not based on SDP approximations of the original problem, such as in [1, 23, 28, 43, 46]. We allow keeping some non-linearities from the original problem as opposed to classical SDP approaches that linearize all constraints. Additionally, independently of how much non-linearity one keeps, the resulting SDP constraints are of the size of the number of uncertain parameters. On the contrary, other approaches use SDPs of at least the size of the number of first- and second-stage variables. The number of uncertain parameters is often substantially smaller than the number of variables. For example, in ACOPF the number of variables depends on the number of buses in the system, uncertainty usually occurs in loads or renewable generators, and not all buses have those. Hence, the size of SDP constraints in our approach could be substantially smaller than in the above-mentioned papers.

Our approach differs from other robust optimization methods where non-linear constraints are linearized, such as [29]. We linearize equalities only, linearize locally, and our approximations are closely related to the original constraints via Taylor series. The results in this paper are also distinct from [32] since in the latter work a robust solution is obtained by iteratively tightening the inequality constraints. In [21], the authors tackle general non-linear optimization problems but use an alternative ARO formulation and thus a distinct solution strategy. Our approach is close in spirit to the approach in [38] which uses local Taylor expansions too. However, the authors of the latter paper work with chance constraints, use full linearizations and address specifically ACOPF. Similarly, the approach in [7] is tailored to linear complementarity problems.

This paper is organized as follows. In Section 2, we present the problem formulation and motivate our solution approach. In Section 3, we describe the proposed algorithms in detail. In Section 5, we evaluate our approach on ACOPF instances. Finally, in Section 6 we present conclusions and discuss directions for future work.

2. Problem formulation and general framework.

We begin this section with 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 $\{1, \ldots, n\}$. For a vector V of length n, we denote the i^{th} entry of V by V_i. 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 $\|\cdot\|$. We say that two vectors (or elements of a vector space) a, b are ε-close to each other if $\|a - b\| \leq \varepsilon$ for some properly defined norm.

In this paper, we deal with *quadratically constrained quadratic problems (QCQP)*. Our approach generalizes to problems of higher degree as we explain later, but we focus on QCQP for simplicity. Now, let n_ζ, n_y, n_x, m_{eq}, and m_{in} be natural numbers and consider semialgebraic sets $S_y \in \mathbb{R}^{n_y}$, $S_x \in \mathbb{R}^{n_x}$ defined by quadratic constraints. Consider *quadratic* mappings $f : \mathbb{R}^{n_y} \to \mathbb{R}$, G :

$\mathbb{R}^{n_\varsigma} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{m_{in}}$, $L_x : \mathbb{R}^{n_\varsigma} \times \mathbb{R}^{n_x} \to \mathbb{R}^{m_{eq}}$ and $L_y : \mathbb{R}^{n_\varsigma} \times \mathbb{R}^{n_y} \to \mathbb{R}^{m_{eq}}$. Now, we can introduce the *two-stage nominal* problem.

PROBLEM 2.1 (Nominal two-stage quadratic optimization problem).

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

Problem (2.1) 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 [11]), water problems (the valve placement problem by [18]) or gas problems (passive gas network feasibility problem studied by [1]). Moreover, problem (2.1) 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 (2.1) if the initial problem was linear. 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 [37, 3]. More information about linear complementarity problems and robust optimization approaches for them can be found in [7].

Remark 2.2. Problem (2.1) 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 $a^2 = b^2, a \geq 0$.

An example of problem (2.1) that is the main use case in this paper is the ACOPF problem. 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 [18] 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 (2.1) under uncertainty, which results in the next formulation.

PROBLEM 2.3 (Two-stage quadratic ARO problem).

$$
\begin{aligned}
z = \inf_{y} \quad & f(y) \\
\text{s.t.} \quad & y \in S_y \\
& \textit{and for any } \zeta \in \Omega \textit{ there exists } x \textit{ such that the following holds:} \\
& L_i(y,\zeta,x) = 0 \textit{ for all } i \in [m_{eq}] && \text{(2.3a)} \\
& G_i(y,\zeta,x) \geq 0 \textit{ for all } i \in [m_{in}] && \text{(2.3b)} \\
& x \in S_x. && \text{(2.3c)}
\end{aligned}
$$

Problem (2.3) is a two-stage ARO problem with 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 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 the second stage for any uncertainty realization $\zeta \in \Omega$. Any solution to problem (2.3) is feasible for the underlying nominal problem and robust against potential uncertainty. We obtain the nominal problem by setting $\zeta = 0$ in problem (2.3). Next, we summarize main assumptions used in this paper.

ASSUMPTION 2.4 (The characteristics of the objective and the constraints).

(a) *S_x, S_y are compact sets.*

(b) *All functions in the equality constraints* (2.3a) *are continuously differentiable.*

(c) *All functions besides the equality constraints* (2.3a) *are polynomials of degree at most two.*
This assumption is for simplicity. In fact, the problem only has to be quadratic in ζ, the requirements for other variables are milder. Our approach would be still applicable if the problem were polynomial of higher degree in y; see Section 4 for more details. The inequality constraints could be general polynomials in x as well. If the inequalities have higher degree in x, we can obtain degree-two polynomials 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 only depends on the number of control and uncertainty variables, i.e., y and ζ.

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

$$
\Omega = \left\{ \zeta \in \mathbb{R}^{n_\zeta} : \zeta^\top \Sigma \zeta + \sigma^\top \zeta + r \geq 0,\ j \in [m_\zeta] \right\}, \tag{2.1}
$$

where Σ is negative semidefinite. Ω has a non-empty interior.
We need assumptions on the degree of ζ and the shape of Ω to efficiently eliminate ζ from the problem using the S-lemma [47]; see Section 4.1. We have chosen an ellipsoidal uncertainty set as the base case since it is convenient for our approach; frequently appears in the literature as being less conservative than, for instance, box uncertainty; and has interpretations from both robust and chance-constrained perspectives; see, e.g., [20, 13].

ASSUMPTION 2.5 (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 (2.3a) as a preprocessing step. If S_x becomes unbounded afterwards, a large ball constraint for x can be added to the problem.

(b) $n_x = m_{eq}$ *and there are no redundant equality constraints in* (2.3a).
The assumption is w.l.o.g. If $n_x \leq m_{eq}$ and there are redundant constraints, they could be detected and eliminated as a preprocessing step. If $n_x > m_{eq}$, then some state variables are free and could be added to the pool of control variables.

(c) 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 problem in control variables y* only, since they represent the actual decisions to implement. The following idea motivates our approach: if we could analytically solve equalities (2.3a) for the second-stage variables x, we would express x as a function of y and ζ obtaining the second-stage decision rule $p : \mathbb{R}^{n_y \times n_\zeta} \to \mathbb{R}^{n_x}$. Then problem (2.3) would be equivalent to the following problem:

PROBLEM 2.6.

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

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 unique analytical expression for the decision rule in the problem. However, such expressions exist on small subsets of $S_y \times S_\zeta$ under known conditions according to the implicit function theorem (see, e.g., [41]). In essence, the approach we suggest restricts the optimization problem to such subsets and replaces the implicit decision rule by its first-order Taylor approximation on each subset. We present the approach in detail in the next sections.

3. Piecewise affine approximations of equality constraints. Our approach looks at piecewise affine approximations of the second-stage decision rules $x = p(y, \zeta)$ in problem (2.6), which are given implicitly via the system of equalities (2.3a), using Taylor series. In addition to its simplicity, the main advantages of the Taylor approximation are its good fit for the original function around the approximation point and the possibility to construct it for an implicit function. For the construction, we need some Jacobians of the system of equalities (2.3a). Denote $L(y, \zeta, x) := \begin{bmatrix} L_1(y, \zeta, x) \\ \dots \\ L_{m_{eq}}(y, \zeta, x) \end{bmatrix}$. For an $(\hat{y}, \hat{\zeta}, \hat{x}) \in S_y \times \Omega \times S_x$, let $J_y^{\hat{y}, \hat{\zeta}, \hat{x}}$, $J_\zeta^{\hat{y}, \hat{\zeta}, \hat{x}}$ and $J_x^{\hat{y}, \hat{\zeta}, \hat{x}}$ be the Jacobians of L with respect to y, ζ, and x, respectively, evaluated at $(\hat{y}, \hat{\zeta}, \hat{x})$. By Assumption 2.4(b), the Jacobian $J_x^{\hat{y}, \hat{\zeta}, \hat{x}}$ is a square matrix of the size $n_{eq} \times n_{eq}$.

