Robust Approximation of Chance Constrained DC Optimal Power Flow under Decision-Dependent Uncertainty

Kevin-Martin Aigner*[1], Jan-Patrick Clarner[2], Frauke Liers[1] and Alexander Martin[1,3]

[1]Friedrich-Alexander University Erlangen-Nürnberg, D-91058 Erlangen, Germany
[2]Zuse Institute Berlin, D-14195 Berlin, Germany
[3]Fraunhofer Institute for Integrated Circuits IIS, D-90411 Nürnberg, Germany

22nd October 2021

Abstract

We propose a mathematical optimization model and its solution for joint chance constrained DC Optimal Power Flow. In this application, it is particularly important that there is a high probability of transmission limits being satisfied, even in the case of uncertain or fluctuating feed-in from renewable energy sources. In critical network situations where the network risks overload, renewable energy feed-in has to be curtailed by the transmission system operator (TSO). The TSO can reduce the feed-in in discrete steps at each network node. The proposed optimization model minimizes curtailment while ensuring that there is a high probability of transmission limits being maintained. The latter is modeled via (joint) chance constraints that are computationally challenging. Thus, we propose a solution approach based on the robust safe approximation of these constraints. Hereby, probabilistic constraints are replaced by robust constraints with suitably defined uncertainty sets constructed from historical data. The ability to discretely control the power feed-in then leads to a robust optimization problem with decision-dependent uncertainties, i.e. the uncertainty sets depend on decision variables. We propose an equivalent mixed-integer linear reformulation for box uncertainties with the exact linearization of bilinear terms. Finally, we present numerical results for different test cases from the Nesta archive, as well as for a real network. We consider the discrete curtailment of solar feed-in, for which we use real-world weather and network data. The experimental tests demonstrate the effectiveness of this method and run times are very fast. Moreover, on average the calculated robust solutions only lead to a small increase in curtailment, when compared to nominal solutions.

Keywords: OR in energy, optimal power flow, chance constrained programming, robust optimization, decision-dependent uncertainty

1 Introduction

The current transition towards sustainable energy places complex demands on the operation of power grids. The proportion of power provided by renewable energies such as solar or wind energy is constantly increasing and it can be assumed that this trend will continue in the future. However, this constitutes a challenge for the operation of electricity power networks. Due to changes in the weather, renewable energy sources are uncertain and subject to strong

*Corresponding author. E-mail address: kevin-martin.aigner@fau.de

fluctuations. To guarantee network stability, curtailment of renewable feed-in often cannot be avoided. This entails using only a fraction of the feed-in to avoid overloading transmission lines. However, curtailment should clearly be minimized as far as possible. In addition, regulating feed-in leads to potential compensation payments from the transmission system operator (TSO) to the energy producers. Based on the current network status and using available short-term weather forecasts, the risk of critical network situations is assessed and appropriate interventions are carried out. As a consequence, it may be necessary to curtail feed-in at some network nodes. In practice, this process is frequently not yet fully automatized, but guided by dispatchers. In this work, we will approach this issue from a mathematical perspective.

The primary tool used to model the operation of electrical power networks is the Optimal Power Flow (OPF) problem. Traditionally, it calculates the minimum-cost production and optimal distribution of electrical power in an electricity network. It is a non-linear non-convex optimization problem [19]. Several methods have been developed for its solution. Due to its computational difficulty, they are predominantly based on local solution methods (e.g. interior point algorithms [18]), on the construction of relaxations [40, 61] or on approximation techniques. Local algorithms often generate solutions for large instances within a relatively short computational time period. However, quality statements cannot always be made for these solutions and they may be far from an optimal solution. The performance also depends on the choice of an initial starting value. It is not guaranteed that these methods always deliver a feasible or (local) optimal solution. However, global methods that lead to solutions with demonstrable quality guarantee are usually computationally very expensive and therefore only applicable for small networks. In order to be able to derive efficient solution approaches, approximations are often applied directly to modeling the underlying laws of physics, see [42] for further information.

One of the widely accepted approximations of OPF is the DC Optimal Power Flow (DC OPF) [21]. This approximation assumes fixed voltage magnitudes and small phase angle differences in the network. The reactive part of electrical power is neglected. This results in a power model consisting of linear constraints. Typically, the objective functions modeling power generation costs are linear or convex quadratic. The resulting optimization problem can therefore be solved efficiently by currently available state-of-the-art solvers, even for large network topologies. In general, solutions generated by DC OPF are not feasible for AC OPF. However, due to its algorithmic tractability, this convex approximation is heavily used and of great value in practice. For many questions, it provides a heuristic solution that can be used and processed by power-system operators.
An additional difficulty in power-flow problems is that electricity networks are largely affected by uncertainties. In particular, feed-in from renewable energy sources is difficult to predict and fluctuates heavily due to weather effects. To guarantee network stability, the treatment of uncertain parameters must be included in the optimization models.

For optimization problems including uncertainty, two principal approaches are adopted. Stochastic optimization traditionally involves stochastic quantities (e.g. expectation value or probability), under the assumption of certain probability distributions. Robust optimization is generally used to protect against all possible realizations of unknown parameters from a predefined uncertainty set. Both paradigms usually lead to models that are hard to solve. Only in a few exceptional cases an algorithmically tractable equivalent reformulation exists. The reason is as the calculation of stochastic quantities may involve multidimensional integration with respect to the underlying probability distributions. Due to the worst-case consideration of uncertain parameters in robust optimization, these problems have a bilevel structure, frequently rendering them computationally intractable. However, for several large classes of problems (e.g. linear, convex or combinatorial), algorithmically tractable robust counterparts can be derived. For a broad overview of these two paradigms, we refer the reader to [9, 27] and [49].

In the power network application, it is important to ensure that there is a sufficiently high probability (chosen beforehand) that all transmission limits are satisfied. This can be modeled with joint chance constraints that enforce the simultaneous satisfaction of several constraints with a predefined probability. From a practical perspective, protection through probabilistic constraints is suitable. Short-term overloads in the electricity network are acceptable. In the event of larger or longer lasting overloads, countermeasures will need to be taken, where a TSO will need to (re-)optimize interventions in order to stabilize the network.
In our model, curtailment limits the output of renewable power production to a specific percentage proportion of the installed power. According to the German Renewable Energy Sources Act, these curtailment options realizes in discrete steps in Germany. Due to the nature of the problem, the model consists of two stages. In the first stage, the nominal network operating solution, including generator output, (discrete) curtailment, power flows and voltage angles, has to be decided before the realization of uncertainty is revealed (here-and-now). After the uncertain parameters manifest themselves, the two-stage variables react to them. In the second stage, the network response to fluctuation ensures that there is a high probability of transmission limits are maintained.

We approximate the probabilistic optimization problem using robust constraints within a robust safe approximation, see [46]. By a suitable choice of the uncertainty set we can ensure that all robust feasible solutions are also feasible for the stochastic optimization problem. The constraints of the robust approximation thus lead to sufficient conditions for the chance constraints being satisfied. Therefore, we are able to make quality statements about the approximate solutions. We derive a mixed-integer reformulation for the robust counterpart using box uncertainty sets. Thus, by solving only one mixed integer optimization program, a robust solution can be calculated that is feasible for the chance constrained problem. We adapt the procedure from [41], where the respective uncertainty sets are computed with the help of the scenario approach [15], which uses available historical data. In this approach, we do not need to make any concrete assumptions on the unknown distribution and can rely on real-world data.

The combination of discrete decision variables for curtailment and the robust safe approximation leads to robust constraints with decision-dependent (or endogenous) uncertainty sets. This means that the shape and size of the uncertainty sets are altered by choices of decision variables. In our application, the uncertainty sets shrink with increasing curtailment.
For linear robust optimization problems with decision-dependent uncertainty, two general frameworks exist in the literature. These introduced reformulations are based on duality arguments with linearizations of bilinear products using mixed-integer linear constraints (big-M) [47] or implication constraints [38]. In both approaches it is generally not possible to derive an equivalent algorithmic tractable reformulation. Only in the special case of binary control of upper bounds in the uncertainty set can an estimation of exact linearizations be given under non-negativity assumptions. In this work, we show that these assumptions can be dropped in the case of decision-dependent box uncertainty sets with discrete linear dependencies. We present an exact reformulation into an algorithmically tractable mixed-integer optimization problem and show that our curtailment model is appropriate in this setting.

For the consideration of decision-dependent uncertainty in stochastic optimization problems, one can distinguish between two types of models in the literature: decisions impacting the timing and realization of information in stochastic processes and decision-dependent probability distributions. For example the authors of [28] propose a mixed-integer disjunctive programming model incorporating decision-dependent information discovery in stochastic processes. Basciftci et al. [6] formulate a mixed-integer linear program for finite processes in capacity expansion problems along with approximation algorithms. In the context of decision-dependent distributions, we refer exemplarily to [5] and [7] where mixed-integer reformulations are used for the operation scheduling in power systems and facility location planning. An overview of recent

studies for this class of problems can be found in [31].
The discrete curtailment of solar feed-in impacts the realization of uncertainty, and therefore this paper is related to stochastic problems with decision-dependent probability distributions.

Contribution The key contributions of our work are:

Chance Constrained DC OPF with Discrete Decision-Dependent Uncertainty: We propose a new stochastic optimization model for DC OPF with chance constraints involving decision-dependent uncertainties and discrete variables.
Robust Approximation: We derive an inner approximation that is a robust optimization problem with decision-dependent uncertainties. Suitable uncertainty sets are constructed from historical data.
Reformulation for Decision-Dependent Uncertainty: We prove that the robust approximation with box uncertainties has an equivalent algorithmically tractable reformulation with the exact linearization of bilinear products.
Computational Results: We provide a computational study on benchmark instances from the NESTA test case archive [22] as well as on a real world problem example. In both cases our approach is able to compute robust solutions with a relatively small increase in costs compared to nominal solutions.

Outline This paper is structured as follows. First, we review the current state of the art and refer to related literature in Section 2. In Section 3 we introduce the two-stage stochastic optimization problem. We start our considerations from the model of [12] and integrate curtailment options, thus obtaining a joint chance constrained DC OPF with discrete decisions. To ensure a tractable problem formulation, Section 4 focuses on a robust safe approximation of the stochastic optimization problem. Suitable uncertainty sets are constructed using the scenario approach, as well as a nearest-neighbour selection of historical data points. For the resulting robust counterpart with decision-dependent box uncertainty sets, we present an equivalent linear reformulation that is algorithmically tractable. We formulate the generalization of this reformulation for general box uncertainties in Section 4.2.1. Numerical results in Section 5 demonstrate the effectiveness of our approach. For selected benchmark instances and for the distribution network of N-ENERGIE GmbH, solutions of the robust approximation can be calculated very quick and efficiently. All solutions fulfill the chance constraints (with a high confidence) and lead to a relatively small cost increase in comparison to the nominal solution. Thus, the robust approximation in combination with the scenario approach is well suited to determine best-possible curtailment of renewables in the DC OPF model.

2 Overview of Optimization under Uncertainty for the DC OPF Problem

In this section, we will review related literature and explain how to classify this work in terms of the state of art of stochastic and robust optimization. One possibility for the stochastic treatment of uncertain parameters in optimization problems is the use of chance constraints. Here it is assumed that the parameter $\omega : \Omega \to \mathbb{R}^n$ is a continuous random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The feasibility of a variable $y \in \mathbb{R}^m$ with respect to a constraint $g : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is ensured with a given probability $1 - \varepsilon \in [0, 1]$ via

$$\mathbb{P}(g(y, \omega) \leq 0) \geq 1 - \varepsilon.$$

Despite their wide usage in modeling, the resulting models are typically hard to solve. Even the combination of linear models with probabilistic constraints, where g is linear (linear chance constraints) often leads to non-convex optimization problems. There are only a few cases

in which a tractable equivalent reformulation is possible [16] or convexity statements for the resulting problems can be made [34, 56].
In many situations, the uncertain parameter ω is contained in several constraints $g_j : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ for $j = 1, ..., J$ and $J \in \mathbb{N}$. For probabilistic constraints, it is necessary to differentiate a model containing several single constraints of the form

$$\mathbb{P}(g_j(y, \omega) \leq 0) \geq 1 - \varepsilon_j \quad \forall j \in \{1, .., J\}$$

with $1 - \varepsilon_j \in [0, 1]$ for $j = 1, .., J$ from joint chance constraints

$$\mathbb{P}\left(g_j(y, \omega) \leq 0 \quad \forall j \in \{1, ..., J\}\right) \geq 1 - \varepsilon.$$

The latter enforces simultaneous fulfillment of all constraints with a certain probability. Single chance constraints allow the separate consideration of the individual inequalities and therefore joint stochastic constraints are more restrictive and represent stronger safety conditions.

There are many successful approaches and applications for models involving (joint) chance constraints that exploits specific problem structures. For example in the cases of finite discrete probability distributions, joint chance constraints can be equivalently rewritten with deterministic mixed-integer constraints, see [24]. Based on that, the authors of [37] studied strong extended formulations and valid inequalities for mixing sets with knapsack constraints for such reformulations. For two-stage stochastic models efficient combinatorial (e.g. [33]) and decomposition (e.g. [1, 39]) algorithms exist in the literature. The mathematical properties of linear joint chance constraints with continuous distributions were examined in [44]. Despite the insightful theory and methods in the literature, in general the resulting optimization models are difficult and challenging. Unfortunately in some cases (especially with continuous multivariate distributions), for these constraints even the evaluation of a potential solution y can be numerically intractable. As in our setting, exact solution approaches are not applicable for models with an unknown probability distributions.

