Power Systems Optimization under Uncertainty: A Review of Methods and Applications

Line A. Roald, David Pozo, Anthony Papavasiliou, Daniel K. Molzahn, Jalal Kazempour, Antonio Conejo

***Abstract*—Electric power systems and the companies and customers that interact with them are experiencing increasing levels of uncertainty due to factors such as renewable energy generation, market liberalization, and climate change. This raises the important question of how to make optimal decisions under uncertainty. This paper aims to provide an overview of existing methods for modeling and optimization of problems affected by uncertainty, targeted at researchers with a familiarity with power systems and optimization. We also review some important applications of optimization under uncertainty in power systems and provide an outlook to future directions of research.**

***Index Terms*—Stochastic optimization**

I. Introduction

In electric power systems, optimization is used for a multitude of tasks, ranging from real-time operation to long-term planning. To make optimal decisions, system operators, generation companies, and consumers rely on a variety of input data for determining parameters in the formulation of a mathematical optimization model that supports their decision-making. Examples of such parameters include forecasts of load and renewable energy, knowledge about future electricity prices, and long-term climate change trends. Unfortunately, many of these parameters are uncertain. For example, forecasting of load and renewable energy generation is impacted by weather forecast uncertainty, and electricity prices are affected by both variations in load and renewable energy generation as well as by the actions of other participants in the electricity market. Future climate change trends are hard to predict because we lack knowledge of how emissions will evolve and how these translate into impacts on the grid.

Although we do not yet know the exact values of these uncertain parameters, we still have to make decisions *now*. For example, utilities and system operators have to decide which generation units to commit during day-ahead planning, before the exact supply of renewable energy is known. The question of how to make optimal decisions in the presence of uncertainty gives rise to the field of *optimization under uncertainty*, which includes stochastic optimization, chance-constrained optimization, robust optimization, and distributionally robust optimization.

In electric power systems, the interest in optimization under uncertainty accelerated rapidly over the past two decades due to the advent of large-scale renewable generation. The presence of renewable energy has drastically increased uncertainty in power systems, with significant impacts on power system operations, electricity markets, and long-term planning. At the same time, methods for optimization under uncertainty have also undergone rapid developments. Importantly, improvements in modeling and solution algorithms have made optimization under uncertainty easier to apply and better suited for scaling to realistic system sizes.

This paper provides an introduction to modeling and optimization under uncertainty for an electric power systems audience along with a review of certain state-of-the-art examples of recent power systems applications. We also discuss remaining challenges and important future directions to motivate research in these areas. One particular group we hope may benefit from this paper is that of graduate students whose research focuses on either electric power systems or optimization and who have some level of familiarity with both areas.

The goal of the paper is not to provide a tutorial or in-depth explanation of any particular method. Instead, we discuss key characteristics as well as advantages and disadvantages of a range of different optimization methods and modeling techniques. Our aim is to introduce our readers to different ways in which we can model and solve an optimization problem under uncertainty, along with sufficient information for judging which approach may be most suitable for their setting. This exposition is supported by references for detailed information on specific methods.

The remainder of the paper is organized in three parts. First, we provide a brief overview of optimization under uncertainty, including a generic problem setup and a discussion of uncertainty modeling in Section II, an overview of different ways to formulate optimization problems under uncertainty in Section III, and discussions regarding tractable reformulations and solution algorithms in Section IV as well as methods for evaluating solution quality in Section V. Second, we review several important and emerging optimization problems under uncertainty in power systems operation. Applications include

Line A. Roald is with the University of Wisconsin - Madison in Madison, Wisconsin, USA (Email: roald@wisc.edu). Her work on this paper is supported by the U.S. National Science Foundation Energy, Power, Control and Networks (EPCN) under award #2045860.
David Pozo is with Skoltech in Moscow, Russia.
Anthony Papavasiliou is with the National Technical University Athens in Athens, Greece.
Daniel K. Molzahn is with Georgia Institute of Technology in Atlanta, Georgia, USA. His work on this paper is supported by the U.S. National Science Foundation Energy, Power, Control and Networks (EPCN) under award #2145564.
Jalal Kazempour is with the Technical University of Denmark in Kgs. Lyngby, Denmark.
Antonio Conejo is with the Ohio State University in Columbus, Ohio, USA.

security-constrained optimal power flow in Section VI, chance-constrained optimal power flow in Section VII, stochastic unit commitment in Section IX, a discussion on power flow modeling in Section XI, stochastic dual dynamic programming in Section XII, consumer energy procurement in Section XIII, and transmission expansion planning in Section X. Finally, we provide a discussion and outlook on existing challenges and promising directions in Section XIV.