THEOREM 3.1 (The implicit function theorem: [41] Theorem 2-12). *Let $L(y, \zeta, x) : \mathbb{R}^{n_y} \times \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_{eq}}$ be a continuously differentiable function. Let $(\hat{y}, \hat{\zeta}, \hat{x}) \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_x}$ be such that $L(\hat{y}, \hat{\zeta}, \hat{x}) = 0$ and the Jacobian $J_x^{\hat{y}, \hat{\zeta}, \hat{x}}$ of L with respect to x at $(\hat{y}, \hat{\zeta}, \hat{x})$ is non-singular. Then there exists an open set $U \subseteq \mathbb{R}^{n_y} \times \mathbb{R}^{n_\zeta}$ with $(\hat{y}, \hat{\zeta}) \in U$, an open set $\hat{S}_x \subset \mathbb{R}^{n_x}$ with $\hat{x} \in \hat{S}_x$, and* *a unique function $p(y,\zeta): U \to \hat{S}_x$ such that $\hat{x} = p(\hat{y},\hat{\zeta})$ and $L(y,\zeta,p(y,\zeta)) = 0$ for all $(y,\zeta) \in U$. Moreover, p is differentiable, and its Jacobian at $(\hat{y},\hat{\zeta},\hat{x})$ equals*

$$-\left(J_x^{\hat{y},\hat{\zeta},\hat{x}}\right)^{-1}\begin{bmatrix} J_y^{\hat{y},\hat{\zeta},\hat{x}} \\ J_\zeta^{\hat{y},\hat{\zeta},\hat{x}} \end{bmatrix}. \tag{3.1}$$

From here on we call the main condition mentioned in Theorem 3.1 *the Implicit Decision Rules (IDR) condition* for brevity:

Definition 3.2 (The Implicit Decision Rules (IDR) condition). *We say that the IDR condition holds for $\hat{y} \in S_y$, $\hat{\zeta} \in \Omega$ if there exists $\hat{x} \in \mathbb{R}^{n_x}$ such that $L(\hat{y},\hat{\zeta},\hat{x}) = 0$ and the Jacobian $J_x^{\hat{y},\hat{\zeta},\hat{x}}$ is non-singular. We say that the IDR condition holds on sets $\hat{S}_y$, $\hat{\Omega}$ if it holds for each $y \in \hat{S}_y$, $\zeta \in \hat{\Omega}$.*

We assume that it is possible to check quickly if the IDR condition holds for given $\hat{y}, \hat{\zeta}$ since one can solve equalities (2.3a) rather efficiently in practice. This paper was inspired by engineering problems, such as ACOPF. For these problems, Newton's methods [50] are usually well-developed, fast, and precise enough to work for realistic problem instances. Theorem 3.1 implies that, if the IDR condition holds for some $(\hat{y},\hat{\zeta})$, a unique local second-stage decision rule $x = p(y,\zeta)$ as in (2.6) exists on some open set around $(\hat{y},\hat{\zeta})$, and the first-order Taylor approximation of this rule is

$$\hat{p}^{\hat{y},\hat{\zeta},\hat{x}}(y,\zeta) := \hat{x} - \left(J_x^{\hat{y},\hat{\zeta},\hat{x}}\right)^{-1}\left(J_y^{\hat{y},\hat{\zeta},\hat{x}}(y-\hat{y}) + J_\zeta^{\hat{y},\hat{\zeta},\hat{x}}(\zeta-\hat{\zeta})\right), \quad \text{where } \hat{x} = p(\hat{y},\hat{\zeta}). \tag{3.2}$$

The above result applies to $\hat{y} \in S_y$ for which the Jacobian with respect to x is non-singular. This condition is in general desirable in applications to dynamical systems, which are the main target of this paper. For such applications, solutions with singular Jacobians may be physically unstable, see, for example, [45].

The idea is to work on small subsets of S_y and Ω and use piecewise affine decision rules based on the above Taylor approximations within subsets. Fix some subset of S_y and $\hat{y} \in \hat{S}_y$. For $M_\zeta > 0$, partition Ω into $\Omega^1,\ldots,\Omega^{M_\zeta}$ and pick some $\hat{\zeta}^1 \in \Omega^1,\ldots,\hat{\zeta}^{M_\zeta} \in \Omega^{M_\zeta}$. If $\hat{S}_y$ is small enough and the IDR condition holds for $\hat{y}$ and $\hat{\zeta}^1,\ldots,\hat{\zeta}^{M_\zeta}$, the implicit decision rules from Theorem 3.1 exist on $\hat{S}_y$ and some open subset of Ω. Now, in problem (2.6), we impose the inequalities for each Ω^k substituting the implicit rules with the corresponding approximation (3.2). The procedure results in an almost standard ARO approximation with piecewise affine decision rules (see problem (3.3) in Algorithm 1). The differences are that (i) y is restricted to a subset, and (ii) the decision rules are known, we do not need to optimize over them. We can solve the approximation and obtain some solution y^*. As the last step, we could check that y^* is feasible for some selected values of the uncertainty that are desirable in practice, for instance, that it is nominally feasible. Then we can repeat the procedure considering other subsets of S_y and choose y^* with the best objective value.

Intuitively, if the subsets of S_y and Ω are small enough, the above approach should be able to discard solutions $\hat{y} \in S_y$ for which no decision rules from Theorem 3.1 exist and select solutions for which these rules exist. We develop this idea in the next corollary.

Corollary 3.3. *Consider an $\hat{y} \in S_y$, a finite cover $\Omega = \bigcup_{k=1}^{M_\zeta} \Omega^k$ and $\hat{\zeta}^1 \in \Omega^1,\ldots,\hat{\zeta}^{M_\zeta} \in \Omega^{M_\zeta}$.*

(a) Let the IDR condition hold for $\hat{y}$ on $\hat{\zeta}^1,\ldots,\hat{\zeta}^{M_\zeta}$, and assume that equalities (2.3a) have no solution for $\hat{y}$ on some $\hat{\Omega} \subseteq \Omega$. If each $\Omega^k, k = 1,\ldots,M_\zeta$ is contained in a ball of radius $\epsilon > 0$, then $\hat{\Omega}$ cannot contain a ball of radius larger than 2ϵ.

(b) Let the IDR condition hold for $\hat{y}$ on Ω. Then there exist a closed ball $\hat{S}_y \subset \mathbb{R}^{n_y}$ with $y \in \hat{S}_y$ and a cover $\Omega = \bigcup_{k=1}^{M_\zeta} \Omega^k$ by finitely many compact subsets such that the unique decision rules from Theorem 3.1 exist on products of $\hat{S}_y$ and each $\Omega^1, \dots, \Omega^{M_\zeta}$.

Proof. We begin by proving part (a). In each Ω^k, $k = 1, \dots, M_\zeta$ there is at least one ζ for which the IDR condition holds and therefore equalities (2.3a) have solutions. Hence $\hat{\Omega}$ does not contain fully one of Ω^k. To cover a ball of radius up to r with balls of radius ϵ in such a way that none of the smaller balls is fully inside the larger ball, we need to have $r \leq 2\epsilon$. Thus we cannot have a violating subset that contains a ball of radius larger than 2ϵ. Next we prove part (b). By Theorem 3.1, for every $\zeta \in \Omega$ there is an open ball $B \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_\zeta}$ with radius ϵ such that $(\hat{y}, \zeta) \in B$ and decision rules $x = p(\hat{y}, \zeta)$ exist on B. Now, consider an open ball B_y around y with radius $\varepsilon/2$ and an open ball B_ζ around ζ with the same radius. Notice that $B_y \times B_\zeta \subset B$. We can cover Ω by the balls B_ζ for every $\zeta \in \Omega$, and by compactness of Ω there will be a finite subcover. Taking a closure of each ball in this subcover and intersecting it with Ω, we obtain a finite compact cover $\Omega = \bigcup_{k=1}^{M_\zeta} \Omega^k$. Now, each of Ω^k corresponds to an open subset of $\mathbb{R}^{n_y}$ that contains $\hat{y}$ for which the decision rules exist by construction. Since there are finitely many such subsets, their intersection results in an open subset of $\mathbb{R}^{n_y}$ that contains $\hat{y}$ and the required closed ball $\hat{S}_y$ around it. ◻

3.1. Approximation algorithm for general problems. We formalize the ideas from the previous section in Algorithm 1, and the approximation guarantee is presented in Theorem 3.4.

Algorithm 1: Piecewise affine approximations of problem (2.3) using Taylor series

---

Input: Problem (2.3), $\hat{y}^i \in S_y^i \subseteq S_y, i = 1, \dots, M_y$ for some M_y, $\hat{\zeta}^k \in \Omega^k \subseteq \Omega$
$k = 1, \dots, M_\zeta$ for some M_ζ, where $\Omega = \bigcup_{k=1}^{M_\zeta} \Omega^k$. All sets are compact.

1 **for** $i = 1, \dots, M_y$ **do**
2 Check that the IDR condition holds for $\hat{y}^i$ and all $\hat{\zeta}^1, \dots, \hat{\zeta}^{M_\zeta}$, denote the corresponding solutions by $\hat{x}^{i,k}$, $k = 1, \dots, M_\zeta$.
3 **if** *The IDR condition fails for $\hat{y}^i$ and some of $\hat{\zeta}^1, \dots, \hat{\zeta}^{M_\zeta}$* **then**
4 set $z_i := \infty$, $, y^{*,i} := \emptyset$
5 **else**
6 Solve Problem 3.3: (Piecewise affine approximation of problem (2.3) on S_y^i)

$$z_i = \inf_y \quad f(y)$$
$$\text{s.t.} \quad y \in S_y^i$$
$$\text{(3.3a)} \qquad \hat{p}^{\hat{y}^i, \hat{\zeta}^k, \hat{x}^{i,k}}(y, \zeta) \in S_x \text{ for all } \zeta \in \Omega^k,\ k \in [M_\zeta]$$
$$\text{(3.3b)} \qquad G_m\left(y, \zeta, \hat{p}^{\hat{y}^i, \hat{\zeta}^k, \hat{x}^{i,k}}(y, \zeta)\right) \geq 0 \text{ for all } \zeta \in \Omega^k,\ m \in [m_{in}],\ k \in [M_\zeta]$$

where $\hat{p}^{\hat{y}^i, \hat{\zeta}^k, \hat{x}^{i,k}}(y, \zeta)$ is defined in (3.2). Let $y^{*,i}$ be the optimal solution to (3.3).
7 Save z_i, $y^{*,i}$ setting $z_i := \infty$, $y^{*,i} := \emptyset$ if problem (3.3) is infeasible.
8 Choose $i^* := \arg\min_{i=1}^{M_y} z_i$, $z^* := z_{i^*}$, $y^* := y^{*,i^*}$.
9 For $\hat{\zeta}^k$, $k = 1, \dots, M_\zeta$, check if y^* is feasible for problem (2.3) restricting Ω to $\hat{\zeta}^k$.
10 Return z^*, y^* if the check in step 9 is positive, return $z = \infty$, $y^* = \emptyset$ otherwise.

The next theorem shows that Algorithm 1 provides approximation guarantees and is able to find solutions robust in a "strong" sense (such that the inequality constraints hold with a margin).