Due to their difficulty, a wide range of approximation techniques has been developed to tackle (joint) chance constraints, such as sampling methods like Monte Carlo methods [35] and the sample average approximation [36]. The application of the sample average approximation of chance constraints has been studied, notably in [48]. The authors analyzed the convergence properties of this approximation when the sample number goes to infinity and showed how to construct good approximate solutions. They demonstrated that sample average approximation can be used to approximate chance constraints and also derived recommendations for practical implementations. These approximations usually grow linear in size with the number of data samples. Therefore, we aim for a data-driven approximation model with constant problem size.

At this point, we want to note that if the probability distribution is unknown or is itself uncertain, models with distributionally robust stochastic constraints are also used in the literature and in practical applications. Typically, the stochastic optimization problem is protected in a worst-case sense against all uncertain distribution functions from a predefined ambiguity set. This set of probability distributions usually contains momentum information [25] or is constructed from historical data [59]. For some cases exact solution approaches for distributionally robust chance constraints are known, e.g. right-hand side or technology vector uncertainty under Wasserstein ambiguity [32]. However, an equivalent algorithmically tractable reformulation for the distributionally robust treatment of (joint) chance constraints only exists under strong modeling assumptions. In our setting, we consider joint chance constraints under decision-dependent uncertainty which we solve approximately. Therefore, the additional complexity by incorporating distributionally robust constraints is out of scope in this work. We refer the reader to [51] for further literature on distributionally robust optimization.

In contrast to stochastic optimization, no probabilistic quantities are considered in classical robust optimization problems. In its basic form, robust optimization is used for a protection in a worst-case sense against all possible realizations of the unknown parameter $\omega \in \mathbb{R}^n$ from a predefined uncertainty set $\mathcal{U} \subseteq \mathbb{R}^n$ with $n \in \mathbb{N}$. The robust feasibility of a variable $y \in \mathbb{R}^m, m \in \mathbb{N}$ with respect to constraint function $g : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is schematically given by

$$\begin{aligned} & g(y,\omega) \leq 0 \quad \forall \omega \in \mathcal{U} \\ \Leftrightarrow \quad & \max_{\omega \in \mathcal{U}} \; g(y,\omega) \leq 0. \end{aligned}$$

This is relevant for problems with a high desired level of security. In robust optimization, prior knowledge of the distribution of uncertain quantities is not necessary.

One approximation technique for chance constrained optimization problems that uses a robust treatment of uncertainty is robust approximation [46]. Hereby, (joint) chance constraints are replaced by robust constraints

$$\max_{\omega \in \mathcal{U}} g_j(y,\omega) \leq 0 \quad \forall j = 1, ..., J,$$

with the usage of a suitable uncertainty set $\mathcal{U}$. In this context, suitable means that

$$\mathbb{P}(\omega \in \mathcal{U}) \geq 1 - \varepsilon.$$

The uncertainty set are constructed such that it has a probability measure of at least of the security level of the initial chance constraint. In this case, it follows

$$\max_{\omega \in \mathcal{U}} g_j(y,\omega) \leq 0 \quad \forall j = 1, ..., J \quad \Rightarrow \quad \mathbb{P}\left(g_j(y,\omega) \leq 0 \quad \forall j \in \{1, ..., J\}\right) \geq 1 - \varepsilon. \tag{1}$$

The main advantage of doing this is the fact that by construction any robust feasible solution is also fulfilling the original probabilistic constraints. However, this is generally not true for the reverse situation. There are solutions that meet the chance constraints and are not feasible for the robust approximation. Therefore, the robust constraint is an inner approximation and may be more conservative than the probabilistic inequality but solution techniques from robust optimization can be applied. Additionally, the realizations of uncertainty for which a robust solution is guaranteed to be feasible are known beforehand as elements of the uncertainty set.

If some form of duality holds, the derivation of an equivalent algorithmically tractable robust counterpart is often possible using duality arguments, see [8] and [10]. Therefore, a feasible solution for the chance constraint model can be computed by solving the tractable reformulation of the robust approximation. Geometries commonly used for the uncertainty set are polyhedral and ellipsoidal sets. For the application of convex duality (e.g. Fenchel duality), more general convex uncertainty sets are also possible. For example, the combination of a robust linear constraint and a polyhedral uncertainty set can easily be reformulated using duality for linear optimization problems. For more details, we refer the reader to [29].

An additional question is how to construct uncertainty sets for the robust approximation in order to obtain sufficient constraints for the chance constraints. With no prior knowledge of the distribution, the uncertainty set can be estimated from data as proposed by [41]. The procedure is based on the scenario approach from stochastic programming, see [15]. Hereby, a certain number of samples is chosen randomly from the data set and the resulting uncertainty set is determined such that it contains all drawn scenarios. If the number of drawn data samples is sufficiently large, quality statements about the probability measure of the constructed set are satisfied. It is possible to ensure that the uncertainty set has the correct size with respect to the unknown probability distribution with a high confidence [17]. Likewise, the authors of [11]

provide data-driven construction schemes for uncertainty sets that imply certain probability guarantees under more specific assumptions on the distribution.

The OPF problem with uncertain parameters is a very active research area. Due to the nature of the problem, two-stage stochastic models are predominantly used to protect against uncertainties. The goal is to calculate an optimal network operating point (first-stage) that stays feasible under uncertainty with a predefined probability. The nominal operating point is a here-and-now decision before the uncertainty manifests. With chosen first-stage variables, wait-and-see decisions have to react to any potential realization of uncertainty. The network reaction to realizations of uncertain parameters is modeled in the second-stage and often contains chance constraints. Due to the non-convexity of the nominal AC OPF, only approximate solutions as in [20, 23, 52] can be achieved for a robust or stochastic treatment of uncertainties.
An essential assumption to handle optimization problems with unknown or uncertain parameters is that it is possible to solve the underlying nominal problem efficiently. This is why the consideration of DC OPF under uncertainties is suitable and of great interest. These optimization problems are usually tackled by reformulating under concrete assumptions of the underlying distribution function or using approximation techniques from stochastic programming. We refer the reader to [12, 54]. In [60], the authors considered the application of so called two-sided chance constraints to the DC OPF. This special class of chance constraints ensure the feasibility of an inequality involving absolute values with a specific probability. They assumed first and second order momentum information about the distribution and enforced the two-sided chance constraint for the worst-case distribution. This belongs to the class of distributionally robust optimization. They were able to provide an equivalent conic reformulation. Their solutions could be calculated efficiently and were fairly stable to input parameters.

Networks with a large proportion of feed-in from renewable energy sources are particularly affected by strong fluctuations. An excessively large feed-in can lead to an overload or even to a failure of the network. These critical situations must be avoided. Therefore, curtailment of renewable power is used in practice to reduce the feed-in of renewable energy sources, maintaining network stability and avoiding transmission line overloads. However, great care should also be taken to ensure that no unnecessarily large amounts of energy are lost. Otherwise, power will be wasted, potentially leading to compensation payments by the electricity network operators. This concept has also been considered in several OPF models. Examples in the literature can be found in [23, 26, 50, 53, 58]. There are two principal types of curtailment, which are usually modeled by continuous decision variables or fixed parameters. The first and more common type of curtailment uses output capacities, which limit the maximum feed-in. This limit cannot be exceeded and any potential power production above the limit is cut off. The second type of curtailment reduces feed-in by a constant. The produced energy is reduced by a fixed value regardless of how high the feed-in amount is. Chance constraints in combination with curtailment are usually tackled by sampling techniques from stochastic optimization already mentioned. To the authors' knowledge, the treatment of discrete curtailment decisions in this context has not yet been studied in the literature.

In the next section, we will introduce the stochastic DC OPF and extend it to include the possibility of discretely controlling the uncertain feed-in.

3 Power System Modeling under Uncertainty

We model the electrical power network as an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ with node set $\mathcal{N}$ and edge set $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$. In the context of power system optimization, nodes are also called buses and edges are also called (transmission) lines. Further, let $\mathcal{N}(k) \subseteq \mathcal{N}$ denote the set of adjacent nodes of $k \in \mathcal{N}$. Without loss of generality, we assume that each bus is connected to a

generation unit (generators) and is also supplied with uncertain power feed-in. Otherwise, the energy production on a bus without generators is set to zero.

3.1 DC Optimal Power Flow

In the following, we introduce the DC approximation of AC optimal power flow as it is typically studied in the literature. The aim in DC OPF is to compute the cost minimal production and transportation of power through an electrical network respecting physical, operational and technical constraints. The following Table 1 displays the notation used for variables and input parameters. The model reads

Symbol	Variable
$P_k^G \in \mathbb{R}$	generator output on bus k
$\theta_k \in \mathbb{R}$	voltage angle on bus k
$p_{kl} \in \mathbb{R}$	power flow on line (k,l)

Symbol	Parameter
$P_k^D \in \mathbb{R}$	power sink/sources on bus k
$b_{kl} > 0$	line susceptance of line (k,l)
$P_k^U \geq 0$	uncertain feed-in on bus k
$P_k^{G,\min} \in \mathbb{R}$	lower generator bounds on bus k
$P_k^{G,\max} \in \mathbb{R}$	upper generator bounds on bus k
$d_{kl}^{\max} > 0$	transmission limits of line (k,l)

Table 1: Notation for decision variables and input parameters

$$
\begin{aligned}
\min_{P^G,\theta,p} \quad & \sum_{k\in\mathcal{N}} c_k(P_k^G) \\
\text{s.t.} \quad & P_k^G + P_k^U - P_k^D = \sum_{l\in\mathcal{N}(k)} p_{kl} && \forall k \in \mathcal{N}, && \text{(2a)}\\
& p_{kl} = b_{kl}(\theta_k - \theta_l) && \forall (k,l) \in \mathcal{L}, && \text{(2b)}\\
& P_k^{G,\min} \leq P_k^G \leq P_k^{G,\max} && \forall k \in \mathcal{N}, && \text{(2c)}\\
& -d_{kl}^{\max} \leq p_{kl} \leq d_{kl}^{\max} && \forall (k,l) \in \mathcal{L}, && \text{(2d)}\\
& P^G, \theta \in \mathbb{R}^{|\mathcal{N}|}, \quad p \in \mathbb{R}^{|\mathcal{L}|}.
\end{aligned}
$$

Decision variables are generator output $P^G \in \mathbb{R}^{|\mathcal{N}|}$, voltage angles $\theta \in \mathbb{R}^{|\mathcal{N}|}$ and power flows $p \in \mathbb{R}^{|\mathcal{L}|}$. Equality constraints (2a)-(2b) model the active power flow, which is determined by Kirchhoff's first law and the power flow equations. The power on each node has to be balanced. This means that on each node the active power production $P_k^G + P_k^U \in \mathbb{R}$ from generators and renewables equals the demand $P_k^D \in \mathbb{R}$ plus the active power sent to adjacent nodes $\sum_{l\in\mathcal{N}(k)} p_{kl} \in \mathbb{R}$. The active power flow on transmission line $(k,l) \in \mathcal{L}$ is the product of voltage angle differences $\theta_k - \theta_l \in \mathbb{R}$ and susceptance $b_{kl} > 0$. The generator output P^G can be continuously controlled within the generator bounds (2c). The resulting power flows are not allowed to exceed transmission limits, as stated in (2d). Further, we assume that there is a reference node $\tilde{k} \in \mathcal{N}$ with $\theta_{\tilde{k}} = 0$.
The objective modeling power generation cost is usually a separable linear or convex quadratic function of P^G. Therefore, (2) is a linear or convex quadratic program and can be solved efficiently with standard techniques.