Part I: Overview of Optimization under Uncertainty

II. Modeling Considerations

A. Notation

In the following, we provide a general overview and some generic examples of our notation (which is defined in more detail below). We use **bold fonts** to denote vectors of decision variables $\mathbf{x}$, $\mathbf{y}$, parameters $\boldsymbol{a}$, , and functions $\boldsymbol{g}(\cdot)$, $\boldsymbol{h}(\cdot)$. We use normal fonts and subscripts x_i, $h_j(\,)$ to refer to individual entries of these vectors. Scalar functions $f(\,)$ and variables q are denoted with normal fonts.

The vector of uncertain parameters in our problem is denoted by . This parameter may be described as a random variable, which is to be distinguished from a decision variable. We will refer to a specific realization of the uncertain parameter as a *scenario* s and a set of many scenarios as a *sample* S. The parameter values and decision variables associated with scenario s are denoted by bold subscript s, e.g., $_s$ and $\mathbf{y}_s$. We also use a number of common mathematical symbols and operators, such as the expectation of random variables.

B. Problem Set-Up

To set up our subsequent discussions, we start from a general formulation of a two-stage optimization problem with uncertain parameters . Two-stage optimization problems can be used to represent many problems that are relevant to power systems, with examples including security-constrained optimal power flow (the first stage optimizes operations prior to any contingency and the second stage represents post-contingency operations) and reserve dimensioning (the first stage decisions regard the amount of reserves to procure, while the second stage determines how the reserves are activated to balance the system). Such models can be formulated as follows:

$$\min_{\mathbf{x},\mathbf{y}} \; f^{\mathrm{F}}(\mathbf{x}) + R_{\mathrm{cost}}\left[f^{\mathrm{S}}(\mathbf{x};\mathbf{y}\;;\;)\right] \tag{1a}$$

s.t.

$$\boldsymbol{h}^{\mathrm{F}}(\mathbf{x}) = 0;\quad \boldsymbol{g}^{\mathrm{F}}(\mathbf{x}) \quad 0; \tag{1b}$$

$$R_{\mathrm{viol}}\begin{bmatrix}\boldsymbol{h}^{\mathrm{S}}(\mathbf{x};\mathbf{y}\;;\;) = 0\\ \boldsymbol{g}^{\mathrm{S}}(\mathbf{x};\mathbf{y}\;;\;) \quad 0\end{bmatrix}: \tag{1c}$$

The challenge of this problem, whose parameters and constraints are defined in more detail below, is that we do not know the value of the uncertain parameters . To obtain a well-defined problem, we introduce risk operators $R_{\mathrm{cost}}[\,]$ and $R_{\mathrm{viol}}[\,]$ that quantify the risk associated with different realizations of . The risk operator $R_{\mathrm{cost}}[\,]$, represents the risk of excessive costs, while the risk operator $R_{\mathrm{viol}}[\,]$ measures the risk of constraint violations.

In this problem, the first-stage decisions represented by decision variables $\mathbf{x}$, also referred to as *here-and-now* decisions, have to be decided *before* the value of is known. The second-stage decision variables $\mathbf{y}$, also referred to as *wait-and-see* decisions, are taken *in response to* the realization of the uncertain parameters . We use subscript $\mathbf{y}$ to highlight the fact that these decisions are taken after the values of the uncertain parameters are known. The first- and second-stage decisions $\mathbf{x}$ and $\mathbf{y}$ must be chosen such that they minimize the sum of the first-stage cost $f^{\mathrm{F}}(\mathbf{x})$ and the second-stage cost $R_{\mathrm{cost}}\left[f^{\mathrm{S}}(\mathbf{x},\mathbf{y}\;,\;)\right]$ and satisfy both the first-stage constraints (1b) and the second-stage constraints (1c). In the second-stage constraint functions $\boldsymbol{h}^{\mathrm{S}}$, $\boldsymbol{g}^{\mathrm{S}}$, we include the uncertain parameters as an input argument to explicitly highlight that the function values depend on the realization of , as is common in parts of the stochastic programming literature [1], [2]. However, it is worth emphasizing that is a parameter of the function, not a decision variable.

To obtain well-defined, tractable formulations for (1), we have to define and evaluate the risk operators $R_{\mathrm{cost}}[\,]$ and $R_{\mathrm{viol}}[\,]$, giving rise to a range of different formulations including stochastic optimization, distributionally robust optimization, robust optimization, and chance-constrained optimization. The choice of formulation depends on what knowledge is available about the uncertain parameters at the time of decision-making and how we would model the impacts of uncertain parameters on our problem. Specific problem formulations and associated solution algorithms will be described in much more detail below.

We note that the above formulation is only one example of a problem with uncertain parameters, and there exist others. Another important class of problems is *multi-stage* stochastic problems, where uncertainty is revealed in multiple steps, such as throughout several hours of the day or across several years. In this situation, we may have the ability to update our wait-and-see decisions $\mathbf{y}$ each time new information about is revealed, while still accounting for the fact that we do not know what future realizations will look like. Important examples of multi-stage stochastic programming problems include unit commitment (where the first-stage variables are the on/off statuses of generators and the generation dispatch is adapted in several stages throughout the day) or long-term planning problems (where decisions on which transmission infrastructure to build need to account for a longer time horizon where changes to the plans can be made at several points in time). Multi-stage stochastic optimization problems pose significant additional challenges relative to the two-stage formulation (1). For the simplicity of discussion, the first part of this paper will focus primarily on two-stage optimization, but will also provide some discussion and references for further reading on multi-stage problems where applicable.