THEOREM 3.4 (Approximation guarantees of Algorithm 1 for problem (2.3)). *Consider a $\delta > 0$, and let $y^* \neq \emptyset$ be a solution returned by Algorithm 1.*

Guarantees limiting infeasibility

(a) There exists $\epsilon > 0$ such that, if S^{i^*} and each subset in the partition $\Omega = \bigcup_{k=1}^{M_\zeta} \Omega^k$ fit in balls of radius ϵ, then the following conditions hold on Ω, except for possibly a subset where no ball of size larger than 2ϵ fits: (i) equality constraints (2.3a) hold, and (ii) each inequality constraint (2.3b), (2.3c) is violated by at most δ.

(b) Let y^* be not feasible for problem (2.3) on $\hat{\Omega} \subset \Omega$ whose Lebesgue measure constitutes at least $\beta < 1$ of that of Ω. If $M_\zeta \geq \frac{\log \theta}{\log(1-\beta)}$ and $\hat{\zeta}_1, \ldots, \hat{\zeta}_{M_\zeta}$ are sampled uniformly i.i.d. on Ω, the probability that y^* passes the check in step 9 of Algorithm 1 is not larger than θ.

Guarantees for feasible solutions

(c) Consider the set $S_y^\delta \subseteq S_y$ of all solutions y to problem (2.3) which (i) satisfy the IDR condition on Ω and (ii) are such that for any $\zeta \in \Omega$ constraints (2.3b), (2.3c) hold strictly with a margin δ (i.e., the inequalities "$\geq \delta$" hold). Then there is a subset of S_y and a finite cover of Ω for which Algorithm 1 will return an optimal $y^* \in S_y^\delta$.

Proof. First, consider statement (a). We know that step 2 in the algorithm did not fail for y^* and the corresponding partition $\Omega = \bigcup_{k=1}^{M_\zeta} \Omega^k$. If ε is small enough, then by Corollary 3.3 (a) there are decision rules $x = p(\hat{y}^{i^*}, \zeta)$ on Ω, except for possibly a subset where no ball of size larger than 2ϵ fits. A small enough ϵ ensures also that $S_y^{i^*}$ is such that the decision rules are valid for both $\hat{y}^{i^*}$ and y^*. Thus part (i) of statement (a) follows. Now, consider inequality constraints (2.3b) (2.3c). By assumption (2.4) (c), these constraints are continuous, and there are finitely many of them. Since we work on compact sets, Taylor approximation of each decision rule $\hat{p}^{\hat{y}^{i^*}, \hat{\zeta}^k, \hat{x}^{i^*,k}}(y^*, \zeta)$ (3.2) is close enough to $p(y^*, \zeta)$ on $S_y^{i^*}$ and Ω^k so that the approximated inequality constraints (3.3b) in problem (3.3) are δ-close to the original inequality constraints (2.3b), (2.3c) in problem (2.3). Thus the original constraints cannot be violated (within the subsets of Ω where the rules exist) by more than δ, so y^* possesses property (ii). To prove statement (b), notice that with i.i.d. uniform sampling, the probability that all constraints of problem (2.3) are satisfied for one sample ζ is at most $1 - \beta$. Hence the probability that they are satisfied for M_ζ samples is at most $(1-\beta)^{M_\zeta}$, and if $M_\zeta \geq \frac{\log \theta}{\log(1-\beta)}$, then the latter probability is not larger than θ. Finally, consider statement (c). By Corollary 3.3, there exist a compact subset of S_y that contains y^* and a compact finite cover of Ω where the implicit decision rules used by Algorithm 1 are valid. By the above argument for part (ii) of statement (a), if we make the subset of S_y and each subset in the cover of Ω small enough, the Taylor approximations are so close to the actual decision rules that the inequality constraints cannot differ from their approximations by more than δ, and thus Algorithm 1 will choose y^* if we restrict the search to subsets from S_y^δ. ∎

To our knowledge, Algorithm 1 is the first ARO algorithm which uses the second-stage policies defined by implicit functions, represented by equalities (2.3a). Theorem 3.4(b) speaks about uniform sampling from Ω, which is a numerically efficient procedure since Ω is an ellipsoid [14].

The main limitation of the algorithm in practice is the need to choose the subsets of S_y and Ω. This is not trivial since inefficient partitioning may lead to long running times and numerically unstable problems. Constructing good subsets is a topic for separate research. For instance, to split Ω, one can use the results from [35]. In this paper, we are especially interested in testing Algorithm 1 as a proof-of-concept for robust ACOPF. This problem has an additional structure: control variables y appear linearly and separately in L. In the next subsection, we introduce a version of Algorithm 1 that uses this property and allows working on larger subsets of S_y.

3.2. Approximation algorithm for problems where L is linear in the control variables. In this subsection, we assume that the control variables y appear linearly and separately in the equality constraints $L(y,\zeta,x)=0$. That is, we can write

$$L(y,\zeta,x) = L_1(x,\zeta) + J_y y, \tag{3.4}$$

where L_1 is a polynomial function and J_y is a constant matrix. As a result, the Jacobian of $L(y,\zeta,x)$ with respect to y is J_y, and other Jacobians do not depend on y. Denote the Jacobians with respect to ζ and x for any y and some $(\hat\zeta,\hat x)$ by $J_\zeta^{\hat\zeta,\hat x}$ and $J_x^{\hat\zeta,\hat x}$, respectively. Then the first-order Taylor approximation (3.2) in $(\hat y,\hat\zeta,\hat x)$ where $L(\hat y,\hat\zeta,\hat x)=0$ simplifies to

$$\begin{aligned}\hat p^{\hat\zeta,\hat x}(y,\zeta) &:= \hat x - \left(J_x^{\hat\zeta,\hat x}\right)^{-1}\left(J_y(y-\hat y) + J_\zeta^{\hat\zeta,\hat x}(\zeta-\hat\zeta)\right)\\ &= \hat x - \left(J_x^{\hat\zeta,\hat x}\right)^{-1}\left(J_y y + J_\zeta^{\hat\zeta,\hat x}(\zeta-\hat\zeta) + L_1(\hat\zeta,\hat x)\right).\end{aligned} \tag{3.5}$$

Observe that the Jacobian $J_x^{\hat\zeta,\hat x}$ and the above Taylor approximation are independent of $\hat y$. That is, the same approximation could be valid for various subsets of S_y as long as the decision rules for these subsets and some $\Omega^k\subseteq\Omega$ lead to the same subset $\hat S_x\subset\mathbb{R}^{n_x}$. Hence it could be beneficial to use subsets of S_x instead of subsets of S_y, reconsidering Algorithm 1 as presented in Algorithm 2.

Algorithm 2: Piecewise affine approximations of problem (2.3) where L is linear in y

---

Input: Problem (2.3), $\gamma>0$, $\hat x^j\in S_x, j=1,\dots,M_x$ for some M_x, $\hat\zeta^k\in\Omega^k\subseteq\Omega$ $k=1,\dots,M_\zeta$ for some M_ζ, where $\Omega = \bigcup_{k=1}^{M_\zeta}\Omega^k$. All sets are compact.

1 Solve Problem 3.6 (Approximation of problem (2.3) where L is linear in y)

$$\begin{aligned} z^* = \inf_y\quad & f(y)\\ \text{s.t.}\quad & y\in S_y\\ & \textbf{For } k=1,\dots,M_\zeta:\\ & \text{At least one of conditions (j) for } j=1,\dots,M_x \text{ holds:}\end{aligned}$$

$$(j):\begin{cases}\text{The IDR condition holds for some } \hat y\in S_y \text{ in } \hat\zeta^k \text{ with solution } \hat x^j,\\ \|\hat p^{\hat\zeta^k,\hat x^j}(y,\hat\zeta^k)-\hat x^j)\|\le\gamma \quad (**)\\ \hat p^{\hat\zeta^k,\hat x^j}(y,\zeta)\in S_x \text{ for all } \zeta\in\Omega^k\\ G_m\left(y,\zeta,\hat p^{\hat\zeta^k,\hat x^j}(y,\zeta)\right)\ge 0 \text{ for all } m\in[m_{in}],\zeta\in\Omega^k,\end{cases} \tag{3.6a}$$

where $\hat p^{\hat\zeta^k,\hat x^j}(y,\zeta)$ is defined in (3.5). Let y^* be the optimal solution to problem (3.6).

2 For $\hat\zeta^k$, $k=1,\dots,M_\zeta$, check if y^* is feasible for problem (2.3) restricting Ω to $\hat\zeta^k$.

3 Return z^*, y^* if the check in step 2 is positive, return $z=\infty$, $y^*=\emptyset$ otherwise.

The next Corollary 3.5 shows that the performance of Algorithm 2 is similar to the one of Algorithm 1.

Corollary 3.5 (Approximation guarantees of Algorithm 2). *Let $y^*\neq\emptyset$ be a solution returned by Algorithm 2, and let γ be the input into Algorithm 2 used in (3.6a) (**).*

(a) The result of Theorem 3.4 (b) holds for y^*, and there is $\varepsilon>0$ such that if $\gamma\le\epsilon$ and $\Omega^1,\dots,\Omega^{M_\zeta}$ fit in balls of radius at most ϵ, the result of Theorem 3.4 (a) holds for y^*.

(b) *Define* $\tau = \max_{\zeta_1,\zeta_2\in\Omega} \|\zeta_1 - \zeta_2\|$. *If* τ *and* γ *are small enough and* $M_\zeta = 1$ *(no need to partition* Ω*), then conditions (i) and (ii) of Theorem 3.4 (a) are satisfied on* Ω.