3.2 Modeling of Uncertainty

In practice, the power production P^U is initially unknown. In addition renewable power, for example, is subject to high fluctuations and is therefore an uncertain parameter. Based on a network operating point that is computed by ignoring uncertainties, a sudden fluctuation of renewable energy can lead to overloads in the electricity network. In the worst case, this can lead

to the failure of network elements owing to cascade effects. To prevent this, the model has to be protected against such fluctuations, and individual feed-in units may have to be regulated. At the same time, the TSO has to ensure that curtailment is as small as possible. In the following we will include these considerations in the optimization model.
As in [12], we represent the uncertain feed-in as the sum of a nominal forecast value $P^F \in \mathbb{R}^{|\mathcal{N}|}$ and a random vector $\omega : \Omega \to \mathbb{R}^{|\mathcal{N}|}$ (defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$) via

$$P_k^U = P_k^F + \omega_k \quad \forall k \in \mathcal{N}. \tag{3}$$

We need to determine a nominal operating solution P^G, θ, p that is feasible for the nominal feed-in value P^F (corresponding to $\omega = 0$). The variables have to fulfill constraints (2a)-(2d) with (3) and $\omega = 0$. We require that there is a high probability that the network reaction to fluctuating feed-in remains feasible. To model the network reaction we define duplicates $P^{G,\omega}, \theta^\omega \in \mathbb{R}^{|\mathcal{N}|}$ and $p^\omega \in \mathbb{R}^{|\mathcal{L}|}$ of decision variables that can be adjusted, depending on the realization of ω. Realizations of ω may lead to a changed distribution of energy in the network. The generators then change their output to $P^{G,\omega}$ in order to balance the total active network power. The other state variables θ^ω, p^ω are adjusted to ensure feasibility of the solution. In the setting of a two-stage stochastic optimization problem, the variables P^G, θ, p refer to first-stage (or here-and-now) decisions. They must be decided for the nominal feed-in value ($\omega = 0$), before the uncertainty is revealed. For fixed first-stage variables, any realization of the uncertain parameter ω leads to a reaction of the network by choosing optimal second-stage (or wait-and-see) variables $P^{G,\omega}, \theta^\omega, p^\omega$.
In the context of OPF, adjustments in the second stage fulfill

$$P_k^{G,\omega} + P_k^F + \omega_k - P_k^D = \sum_{l \in \mathcal{N}(k)} p_{kl}^\omega \qquad \forall k \in \mathcal{N}, \tag{4a}$$

$$p_{kl}^\omega = b_{kl}(\theta_k^\omega - \theta_l^\omega) \quad \forall (k,l) \in \mathcal{L}. \tag{4b}$$

The reaction of generator output $P^{G,\omega}$ in an electrical network is given by the Automatic Generation Control [14]. This adjustment policy defines how the generation units respond to the total power imbalance $\Omega := \sum_{k \in \mathcal{N}} \omega_k$ induced by the realization of ω. Here, we follow the modeling of [12]. To this end, the mismatch Ω is divided among all generators according to participation factors $\alpha_k \geq 0, k \in \mathcal{N}$ such that $\sum_{k \in \mathcal{N}} \alpha_k = 1$. For each individual generator at node k, the participation factor α_k determines what proportion of Ω the generator compensates. The second-stage generator output depends on the uncertainty and on the values of the first-stage variables. It reads

$$P_k^{G,\omega} = P_k^G - \alpha_k \Omega \quad \forall k \in \mathcal{N}. \tag{5}$$

In the context of two-stage optimization under uncertainty, the question of how the individual stages are coupled arises. This is often modeled by decision rules. The second-stage decisions are written as a function of the uncertainty and first-stage variables as an ansatz function. Not only can decision rules be used to model problem-specific relationships of different stages, they can also be used as modeling simplifications to solve problems that would otherwise not be algorithmically tractable. For (affine) linear decision rules, this functional relationship is linear and so leads to algorithmically tractable reformulations. Therefore, (5) can be interpreted as a special affine decision rule. The values for participation factors α can either be decided by the optimizer, or may be fixed and given by the TSO. In short-term planning, they are often fixed. If they can be decided by the optimizer, they are decided in the first-stage and integrated into the set of here-and-now decision variables. For problems with a longer time horizon (e.g. network expansion, generator design) the factors can indeed be used for optimization. In the following sections, participation factors are considered as continuous first-stage variables.

The authors of [12] showed that second-stage variables $P^{G,\omega}, \theta^\omega, p^\omega$ are uniquely determined by (4)-(5). Hereby, $\theta_{\tilde{k}}^\omega = \theta_{\tilde{k}} = 0$ is assumed for a reference node $\tilde{k} \in \mathcal{N}$. For fixed first-stage

variables, the authors proved the existence of a unique solution of equation system (4) for any realization of ω. In the following section we will extend this result so that it also can be applied to the curtailment model.

3.3 Modeling of Curtailment

So far, in (2) the full renewable power production is fed into the network. However, if the fluctuations are too large, the network stability cannot be guaranteed for a nominal feed-in that ignores uncertainties. In practice, if necessary, the feed-in is curtailed to a certain percentage of the installed capacity $P_k^{\text{inst}}, k \in \mathcal{N}$. The installed capacity is the intended full-load sustained energy production on each node. To do this, we introduce additional first-stage decision variables $\beta_k \in \mathcal{S}_k \subset [0,1]$, $k \in \mathcal{N}$, which we refer to as curtailment factors from the discrete set $\mathcal{S}_k$ of curtailment options. In practice, sets with few discrete levels are common. Sets of such discrete steps may be used are for example $\{0, 0.3, 0.6, 1\}$ or $\{0, 0.1, ..., 0.9, 1\}$ at each node. At a node $k \in \mathcal{N}$, the power fed into the network cannot exceed $\beta_k P_k^{\text{inst}}$. Any potential feed-in above this value is cut off. This changes uncertain feed-in (3) to

$$P_k^U(\beta_k, \omega_k) = \min(P_k^F + \omega_k, \beta_k P_k^{inst}) = \begin{cases} P_k^F + \omega_k & \text{if} \quad P_k^F + \omega_k \leq \beta_k P_k^{inst}, \\ \beta_k P_k^{inst} & \text{otherwise.} \end{cases}$$

To integrate curtailment into the two-staged modeling we split the curtailed feed-in into the nominal part (ω=0) and a curtailed uncertain part.

Proposition 3.1. *The curtailed uncertain feed-in $P_k^U(\beta_k, \omega_k)$ can be rewritten as*

$$P_k^U(\beta_k, \omega_k) = \underbrace{\min(P_k^F, \beta_k P_k^{inst})}_{\text{curtailed forecasted value (nominal part)}} + \underbrace{h_k(\beta_k, \omega_k)}_{\text{curtailed fluctuation (uncertainty)}},$$

where

$$h_k(\beta_k, \omega_k) := \min(P_k^F + \omega_k, \beta_k P_k^{inst}) - \min(P_k^F, \beta_k P_k^{inst}).$$

The introduced curtailment function h_k is piecewise linear and increasing in ω_k for every $k \in \mathcal{N}$. Rather like to Section 3.2, we next show that first-stage operating solution P^G, θ, p together with curtailment values $\beta_k \in \mathcal{S}_k$, $k \in \mathcal{N}$ are feasible for the nominal feed-in value $\min(P_k^F, \beta_k P_k^{inst})$ (corresponding to $\omega = 0$).
Realizations of the uncertainty ω lead to a power imbalance in the network. The total power generation mismatch with curtailed feed-in fluctuation reads $\widetilde{\Omega} := \sum_{k \in \mathcal{N}} h_k(\beta_k, \omega_k)$, which defines the second-stage generator output in the Automatic Generation Control scheme as

$$P_k^{G,\omega} = P_k^G - \alpha_k \widetilde{\Omega} \tag{6}$$

for every $k \in \mathcal{N}$ and fixed first-stage. The adjustment of voltage angles θ_k^ω and flows p_{kl}^ω are implicitly given as the solution of the power flow equations

$$P_k^{G,\omega} + \min(P_k^F, \beta_k P_k^{inst}) + h_k(\beta_k, \omega_k) - P_k^D = \sum_{l \in \mathcal{N}(k)} p_{kl}^\omega \qquad \forall k \in \mathcal{N}, \tag{7a}$$

$$p_{kl}^\omega = b_{kl}(\theta_k^\omega - \theta_l^\omega) \quad \forall (k,l) \in \mathcal{L}. \tag{7b}$$

For the reference angle node $\tilde{k} \in \mathcal{N}$, we assume that $\theta_{\tilde{k}}^\omega = \theta_{\tilde{k}} = 0$. Analogous to [12], we now show that the second stage variables are uniquely determined by the assignment of first stages and by realizations of ω.

Proposition 3.2. *Second stage variables $P^{G,\omega}, \theta^\omega, p^\omega$ are uniquely given by* (6), (7) *for fixed values of first-stage decisions $\alpha, \beta, P^G, \theta, p$ and fixed realization of ω.*

Proof. Combining the equality constraints (7a) and (7b), we obtain the system

$$B\theta^{\omega} = P^{G,\omega} + P^{U} - P^{D}, \tag{8}$$

where $B \in \mathbb{R}^{|\mathcal{N}|\times|\mathcal{N}|}$ with

$$B_{k,l} = \begin{cases} -b_{kl} & (k,l) \in \mathcal{L}, \\ \sum_{j\in\mathcal{N}(k)} b_{kj} & k = l, \\ 0 & \text{otherwise.} \end{cases}$$

Matrix B is the graph Laplacian for the graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ with edges weighted by the vector of line susceptances $b \in \mathbb{R}^{|\mathcal{L}|}$. For a broader overview of the theory of such matrices, see [45]. It is known that for connected graphs $\mathcal{G}$ we have $rank(B) = |\mathcal{N}| - 1$. As the reference node $\tilde{k}$ with $\theta_{\tilde{k}} = \tilde{\theta}_{\tilde{k}} = 0$ is given and the sum of components on the right-hand side of the system (8) equals zero, the solution θ can be given explicitly and is unique. Thus, there exists a matrix $\tilde{B} \in \mathbb{R}^{|\mathcal{N}|\times|\mathcal{N}|}$ such that

$$\theta^{\omega} = \tilde{B}(P^{G,\omega} + P^{U} - P^{D}). \tag{9}$$

The matrix $\tilde{B}$ can be calculated by striking row $\tilde{k}$ and column $\tilde{k}$ of B, inverting the resulting submatrix and inserting a row and a column with zeros at the same index $\tilde{k}$, see [12].
To prove that the right-hand sides components add up to zero, we calculate

$$\begin{aligned}
&\sum_{k\in\mathcal{N}} \left(P_k^G - \alpha_k\widetilde{\Omega} + \min(P_k^F, \beta_k P_k^{inst}) + h_k(\beta_k,\omega_k) - P_k^D\right) \\
&= \sum_{k\in\mathcal{N}} \left(P_k^G + \min(P_k^F, \beta_k P_k^{inst}) - P_k^D\right) - \widetilde{\Omega}\sum_{k\in\mathcal{N}} \alpha_k + \sum_{k\in\mathcal{N}} h_k(\beta_k,\omega_k) \\
&= \sum_{k\in\mathcal{N}} \left(P_k^G + \min(P_k^F, \beta_k P_k^{inst}) - P_k^D\right) \\
&= \sum_{k\in\mathcal{N}}\sum_{l\in\mathcal{N}(k)} p_{kl} = \sum_{k\in\mathcal{N}}\sum_{l\in\mathcal{N}(k)} b_{kl}(\theta_k - \theta_l) = 0.
\end{aligned}$$

The last equality holds as every line appears in the summation twice with a different sign and hence the summands cancel out. □

Equation (9) can be rewritten by computing

$$\begin{aligned}
\theta_l^{\omega} &= \sum_{k\in\mathcal{N}} \tilde{B}_{lk}\left(P_k^{G,\omega} + P_k^U - P_k^D\right) \\
&= \sum_{k\in\mathcal{N}} \tilde{B}_{lk}\left(P_k^G - \alpha_k\tilde{\Omega} + \min(P_k^F, \beta_k P_k^{inst}) + h_k(\beta_k,\omega_k) - P_k^D\right) \\
&= \sum_{k\in\mathcal{N}} \tilde{B}_{lk}\left(P_k^G + \min(P_k^F, \beta_k P_k^{inst}) - P_k^D\right) + \sum_{k\in\mathcal{N}} \tilde{B}_{lk}\left(-\alpha_k\tilde{\Omega} + h_k(\beta_k,\omega_k)\right) \\
&= \theta_l + \sum_{k\in\mathcal{N}} \tilde{B}_{lk}\left(h_k(\beta_k,\omega_k) - \alpha_k\tilde{\Omega}\right).
\end{aligned}$$

For fixed first-stage variables, (6) defines the reaction of $P^{G,\omega}$ to any realization of the uncertainty as a function of ω. Combining this with (9), the assignment of θ^{ω} is given. Together with equation (7b), the power flow reacting to ω can be computed. Since all second stage variables are determined by the realization of the uncertainty, $P^{G,\omega}, \theta^{\omega}, p^{\omega}$ are themselves random variables. Thus we can drop the structure of a two-stage optimization problem since no decisions are made in the second stage. Ordinarily, from an algorithmic point of view this is easier to handle than general two-stage stochastic problems as no additional optimization problems need to be solved at the second stage.

3.4 Chance Constrained DC Optimal Power Flow with Discrete Curtailment

By construction, generator output $P_k^{G,\omega}$ and flows p_{kl}^{ω} are random variables that depend on the realization of the uncertainty ω and on the values of first-stage variables. We wish the automatic generation control scheme to yield solutions that satisfy generator limits of type (2c) with a probability of at least $1-\varepsilon_1 \in [0,1]$ as well as transmission capacities (2d) with a probability of at least $1-\varepsilon_2 \in [0,1]$. As in [12, 52], the parameter $1-\varepsilon_1$ for the generator bounds will be chosen close to 1. The reason is that generator limits are usually impossible to exceed in practice and a probabilistic two-stage model should prevent these violations.
We model this requirement by two joint chance constraints in order to guarantee network stability. This means that the desired compliance probabilities for all generators and transmission limits are simultaneously met. As an alternative, one could also choose $\varepsilon_1 = \varepsilon_2$ or integrate both set of constraints in one joint chance constraint. We could also think of using individual chance constraints, as they are easier to handle from an algorithmic perspective. However, they secure each line individually against failure. In contrast, from a practical point of view, the usage of two joint chance constraints is much more appropriate, as it ensures that the entire network is protected against overloading with a certain probability and ensures with an even higher probability that the second-stage solutions do not violate generator bounds. Combining all model elements of the previous sections, we formulate the joint chance constrained DC OPF problem with discrete curtailment as