C. Common Uncertainty Sources

There are many different sources of uncertainty in power system optimization problems that can be captured by the uncertain parameters . Some examples include:

Renewable energy generation, where can represent either the amount of power provided by solar or wind generators, a deviation from a forecasted value, or an uncertain upper bound on the total generation available.
Component outages, i.e., whether or not a transmission line or a generator is experiencing a failure, where can represent whether or not a certain outage has occurred.
Price of electricity, where can represent the price a consumer has to pay or a producer receives for their electricity.
Precipitation, where may represent the water inflow to a hydropower reservoir.
Ambient conditions, where represents, e.g., uncertainty regarding the future temperature that may impact the total load or generation capacities.
Occurrence of extreme weather, where may represent the path of a hurricane or the location of a wildfire ignition.

An important distinction between different uncertainty sources is whether the uncertainty is continuous or discrete. Uncertain parameters that can take on any value in a range (e.g., renewable generation, load demands, and prices) are well represented as *continuous random variables* with a continuous probability distribution, typically centered around a forecast value. These variables can take on an infinite number of possible values, which is referred to as *infinite support*. For instance, normally distributed random variables belong to this group. We note that such continuous random variables can still have *bounded support*, i.e., they may be limited to a finite range. A uniformly distributed random variable has infinite, but bounded support. Uncertain parameters that represent discrete events (e.g., component outages and wildfire ignitions) are represented as *discrete random variables* with an associated discrete probability distribution. If there is a finite number of possible realizations of those variables, we refer to them as having *finite support*.

Furthermore, it is worth noting that the probability distribution and the best way to represent a source of uncertainty can change depending on the considered time scale. For example, short-term variations in renewable generation may be well represented by a Gaussian distribution, while a long-term distribution for a renewable energy plant typically follows a Weibull distribution. In long-term planning, it may be necessary to focus more on extreme scenarios to assess system adequacy rather than the full distribution. In addition to varying with the considered time scale, the characterization of uncertainty can also depend on who is the decision maker. For example, one generation company does not know the bids of other generation companies and perceive those as being uncertain, but all bids are known and deterministic for the system operator.

D. Uncertainty Propagation

The impacts of variations in the uncertain parameters will propagate through our model and impact the second-stage cost $R_{\text{cost}}\left[f^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ ,\ )\right]$ as well as constraints (1c). For example, generators that participate in balancing the system as the renewable generation fluctuates may need to produce more or less power, thus impacting the system cost. Further, fluctuations in renewable generation will cause power flows and voltage magnitudes to change as well, leading to possible constraint violations. Modeling these effects can be very challenging as it requires the simultaneous consideration of multiple uncertainty sources and how they combine.

It is thus helpful to make a distinction between *input uncertainty*, i.e., the uncertain parameters that are the source of variability represented by such as renewable energy generation, and the *output uncertainty*, i.e., the quantities represented by $f^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ ,\ )$, $h^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ ,\ )$, and $g^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ ,\ )$ such as power generation, power flows, etc. Often times, it is comparatively simpler to quantify or forecast the input uncertainty , e.g., through historical data or probabilistic forecasting methods. However, even if we have a perfect estimate of the uncertainty from each source, it can be challenging to determine how the uncertainty from many different sources combines and propagates. This is true even for linear models, and uncertainty propagation is even more challenging for models that incorporate non-linear equality constraints, such as the alternating current (AC) power flow equations. When choosing an uncertainty representation, we need to consider methods that allow us to accurately represent both the input uncertainty and the output uncertainty $f^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ ,\ )$, $g^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ ,\ )$, and $h^{\mathsf{S}}(\mathbf{x},\mathbf{y}\ \ )$.

E. Uncertainty Representation

The choice of uncertainty representation is often constrained by practical limitations on access to uncertainty data. When determining what kind of representation to use, we need to determine 1) what we know about the probability distribution of the uncertain parameters , 2) who has (or can get) access to uncertainty data, and 3) how much data we can expect to access (e.g., will there be limited or unlimited access to scenarios?). Depending on the answers to these questions (as well as questions related to the model formulation and solution method which will be discussed in Sections III and IV), there are different ways in which we can represent the uncertain parameters in an optimization problem. Fig. 1 illustrates a two-dimensional uncertainty distribution and some common uncertainty representations, which are further discussed below.