Proof. We begin by proving statement (a). The result of Theorem 3.4 (b) holds by step 2 of Algorithm 2, which coincides with step 9 of Algorithm 1. Now, move to the result of Theorem 3.4 (a). Denote the set of all solutions $y \in S_y$ to (2.3a) for some $\hat{\zeta}^k,\ k = 1,\dots,M_\zeta$ and $\hat{x}^j, j = 1,\dots,M_x$ by $S_y^{j,k}$. In each condition (j) in problem (3.6), we verify that the IDR condition holds for some $y \in S_y^{j,k}$ and $\hat{\zeta}^k$ resulting in the second-stage solution $\hat{x}^j$. Thus the IDR condition holds for all $y \in S_y^{j,k}$ in $\hat{\zeta}^k$ since the Jacobian $J_x^{\hat{\zeta}^k,\hat{x}^j}$ does not depend on y. If the verification is successful, a second-stage rule exists for an open ball around any $y \in S_y^{j,k}$ and $\hat{\zeta}^k$. Since $S_y^{j,k}$ is compact, it has a finite cover by those balls. Consider the union of the balls in this cover, which is an open set containing $S_y^{j,k}$, and look at the complement of this union. The distance between this complement and $S_y^{j,k}$ is positive since they are disjoint, $S_y^{j,k}$ is compact and the complement is closed. Hence, there is some $\epsilon > 0$ such that if $\gamma \le \epsilon$, then the restriction on the Taylor approximation in (3.6a) (**) pushes y^* to be in the earlier constructed open cover of $S_y^{j,k}$, so there exist second-stage rules for y^* around $\hat{\zeta}^k$. Moreover, we can adjust γ and and the size of Ω^k to make sure that the Taylor approximations of the inequality constraints are δ-close to the original constraints, for any $\delta > 0$. These are the conditions we used to prove Theorem 3.4 (a), thus the corresponding result holds for y^*. Finally, statement (b) follows from the proof of statement (a) under the following consideration. If $M_\zeta = 1$, then statement (a) applies to the whole $\Omega = \Omega^1$. Thus, if τ and γ are small enough and we obtain $y^* \neq \emptyset$, the conditions (i) and (ii) of Theorem 3.4 (a) are satisfied on the whole Ω. ∎

Algorithm 2 could be more efficient than Algorithm 1 since it could capture more values of $y \in S_y$, even if it looks around only one solution $\hat{x}$. However, we cannot guarantee that a strictly feasible solution y^* as defined in Theorem 3.4 (c) will be found under some conditions as we cannot control y as precisely as in Algorithm 1. Also, to say that *at least* one of the conditions (j) in (3.6a) holds for each Ω^k, one may have to introduce additional binary control variables.

By Corollary 3.5 (b), if Ω is small enough for the chosen approximation error, it suffices to find one good small subset of S_x, and there is no need to partition Ω. The nominal problem (2.1) for which we want to find a robust solution is usually feasible and, moreover, its optimal solution is usually "stable" in the sense that the Jacobian in (3.6a) is non-singular. Thus, for small uncertainty sets, one could start from some nominal feasible control solution $\hat{y}^1$ at hand (e.g., the original optimal solution), find the corresponding state variables solution $\hat{x}^1$ such that $L(\hat{y}^1, 0, \hat{x}^1) = 0$ and restrict Algorithm 2 to work around $\hat{x}^1$ as in (**) in (3.6a). If the algorithm finds a solution $\hat{y}^2 \neq \hat{y}^1$, one could set $\hat{x}^2$ such that $L(\hat{y}^2, 0, \hat{x}^2) = 0$ and repeat the procedure trying to find a better solution. Notice that a similar procedure would be valid for Algorithm 1 too, by Theorem 3.4 (a), but we would construct subsets around y^1, y^2 in S_y instead of subsets around x^1, x^2 in S_x.

Now, we have constructed two algorithms, one is suitable for general problems (2.3) (Algorithm 1), and one is tailored to problems (2.3) linear in control variables (Algorithm 2). To solve the optimization problems (3.3) and (3.6) in the algorithms, we need to deal with quadratic inequality constraints under ellipsoidal uncertainty, which is the topic of the next section.

4. Quadratic inequality constraints under uncertainty.

In this section, we solve problem (3.3) in Algorithm 1 or problem (3.6) in Algorithm 2, depending on which algorithm is used. Our first step is to eliminate ζ from (3.3b), (3.3a) and (3.6a). The constraints in question have the following general form:

$$h(y,\zeta) \geq 0 \text{ for all } \zeta \in \Omega. \tag{4.1}$$

For instance, we can have $h(y,\zeta) := G_i\left(y,\zeta,\hat{p}^{\hat{y},\hat{\zeta},\hat{x}}(y,\zeta)\right)$ for some $i \in [m_{in}]$. Since (4.1) 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 (4.1) that contains no ζ. In this way, we eliminate the uncertainty ζ from the problems in question. After that, we stop treating y as a parameter, and y becomes the only variable left in the reformulated problems (3.3) and (3.6), which we solve in subsection 4.2.

The rest of this section is needed if the initial problem (2.3) contains *inequalities* that are non-linear in y or ζ. If it turns out that problem (3.3) in Algorithm 1 or problem (3.6) in Algorithm 2 is linear in y and ζ, then classical robust optimization techniques apply to the corresponding problem. In this case one can use a standard reformulation of a linear constraint under the ellipsoidal uncertainty (see,e.g., [5]) and rewrite the problem as a second-order cone program. Finally, from here on, we focus on problem (3.6) for brevity, but the method is valid for problem (3.3) as well.

4.1. Eliminating the uncertainty from the problem. Under Assumption 2.4, all inequalities in problem (2.3) are of degree at most two in x and ζ, hence so is h in (4.1):

$$h(y,\zeta) := \zeta^\top A\zeta + (y^\top B + b^\top)\zeta + c^\top y + d + g(y) \geq 0, \tag{4.2}$$

for some given parameters $A,\ B,\ b,\ c,\ d$, and a function g which contains all monomials that are non-linear in y. The precise values of the parameters and the form of g depend on the functions in (3.6a) and can be obtained directly from those functions. Recall that Ω has the form (2.1). Such a combination of h and Ω allows reformulating (4.1) and eliminating the uncertainty from it due to the *S-lemma*.

Proposition 4.1 ([47]). *Let $h, g \in \mathbb{R}_2[\zeta]$ and suppose there is $\zeta \in \mathbb{R}^{n_\zeta}$ such that $g(\zeta) < 0$. Then the following two statements are equivalent:*

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

S-lemma is well known in robust optimization (see [4]) but is usually applied to convex problems in y,ζ while we use it for a general quadratic constraint (4.1).

Proposition 4.2. *Constraint (4.1) with h as in (4.2) and the uncertainty set (2.1) holds if and only if there exist λ, γ such that*

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

where Σ, σ and r are the parameters from (2.1), and other parameters are defined in (4.2).

Proof. We have

$$
\begin{aligned}
& h(y,\zeta) \geq 0 \text{ for all } \zeta \in \Omega \\
\overset{\text{Prop. 4.1}}{\Longleftrightarrow} & \zeta^\top A\zeta + (y^\top B + b^\top)\zeta + c^\top y + d + g(y) \geq 0 \text{ for all } \zeta \in \mathbb{R}^{n_\zeta} : \zeta^\top \Sigma\zeta + \sigma^\top \zeta + r \leq 0 \\
\Longleftrightarrow & \zeta^\top A\zeta + (y^\top B + b^\top)\zeta + c^\top y + d + g(y) - \lambda(-r - \zeta^\top \Sigma\zeta - \sigma^\top\zeta) = \begin{bmatrix} 1 \\ \zeta \end{bmatrix}^\top S \begin{bmatrix} 1 \\ \zeta \end{bmatrix}, \\
& S \succeq 0, \ \lambda \geq 0 \\
\Longleftrightarrow & \begin{bmatrix} \gamma + c^\top y + d + \lambda r & \frac{1}{2}(y^\top B + b^\top + \lambda\sigma^\top) \\ \frac{1}{2}(B^\top y + b + \lambda\sigma) & \lambda\Sigma + A \end{bmatrix} \succeq 0, \ g(y) = \gamma, \ \lambda \geq 0.
\end{aligned}
$$

∎

If some inequalities in problem (2.3) are linear in x and ζ, then h in (4.1) is linear in ζ as well. In this case, the constraint considered in this section simplifies to the classical linear robust constraint. It is well known how to eliminate uncertainty from such constraints; see, e.g., [4]. Each linear constraint under the ellipsoidal uncertainty in (2.1) will become a second-order cone constraint after eliminating the uncertainty.

4.2. Alternating projections algorithm. In Section 4.1, we eliminate ζ from constraints (3.6a) in problem (3.6) and add auxiliary control variables (e.g., λ, γ in Proposition 4.2). By eliminating ζ, we reformulate problem (3.6). This reformulation can be written in a general way as

PROBLEM 4.3.

$$
\begin{aligned}
&(4.3a) && z_j = \inf_y \quad f(y) \\
&(4.3b) && \text{s. t.} \quad H(y) = \gamma \\
&(4.3c) && F \left[\begin{smallmatrix} y \\ \gamma \end{smallmatrix} \right] \in \mathcal{C},
\end{aligned}
$$

where γ are auxiliary control variables as in Proposition 4.2, 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).

Under Assumption 2.4, the equality constraints are quadratic, the objective function is convex with degree at most two, and $\mathcal{C}$ is the positive semideninite cone. The approach we use next is applicable for general polynomial equality constraints and semialgebraic cones $\mathcal{C}$, therefore we introduce the formulation above.

Problem (4.3) would be a classical conic problem for which many solvers exist if we did not have those non-linear equality constraints. Hence, we first relax those complicating equalities and obtain a solution feasible for (4.3c). Then we iteratively transform it into a solution feasible for the whole problem (4.3). Notice that problem (4.3) cannot be unbounded as the feasible set of its nominal problem is compact by Assumption 2.4. 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., SDP relaxations or the reformulation-linearization technique as in [40]. The lower-bound relaxation might provide a solution that is infeasible for (4.3b). 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 (4.3). Define two sets:

$$(4.3) \qquad \mathcal{A} := \{(y,\gamma) : (4.3c) \text{ holds}, f(y) \leq z^U\},$$

$$(4.4) \qquad \mathcal{B} := \{(y,\gamma) : (4.3b) \text{ holds}\}.$$

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 (4.3b). We use two subsets of the variables to speed up the algorithm by skipping the iterations in which only y^c changes.