$$
\begin{aligned}
\min_{\alpha,\beta,P^G,\theta,p} \quad & \sum_{k\in\mathcal{N}} f_k(P_k^G,\beta_k) \\
\text{s.t.} \quad & P_k^G + \min(P_k^F, \beta_k P_k^{inst}) - P_k^D = \sum_{l\in\mathcal{N}(k)} p_{kl} && \forall k\in\mathcal{N}, && \text{(10a)}\\
& p_{kl} = b_{kl}(\theta_k - \theta_l) && \forall (k.l)\in\mathcal{L}, && \text{(10b)}\\
& P_k^{G,\min} \le P_k^G \le P_k^{G,\max} && \forall k\in\mathcal{N}, && \text{(10c)}\\
& p_{kl} \le d_{kl}^{\max} && \forall (k,l)\in\mathcal{L}, && \text{(10d)}\\
& \mathbb{P}\left(P_k^{G,\min} \le P_k^{G,\omega} \le P_k^{G,\max} \quad \forall k\in\mathcal{N}\right) \ge 1-\varepsilon_1, && && \text{(10e)}\\
& \mathbb{P}\left(-d_{kl}^{max} \le p_{kl}^{\omega} \le d_{kl}^{max} \quad \forall (k,l)\in\mathcal{L}\right) \ge 1-\varepsilon_2, && && \text{(10f)}\\
& \alpha_k \ge 0 \quad \forall k\in\mathcal{N}, \quad \sum_{k\in\mathcal{N}} \alpha_k = 1, \\
& \alpha, P^G, \theta \in \mathbb{R}^{|\mathcal{N}|}, \quad p\in\mathbb{R}^{|\mathcal{L}|}, \quad \beta_k \in \mathcal{S}_k \quad \forall k\in\mathcal{N},
\end{aligned}
$$

where $P_k^{G,\omega}$ and p_{kl}^{ω} are determined by (6), (7a) and (7b).
Constraints (10a)-(10d) state the feasibility of here-and-now decisions corresponding to the nominal curtailed power feed-in. The chance constraints (10e) and (10f) ensure there is a high probability of the network stability constraints are met even under fluctuation. The objective should include the original production costs $\sum_{k\in\mathcal{N}} c_k(P_k^G)$ of (2). At the same time, we aim for a minimal curtailment of discretely controllable units. This can be modeled with the minimization of $\sum_{k\in\mathcal{N}}(1-\beta_k)P_k^{inst}$. This ensures that the curtailment levels β are selected as high as possible in order to curtail as little power as possible. If it is necessary to limit the feed-in, this should be done in a cost minimal way. A weighting of these discrete steps is performed in our model with the installed capacity at the considered network node. This results in a minimal curtailment of the installed capacity.
Other choices of objective functions for optimization problem (10) can be made. A minimization of the expected value for curtailment or other risk measures usually used in stochastic optimization are possible. For example, the objective function $\sum_{k\in\mathcal{N}} \max(0, P_k^F - \beta_k P_k^{inst})$ minimizes the nominal curtailment of forecast feed-in P^F. The disadvantage of such a function is that it can lead to multiple cost optimal assignments for β_k. In this example every feasible β_k with

$\beta_k P_k^{inst} \geq P_k^F$ has the same objective value of 0 and cannot be further minimized. In order to select the largest possible value for this type of adjustment, we decided to use objective $\sum_{k\in\mathcal{N}}(1-\beta_k)P_k^{inst}$. We use a weighting parameter $\lambda \in [0,1]$ to set up a function containing both features, writing

$$f_k(P_k^G, \beta_k) := \lambda c_k(P_k^G) + (1-\lambda)(1-\beta_k)P_k^{inst}$$

for all $k \in \mathbb{R}^{|\mathcal{N}|}$.

Ordinarily, optimization problem (10) is algorithmically challenging. Even under concrete assumptions for the probability distribution this problem would usually be nonlinear and non-convex. Even testing feasibility of a solution cannot be done efficiently. The evaluation of constraint (10e) or (10f) involves the multi-dimensional integration of random vectors. As an additional difficulty, it contains decision-dependent uncertainty in from of discrete variables that impact the realizations of uncertain parameters. Therefore, approximation techniques are suitable for generating (approximate) solutions to this optimization problem.

In the next section, we will derive an algorithmically tractable reformulation for the robust approximation.

4 Robust Approximation of the Chance Constrained OPF

The optimization problem (10) introduced in Section 3.4 cannot be solved easily. Some approximation techniques for joint chance constraints exist in the literature, e.g. scenario based sampling approaches [57] or solving relaxations [4]. These methods often generate solutions that are approximately feasible for the probabilistic constraints. The methodology we chose to solve optimization problem (10) is the robust approximation of chance constraints, see [46]. The probabilistic constraints (10e) and (10f) are replaced by a worst-case protection via robust constraints using uncertainty sets $\mathcal{U}_1, \mathcal{U}_2 \subseteq \mathbb{R}^{|\mathcal{N}|}$, respectively. For each joint chance constraint, we use a separate uncertainty set because the probability levels can differ ($\varepsilon_1 < \varepsilon_2$). To guarantee the validity of the resulting approximation, we require that

$$\mathbb{P}(\omega \in \mathcal{U}_1) \geq 1-\varepsilon_1, \quad \mathbb{P}(\omega \in \mathcal{U}_2) \geq 1-\varepsilon_2. \tag{11}$$

The robust approximation of (10) is given by the following robust optimization problem

$$\begin{aligned}
\min_{P^G,\alpha,\beta,\theta,p} \quad & \sum_{k\in\mathcal{N}} f_k(P_k^G, \beta_k) \\
\text{s. t.} \quad & P_k^G + \min(P_k^F, \beta_k P_k^{inst}) - P_k^D = \sum_{l\in\mathcal{N}(k)} p_{kl} && \forall k \in \mathcal{N}, \\
& p_{kl} = b_{kl}(\theta_k - \theta_l) && \forall (k,l) \in \mathcal{L}, \\
& P_k^{G,\min} \leq P_k^G \leq P_k^{G,\max} && \forall k \in \mathcal{N}, \\
& p_{kl} \leq d_{kl}^{\max} && \forall (k,l) \in \mathcal{L}, \\
& \max_{\omega\in\mathcal{U}_1} P_k^{G,\omega} \leq P_k^{G,\max}, \quad \min_{\omega\in\mathcal{U}_1} P_k^{G,\omega} \geq P_k^{G,\min} && \forall k \in \mathcal{N}, && \text{(12a)} \\
& \max_{\omega\in\mathcal{U}_2} p_{kl}^{\omega} \leq d_{kl}^{max}, \quad \min_{\omega\in\mathcal{U}_2} p_{kl}^{\omega} \geq -d_{kl}^{max} && \forall (k,l) \in \mathcal{L}. && \text{(12b)} \\
& \alpha_k \geq 0 \quad \forall k \in \mathcal{N}, \quad \sum_{k\in\mathcal{N}} \alpha_k = 1, \\
& P^G, \theta \in \mathbb{R}^{|\mathcal{N}|}, \quad p \in \mathbb{R}^{|\mathcal{L}|}, \quad \beta_k \in \mathcal{S}_k \quad \forall k \in \mathcal{N},
\end{aligned}$$

where $P_k^{G,\omega}$ and p_{kl}^{ω} are determined by (6), (7a) and (7b).
Because of implication (1) one can verify that under assumption (11) every feasible solution of

this robust optimization problem is also feasible for (10). This is because a robust solution will stay feasible for every possible fluctuation from the constructed uncertainty sets. Therefore, the probability of fulfilling these constraints is given by the probability measure of the uncertainty sets. By construction via (11), they have a sufficient size. An additional advantage of a robust approximation is that one knows the realizations of uncertainty for which robust solutions are guaranteed to be feasible as the elements of the uncertainty set.
To note, there may be solutions that satisfy the chance constraints (10e) and (10f) that are, however not included in the feasible set of the robust approximation. For this reason, the robust approximation belongs to so-called safe (or inner) approximation techniques for solving chance constrained optimization problems. The feasible set of problem (10) is reduced, while the required probability properties are preserved. In some cases, the conservatism of the solutions increases when compared to the original probabilistic formulation. The solution quality strongly depends on the uncertainty sets $\mathcal{U}_1, \mathcal{U}_2$ satisfying (11). The larger the set, the stronger the reduction of the feasible set, which can also lead to unnecessarily conservative and costly solutions. To avoid this, attention must be paid to the construction of uncertainty sets. Appropriate choices leading to only a mild increase in conservatism will be addressed in the next section.

4.1 Construction of Uncertainty Sets from Historical Data

An obvious question that must be addressed is how to construct appropriate uncertainty sets $\mathcal{U} \subseteq \mathbb{R}^{|\mathcal{N}|}$ for a given probability level $1-\varepsilon \in [0,1]$. The uncertainty sets must be large enough to guarantee the claimed probabilistic assumption

$$\mathbb{P}(\omega \in \mathcal{U}) \geq 1-\varepsilon. \tag{13}$$

At the same time, it is necessary to ensure that the sets are not unnecessarily large. In this section, we explain two means of accomplishing this task for the case where historical data is available. Under concrete assumptions for the distribution of an uncertain vector ω, uncertainty sets can be computed with classical confidence regions or with alternative statistical methods (See, e.g., [3, 11, 43]).
We do not make specific assumptions about the probability distribution of the uncertain vector ω with realizations in $\mathbb{R}^{|\mathcal{N}|}$. In order to achieve an efficient reformulation, we fix the geometry of the sets for the robust approximation to box uncertainty sets

$$\mathcal{U} = [l, u] \subseteq \mathbb{R}^{|\mathcal{N}|},$$

where $l, u \in \mathbb{R}^{|\mathcal{N}|} \cup \{\pm\infty\}$ denote the lower and upper bounds of the box, respectively. This special geometry enables us to derive an equivalent linear reformulation for the robust approximation of the chance constrained DC OPF (10). In the next two subsections, we explain two different methodologies to construct such an uncertainty set $\mathcal{U}$ for a given probability. These construction methods can also be applied for more general convex geometries of uncertainty sets. Polyhedral or ellipsoidal sets are most common in robust optimization. A fixed geometry restricts the modeling possibilities and may lead to more conservative robust solutions as in comparison to other geometries. However in this work, we focus on box uncertainties because this geometry is suitable for the further computations in Section 4.2.

4.1.1 Scenario Approach for the Construction of Uncertainty Sets

First, we apply the idea of [41] and formulate the estimation of a convex uncertainty set as an auxiliary probabilistic optimization problem. For this problem with chance constraints we apply the scenario approach [15] from stochastic programming. The latter replaces chance constraints by finite constraints given by samples drawn randomly from the unknown distribution.

Instead of ensuring feasibility with a certain probability, the resulting scenario program enforces feasibility for all drawn realizations of uncertainty. If the number of samples is sufficiently large, quality statements can be made about the computed uncertainty set.
We want to mention that one can think of associating the scenario approach directly to the chance constrained model (10). Unfortunately, this leads to very large mixed-integer optimization problems as the size of the scenario program grows linear with number of data samples. In contrast, the reformulated robust approximation can be solved efficiently. Therefore, the association of the scenario approach to the robust approximation is preferable in our setting.

The auxiliary optimization problem in its general form consists of a chance constraint modeling the enclosure of a probability mass of at least $1-\varepsilon \in (0,1)$ via constraint (13). At the same time, this problem should aim for an uncertainty set of minimal size. The auxiliary problem is given schematically by

$$\begin{aligned} \min_{\mathcal{U}\subseteq\mathbb{R}^{|\mathcal{N}|}} \quad & size(\mathcal{U}) \\ \text{s.t.} \quad & \mathbb{P}(\omega \in \mathcal{U}) \geq 1-\varepsilon, \end{aligned}$$

where function $size$ with $size(\mathcal{U}) \in \mathbb{R}$ models the size of the uncertainty set (e.g. volume, diameter, perimeter). In order to apply the scenario approach, the uncertainty set must be parameterized in finite dimensional variables. Furthermore, the objective function and the constraints modeling "$\omega \in \mathcal{U}$" involved in the chance constraint must be convex, see Theorem 1 in [41]. The scenario approach can be used to compute more general convex uncertainty sets if the auxiliary problem fulfills these convexity assumptions and works also for possibly unbounded random variables. To estimate box uncertainties $\mathcal{U} = [l,u] \subseteq \mathbb{R}^{|\mathcal{N}|}$, we use the probabilistic optimization problem

$$\min_{l,u\in\mathbb{R}^{|\mathcal{N}|}} \quad \sum_{k\in\mathcal{N}} (u_k - l_k) \tag{15a}$$

$$\text{s.t.} \quad \mathbb{P}(l \leq \omega \leq u) \geq 1-\varepsilon. \tag{15b}$$

To control the set's size, we minimize the sum of interval lengths in every dimension. In contrast, if minimization of the box volume were used instead, this would lead to a non-convex objective. In this case, the scenario approach is no longer applicable. Although the solution of (15) does not necessarily minimize the box volume, the solution of the following scenario program does. This is why this choice of objective is suitable. We further explain this after introducing the scenario program.
Assume $N > 0$ samples are drawn randomly from an unknown probability distribution. Instead of (15b), in the scenario approach we must ensure that these N scenarios are included in the uncertainty set. The resulting scenario program for computing $\mathcal{U} = [l,u]$ is given by

$$\min_{l,u\in\mathbb{R}^{|\mathcal{N}|}} \quad \sum_{k\in\mathcal{N}} (u_k - l_k) \tag{16a}$$

$$\text{s.t.} \quad l \leq \omega^i \leq u \quad \forall i = 1,\dots,N. \tag{16b}$$

The solution of this optimization problem can be written explicitly as (l^*,u^*), where $l_k^* := \min_{i=1,\dots,N}(\omega_k^i)$ and $u_k^* := \max_{i=1,\dots,N}(\omega_k^i)$ for every vector component k. The set $\mathcal{U}^* := [l^*,u^*]$ also minimizes the volume over all sets enclosing samples $\omega^i, i = 1,...,N$. Although the solution of problem (15) generally does not calculate boxes with minimal volume, this does results for the scenario expanded problem (16).
If N is large enough, probabilistic guarantees can be made. The next Theorem adapted from [17] is valid for more general chance constrained optimization problems involving convex functions. We formulate the mathematical statement specifically applied to (15) and (16).