Certainty Equivalent: While it may seem trivial, a common approach to dealing with uncertainty is to simply replace the uncertain parameters by a single best guess. This gives rise to the so-called *certainty equivalent problem*. A common choice is to replace the uncertain parameters by their expected values (also referred to as the mean) $= \mathsf{E}_P[\ ]$, where $\mathsf{E}_P[\ ]$ is the expectation operator with respect to the probability distribution P of the uncertain parameters . Alternatively, it is possible to consider the *mode* of the distribution m, which represents the most likely realization (note that m for general distributions). The expected value and mode of the distribution is shown in red and orange in Fig. 1 (a).

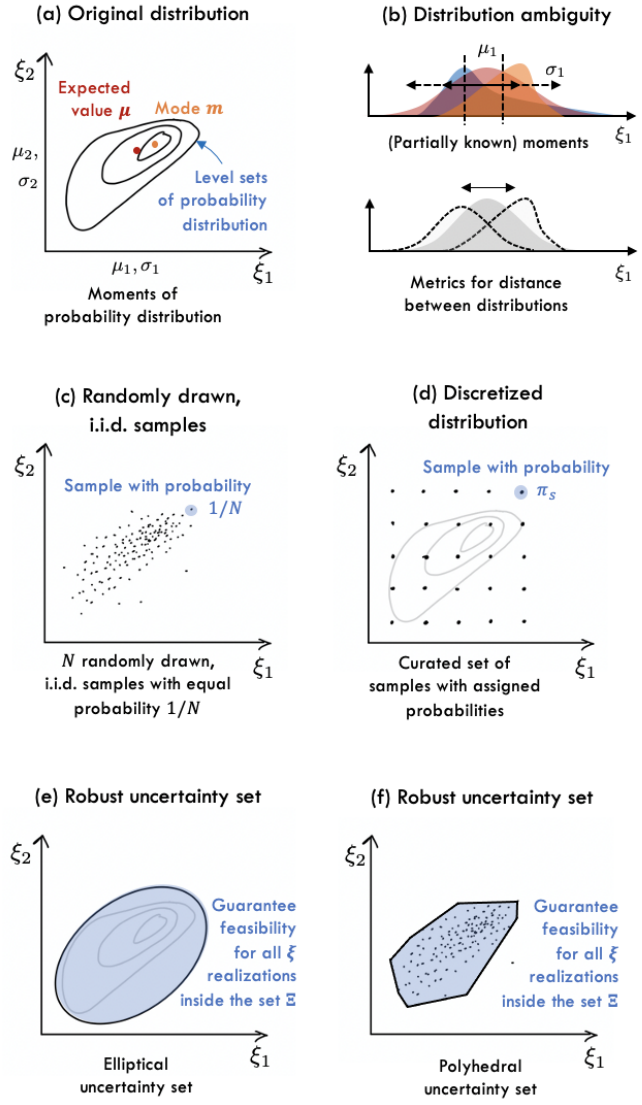

Figure 1: Examples of uncertainty representations for a two-dimensional vector $= f_1; _2g$ of uncertain parameters. (a) Original uncertainty distribution, with the mean $= f_1; _2g$, mode $\boldsymbol{m} = fm_1; m_2g$, and level sets marked. (b) Illustration of distributional robustness for the scalar random variable $_1$, including (partial) knowledge of the first and second moments (top) and metrics that limit the distance from a known distribution (bottom). (c) Example of randomly drawn, *i.i.d.* samples where each sample has probability $1=N$. (d) Example probability distribution represented through samples on a grid, each with an assigned probability $_s$. (e) Example of an elliptical robust uncertainty set. (f) Example of a polyhedral uncertainty set.

In many practical applications, solving the certainty equivalent problem may yield good results and is often significantly easier than formulating and solving an optimization problem under uncertainty. This case can therefore provide an important benchmark to determine whether using a more complex representation is worthwhile.

Perfect Information Model: Another option is to solve a *perfect information model*, which assumes that we are able to perfectly forecast . This model replaces the uncertain parameters by their actual realizations, thus providing the best possible solution we can get. In general, this is not a realistic model (in particular, if we were able to perfectly forecast , we would no longer need to consider (1) to be an optimization problem under uncertainty), but the perfect information model can still provide a helpful benchmark for our problem.

Probability Distribution: In certain special cases, it is possible to work directly with the probability distribution of . For example, assume that follows a multivariate normal distribution and appears linearly in our problem. In this case, all other quantities can be interpreted as a weighted linear combinations of , and will also follow a normal distribution. Another example is if has finite support (i.e., a discrete distribution with a finite number of possible outcomes) and the probability associated with each outcome is known, in which case we can investigate each realization separately.