THEOREM 4.4. *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 2.4 and Theorem 7.3 in [15] 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 [15], 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. 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 $f(\hat{y})$. ∎

In the next subsection, we demonstrate the performance of the combination of Algorithm 2 with Algorithm 3 to solve problem (2.3) on the ACOPF. The experiments show that in the majority of tested cases this combination of algorithms finds a robustly feasible solution within 15 minutes of computational time for small uncertainty sets.

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 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 [39], many of which use the DC power flow approximation; see, e.g., [44, 8] 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,

Algorithm 3: Alternating projections algorithm for problem (4.3)

Input: small $tol > 0$, upper bound $z^U \geq z_j$, $(\nu_i)_{i=1}^N \in (0, 1]$, $N \geq 1$
Solve a lower bound relaxation of problem (4.3) and denote its solution by (y_0, γ_0) and its objective value by z^L;
if *the lower bound relaxation is infeasible,* **then**
  Lower-bound relaxation is infeasible, stop, return "Problem (4.3) infeasible"
Set $i = 0$ and choose (y_1, γ_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 $\mathcal{B}$: $(y_1, \gamma_1) := (y_0, H(y_0))$;
  Project on $\mathcal{A}$: $(y_0, \gamma_0) := \arg\min_{(y,\gamma)\in\mathcal{A}} \|(y^{nc}, \gamma) - (y_1^{nc}, \gamma_1)\|$;
  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 $y_0^c = \arg\min_{(y_0^{nc}, y^c, \gamma_0)\in\mathcal{A}} f\,(y_0^{nc}, y^c)$;
    Save $(y^*, \gamma^*) := (y_0^{nc}, y_0^c, \gamma_0)$ as the current best feasible solution;
    Try to find a feasible solution with a better objective value:
    Update $z^U := \nu_i f(y^*) + (1 - \nu_i) z^L$ (to decrease the upper bound);
    Update $\mathcal{A} := \mathcal{A} \cap \{(y, \gamma) : f(y) \leq z^U\}$ (to work with lower objective values);
    Adjust $(y_1, \gamma_1) := \frac{1}{tol}(y_0, \gamma_0)$ (to proceed with the while loop)
if *No feasible solution is obtained,* **then**
  Problem (4.3) could be infeasible, return "Inconclusive, out of iterations"
else
  Return "The best obtained solution is (y^*, γ^*)"

such as the work in [34] and [38], or convex relaxations, such as the work in [43]. 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. Since power flow approximations and relaxations can introduce substantial errors relative to the non-linear AC power flow equations [36, 16, 2], 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 [29]. 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 [12] or convex relaxations as in [28]. 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 [32] 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 [26] that uses so-called "convex restriction" techniques (see [25]) 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.

Accordingly, new computational methods such as the framework proposed in this paper are needed to address power system optimization problems that model both uncertain power injections from renewable generators and non-linearities from the AC power flow equations. The proposed framework would most naturally fit into a system operator's generation scheduling processes as part of near-real-time (five-minute to hourly) generator setpoint computations via optimal power flow algorithms, which is a key application at the heart of power system operations. Since optimization in this case is repeated frequently during the day, the uncertain deviations in loads or generation are likely to be of small size. However, they can lead to infeasibility of the nominal solutions, thus an approach is needed that takes the uncertainty into account. This situation corresponds well with the result in Corollary 3.5 (b) about small uncertainty sets. Extensions to incorporate binary variables would facilitate other applications, such as solving unit commitment problems for day-ahead scheduling that models generators' start up and shut down decisions. Thus, developing algorithms suitable for such extensions is an important direction for future work.

5.1. Robust ACOPF formulation. Exploiting the polynomial representation of the AC power flow equations, we next apply the approach described in this paper to the robust ACOPF problem, beginning with our notation and the problem formulation. Consider a power network with the set of buses $N = \{1, \dots, n\}$ 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^d_k and Q^d_k, 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 [50]. 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^r_k and Q^r_k, respectively, for all $k \in N$. In the nominal ACOPF without uncertainty, P^r_k and Q^r_k are known and fixed. Next we define the ACOPF problem as a quadratic optimization problem.

Problem 5.1.

$$
\begin{aligned}
z^{nom} = \inf_{x, P^g, Q^g} \quad & \sum_{k \in G} c^2_k (P^g_k)^2 + c^1_k P^g_k + c^0_k \\
\text{s.t.} \quad & P^{\min}_k \le P^g_k \le P^{\max}_k && \textit{for all } k \in N \\
& Q^{\min}_k \le Q^g_k \le Q^{\max}_k && \textit{for all } k \in N \\
& (V^{\min}_k)^2 \le x_k^2 + x_{k+n}^2 \le (V^{\max}_k)^2 && \textit{for all } k \in N \\
& tr(Y_{lm} x x^\top) \le S^{\max}_{lm} && \textit{for all } \{lm\} \in E \\
& P^g_k + P^r_k = P^d_k + tr(Y_k x x^\top) && \textit{for all } k \in N \\
& Q^g_k + Q^r_k = Q^d_k + tr(\bar{Y}_k x x^\top) && \textit{for all } k \in N \\
& x_{s+n} = 0,
\end{aligned}
$$

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 $P^r_k = \bar{P}^r_k + \zeta_k$, 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

$$
Q^r_k = \bar{Q}^r_k + \gamma_k \zeta_k, \ \gamma_k := \begin{cases} 0 & \text{if } P^d_k = 0, \\ \frac{\sqrt{1 - \cos^2 \phi_k}}{\cos \phi_k} & \text{otherwise.} \end{cases}
$$

Without loss of generality, we let $\bar{P}^r_k = \bar{Q}^r_k = 0$, otherwise one can adjust the loads P^d_k and Q^d_k. 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^g_k and an adjustment in generation due to the uncertainty:

$$P^g_k - \alpha_k \left(\sum_{i=1}^n \zeta_i - \delta \right), \ \alpha_i \geq 0, \ \sum_{i=1}^n \alpha_i = 1, \tag{5.1}$$

Thus, when introducing uncertainty to problem (5.1), we replace P^g_k in this problem with (5.1). We note that this model represents the steady-state behavior of widely used automatic generation control (AGC) (see [22]) and is adopted in many robust and stochastic OPF formulations, e.g., those used by [43], [38], and [32]. 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 $V^g_k := x_k^2 + x_{k+n}^2$ 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^g_k and squared voltage magnitudes V^g_k are set by the operator while the reactive power Q^g_k 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^g_k, where k belongs to the set of PV buses, and V^g_k, where k belongs to the union of PV and reference buses.

Now we define the problem in the same form as (2.3) to more easily use the results from the earlier sections. This yields the following:

$$f(t, P^g, V^g) = \sum_{k \in G \setminus \{s\}} c^2_k (P^g_k)^2 + c^1_k P^g_k + c^0_k + c^2_s t^2 + c^1_s t + c^0_s, \tag{5.2}$$

$$\begin{aligned} S_y = \big\{ &(P^g, V^g) : P^{\min}_k \leq P^g_k \leq P^{\max}_k \ \text{ for all } k \in G \setminus \{s\}, \\ &(V^{\min}_k)^2 \leq V^g_k \leq (V^{\max}_k)^2 \ \text{ for all } k \in G, \ P^{\min}_s \leq t \leq P^{\max}_s \big\}, \end{aligned} \tag{5.3}$$

$$L(P^g, V^g, \zeta, x) = \begin{bmatrix} P^d_k + \operatorname{tr}(Y_k x x^\top) - \zeta_k - P^g_k + \alpha_k \sum_{i=1}^n \zeta_i & \text{for all } k \in G \setminus \{s\} \\ P^d_k + \operatorname{tr}(Y_k x x^\top) - \zeta_k + \alpha_k \sum_{i=1}^n \zeta_i & \text{for all } k \in N \setminus G \\ Q^d_k + \operatorname{tr}(\bar{Y}_k x x^\top) - \gamma_k \zeta_k & \text{for all } k \in N \setminus G \\ x_k^2 + x_{k+n}^2 - V^g_k & \text{for all } k \in G \\ x_{s+n} & \end{bmatrix}, \tag{5.4}$$

$$G(t,P^g,V^g,\zeta,x) = \begin{bmatrix} -P_s^{\min} + P_s^d + \operatorname{tr}(Y_s x x^\top) - \zeta_s + \alpha_k \sum_{i=1}^n \zeta_i & \\ t - P_s^d - \operatorname{tr}(Y_s x x^\top) + \zeta_s - \alpha_k \sum_{i=1}^n \zeta_i & \\ -Q_k^{\min} + Q_k^d + \operatorname{tr}(\bar{Y}_k x x^\top) - \gamma_k \zeta_k & \text{for all } k \in G \\ Q_k^{\max} - Q_k^d - \operatorname{tr}(\bar{Y}_k x x^\top) + \gamma_k \zeta_k & \text{for all } k \in G \\ -\operatorname{tr}(Y_{lm} x x^\top) + S_{lm}^{\max} & \text{for all } \{lm\} \in E \end{bmatrix}, \tag{5.5}$$

$$S_x = \left\{ x \in \mathbb{R}^{2n} : (V_k^{\min})^2 \le x_k^2 + x_{k+n}^2 \le (V_k^{\max})^2 \text{ for all } k \in N \setminus G \right\} \tag{5.6}$$

In the next subsection, we run numerical experiments solving problem (2.3) with the inputs defined above and the instances from MATPOWER [50].