Theorem 4.1 (Scenario Approach, see [17]). *Let* $1-\varepsilon \in (0,1)$. *The optimal solution* (l^*, u^*) *of* (16) *fulfills the chance constraint of* (15b) *with a probability of at least* $1-\delta$ *if* N *is chosen such that*

$$\sum_{j=0}^{2|\mathcal{N}|-1} \binom{N}{j} \varepsilon^j (1-\varepsilon)^{N-j} \leq \delta.$$

The necessary number of samples N for a given confidence level $1-\delta$ is given implicitly. An explicit sufficient condition was derived in [2] and reads as

$$N \geq \left\lceil \frac{1}{\varepsilon} \frac{e}{e-1} \left(2|\mathcal{N}| - 1 + \ln \frac{1}{\delta} \right) \right\rceil. \tag{17}$$

4.1.2 Empirical Quantiles

A heuristic way to construct uncertainty sets from historical data is to use the empirical distribution. In order to do this, we select $(1-\varepsilon) \cdot 100\%$ of all samples in the multi-dimensional data set and enclose them by a smallest box $\mathcal{U} = [l, u] \subseteq \mathbb{R}^{|\mathcal{N}|}$ that contains all these samples. The corresponding samples are chosen as the nearest neighbours to the nominal scenario $0 \in \mathbb{R}^{|\mathcal{N}|}$ measured in the infinity norm $\|\cdot\|_\infty$ of $\mathbb{R}^{|\mathcal{N}|}$. Different geometries and vector norms are also applicable here. It is only necessary to make sure that the uncertainty set contains the scenarios. At the same time, the set should be constructed to be as small as possible to avoid unnecessarily conservative solutions.
Although we cannot make quality statements about the size of the uncertainty for an unknown distribution, the resulting uncertainty set has an empirical probability measure of $1-\varepsilon$ with respect to the empirical distribution function induced from the historical data. According to the law of large numbers, the empirical distribution function converges towards the actual function when the number of data points goes to infinity, and so does the probability measure of the uncertainty set computed with the nearest neighbour algorithm.

4.2 Equivalent Reformulation into Tractable Robust Counterpart

In this section, we introduce an algorithmically tractable reformulation for constraints of the form (12). To do this, we first derive an equivalent reformulation for decision-dependent box uncertainties with discrete linear decision dependencies. We show afterwards that problem (12) fits in this setting with box uncertainties and give the resulting reformulation. Since the discrete curtailment of the feed-in restricts possible realizations of uncertainty, the robust optimization problem can be rewritten with decision-dependent uncertainty sets. The box bounds can be expressed in linear dependence of the curtailment levels. In this application, curtailing feed-in in discrete steps reduces the uncertainty.

4.2.1 Reformulation for General Endogenous Box Uncertainties

Let $y \in Y \subset \mathbb{R}^{n_1}$ denote bounded continuous variables and $z \in Z \subset \mathbb{R}^{n_2}$ denote bounded discrete variables of a mixed-integer linear optimization problem with $n_1, n_2 \in \mathbb{N}$. We consider decision-dependent robust constraints as semi-infinite constraints that we subsequently rewrite as a finite maximization problem:

$$a(y,z)^\top \omega - b(y,z) \leq 0 \quad \forall \omega \in \{\omega \in \mathbb{R}^n \mid \tilde{l}(z) \leq \omega \leq \tilde{u}(z)\}$$

$$\Leftrightarrow \left\{ \begin{array}{ll} \max\limits_{\omega \in \mathbb{R}^n} & a(y,z)^\top \omega - b(y,z) \\ \text{s.t.} & \tilde{l}(z) \leq \omega \leq \tilde{u}(z) \end{array} \right\} \leq 0, \tag{18}$$

where $a, b : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^n$ are linear in both arguments and $\tilde{l}, \tilde{u} : \mathbb{R}^{n_2} \to \mathbb{R}^n$ are linear functions. Further, we assume that $\tilde{l}(z) \leq \tilde{u}(z)$ for all $z \in Z$, as otherwise the uncertainty set could be empty. For each fixed assignment of the discrete variable, this forms a classical robust constraint with a box uncertainty set. Therefore, linear duality applies to the maximization problem. Further modeling and similar reformulations of robust constraints with decision-dependent uncertainty can be found in [47] and [38]. However, it is generally not guaranteed in these publications whether the reformulations are equivalent. Therefore we clarify this issue in our setting with the next theorem and following remarks.

Theorem 4.2. *Assume that a is bounded from above, i. e. there exists an $\overline{a} \in \mathbb{R}^n$ such that $a(y, z) \leq \overline{a}$ holds component-wise for all feasible $y \in Y \subset \mathbb{R}^{n_1}$ and $z \in Z \subset \mathbb{R}^{n_2}$. Then, robust constraints of the form* (18) *are equivalent to*

$$(\tilde{u}(z) - \tilde{l}(z))^\top \zeta + a(y, z)^\top \tilde{l}(z) - b(y, z) \leq 0, \tag{19a}$$
$$\zeta \geq a(y, z), \tag{19b}$$
$$\hat{a} \geq \zeta \geq 0, \tag{19c}$$

where $\zeta \in \mathbb{R}^n$ denotes a (dual) vector and $\hat{a} := \max(\overline{a}, 0) \in \mathbb{R}^n$.

Proof. As a first step, we reformulate the maximization problem on the left-hand side of (18) by introducing shifted variables $\tilde{\omega} = \omega - \tilde{l}(z) \in \mathbb{R}^n$. We then obtain the optimization problem

$$\begin{aligned} \max_{\tilde{\omega} \in \mathbb{R}^n} \quad & a(y, z)^\top (\tilde{\omega} + \tilde{l}(z)) - b(y, z) \\ \text{s. t.} \quad & 0 \leq \tilde{\omega} \leq \tilde{u}(z) - \tilde{l}(z). \end{aligned}$$

Introducing the dual variable $\zeta \in \mathbb{R}^n$, the dual program reads as

$$\begin{aligned} \min_{\zeta \in \mathbb{R}^n} \quad & (\tilde{u}(z) - \tilde{l}(z))^\top \zeta + a(y, z)^\top \tilde{l}(z) - b(y, z) \\ \text{s. t.} \quad & \zeta \geq a(y, z), \\ & \zeta \geq 0. \end{aligned}$$

As $\tilde{u}(z) - \tilde{l}(z) \geq 0$ for all z, the optimal solution of the dual program is $\zeta^* = \max(a(y, z), 0)$. The equality holds component-wise. We use the upper bound $\overline{a}$ of $a(y, z)$ to derive an upper bound for the optimal solution ζ^*. It holds that

$$\zeta^* = \max\left(a\left(y, z\right), 0\right) \leq \max(\overline{a}, 0) = \tilde{a}$$

for all y, z. Hence, we can insert this bound in the dual optimization problem without cutting off the optimal solution. The minimum can be dropped because this does not affect the inequality of (18). □

To retrieve a tractable reformulation, constraints (19a)-(19c) need to be linearized. This can be accomplished using mixed-integer linear constraints. Due to the linearity of $\tilde{l}$ and $\tilde{u}$, the term $(\tilde{u}(z) - \tilde{l}(z))^\top \zeta$ can be rewritten as

$$(\tilde{u}(z) - \tilde{l}(z))^\top \zeta = \left(A^{\tilde{u}} - A^{\tilde{l}})^\top z + (b^{\tilde{u}} - b^{\tilde{l}})\right)^\top \zeta = \sum_{i=1}^{n_2} \sum_{j=1}^{n} K_{ij} z_i \zeta_j + \sum_{j=1}^{n} k_i \zeta_i$$

for some $A^{\tilde{u}}$, $A^{\tilde{l}} \in \mathbb{R}^{n_2 \times n}$, $b^{\tilde{u}}$, $b^{\tilde{l}} \in \mathbb{R}^n$ and some constants $K_{ij}, k_i \in \mathbb{R}$ for $i \in \{1, ..., n_2\}$ and $j \in \{1, ..., n\}$. This contains bilinear products $z_i \zeta_j$ of the discrete variable z and the continuous variable ζ. Such bilinear products can be equivalently linearized with mixed-integer linear constraints if the continuous variable is bounded. In Theorem 4.2, we are able to prove

bounds for the continuous dual variable. To carry out the linearization one substitute the discrete variable as sum of auxiliary binaries combined with special ordered sets constraints (of type one). For the resulting bilinear terms one can apply the McCormick inequalities which are known to be exact in this setting. For more details about the linearization of bilinear expressions involving discrete variables, see [13] and [55]. Due to the linearity of the bounded term $a(y,z)$, the expression $a(y,z)^\top \tilde{l}(z)$ can also be linearized. Hence, a full linearization of (19) with mixed-integer linear constraints can be carried out.
Theorem 4.2 is used in the next section to reformulate the robust optimization problem (12).

4.2.2 Reformulation for Robust DC OPF

We now show that (12) can be written as a robust optimization problem with decision-dependent uncertainties. We then use Theorem 4.2 to obtain an equivalent algorithmically tractable reformulation.
The robust constraints in (12) are given by

$$\begin{aligned}
&\max_{\omega\in\mathcal{U}_1} P_k^{G,\omega} \le P_k^{G,\max}, \quad \min_{\omega\in\mathcal{U}_1} P_k^{G,\omega} \ge P_k^{G,\min} \qquad \forall k\in\mathcal{N},\\
&\max_{\omega\in\mathcal{U}_2} p_{kl}^{\omega} \le d_{kl}^{max}, \quad \min_{\omega\in\mathcal{U}_2} p_{kl}^{\omega} \ge -d_{kl}^{max} \qquad \forall (k,l)\in\mathcal{L},
\end{aligned}$$

where $P_k^{G,\omega}$ and p_{kl}^{ω} are determined by (6), (7a) and (7b). In Section 3.3, we derived explicit formulas for $P^{G,\omega}, p^{\omega}$ given by

$$\begin{aligned}
P_k^{G,\omega} &= P_k^G - \alpha_k \sum_{k\in\mathcal{N}} h_k(\beta_k,\omega_k),\\
p_{kl}^{\omega} &= b_{kl}(\theta_k^{\omega} - \theta_l^{\omega}),\\
\theta_k^{\omega} &= \theta_k + \sum_{l\in\mathcal{N}} \tilde{B}_{kl}\left(h_l(\beta_l,\omega_l) - \alpha_l \sum_{m\in\mathcal{N}} h_m(\beta_m,\omega_m)\right).
\end{aligned}$$

To see that every robust constraint in (12) can be written in form of (18), we set $y=(P^G,\alpha,\theta)$ and $z=\beta$. Further, we define the functions $a,b:\mathbb{R}^{3|\mathcal{N}|}\times\mathbb{R}^{|\mathcal{N}|}\to\mathbb{R}^{|\mathcal{N}|}$ corresponding to Table 2 for each robust constraint, where $\mathbf{1}\in\mathbb{R}^{|\mathcal{N}|}$ denotes a vector consisting of ones, $\check{b}^{kl} = b_{kl}\left(\tilde{B}_{\cdot,k}-\tilde{B}_{\cdot,l}\right)$, $\check{B}^{kl} = b_{kl}\mathbf{1}\left(\tilde{B}_{\cdot,k}-\tilde{B}_{\cdot,l}\right)^\top$ with $\tilde{B}$ being the inverted graph Laplacian as in Section 3.3.

constraint	$a(y,z)$	$b(y,z)$
generation (12a) lower bound	$\alpha_k\mathbf{1}$	$P_k^G - P_k^{G,\min}$
generation (12a) upper bound	$-\alpha_k\mathbf{1}$	$P_k^{G,\max} - P_k^G$
transmission (12b) lower bound	$\check{B}^{kl}\alpha - \check{b}^{kl}$	$b_{kl}(\theta_k-\theta_l) - d_{kl}^{\max}$
transmission (12b) upper bound	$\check{b}^{kl} - \check{B}^{kl}\alpha$	$d_{kl}^{\max} - b_{kl}(\theta_k-\theta_l)$

Table 2: Decision-dependent constraints

Using these definitions, the robust constraints (12a) and (12b) can be written in the form of

$$\left\{\begin{array}{ll} \max\limits_{\omega\in\mathbb{R}^{|\mathcal{N}|}} & \sum\limits_{k\in\mathcal{N}} a_k(y,z)h_k(\beta_k,\omega_k) - b(y,z)\\ \text{s.t.} & l\le\omega\le u \end{array}\right\} \le 0, \tag{20}$$

To see that these constraints are decision-dependent in the sense that the uncertainty set changes with decisions made in advance as in (18), we derive an equivalent reformulation of (20).

Lemma 4.3. *Robust constraints as in* (20) *can be restated equivalently as*

$$\left\{ \begin{array}{ll} \max\limits_{\omega \in \mathbb{R}^{|\mathcal{N}|}} & a(y,z)^\top \omega - b(y,z) \\ s.\,t. & l \leq \omega \leq u \\ & \hat{l}(\beta) \leq \omega \leq \hat{u}(\beta) \end{array} \right\} \leq 0, \tag{21}$$

where $\hat{l}_k(\beta) = \min(l_k + P_k^F - \beta_k P_k^{inst}, 0)$ *and* $\hat{u}_k(\beta) = \max(\beta_k P_k^{inst} - P_k^F, 0)$ *for* $k \in \mathcal{N}$.