Distributionally Robust Representation: We frequently may not have access to the *true* probability distribution P of our uncertain parameters . This may be because the distribution is unknown, or because it is numerically intractable to work directly with the distribution. Instead, we may identify a *family* of distributions that the uncertain parameters may belong to, with the size of the family representing the underlying *ambiguity* of the uncertain parameters. We typically refer to this family of distributions as the *ambiguity set* A. A distributionally robust representation of the uncertain parameters considers all distributions $P \; 2 \; A$ as possible distributions. There are two main types of ambiguity sets that have been widely used in the literature, namely *moment-based* and *metric-based* ambiguity sets, as illustrated in Fig. 1 (b). Both sets can be built upon empirical (historical) data; thus, a distributionally robust representation of the uncertainty might be seen as a *data-driven* model.

The moment-based approaches estimate the moment information, e.g., mean (the first-order moment) and covariance (the second-order moment), from the empirical data and build an ambiguity set containing all distributions with the same moments [3], [4]. One can enlarge the resulting ambiguity set (to be more ambiguity averse) by allowing the values of moments to be inexact, i.e., by adding more distributions whose moments are close but not necessarily identical to those estimated from the empirical data [5]. In contrast, one can shrink the ambiguity set (to be less ambiguity averse) by limiting the distributions that are included in the set, e.g., by adding unimodality constraints that exclude all distributions with more than one spike [6]–[8].

The metric-based approaches exploit probabilistic distance metrics (mainly Wasserstein distance [9], which measures the distance of two given distributions) and include any distribution in the ambiguity set whose distance to the empirical one is less than or equal to a given positive value [10]. This value allows the user to adjust its ambiguity aversion, such that assigning a higher value results in a larger ambiguity set and therefore a more conservative solution, and vice versa.

Independent Identically Distributed Samples: In many situations, we may have access to scenarios that represent possible realizations (arising from, e.g., historical data or probabilistic forecasts). If we assume that these samples $_s$ are *independent and identically distributed (i.i.d.)*, each scenario $_s$ represents one realization of the uncertain parameters, and all scenarios have the same probability of occurrence $1/N$. If there is a higher probability that the uncertain parameter takes on values in a certain range, this will be reflected via more realizations $_s$ in that range, as shown in Fig. 1 (c). Thus, a benefit of using *i.i.d.* samples is that we do not need to explicitly model the distribution or make (possibly restrictive) choices in how to discretize the distribution and assign values to individual realizations. The use of *i.i.d.* samples is a purely data-driven method, though it requires the access to a (possibly large) set of representative scenarios.

A drawback of using *i.i.d.* scenarios is that an accurate representation of the probability distribution may require consideration of a large number of samples N. Furthermore, *i.i.d.* data is obtained using random sampling, and if we draw a new sample set S, we may obtain a different solution to our optimization problem. As a result, not only the input data, but also the solution is random. Finally, it can be hard to obtain samples that are truly *i.i.d.*. For example, historical data obtained from a time series is typically not *i.i.d.* as the uncertainty realization in one time step is often not independent of the uncertainty realization in the next time step.

Discretized Probability Distribution: Another way to represent a probability distribution is to discretize it into a set of samples $_s$ with *explicitly assigned* probabilities $_s$. Fig. 1 (d) shows an example of a distribution that has been discretized with values on an equidistant grid[1]. The key difference between a discretized distribution and using *i.i.d.* samples is that the scenarios and associated probabilities in the discretized case are designed to reflect the underlying distribution as closely as possible. This can reduce the number of samples compared with the *i.i.d.* case, but the number of required scenarios may still grow quickly with the dimension of the vector . Furthermore, the accuracy of this approach depends on our ability to create a discretization with scenarios and probabilities that accurately represents the underlying distribution.

Robust Uncertainty Set: In certain situations, we may not have access to sufficient data in order to estimate the distribution of or may simply desire to ensure that our solutions are safe across a range of possible realizations. In this case, we define as any realization within a given *uncertainty set* . The uncertainty set can be represented using a set of scenarios $_s$, which may include all possible realizations if has finite support or be a set of scenarios that approximates a continuous distribution. Another option for representing the uncertainty set is to allow continuous variations of but constrain these variations to lie within a predetermined set. Common representations of uncertainty sets include elliptical sets [11] (illustrated in Fig. 1 (e)), box-constrained sets, budgeted uncertainty sets [12], and more general polyhedral sets, typically generated using data-driven methods [13]. In some cases, the uncertainty set may be chosen to truly represent all possible realizations of , such as if the probability distribution has bounded support (i.e., can only vary within finite bounds). In other cases, the uncertainty set may be defined to contain a certain probability mass or (in practical settings) to simply contain all realizations of which we would like to safeguard our solution against.

An important aspect of using a robust uncertainty representation is that we disregard variations in the probability of occurrence for different realizations of within the uncertainty set. It is therefore worth noting that a larger uncertainty set which contains more realizations of leads to a safer, but also more costly solution.