5.2. Numerical results. In this section, we implement Algorithm 2 and compare it with several algorithms used in the literature. All computations are done using MATLAB R2021a and Yalmip (see [27]) on a computer with the processor Intel® Core® i7-8665U CPU @ 1.90GHz and 16 GB of RAM. Semidefinite programs are solved with MOSEK, Version 9.3.21 [33].

In the experiments, we need to choose the vector α and the parameters defining the ellipsoid Ω. We assume that all uncertainty balances out on the reference bus since this is the default setup of MATPOWER. Therefore $\alpha_s = 1$ and $\alpha_k = 0$ for $k \neq s$.

We implement Algorithm 2 with the following inputs:

- We do not split the uncertainty set Ω and set $\hat{\zeta} = 0$.
- We consider only one $\hat{x}$: the initial feasible solution to the solution of the nominal problem (5.1) provided by MATPOWER [50]. For constraint (**) in Algorithm 2, we set $\gamma = \sqrt{\|x_{j-1}\|/10}$ for $|N| < 30$, and $\gamma = \sqrt{\|x_{j-1}\|/30}$ for $|N| \ge 30$ since the solution norm grows with the instance size.
- We reformulate problem (3.6) as problem (4.3) (see Proposition 4.2).
- To solve problem (4.3), we use Algorithm 3 with $f_0 = 1 \times 10^5$, $tol = 1 \times 10^{-4}$, $N = 100$, $\nu_i = 1$ for all $i \in [N]$.
- To compute the lower bound in Algorithm 3, we use the classical SDP relaxation of the quadratic problem (4.3), linearizing all non-convex quadratic terms in the control variables and obtaining an SDP constraint of the size $n_y + 1$.
- In the ACOPF problem, some matrices A in (4.2) are negative semidefinite. We do not have to project on the corresponding equality constraints (4.3b). Instead, we replace the equality sign by "$\le$" and add the resulting constraints to the definition of $\mathcal{A}$ in (4.3).
- We do the posterior check in step 2 of Algorithm 2 for $\hat{\zeta} = 0$, i.e., we check that the obtained solution is nominally feasible.

Setting $\nu_i = 1$ implies that we do not try to improve the objective value after finding the first feasible solution. We considered setting $\nu_i < 1$, however, for all 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 or close to feasible solution to problem (4.3). Therefore, we decided to stop at the first obtained feasible solution by setting $\nu_i = 1$ for all $i \in [N]$.

As benchmarks, we use three other approaches. First, we consider the nominal solution found by MATPOWER to see how robust or not it is. Next, we implement the classical DCOPF relaxation of ACOPF, which is a linear problem. We follow the approach from [8]. In this paper, the parameters α are optimized while in our case they are fixed. We can eliminate all equalities and second-stage variables in DCOPF, thus obtaining a standard linear robust optimization problem under ellipsoidal uncertainty. Finally, we implement an approach based on the classical SDP relaxation of ACOPF [24]. Without the uncertainty, this relaxation usually provides exceptionally good approximations. A number of papers in the literature (e.g., [43, 46]) implement an ARO version of this relaxation by replacing the nominal positive semidefinite matrix in the relaxation by a linear matrix function of the uncertainty; see the earlier mentioned papers for details. Since the initial relaxation is linear, except for the SDP constraint, after substituting the linear decision rule, all equalities can be eliminated and the standard linear robust approach applies. To ensure approximate feasibility of the uncertain SDP constraint, one usually imposes this constraint for many enough realizations of the uncertainty. We followed the approach in [43] and imposed the SDP constraint for the largest possible value of each uncertainty. We have also tried using the sampling approach to ensure feasibility of the constraint with high probability, as in [46]. However, even for small cases this required many scenarios, resulted in too many SDP constraints, and lead to numerically unstable problems even if we used row generation to add the constraints.

For all algorithms, we limit the running time to 15 minutes (for our approach, 15 minutes from the moment after obtaining the SDP lower bound). DCOPF and Taylor approach do not reach this time limit in the presented cases. For the SDP relaxation, this limit is reached for the large cases (57 and 118 buses), therefore this approach is omitted in the corresponding table.

The two benchmark models we consider have affine constraints and use a chance constraints framework. However, to obtain the final reformulations, both models exploit the analogy between the ellipsoidal uncertainty and chance constraints for normally distributed random variables and affine constraints. Therefore, we also use this analogy to conduct experiments. Namely, we say that the uncertainty vector ζ is a random variable with the mean zero and the vector of standard deviations is equal to $w\%$ of the initial load. Let ζ be normally distributed, and let W be a variance-covariance matrix of ζ. Let P be such that $W = PP^\top$ (obtained, e.g., via Cholesky decomposition) and $q_{1-\varepsilon}$ be the $(1-\varepsilon)$-percentile of the standard normal distribution. Then making one affine constraint robust against the uncertainty set

$$\Omega := \{\zeta : \|P^{-1}\zeta\| \leq q_{1-\varepsilon}\} \tag{5.7}$$

is equivalent to requiring this constraint to hold with probability at least $1-\varepsilon\%$ for ζ. This is a well-known result which follows from classical robust optimization (see, e.g., [4]) applied to a linear constraint with ellipsoidal uncertainty and a straightforward reformulation of a chance constraint using the cumulative probability function.

We allow the uncertainty to occur at all buses with positive active power loads. This is easy 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 the variance-covariance matrix W. For the first option, we set $W = \text{Diag}\,(wP^d_k)$, where Diag is the operator that creates a diagonal matrix from a vector. We let w vary 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 random correlations among the uncertainties. Namely, we generate a random positive definite matrix and scale it such that the standard deviations are equal to wP^d_k. We relate our uncertainty to 95% chance constraints using $\Sigma = {P^{-1}}^\top P^{-1}$, $\rho = 0$ and $r = -1.65^2$ in (2.1). 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.

Choosing the probability of violations of the chance constraints is not trivial. The number of constraints in the problem is large even for small cases. Thus, when we restrict the probability that each chance constraint is violated, the probability that *at least* one constraint is violated tends to one when the number of constraints grows. Formally, to obtain the 5%-safe radius of the ellipsoid for m constraints, we can just require all chance constraints to hold with probability $1 - \varepsilon/m$. However, in practice this approach is usually too conservative, and thus we choose the typical 5% level for each constraint and investigate how robust the solutions are for this setup.

Another reason to connect robust constraints to a probabilistic setup is the property that the volume of the ball with a constant radius tends to zero when the dimension of the ball grows. Thus, for larger cases, it would be hard to generate a large enough sample of uncertainties within the unit ball while we can easily generate uncertainties from a normal distribution.

To evaluate the performance of the algorithms, we generate ζ according to the normal distribution around zero with the covariance matrix W. In particular, we first generate a standard normal vector z and then use $\zeta = Pz$, where P comes from the earlier mentioned factorization $W = PP^\top$. We generate 1000 values of the uncertainly including the nominal case where $\zeta = 0$.

The computational results are presented in Tables 1–3. Tables 1 and 2 cover power networks with fewer than 30 buses, and Table 3 shows the results for larger power networks. Our main indicator is the share of experiments where the solution provided by the corresponding approach had some infeasibilities. We round all infeasibilities down to 1×10^{-3} per unit since this is precise enough for our application. We also checked more precise rounding, and the differences in the tables were minor. In all presented instances, the equality constraints were always feasible.

For each case, we present all levels of uncertainty where at least one model has at most 10% of constraint violations. As we mentioned earlier, it is not obvious which performance should be considered good, especially for large cases. A perfect performance would be to have less than 5% of experiments with constraints violations since this is the violation possible in our setup for a single constraint. However, given that even the smallest case we present has more than ten constraints, we choose 10% as an acceptable violation.

We do not present typical Matpower smallest cases with three buses (LMBM3) and five buses (WB5) since for the first case even the nominal solution is robustly feasible for moderate values of the uncertainty, and for the second case none of the models is able to obtain a robustly feasible solution. The latter test case is known to be especially challenging in various contexts (see [31]).

In all tables, the first column denotes the values of the uncertainty as a fraction of load, mentioned earlier as w, in percent. The second column indicates the evaluated model. The first two model names speak for themselves, the third name "SDP" means the model from [43], and the name "Taylor" indicates the approach from this paper. The third column shows average estimated objective values among all instances. In general, robust optimization approaches do not aim to optimize the average objective, so the second column is just a rough indicator of the performance of each approach in terms of the objective value. We find this indicator sufficient since our primary goal is not to compare the objectives but rather to ensure feasibility of the constraints. The fourth column shows the running time of each benchmark algorithm required to find a solution used for the experiments. We do not include the input construction time for each test case since all benchmarks use roughly the same inputs. The fifth column highlighted by grey shows our main indicator—the share of all experiments where the solution provided by the corresponding approach had some infeasibilities. The last two columns show the maximum number of violated constraints among all experiments: "PQ" indicates violations of active and reactive power bounds in one test case and "VI" indicates violations of voltage and current flow bounds in one test case. Small violations of the "PQ" constraints could lead to more serious problems than small violations of the "VI" constraints, therefore we have separated these indicators.

A hyphen indicates that the corresponding model delivered no solution (the resulting problem was infeasible). The bold font in the "Model" column indicates the best model: the one which is most robust and runs in the shortest time. We have also marked by the bold font in the "Constraint violations" column all best performing models in terms of feasibility.