Proof. We need to show that for every fixed curtailment option β and for all components $k \in \mathcal{N}$, the sets

$$\begin{aligned} M_1 &:= \{h_k(\beta_k, \omega_k) \mid l_k \leq \omega_k \leq u_k\} \subseteq \mathbb{R}, \\ M_2 &:= \{\omega_k \mid \hat{l}_k(\beta) \leq \omega_k \leq \hat{u}_k(\beta) \ \wedge \ l_k \leq \omega_k \leq u_k\} \subseteq \mathbb{R} \end{aligned}$$

are equal. Since $h_k(\beta_k, \cdot)$ is monotonically increasing for all fixed β_k, it holds that

$$M_1 = \{\omega_k \mid h_k(\beta_k, l_k) \leq \omega_k \leq h_k(\beta_k, u_k)\}.$$

For the following steps, we recall that $h_k(\beta_k, \omega_k) = \min(P_k^F + \omega_k, \beta_k P_k^{inst}) - \min(P_k^F, \beta_k P_k^{inst})$.

If $\beta_k P_k^{install} - P_k^F \leq l_k \leq 0 \leq u_k$, then $M_1 = \{\omega_k \mid 0 \leq \omega_k \leq 0\} = M_2$.
If $l_k \leq \beta_k P_k^{inst} - P_k^F \leq 0 \leq 0$, then $M_1 = \{\omega \mid l_k - \beta_k P_k^{inst} + P_k^F \leq \omega_k \leq 0\} = M_2$.
If $l_k \leq 0 \leq \beta_k P_K^{inst} - P_k^F \leq u_k$, then $M_1 = \{\omega_k \mid l_k \leq \omega_k \leq \beta_k P_k^{inst} - P_k^F\} = M_2$.
If $l_k \leq 0 \leq u_k \leq \beta_k P_k^{inst} - P_k^F$, then $M_1 = \{\omega_k \mid l_k \leq \omega_k \leq u_k\} = M_2$. □

Defining $\tilde{l}(\beta) = \max(l, \hat{l}(\beta))$ and $\tilde{u}(\beta) = \min(u, \hat{u}(\beta))$ component-wise, it is apparent that (21) and (18) are equivalent. Lastly, to obtain a tractable reformulation of the robust DC OPF problem (10), Theorem (4.2) needs to be applied to (21).

The terms $\tilde{l}(\beta)$ and $\tilde{u}(\beta)$ only contain linear, minimum and maximum expressions depending on a discrete variable. Therefore, these terms can be linearized by introducing auxiliary variables and additional linear constraints. Combining duality and linearization techniques results in a mixed-integer reformulation of (12) with linear constraints. This problem formulation can ultimately be solved with a standard MIP solver, e.g. Gurobi [30]. As result of the optimization we obtain a global optimal solution to (12).

In the next section we test our model and show its practical suitability by means of numerical tests.

5 Experimental Results

For our numerical experiments we calculate robust curtailment strategies for different test examples and analyze their quality. The basis for this is real-world data from the network operator N-ERGIE GmbH (NNG) and the German Weather Service (DWD). We test our optimization models on selected benchmark instances and on an excerpt of the real distribution network of NNG. In this section we first explain the available data and how we extract information for our optimization approach. We then calculate uncertainty sets from this data. Subsequently, we consider different power grids and calculate robust solutions for them. An analysis of these results shows the practical applicability of our methods.

5.1 Generation of Uncertainty Sets

Our experimental results are based on historical solar feed-in data covering the months May to July 2015-2018 between 8 a. m. and 6 p. m., provided by the electrical distribution network operator N-ERGIE Netz GmbH, and the corresponding radiation forecast data, provided by the Deutscher Wetterdienst (DWD). The feed-in data consists of quarter-hourly averaged produced solar power (in MW) for 11 different grid nodes. All feed-in values are given relative to the installed capacity. For each feed-in value at every node, we match the corresponding radiation forecast (in $\mathrm{kJ/m^2}$). To obtain data on solar feed-in forecast values and errors for parameters in

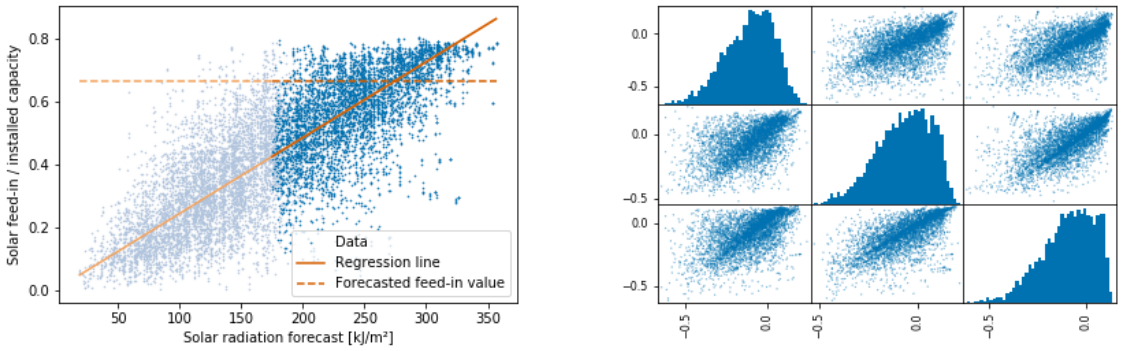

Figure 1: Obtaining feed-in forecast error data for a feed-in point with a solar radiation forecast of 275 $\mathrm{kJ/m^2}$ (left), resulting scatter plots for three nodes (right).

(3), we use a linear regression model, which we fit to the data from the year 2015 for each node independently. For a given solar radiation forecast, the regression delivers a feed-in forecast value P_k^F for each node k. To generate forecast error data, we use all data points in the years 2016 to 2018 with a maximum solar radiation forecast deviation of 100 $\mathrm{kJ/m^2}$ for every node and take the difference of the corresponding feed-in values and the solar feed-in forecast by the regression model.

Figure 1 shows an example of the procedure for obtaining solar feed-in forecast error data for one feed-in point, as well as the resulting scatter plots for three grid nodes. The graph on the left shows that solar radiation forecast of DWD and measured solar feed-in are linearly correlated. In order to simulate a sunny day in the numerical tests, we choose the solar radiation at 275 $\mathrm{kJ/m^2}$. The linear regression model yields a certain feed-in forecast P^F for this value, marked by the dotted line. To obtain samples for the fluctuation ω, we calculate the distance from measured points to P^F. This results in the middle picture of the scatter plot on the right hand side. This procedure is performed for all 11 solar feed-in nodes. From the graph on the right hand side it can also clearly be seen that our data is not normally distributed. Therefore, this assumption so frequently made in stochastic optimization cannot be used.

We calculate uncertainty sets for the robust approximation of chance constraints using the scenario approach from Section 4.1.1. This method depends on randomly drawn samples and every usage of the scenario approach can result in a slightly different uncertainty set. For our numerical results, we use the average uncertainty set resulting from applying the scenario approach 100,000 times. Thus our numerical results are reproducible because the average uncertainty set no longer changes. In addition, we determine empirical uncertainty sets respective to the empirical distribution given by the data set by finding the necessary number of nearest neighbours to the nominal scenario $0 \in \mathcal{R}^{|\mathcal{N}|}$ in the infinity norm and enclosing them with a box as described in Section 4.1.2.
Figure 2 shows the box measure of such calculated uncertainty sets for different probability levels $1 - \varepsilon$. In practice, a security probability of about 90% is usually required.The size is com-

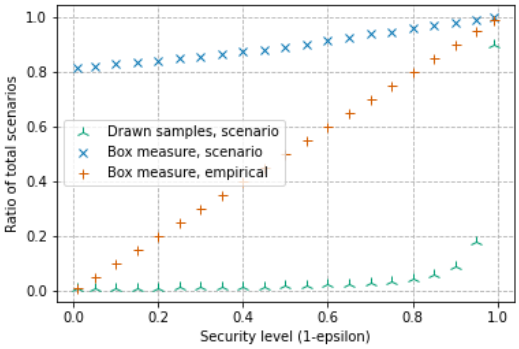

Figure 2: Size and number of drawn samples for calculated uncertainty sets with varying security levels $(1-\varepsilon)$ for the scenario approach with $\delta = 0.01$ and empirical uncertainty sets

puted by counting the number of scenarios inside the set, relative to all data points. Therefore, the size of uncertainty sets shown in Figure 2 equals the empirical probability measure. The amount of data after filtering for a solar radiation forecast of 275 kJ/m^2 as shown in Figure 1 is about 4000. In addition, the number of data points required for the confidence level of $1-\delta = 0.99$ for the scenario approach is displayed. The confidence level $1-\delta$ describes the probability that the number of samples drawn for the scenario approach is sufficient to determine a correct uncertainty set, see Theorem 4.1.
It can be seen that the scenario approach generates larger uncertainty sets than the nearest neighbour algorithm. In return, the scenario approach provides quality statements that must be paid in the form of larger sets. For small probability levels, the scenario approach is much more conservative in the determination of uncertainty sets corresponding to the resulting box size. Thus, for a required security of 20%, it generates a box measure of more than 80%. This is mainly due to the high required confidence probability of 99%. However, one can also see from this graph that the difference between the two box measures becomes smaller as the required security level increases. From a practical point of view, probability levels greater than 80% are more relevant than smaller values. According to equation (17), the number of samples that are sufficient for the scenario approach increases proportionally to $\frac{1}{\varepsilon}$. But relative to the total amount of data points, this number is comparatively small. Even for a desired protection with probability of 90%, less than 20% of all samples (about 800) have to be chosen.

5.2 Robust Curtailment on Electricity Networks

More important than the size of the computed uncertainty sets is the quality of solutions obtained by the robust approximation using these sets by solving (12). For all tests we use the solar feed-in data as previously described, which is real data obtained for a realistic network topology for which we will evaluate our methods later in this section.

In addition to the real network below, we perform numerical experiments on three instances from the NESTA test case archive [22]: *case9-wscc–api*, *case39-epri–api*, *case118-ieee–api*. We considered *Active Power Increase* (api) instances because they are constructed with binding transmission line capacities. Other instances often contain very high or no transmission limits. Since we robustify the DC OPF to avoid overload, these instances are already robust. More network manipulation would be necessary to generate additional suitable test cases. We choose these three instances in particular because they belong to different size classes and have linear generator production costs. Results for additional topologies would be comparable because for every instance we need to use the same solar feed-in data.

The solution approach was also tested on a real subnetwork of the N-ERGIE (NNG) Netz GmbH

distribution network, together with its corresponding real solar feed-in data.
The codes are written in the Python language with Gurobi [30] as a MIP solver. For all experiments, we used a 64-bit computer with Intel Corei4 CPU 3.50 GHz processor and 32 GB RAM. These results are presented in a subsequent section.

We choose $\beta_k \in \{0, 0.3, 0.6, 1\}$ as the set of curtailment options. These are the common curtailment levels used by NNG for the feed-in management of solar power. We implement these discrete variables as sum of auxiliary binaries combined with special ordered set constraints of type one. In our numerical tests we fix the probability level $1 - \epsilon_1 = 0.99$ for the probabilistic generator constraints as the violation of generator limits is usually impossible in practice and should be avoided in the model. This parameter was not set to 1 because the scenario approach only works for probability levels $1 - \varepsilon < 1$. The constraints we focus on in our numerical experiments are transmission limits (10f) affected by parameter ε_2. For the scenario approach, we again set the confidence level $1 - \delta = 0.99$. We omit tests with varying confidence level. This would lead to smaller uncertainty sets, but the quality of these sets also decreases. There would no longer be a 99% confidence probability that the calculated solutions meet the chance constraints.

To weight generator costs as approximately similar to curtailment costs, we choose a heuristic value for the cost weighting parameter λ. Full curtailment of all solar plants should be as expensive as the same amount of power produced by generators, i. e. $\lambda \sum_{k \in \mathcal{N}} c_k (P^{\text{total}}/|\mathcal{N}|) = (1 - \lambda) P^{\text{total}}$, where $P^{\text{total}} = \sum_{k \in \mathcal{N}} P_k^{\text{inst}}$. Further experiments showed that the choice of λ only affects our results for $\lambda = 0$ and $\lambda = 1$. For these values the solutions are not useful. Either generator production costs or curtailment costs are ignored. This leads to unrealistic solutions in practical terms. For the considered instances the weighting factor is always approximately $\lambda \approx 0.6$. However, the following results turn out to be independent of this parameter as long as $0 < \lambda < 1$.

The instances from the NESTA test case archive do not yet contain any solar feed-in. To examine the effectiveness of our methods, we add solar plants to some network nodes. Of course, for the NNG network it is not necessary to add additional solar feed-in because the solar data was measured in this network.
Based on our data, we scale the installed solar power such that the total solar feed-in forecast equals a certain percentage of the total demand. We use a scaling factor $\gamma > 0$ such that $\sum_{k \in \mathcal{N}} p_k^F = \gamma \sum_{k \in \mathcal{N}} p_k^D$. Computational experiments show that a good choice for the demonstration of numerical results for the NESTA instances is $\gamma = 1.5$ and for the NNG network is $\gamma = 2.5$. It turns out that smaller utilization than the chosen values (e.g. $\gamma = 1$) leads to no curtailment and no overload. Larger values for γ result in critical network situations where a complete curtailment of feed-in can be necessary. We thus do not report details about these alternate results.
We add two solar plants in the 9-bus instance and 11 solar plants in the other two benchmark instances, evenly distributed among the nodes.