Although there are important distinctions between these different ways of representing uncertainty, they are also closely related. For example, if we know the probability distribution of the underlying uncertainty, we may choose to generate *i.i.d.* scenarios from this distribution instead of directly representing the distribution in our problem. Alternatively, if we have access to many scenarios, we can estimate the probability distribution from this data or use this data to define an uncertainty set. Furthermore, distributionally robust methods leverage combinations of probability distributions, scenarios, and uncertainty sets. It is also worth noting that obtaining or generating realistic scenarios for the uncertain parameters can be a significant challenge. This is particularly true if the entries of are not independent random variables, but have complicated dependencies that are not well described using linear correlation (and may require, e.g., the use of copulas [14]. For instance, if the uncertain random variables are time series, multivariate time series models can be used to generate samples that are correlated in time and space.

III. Optimization Under Uncertainty

As discussed above, the presence of uncertain parameters implies that a given choice of first-stage decisions $\mathbf{x}$ may give rise to a range of different outcomes in the second stage. When optimizing under uncertainty, we therefore have to decide how we both measure and manage the risk associated with this range of second-stage outcomes, i.e., how we define the risk operators $R_{\text{cost}}[\,]$ and $R_{\text{viol}}[\,]$ in (1). In the following, we discuss some common formulations of optimization problems under uncertainty and associated metrics of risk. We first discuss formulations that explicitly consider

[1]We note that the use of an equidistant grid is only for illustrative purposes, as in reality it is common to use techniques such as clustering to design the set of uncertainty scenarios.

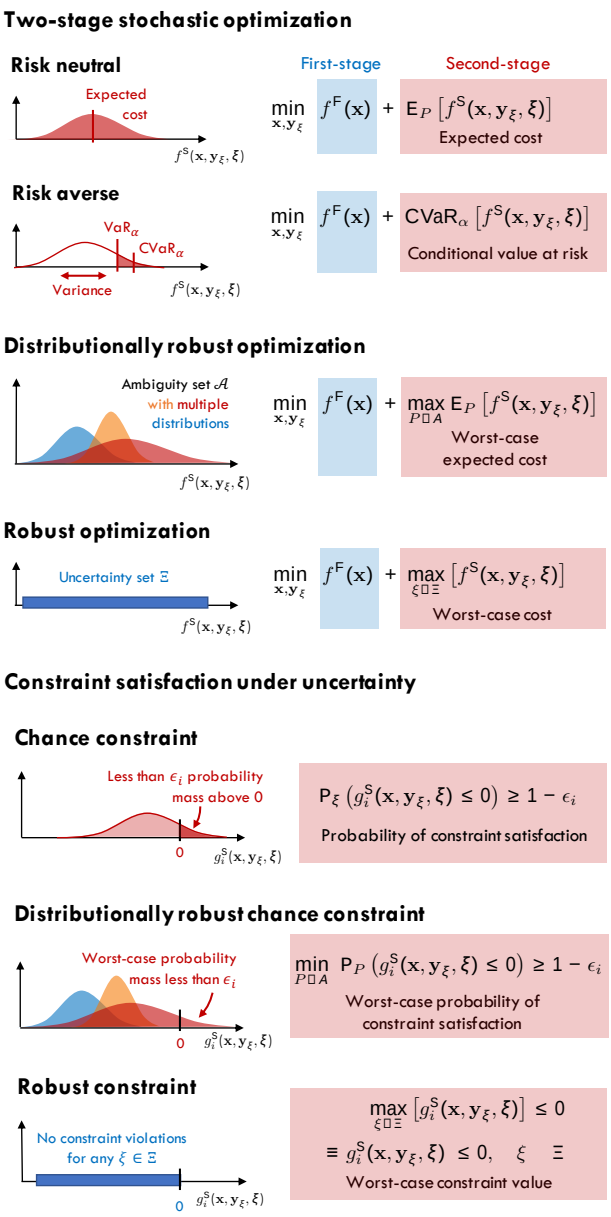

Figure 2: An overview of formulations for power system optimization under uncertainty.

the impact of uncertainty on cost, specifically risk-neutral and risk-averse versions of two-stage stochastic optimization problems, distributionally robust formulations, and robust min-max formulations. We then discuss formulations that focus on providing guarantees of constraint satisfaction despite different realizations of , including chance constraints, distributionally robust constraints, and robust constraints. An overview of the different types of optimization problems under uncertainty is given in Fig. 2.