Table 1
Results for instances up to 30 buses, without correlation. *All objective values are divided by 100 in comparison to the original data.*

Uncertainty, % of load	Model	Average objective	Time, sec.	Constraint violations, %	Max # violations, per exper., PQ	Max # violations, per exper., VI
case 6ww, 6 buses						
1	Nominal	31.3	0.0	43.9	0	2
	DCOPF	-	-	-	-	-
	SDP	31.4	28.4	22.1	0	2
	Taylor	31.6	69.4	**0.0**	0	0
case 9, 9 buses						
1	Nominal	53.0	0.0	0.7	0	2
	DCOPF	53.2	9.1	**0.0**	0	0
	SDP	53.0	22.5	100.0	0	2
	Taylor	53.3	57.6	**0.0**	0	0
5	Nominal	53.2	0.0	35.7	0	2
	DCOPF	53.5	13.8	**0.0**	0	0
	SDP	53.2	30.5	100.0	0	3
	Taylor	53.4	72.7	**0.0**	0	0
10	Nominal	53.5	0.0	43.9	0	2
	DCOPF	54.4	12.6	0.1	1	0
	SDP	53.6	23.8	87.5	0	3
	Taylor	53.7	70.4	**0.0**	0	0
20	Nominal	54.9	0.0	48.5	1	5
	DCOPF	55.5	12.1	3.4	1	0
	SDP	55.0	29.4	99.8	1	6
	Taylor	55.0	74.4	**1.0**	1	0
30	Nominal	57.1	0.0	51.3	1	6
	DCOPF	-	-	-	-	-
	SDP	57.4	26.1	97.8	1	6
	Taylor	57.2	68.1	**7.1**	1	2
case 30, 30 buses						
1	Nominal	6.1	0.1	100.0	0	2
	DCOPF	-	-	-	-	-
	SDP	5.8	178.6	31.7	0	2
	Taylor	5.9	177.2	**1.7**	0	2

In most cases, the nominal solution is not robustly feasible, thus considering the uncertainty is important. The results show that our approach provides a robust solution in reasonable time even for the instance with 118 buses. By Corollary 3.5, the approach we implemented is guaranteed to be robust only for small uncertainty sets, which is indeed reflected in the numerical results. For all presented cases the "Taylor" solution is robust in more than 95% of the experiments when the uncertainty is 1% of the loads while this is not the case for the nominal solutions and other approaches. The second-best approach in terms of robustness is "DCOPF". It works for some small

Table 2
Results for instances up to 30 buses, with random correlation. *All objective values are divided by 100 in comparison to the original data.*

Uncertainty, % of load	Model	Average objective	Time, sec.	Constraint violations, %	Max # violations, per exper., PQ	Max # violations, per exper., VI
case 6ww, 6 buses						
1	Nominal	31.3	0.1	43.9	0	2
	DCOPF	-	-	-	-	-
	SDP	31.4	80.9	25.3	0	2
	Taylor	31.6	176.9	**0.0**	0	0
5	Nominal	31.4	0.1	47.0	0	3
	DCOPF	-	-	-	-	-
	SDP	31.4	80.8	37.6	0	3
	Taylor	32.0	177.0	**4.8**	0.0	2.0
case 9, 9 buses						
1	**Nominal**	53.0	0.0	**0.0**	0	0
	DCOPF	53.2	33.9	**0.0**	0	0
	SDP	53.0	81.7	100.0	0	2
	Taylor	53.3	201.5	**0.0**	0	0
5	Nominal	53.1	0.0	23.3	0	2
	DCOPF	53.3	33.5	**0.0**	0	0
	SDP	53.1	81.8	100.0	0	3
	Taylor	53.3	199.5	**0.0**	0	0
10	Nominal	53.2	0.0	37.6	0	3
	DCOPF	53.8	34.2	**0.0**	0	0
	SDP	53.3	81.7	100.0	0	4
	Taylor	53.5	201.6	**0.0**	0	0
20	Nominal	54.0	0.0	47.8	1	5
	DCOPF	55.4	34.0	2.6	1	0
	SDP	54.2	82.0	0.7	1	0
	Taylor	54.1	201.0	**0.1**	1	0
30	Nominal	55.1	0.0	53.7	1	5
	DCOPF	55.5	33.9	6.1	1	1
	SDP	55.1	82.3	68.6	1	5
	Taylor	55.2	201.0	**2.1**	1	1
40	Nominal	56.7	0.0	57.8	1	5
	DCOPF	-	-	-	-	-
	SDP	56.8	82.0	93.0	1	6
	Taylor	56.8	201.5	**6.5**	1	2
case 30, 30 buses						
Infeasible for all robust approaches.						

instances but fails for larger ones. DCOPF is an accurate approximation under quite restrictive assumptions and seems to become less precise under uncertainty. Finally the “SDP” model provides most violations among the three robust models. However, if we look deeper into these violations, when the uncertainty sets are small, the “SDP” solution usually leads to few small violations of “VI” constraints. This could be expected since the voltages are approximated in the SDP relaxation and

Table 3

Results for instances larger than 30 buses, without correlation*. All objective values are divided by 100 in comparison to the original data. The SDP approach results in too-large-to-solve problems for these instances, so we do not mention this approach in the table.*

Uncertainty, % of load	Model	Average objective	Time, sec.	Constraint violations, %	Max # violations, per exper., PQ	Max # violations, per exper., VI
case 57, 57 buses						
1	Nominal	417.4	0.1	70.6	2	1
	DCOPF	418.5	43.0	100.0	2	1
	Taylor	426.8	467.2	**0.0**	0	0
case 118, 118 buses						
1	Nominal	1296.7	0.2	99.5	9	0
	DCOPF	1315.6	94.9	100.0	21	0
	Taylor	1301.3	830.0	**1.1**	1	0

these approximations could become looser under uncertainty. Adding some tightening constraints to the "SDP" model could potentially improve the robustness of that approach.

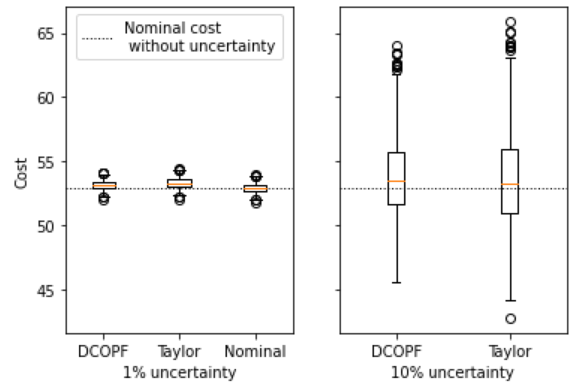
As to the running times, our approach is the slowest for small test cases since it is an iterative approach due to Algorithm 3. However, when our approach found robust solutions, it could usually do that within several iterations of the alternating projections Algorithm 3. Longer running times for the two largest instances indicate that the alternating projections algorithm went through many iterations and could not converge, thus a robust solution may not exist for the considered subset of state variables. For the two largest cases, the SDP approach is the slowest, and it is in fact too large to solve (e.g., for case 118 it includes 100 SDP constraints of the size 236×236).

Finally, our approach often provides somewhat higher average objective than others. Given that the approach is more robust than the others, the difference in the average objective is minimal. To understand the performance with respect to the objective better, we looked at Case 9 with correlations, which is robustly feasible for three of four models for the 1%-uncertainty and robustly feasible for two models for the 10%-uncertainty. Below we show the box plots of the objective values of all robustly feasible approaches. The two box plots use same interval of possible objective values in the y-axes for comparability. We see that the "Taylor" approach tends to have higher cost in both cases, and it also results in larger standard deviations for the larger uncertainty, meaning that very low costs are also more common than in DCOPF. The worst-case objective value of our algorithm in the experiments is a little larger than for DCOPF, and the algorithm does not find the DCOPF solution. To find more solutions, such as the DCOPF solution when it is robust, it could be useful to enlarge the search space of the algorithm.

Finally, we see that adding correlations to the uncertainty ellipsoid does not substantially change the results for small cases. However, when the instance size grows, no robust solutions could be found by any model for "case30", "case57", and "case118", even for the lowest values of the uncertainty. Therefore, we do not provide results with correlations for larger cases. Further investigation is needed to see why exactly this situation occurs and to possibly tune the models to work for larger cases with correlated uncertainties.

Our evaluations are equally conservative for all models. First, we say that more than 10% of experiments with constraint violations is a bad performance while we have many constraints and the analogy between the normal distribution and our ellipsoid is only valid for one constraint. The

Fig. 1. *Box plot of the objective value realizations for models without constraint violations for Case 9 with correlations,* 1% *and* 10% *uncertainty.*



5%-safe radius of the ellipsoid for many constraints would be much larger, but using this radius would be too conservative. Second, we do not deeply analyze magnitudes of violations. By the experiment setup they are larger than 1×10^{-3} per unit, but the importance of such violations can depend on the type of constraint and the details of the test case. Finally, we did not fine-tune the algorithm to each specific case and allowed for as many power injection uncertainties as possible. The performance could improve if the algorithms are fine-tuned, especially for large instances.

6. Conclusions and directions for future research. In this paper, we propose a framework to obtain approximately feasible solutions to quadratic ARO problems with equalities, which implicitly define the second-stage decision rules for the state variables. We replace the implicit decision rules by their explicit piecewise affine approximations. As a result, we can eliminate the state variables from the problem and replace the original ARO problem by a sequence of classical quadratic problems with additional tractable conic constraints. Since, a generally efficient algorithm does not exist yet for the latter problems, we design an alternating projections algorithm that converges to a local optimum of the problem. For any $\varepsilon > 0$, if the piecewise affine approximations are fine enough, the suggested algorithm guarantees that the second-stage equality constraints are satisfied and the inequality constraints are not violated by more than ε on "large" subsets of uncertainty, where "large" is defined in Theorem 3.4 (a), (b). The feasibility of the second-stage equalities is rarely addressed in the literature, thus we consider analysing and ensuring such feasibility an important contribution to the existing research. We suggest two versions of the algorithm, Algorithm 1 for general problems and Algorithm 2 for problems linear in the control variables.