To investigate the quality of solutions of (12), we calculate the objective value of the robust formulation relative to the nominal problem ($1 - \varepsilon_2 = 0$) and an empirical constraint satisfaction probability (or adherence) by Monte Carlo experiments counting samples from the data set not leading to any constraint violation. The relative cost comparison of the robust solution to the nominal one can be interpreted as the cost of robust safe approximation. Also the optimal solution of the chance constraint optimization problem would lead to more costly solutions in comparison to the nominal problem. Since we cannot calculate this solution, we compare the result of the robust safe approximation with the nominal solution.

The left hand side of Figure 3 shows the percentage cost increase relative to the nominal solution for different probability levels. On the right side, we see the calculated empirical adherence

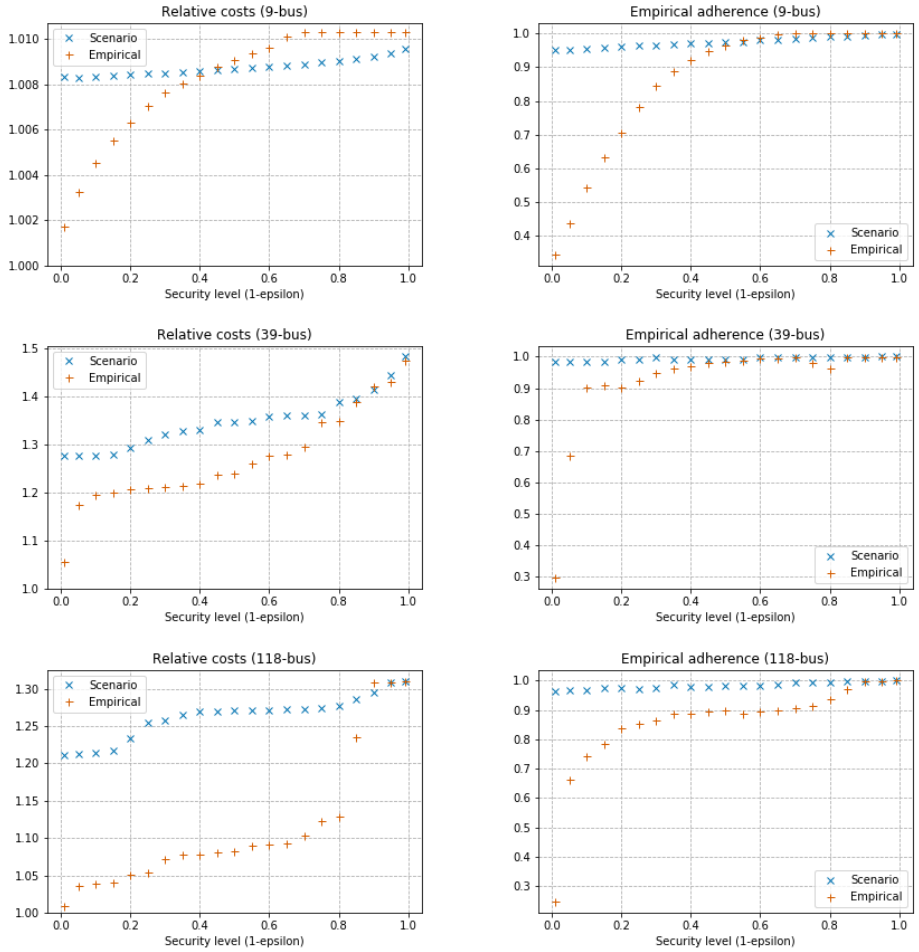

Figure 3: Objective value of the robust formulation relative to the nominal problem (left), empirical security level (right) for uncertainty sets calculated by the scenario approach and empirically for different values of ε_2 for the 9-bus, 39-bus and 188-bus instances.

probabilities of the robust solutions. These are calculated by hand by counting the number of scenarios which does not lead to an overload for the robust solution. Every row belongs to one of the three benchmark instances. Every plot displays the results for the scenario approach and for empirical uncertainty sets.

Considering the 9-bus network, one can see that the scenario approach for small security levels generates slightly more costly solutions. However, above the probability of 0.4, solutions using the empirical uncertainty set are more expensive. However, the cost increases are very small for this test example. Even in the case of a 99% protection, the solutions are only about 1% more expensive than the nominal solution. Looking at the profile of the empirical probability for transmission limit adherence, the conservatism of the scenario approach for small security levels is again striking. The uncertainty sets calculated by this approach lead to very robust solutions in comparison to the empirical set. But here again, the trend is reversed with a desired probability level greater than 0.50 that is also more realistic in practical applications. The next two rows, which consider the 39-bus and 118-bus instances, look similar. The scenario approach starts more conservatively regarding costs and empirical protection for small desired adherence probability. From the practically relevant required security levels of 0.9 and above, no significant difference in performance can be seen between the two uncertainty sets.

The graphs on the right hand side indicate that the scenario approach leads to solutions that are protected against uncertainties. This is also accompanied by the size of the uncertainty

quantity considered in Figure 2. For higher levels of security, the difference in adherence is no longer significantly large. The non-monotonic behavior of the empirical adherence probability is explained as no guarantees can be made for samples outside the uncertainty set. The cost trend is similar. For low security levels the scenario approach leads to more expensive solutions. However, the difference decreases with increasing security levels. In some cases the difference is reversed.
In conclusion, the results show that the scenario approach for the considered instances and the underlying database is a good choice. Although larger uncertainty sets are calculated on average, they are well suited for robust approximation and generate good solutions especially for practical relevant values for $1 - \varepsilon_2$. To produce these results, a mixed-integer optimization problem was solved for each network and each security level. The computing times depend strongly on the network size, but vary slightly for different security levels. The computing times are very low. Indeed, the average computation times for the individual examples in seconds are 0.03s (9-bus), 85.8s (39-bus) and 1143.7s (118-bus), with a maximal run time of 0.03s (9-bus), 140.4s (39-bus) and 2393.6s (118-bus).

Results for Historical Data and a Realistic Electricity Network The following numerical results consider the real-world distribution network of the N-ERGIE Netz GmbH. The 34-bus subnetwork of the N-ERGIE Netz GmbH slightly differs from the previous test instances. This network contains (slack-) generators on boundary nodes (denoted with + in Figure 4). On these nodes it is possible both to insert power into the network and to extract power from it. There are no costs affiliated with these interactions. Hence, there are no generator production costs and different values for the cost weighting parameter λ result in equivalent objective functions. The presented results are therefore identical for $\lambda \in [0, 1)$. For $\lambda = 0$, the results are not realistic due to a missing objective. Moreover, the participation factors are fixed values given by the TSO ($\alpha_{31} = \alpha_{34} = 0.05$, $\alpha_{32} = \alpha_{33} = 0.45$). Symbol $*$ in Figure 4 indicates the position of solar feed-in points. We vary the utilization factor γ and observe the relative curtailed power

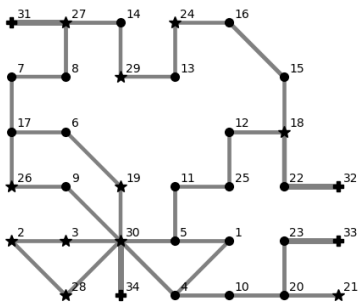

Figure 4: Sketch of N-ERGIE subnetwork

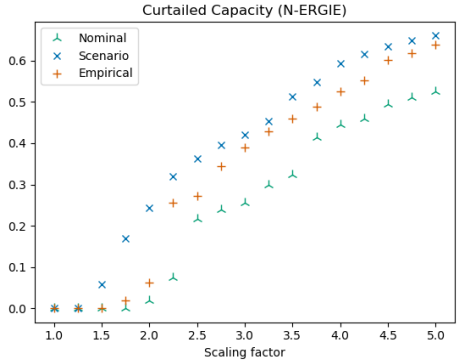

Figure 5: Varying utilization paramater

in Figure 5 for a security level of 0.9. Naturally, the amount of curtailed capacity increases in proportion to the amount of power produced. It can be observed that the quality of solutions produced using the scenario approach is similar to those obtained using empirical uncertainty sets. On average, the scenario approach results in an curtailment capacity increase of about 13% to the non-robust solution. However, the solutions generated by empirical uncertainty sets differ from this by only 0.06% curtailed capacity.

Finally, we compare as in Figure 3 the quality of solutions varying $1 - \varepsilon_2$ while $\gamma = 2.5$ via Monte Carlo simulation. In Figure 6 the same characteristics can be observed as in Figure 3. To calculate these results, again the run times are very small, on average 1.4s with a maximal run time of 3.0s. Solving each MIP took about the same amount of time.

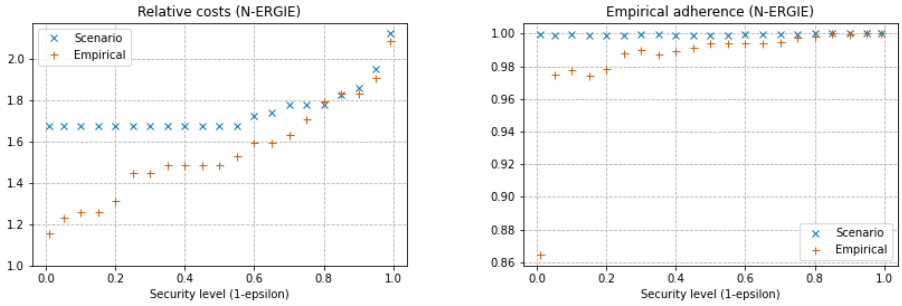

Figure 6: Objective value of the robust formulation relative to the nominal problem (left), empirical security level (right) for uncertainty sets calculated by the scenario approach and empirically for different values of ε_2 for the N-ERGIE instance.

For this real-world instance, the scenario approach is also a high-quality method to generate uncertainty sets for the robust approximation. Due to the short computing times, our methodology can also be used to optimize larger real-world networks. For our networks, the cost increase due to the conservatism of the robust approximation is relatively small for realistic values of $1 - \varepsilon_2$.

Overall System Reliability If one is interested in a robust approximation for the probabilistic DC OPF (10) wit one joint chance constraint

$$\mathbb{P}\begin{pmatrix} P_k^{G,\min} \leq P_k^{G,\omega} \leq P_k^{G,\max} & \forall k \in \mathcal{N}, \\ -d_{kl}^{max} \leq p_{kl}^{\omega} \leq d_{kl}^{max} & \forall (k,l) \in \mathcal{L} \end{pmatrix} \geq 1 - \varepsilon, \tag{22}$$

instead of (10e) and (10f) to ensure a overall system reliability of at least $1 - \varepsilon \in (0, 1)$, one has to construct only one uncertainty set $\mathcal{U} \subseteq \mathbb{R}^{|\mathcal{N}|}$ with $\mathbb{P}(\omega \in \mathcal{U}) \geq 1 - \varepsilon$. The robust approximation results in the problem (12) with $\mathcal{U} = \mathcal{U}_1 = \mathcal{U}_2$ and can again be equivalently reformulated and solved as a tractable mixed-integer linear problem with techniques presented in Section 4.2. However, the results in our numerical experiments do not change significantly for this setting. The cost of robust protection and the empirical adherence probability is exemplarily displayed in Figure 7 for the NESTA instance *case9-wscc–api* and varying probability level $1 - \varepsilon$ for the one joint probabilistic constraint. In comparison to the plots in Figure 3 about this testcase for two joint chance constraints (with $1 - \varepsilon_1 = 0.99$), one can observe very similar characteristics although the robust solutions for the problem with one chance constraint are slightly less conservative. One potential reason is that the robust solutions in the considered instances have binding transmission limits and they are more safety-critical than the generator bounds. The second-stage variables for the power production from generators violate the operational limits only in extreme cases. Therefore, the impact of integrating generator bounds into one joint chance constraint instead of using a separate joint chance constraint with high security level is small in our experiments. This shows that our robust approximation together with the uncertainty set design is also suitable for one joint chance constraint ensuring an overall system reliability.

In addition, one can also guarantee the overall system reliability of robust solutions for two separate chance constraints 10e and (10f), if $\mathcal{U}_2 \subseteq \mathcal{U}_1$ holds in (12). In all our numerical experiments with a security level of $1 - \varepsilon_1 = 0.99$ for generator bounds and $1 - \varepsilon_2 \in \{0.01, 0.05, 0.1, ..., 0.95, 0.99\}$ for transmission limits, we observed that our uncertainty sets resulting from the scenario approach (with confidence $1 - \delta = 0.99$) fulfill $\mathcal{U}_2 \subseteq \mathcal{U}_1$. We want to note, that this does not need to hold true in general for all parameters $1 - \varepsilon_1 \geq 1 - \varepsilon_2$ and other uncertainty set design strategies. However, in the case that $\mathcal{U}_2 \subseteq \mathcal{U}_1$ holds, any solution of the robust approximation (12) yields the overall system reliability (22) with confidence of at least $1 - \delta = 0.99$ in

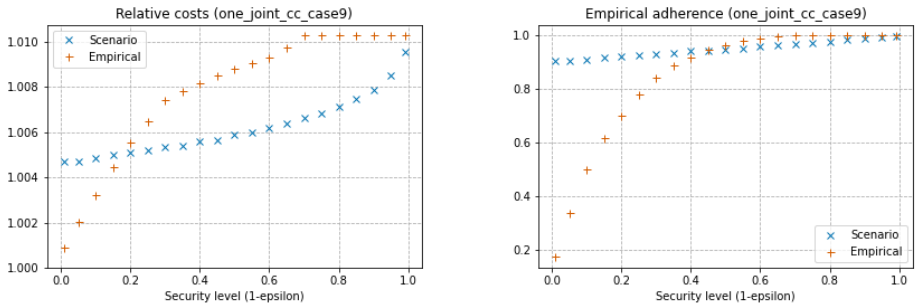

Figure 7: Optimal objective value of the robust approximation relative to the nominal problem (left), empirical security level (right) for uncertainty sets calculated by the scenario approach and empirically for different values of $1 - \varepsilon_2$ for *case9-wscc–api* with one joint chance constraint(22).

addition to feasibility for the two joint chance constraints (10e) and (10f). The reason is that the robust protection in (12) ensures feasibility of second stage variables for all operational limits simultaneously for all realizations $\omega \in \mathcal{U}_2$ if $\mathcal{U}_2 \subseteq \mathcal{U}_1$. Inequality (22) follows then from the fact that $\mathbb{P}(\omega \in \mathcal{U}_2) \geq 1 - \varepsilon_2$ with confidence probability of at least $1 - \delta = 0.99$.