A. Risk-Neutral Two-Stage Stochastic Optimization

A common formulation of two-stage optimization problems minimizes the total *expected* cost. The first-stage cost $f^{\mathrm{F}}(x)$ is known (i.e., deterministic) and the second-stage cost is expressed as the expected value of $f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$, giving rise to the following formulation:

$$\min_{\mathbf{x},\mathbf{y}}\ f^{\mathrm{F}}(\mathbf{x}) + \mathrm{E}_P\left[f^{\mathrm{S}}(\mathbf{x}; \mathbf{y}\ ;\ )\right] \tag{2a}$$

s.t.

$$\boldsymbol{h}^{\mathrm{F}}(\mathbf{x}) = 0; \quad \boldsymbol{g}^{\mathrm{F}}(\mathbf{x}) \quad 0; \tag{2b}$$

$$\boldsymbol{h}^{\mathrm{S}}(\mathbf{x}; \mathbf{y}\ ;\ ) = 0; \tag{2c}$$

$$\boldsymbol{g}^{\mathrm{S}}(\mathbf{x}; \mathbf{y}\ ;\ ) \quad 0: \tag{2d}$$

This formulation is called a *risk-neutral* formulation, as it treats costs above and below the expected value equally. Note that the formulation assumes that second-stage constraints (2c), (2d) are enforced for all realizations for the chosen solution $\mathbf{x}, \mathbf{y}$.

It is important to note that the risk-neutral formulation above is generally *not* the same as the certainty equivalent problem, which reduces the stochastic optimization problem to a deterministic problem by replacing with its expected value $\mathrm{E}_P[\ ]$. Specifically, $\mathrm{E}_P\left[f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}\ ,\ )\right]$ $f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}\ , \mathrm{E}_P[\ ])$.

B. Risk-Averse Two-Stage Stochastic Optimization

The risk-neutral formulation of the two-stage stochastic optimization problem which minimizes the expected cost may not always be desirable, as certain unfavorable outcomes can have a disproportionately large impact, e.g., a large loss could harm future prospects of a company or a large blackout can cause unacceptably high societal costs [15]. Problem formulations that specifically focus on limiting negative impacts of uncertainty realizations are referred to as *risk-averse* formulations. We briefly describe a few common risk metrics:

Standard Deviation: One classical way of limiting risk is to restrict the variance of the cost (or profit). In this framework, originally developed for financial portfolio selection [16], a larger variance is considered to be more risky and thus undesirable. The goal is to choose a solution that minimizes both the expected value and the variance of the cost. This multi-objective problem can be expressed as a trade-off in the objective function, or minimize the expected cost subject to a variance constraint or vice versa (i.e., minimize variance subject to a constraint on cost). A problem with this approach is that limiting the standard deviation also limits the possibility of getting better-than-expected outcomes (i.e., lower cost or higher profit).

Value-at-Risk: Another common risk metric is Value-at-Risk with probability level ϵ (VaR_ϵ), which bounds the largest cost that will occur with more than ϵ probability. This metric is mathematically defined as the $1\ \ \epsilon$-quantile of the cost function $f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$. While VaR_ϵ is widely used as a risk metric in finance, a common criticism of VaR_ϵ is that it does not provide any information about the size of costs beyond the $(1\ \ \epsilon)$ quantile, i.e., it counts how many times the cost exceeds the threshold, not by how much. Furthermore, VaR_ϵ is non-convex in general and does not satisfy the properties of a coherent risk measure (i.e., it is not sub-additive) [17].

Conditional Value-at-Risk: The Conditional-Value-at-Risk ($\text{CVaR}_{1-\epsilon}$), also referred to as the Expected Shortfall, is a convex and coherent risk measure [17] which mitigates some of these drawbacks of VaR. Specifically, $\text{CVaR}_{1-\epsilon}$ is defined as the expected value of all realizations of a random variable above the $1-\epsilon$ quantile and is more convenient for optimization due to its convexity properties that enable numerically efficient reformulations [18], [19]. In general, CVaR_{ϵ} represents an upper bound (which is frequently loose) on VaR_{ϵ}, i.e., $\text{CVaR}_{\epsilon} \geq \text{VaR}_{\epsilon}$.

The above risk metrics can be used to formulate a risk-averse version of the two-stage stochastic optimization problem where the expected value in the objective function (2a) is either replaced by or combined with the standard deviation, $\text{VaR}_{1-\epsilon}$, or $\text{CVaR}_{1-\epsilon}$. It is worth noting that such risk-averse formulations focus on limiting cost in the tail of the distribution, i.e., avoiding a large cost in a small number of scenarios, at the expense of a higher expected cost.

C. Distributionally Robust Optimization

In an ambiguous environment where the probability distribution is not known, distributionally robust optimization can be used to minimize the expected cost associated with the worst-case distribution within the ambiguity set. Since the worst-case expected cost depends both on our solution $\mathbf{x}$, $\mathbf{y}$ and the probability distribution of ξ, it is not possible to determine the worst-case solution a priori (i.e., before we solve the optimization problem). Instead, distributionally robust optimization methods endogenously identify the worst-case distribution within the ambiguity set and minimize the expected cost for this worst-case distribution. This problem can be defined as

$$\min_{\mathbf{x},\mathbf{y}} \quad f^{\mathsf{F}}(\mathbf{x}) + \max_{P \in A} \mathsf{E}_P \left[f^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) \right] \tag{3a}$$

s.t.

$$\boldsymbol{h}^{\mathsf{F}}(\mathbf{x}) = 0, \quad \boldsymbol{g}^{\mathsf{F}}(\mathbf{x}) \leq 0, \tag{3b}$$

$$\boldsymbol{h}^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) = 0, \tag{3c}$$

$$\boldsymbol{g}^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) \leq 0, \tag{3d}$$

where the objective function (3a) determines the worst-case distribution P among all pre-defined distributions within the ambiguity set A and minimizes the expected cost against such a worst-case distribution. For distributionally robust problems, we note that a larger ambiguity set results in a more conservative dispatch, and vice versa.