We implement the algorithm for ACOPF problems with uncertainty in loads and simulate the uncertainty to evaluate the performance of the algorithm in comparison to three known benchmarks: nominal solution, robust DCOPF and robust SDP relaxation. The solutions provided by our approach are robustly feasible for small uncertainty sets and for cases with up to 118 buses. Moreover, the solutions are substantially more robust than the benchmarks. The algorithm also performs well in terms of the objective function values. The experiments show that DCOPF and SDP approximations work well for problems without uncertainty, but possible inexactness of these methods is magnified after adding the uncertainty. As a result, they are less robust in an ex-post experiment. In contrast, our approach yields more robust results in the experiments. The good performance can be explained by the fact that our approach preserves more non-linearities from the original problem. Notice that the size of the final problem (i.e., problem (3.6)) which we solve depends on the number of sources of uncertainty. Hence, even for larger networks, the approach could work well for systems with uncertain power injections at a limited subset of the buses.

We conclude by analyzing the limitations of our algorithms and suggesting directions for further research. First, the non-linearities which we keep in the problem have drawbacks. Namely, when DCOPF can find a robust solution, it does so substantially faster than the approach we propose. At the last step of our approach, after removing the state and uncertainty variables, we need to solve a problem with quadratic and SDP constraints. We proposed the alternating projections in Algorithm 3 for this purpose. This algorithm only finds locally optimal solutions, may take a number of iterations to converge, and each iteration solves an SDP problem, which makes the procedure rather slow for larger instances. However, this approach is still relevant since no generally efficient algorithms exist for such problems yet, and our Algorithm 3 contributes to the development of quadratic optimization with SDP constraints. In the future, one could use an alternative algorithm or a relaxation to solve the obtained problem with quadratic and SDP constraints.

Next, our algorithms can exploit various subsets of uncertainty, control and/or state variables. However, the actual implementation considered small uncertainty sets without partitioning them, and we restricted the search to one subset of state variables around the initial optimal solution. As expected, for larger values of the uncertainty in our numerical experiments, the algorithm becomes imprecise or could not find a feasible solution. A natural remedy to increase precision and find additional feasible solutions would be to consider several subsets of the state variables as described in subproblem (3.6). Next, to work with larger uncertainty sets, we could combine Algorithm 1 and Algorithm 2 with the approach proposed in [35] to split larger uncertainty sets.

Third, some improvements could be done in particular for ACOPF problems. These problems possess much sparsity, and it is to a large extent preserved under the transformation (3.2). Therefore, 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, especially those incorporating binary variables to enable application to unit commitment problems for longer-term generator scheduling.

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. We also greatly thank the referees, the Associate Editor, and the Editor for their thorough and thoughtful comments that helped us improve the quality of the paper. Finally, we are grateful to Roweno J.R.K. Heijmans and Juan C. Vera for fruitful discussions and advice about the paper.

REFERENCES

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

[2] K. Baker, *Solutions of DC OPF are never AC feasible*, in Proceedings of the Twelfth ACM International Conference on Future Energy Systems, 2021, pp. 264–268.

[3] B. Baumrucker and L. T. Biegler, *MPEC strategies for cost optimization of pipeline operations*, Computers & chemical engineering, 34 (2010), pp. 900–913.

[4] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton Series in Applied Mathematics. Princeton University Press, 2009.

[5] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, *Adjustable robust solutions of uncertain linear programs*, Mathematical Programming, A (2004), pp. 351–376.

[6] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, *Adaptive robust optimization for the security constrained unit commitment problem*, IEEE Transactions on Power Systems, 28 (2012), pp. 52–63.

[7] C. Biefel, F. Liers, J. Rolfes, and M. Schmidt, *Affinely adjustable robust linear complementarity problems*, SIAM Journal on Optimization, 32 (2022), pp. 152–172.

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

[9] D. Bienstock and N. Özbay, *Computing robust basestock levels*, Discrete Optimization, 5 (2008), pp. 389–414.

[10] D. Bienstock and A. Verma, *Strong NP-hardness of AC power flows feasibility*, Operations Research Letters, 47 (2019), pp. 494–501.

[11] C. Bingane, M. F. Anjos, and S. Le Digabel, *Tight-and-cheap conic relaxation for the optimal reactive power dispatch problem*, IEEE transactions on power systems, 34 (2019), pp. 4684–4693.

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

[13] W. Chen, M. Sim, J. Sun, and C.-P. Teo, *From CVaR to uncertainty set: Implications in joint chance-constrained optimization*, Operations research, 58 (2010), pp. 470–485.

[14] J. Dezert and C. Musso, *An efficient method for generating points uniformly distributed in hyperellipsoids*, in Proceedings of the Workshop on Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom, vol. 7, 2001.

[15] D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis, *Transversality and alternating projections for nonconvex sets*, Foundations of Computational Mathematics, 15 (2015), pp. 1637–1651.

[16] K. Dvijotham and D. K. Molzahn, *Error bounds on linear power flow approximations: A convex relaxation approach*, in IEEE 55th Annual Conference on Decision and Control (CDC), Las Vegas, NV, USA, December 2016, pp. 2411–2418.

[17] A. Georghiou, A. Tsoukalas, and W. Wiesemann, *A primal–dual lifting scheme for two-stage robust optimization*, Operations Research, 68 (2020), pp. 572–590.

[18] B. Ghaddar, M. Claeys, M. Mevissen, and B. J. Eck, *Polynomial optimization for water networks: Global solutions for the valve setting problem*, European Journal of Operational Research, 261 (2017), pp. 450 – 459.

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

[20] F. Golestaneh, P. Pinson, R. Azizipanah-Abarghooee, and H. B. Gooi, *Ellipsoidal prediction regions for multivariate uncertainty characterization*, IEEE Transactions on Power Systems, 33 (2018), pp. 4519–4530.

[21] N. M. Isenberg, P. Akula, J. C. Eslick, D. Bhattacharyya, D. C. Miller, and C. E. Gounaris, *A generalized cutting-set approach for nonlinear robust optimization in process systems engineering*, AIChE Journal, (2021), p. e17175.

[22] N. Jaleeli, L. S. VanSlyck, D. N. Ewart, L. H. Fink, and A. G. Hoffmann, *Understanding automatic generation control*, IEEE Transactions on Power Systems, 7 (1992), pp. 1106–1122.

[23] J. B. Lasserre, *Tractable approximations of sets defined with quantifiers*, Mathematical Programming, 151 (2015), pp. 507–527.

[24] J. Lavaei and S. H. Low, *Zero duality gap in optimal power flow problem*, IEEE Transactions on Power Systems, 27 (2012), pp. 92–107.

[25] D. Lee, H. D. Nguyen, K. Dvijotham, and K. Turitsyn, *Convex restriction of power flow feasibility sets*, IEEE Transactions on Control of Network Systems, 6 (2019), pp. 1235–1245.

[26] D. Lee, K. Turitsyn, D. K. Molzahn, and L. Roald, *Robust AC optimal power flow with convex restriction*, IEEE Transactions on Power Systems, 36 (2021), pp. 4953–4966.

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

[28] Á. Lorca and X. A. Sun, *The adaptive robust multi-period alternating current optimal power flow problem*, IEEE Transactions on Power Systems, 33 (2018), pp. 1993–2003.

[29] R. Louca and E. Bitar, *Robust AC optimal power flow*, IEEE Transactions on Power Systems, 34 (2019), pp. 1669–1681.

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

[31] D. K. Molzahn, *Computing the feasible spaces of optimal power flow problems*, IEEE Transactions on Power Systems, 32 (2017), pp. 4752–4763.

[32] D. K. Molzahn and L. A. Roald, *Towards an AC optimal power flow algorithm with robust feasibility guarantees*, in 20th Power Systems Computation Conference (PSCC), 2018, pp. 1–7.

[33] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 8.0.0.81.*, 2019, https://www.mosek.com/documentation/.

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

[35] K. Postek and D. d. Hertog, *Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set*, INFORMS Journal on Computing, 28 (2016), pp. 553–574.

[36] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, *Usefulness of DC Power Flow for Active Power Flow Analysis*, in IEEE Power Engineering Society General Meeting, 2005, pp. 454–459.

[37] A. U. Raghunathan and L. T. Biegler, *Mathematical programs with equilibrium constraints (MPECs) in process engineering*, Computers & chemical engineering, 27 (2003), pp. 1381–1392.

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

[39] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo, *Power systems optimization under uncertainty: A review of methods and applications*, to appear in Electric Power Systems Research, presented at the 22nd Power Systems Computation Conference (PSCC 2022).

[40] H. D. Sherali and C. H. Tuncbilek, *A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique*, Journal of Global Optimization, 2 (1992), pp. 101–112.

[41] M. Spivak, *Calculus on Manifolds. A Modern Approach to Classical Theorems of Advanced Calculus.*, Addison-Wesley Publishing Company, 1995.

[42] A. Stuhlmacher and J. L. Mathieu, *Water distribution networks as flexible loads: A chance-constrained programming approach*, in 21st Power Systems Computation Conference (PSCC), 2020, pp. 1–8.

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

[44] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, *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 (2013), pp. 3885–3896.

[45] Y. Wang, L. Da Silva, and W. Xu, *Investigation of the relationship between ill-conditioned power flow and voltage collapse*, IEEE Power Engineering Review, 20 (2000), pp. 43–45.

[46] W. Xie and S. Ahmed, *Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation*, IEEE Transactions on Power Systems, 33 (2018), pp. 1860–1867.

[47] V. A. Yakubovich, *S-procedure in nonlinear control theory*, Vestnik Leningrad University, 4 (1977), pp. 73–93. english translation; original Russian publication in Vestnik Leningradskogo Universiteta, Seriya Matematika (1971), 62–77.

[48] İ. Yanikoğlu, B. L. Gorissen, and D. den Hertog, *A survey of adjustable robust optimization*, European Journal of Operational Research, 277 (2019), pp. 799–813.

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

[50] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, *MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education*, IEEE Transactions on Power Systems, 26 (2010), pp. 12–19.