6 Extension of the Proposed Methods in Applications

The reformulation proposed in this paper is independent of the underlying data. To be applied to other network situations or for e.g. day-ahead planning or network expansion planning, a TSO will need to construct a feed-in forecast. This could simply be a linear regression model to generate a feed-in forecast based on historical data, as was used here. Any further improvement on forecasting weather dependent feed-in values can be considered. Depending on the concrete application, this leads to a multitude of options, for example the use of advanced stochastic modeling or techniques from machine learning. In this work, we focused on solar radiation forecast of DWD and feed-in values to generate the solar power feed-in forecast. The integration of additional aspects in the form of additional data could give more precise predictions and will lead to new research questions. For example, temperature or air pressure measurements could be included to improve the forecast of solar feed-in. The consideration of different energy sources would also be easy to integrate.

In order to apply our method to electricity network operation, our approach could be implemented as follows. Based on the data set and a forecast value, an uncertainty set around the forecast feed-in must be constructed for a chosen probability level. In this work, we presented two possibilities for achieving this, one is represented by the scenario approach, the other by empirical quantiles using the nearest neighbour algorithm. Both compute such uncertainty sets without prior knowledge of the underlying distributions. Further knowledge or several concrete assumptions about the probability of fluctuations around the forecast value could make this task easier. A TSO could also pick concrete scenarios directly from historical data and construct problem-specific uncertainty sets.

Our reformulations enable the TSO to transform the robust problem (12) into an equivalent mixed-integer linear optimization problem that can be solved with standard state-of-the-art software.

7 Conclusion

In this paper, we extended the chance constrained DC Optimal Power Flow problem with the possibility to curtail uncertain feed-in in discrete steps. We approximated the chance constraints with robust constraints. We constructed the required uncertainty sets from histor-

ical data. After further reformulation we ended up with a robust optimization problem with decision-dependent uncertainty sets. In the case of discrete linear dependent box uncertainties we were able to provide an equivalent mixed-integer linear reformulation. Our numerical results demonstrated the applicability of our model and its reformulation.

Future research could add further features and investigate questions arising from the application, for example adding optimal transmission switching under uncertainty or including storage elements. From a mathematical point of view, it would be interesting to study different geometries for decision-dependent uncertainty sets. The major challenge is to find assumptions where an equivalent reformulation for such problems is possible.

8 Acknowledgments

We are grateful to Rainer Bäsmann for many fruitful discussions on the operation of electricity networks. We also thank "Deutscher Wetterdienst", especially Bernhard Reichert, and "N-ERGIE Netz GmbH" for providing data used in this publication. Furthermore, we want to thank the two anonymous reviewers whose comments helped to improve and clarify this manuscript.

This research has been funded by the Federal Ministry of Education and Research of Germany (grant 05M18WEB). We would like to thank the "Deutsche Forschungsgemeinschaft" (DFG) for their support within projects A05, B06, B07, and Z01 of the Sonderforschungsbereich/Transregio 154 "Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks". This research has been performed as part of the Energie Campus Nürnberg and is supported by funding of the Bavarian State Government.

References

[1] Adam, L., Branda, M., Heitsch, H., and Henrion, R. (2020). Solving joint chance constrained problems using regularization and benders' decomposition. *Annals of Operations Research*, 292:683 – 709.

[2] Alamo, T., Tempo, R., and Luque, A. (2010). On the sample complexity of randomized approaches to the analysis and design under uncertainty. In *Proceedings of the 2010 American Control Conference*, pages 4671–4676.

[3] Altman, D., Machin, D., Bryant, T., and Gardner, M. (2013). *Statistics with Confidence: Confidence Intervals and Statistical Guidelines*. Wiley.

[4] Baker, K. and Toomey, B. (2017). Efficient relaxations for joint chance constrained AC optimal power flow. *Electric Power Systems Research*, 148:230 – 236.

[5] Basciftci, B., Ahmed, S., and Gebraeel, N. (2020). Data-driven maintenance and operations scheduling in power systems under decision-dependent uncertainty. *IISE Transactions*, 52(6):589–602.

[6] Basciftci, B., Ahmed, S., and Gebraeel, N. Z. (2019). Adaptive two-stage stochastic programming with an application to capacity expansion planning. *arXiv:1906.03513*.

[7] Basciftci, B., Ahmed, S., and Shen, S. (2021). Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research*, 292(2):548–561.

[8] Ben-Tal, A., Den Hertog, D., and Vial, J.-P. (2015). Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical programming*, 149(1-2):265–299.

[9] Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust optimization*, volume 28. Princeton University Press.

[10] Ben-Tal, A. and Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88:411–424.

[11] Bertsimas, D., Gupta, V., and Kallus, N. (2013). Data-driven robust optimization. *Mathematical Programming*, 167.

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

[13] Bisschop, J. (2006). *AIMMS - Optimization Modeling*. Lulu.com.

[14] Borkowska, B. (1974). Probabilistic load flow. *IEEE Transactions on Power Apparatus and Systems*, PAS-93(3):752–759.

[15] Calafiore, G. and Campi, M. (2005). Uncertain convex programs: Randomized solutions and confidence levels. *Mathematical Programming*, 102:25–46.

[16] Calafiore, G. C. and El Ghaoui, L. (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, 130(1):1–22.

[17] Campi, M. C. and Garatti, S. (2008). The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230.

[18] Capitanescu, F., Glavic, M., Ernst, D., and Wehenkel, L. (2007). Interior-point based algorithms for the solution of optimal power flow problems. *Electric Power Systems Research*, 77(5):508 – 517.

[19] Carpentier, J. (1962). Contribution a l'etude du dispatching economique. *Bull. Soc. Francaise Electricians*, 8:431–447.

[20] Chen, J., Wu, Q., Zhang, L., and Wu, P. (2017). Multi-objective mean–variance–skewness model for nonconvex and stochastic optimal power flow considering wind power and load uncertainties. *European Journal of Operational Research*, 263(2):719–732.

[21] Christie, R. D., Wollenberg, B. F., and Wangensteen, I. (2000). Transmission management in the deregulated environment. *Proceedings of the IEEE*, 88(2):170–195.

[22] Coffrin, C., Gordon, D., and Scott, P. (2014). Nesta, the NICTA energy system test case archive. *arXiv:1411.0359*.

[23] Dall'Anese, E., Baker, K., and Summers, T. (2017). Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Transactions on Power Systems*, 32(5):3427–3438.

[24] Dentcheva, D., Prékopa, A., and Ruszczyński, A. (2000). Concavity and efficient points for discrete distributions in stochastic programming. *Mathematical Programming*, 89:55–77.

[25] Ding, K. (2014). Distributionally Robust Joint Chance Constrained Problem under Moment Uncertainty. *Journal of Applied Mathematics*, 2014(SI24):1 – 8.

[26] Ding, T., Huang, C., Bo, R., Li, R., Yang, Z., Li, F., and Sun, H. (2015). Interval arithmetic based optimal curtailment for infeasible sced considering wind power uncertainties. In *2015 IEEE Power & Energy Society General Meeting*, pages 1–5.

[27] Gabrel, V., Murat, C., and Thiele, A. (2014). Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3):471–483.

[28] Goel, V. and Grossmann, I. E. (2006). A class of stochastic programs with decision dependent uncertainty. *Mathematical programming*, 108(2):355–394.

[29] Gorissen, B. L., İhsan Yanıkoğlu, and [den Hertog], D. (2015). A practical guide to robust optimization. *Omega*, 53:124 – 137.

[30] Gurobi Optimization, L. (2020). Gurobi optimizer reference manual. http://www.gurobi.com.

[31] Hellemo, L., Barton, P., and Tomasgard, A. (2018). Decision-dependent probabilities in stochastic programs with recourse. *Computational Management Science*, 15.

[32] Ho-Nguyen, N., Kılınç-Karzan, F., Küçükyavuz, S., and Lee, D. (2020). Distributionally robust chance-constrained programs with right-hand side uncertainty under wasserstein ambiguity. *arXiv:2003.12685*.

[33] Hong, X., Lejeune, M. A., and Noyan, N. (2015). Stochastic network design for disaster preparedness. *IIE Transactions*, 47(4):329–357.

[34] Jagannathan, R. (1974). Chance-constrained programming with joint constraints. *Operations Research*, 22(2):358–372.

[35] Kalos, M. and Whitlock, P. (2008). *Monte Carlo Methods*. Wiley.

[36] Kleywegt, A. J., Shapiro, A., and Homem-de Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

[37] Küçükyavuz, S. (2012). On mixing sets arising in chance-constrained programming. *Mathematical Programming*, 132:31–56.

[38] Lappas, N. H. and Gounaris, C. E. (2018). Robust optimization for decision-making under endogenous uncertainty. *Computers & Chemical Engineering*, 111:252–266.

[39] Liu, X., Küçükyavuz, S., and Luedtke, J. (2016). Decomposition algorithms for two-stage chance-constrained programs. *Mathematical Programming*, 157:219–243.

[40] Low, S. H. (2014). Convex relaxation of optimal power flow—part i: Formulations and equivalence. *IEEE Transactions on Control of Network Systems*, 1(1):15–27.

[41] Margellos, K., Goulart, P., and Lygeros, J. (2014). On the road between robust optimization and the scenario approach for chance constrained optimization problems. *IEEE Transactions on Automatic Control*, 59(8):2258–2263.

[42] Mary, A., Cain, B., and O'Neill, R. (2012). History of optimal power flow and formulations. *Fed. Energy Regul. Comm.*, 1:1–36.

[43] Mathew, T. and Kasala, S. (1994). An exact confidence region in multivariate calibration. *Ann. Statist.*, 22(1):94–105.

[44] Miller, B. L. and Wagner, H. M. (1965). Chance constrained programming with joint constraints. *Operations Research*, 13(6):930–945.

[45] Molitierno, J. (2016). *Applications of Combinatorial Matrix Theory to Laplacian Matrices of Graphs*. Discrete Mathematics and Its Applications. CRC Press.

[46] Nemirovski, A. (2012). On safe tractable approximations of chance constraints. *European Journal of Operational Research*, 219(3):707 – 718. Feature Clusters.

[47] Nohadani, O. and Sharma, K. (2018). Optimization under decision-dependent uncertainty. *SIAM Journal on Optimization*, 28(2):1773–1795.

[48] Pagnoncelli, B. K., Ahmed, S., and Shapiro, A. (2009). Sample average approximation method for chance constrained programming: theory and applications. *Journal of optimization theory and applications*, 142(2):399–416.

[49] Prékopa, A. (1995). *Stochastic Programming*. Mathematics and Its Applications. Springer Netherlands.

[50] Qiu, F. and Wang, J. (2014). Chance-constrained transmission switching with guaranteed wind power utilization. *IEEE Transactions on Power Systems*, 30(3):1270–1278.

[51] Rahimian, H. and Mehrotra, S. (2019). Distributionally robust optimization: A review. *arXiv:1908.05659*.

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

[53] Roald, L., Misra, S., Chertkov, M., Backhaus, S., and Andersson, G. (2016). Chance constrained optimal power flow with curtailment and reserves from wind power plants. *arXiv:1601.04321*.

[54] Roald, L., Oldewurtel, F., Van Parys, B., and Andersson, G. (2015). Security constrained optimal power flow with distributionally robust chance constraints. *arXiv:1508.06061*.

[55] Sherali, H. D. and Adams, W. P. (2013). *Reformulation–Linearization Techniques for Discrete Optimization Problems*, pages 2849–2896. Springer New York, New York, NY.

[56] van Ackooij, W. (2015). Eventual convexity of chance constrained feasible sets. *Optimization*, 64(5):1263–1284.

[57] Vrakopoulou, M., Margellos, K., Lygeros, J., and Andersson, G. (2013). Probabilistic guarantees for the N-1 security of systems with wind power generation. In *Reliability and risk evaluation of wind integrated power systems*, pages 59–73. Springer.

[58] Wang, Q., Guan, Y., and Wang, J. (2011). A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE Transactions on Power Systems*, 27(1):206–215.

[59] Xie, W. (2020). On distributionally robust chance constrained programs with wasserstein distance. *arXiv:1806.07418*.

[60] Xie, W. and Ahmed, S. (2018). Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. *IEEE Transactions on Power Systems*, 33(2):1860–1867.

[61] Zohrizadeh, F., Josz, C., Jin, M., Madani, R., Lavaei, J., and Sojoudi, S. (2020). A survey on conic relaxations of optimal power flow problem. *European Journal of Operational Research*, 287(2):391–409.