D. Robust Optimization

In some situations, we may not have access to sufficient data in order to determine the distribution of ξ or we may face a situation where an undesirable outcome is truly catastrophic (e.g., a company goes bankrupt or we experience a total system blackout). In this situation, it might be desirable to limit the impact of the worst-case outcomes, regardless of how (un)likely they are. Formulations that minimize the worst-case cost across all possible realizations within the uncertainty set give rise to robust optimization problems.

The min-max formulation of a robust optimization problem minimizes the combination of the first-stage and worst-case second-stage costs, defined as the maximum cost over the uncertainty set $\xi \in \Xi$, i.e.,

$$\min_{\mathbf{x},\mathbf{y}} \quad f^{\mathsf{F}}(\mathbf{x}) + \max_{\xi \in \Xi} \left[f^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) \right] \tag{4a}$$

s.t.

$$\boldsymbol{h}^{\mathsf{F}}(\mathbf{x}) = 0, \quad \boldsymbol{g}^{\mathsf{F}}(\mathbf{x}) \leq 0, \tag{4b}$$

$$\boldsymbol{h}^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) = 0, \tag{4c}$$

$$\boldsymbol{g}^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) \leq 0. \tag{4d}$$

It is worth noting that problems of the form (4) often assume *relatively complete recourse*, i.e., that for any feasible choice of the first-stage decision variables $\mathbf{x}$ and uncertainty realization ξ there exists a choice of $\mathbf{y}$ that satisfies the second-stage constraints (4b). For example, relatively complete recourse for linearized direct current (DC) power flow constraints can be ensured by allowing load shedding and renewable energy curtailment while also setting generator lower bounds to zero. Under these conditions, (4b) are often only explicitly enforced for a small set of scenarios, including the worst-case one. However, it is important to be aware that not all problems satisfy relatively complete recourse. In particular, it can be challenging to guarantee that there exists a feasible solution to a set of non-linear equality constraints (such as the AC power flow constraints) across a range of realizations for ξ.

E. Limiting Risk of Constraint Violations

Many optimization problems under uncertainty aim to identify first- and second-stage variables $\mathbf{x}$ and $\mathbf{y}$ that explicitly limit the risk of constraint violations, represented in (1) with the risk operator $R_{\text{viol}}[\,]$. We discuss three common methods to limit risk of constraint violations, namely chance-constrained, distributionally robust, and robust constraint satisfaction.

1) Chance-Constrained Optimization: Chance constraints, also commonly referred to as probabilistic constraints, require that the constraints are satisfied with a minimum probability of $1-\alpha$. We can distinguish between two main types of chance-constrained problems. A problem with *single chance constraints* enforces that the individual violation probability of each constraint i should not exceed ϵ_i:

$$\mathsf{P}\left(h_j^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) = 0\right) \geq 1 - \epsilon_j \quad \forall j = 1, \ldots, m \tag{5a}$$

$$\mathsf{P}\left(g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) \leq 0\right) \geq 1 - \epsilon_i \quad \forall i = 1, \ldots, n. \tag{5b}$$

A *joint chance constraint* requires that the joint violation probability, i.e., the probability that any of the constraints is violated, should not exceed ϵ:

$$\mathsf{P}_{\xi}\begin{pmatrix} \boldsymbol{h}^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) = 0 \\ \boldsymbol{g}^{\mathsf{S}}(\mathbf{x}, \mathbf{y}, \xi) \leq 0 \end{pmatrix} \geq 1 - \epsilon. \tag{6}$$

The use of single versus joint chance constraints is a modeling choice, i.e., it is not necessarily that one choice is better or worse than the other. While single chance constraints may allow a user to control the risk associated with individual constraints, a joint chance constraint provides stronger security guarantees for the overall system.

It is worthwhile to note that we have included equality constraints $h_i^S(\mathbf{x}, \mathbf{y}, \xi) = 0$ among the chance constraints in (5) and (6). Enforcing equality constraints is very important to the optimization of the electric grid, where equality constraints are used to represent the physical operation of the system. For example, it is necessary to ensure that the power system always remains balanced (i.e., the sum of all power injections minus the network losses exactly equals zero) despite possible variations in load demands and renewable generation. If such basic physical conditions are not satisfied, the mathematical model may no longer be valid[2]. As a result, it is common to use a very low violation probability for the equality constraints (or enforce the equality constraints as robust constraints, as further described below).

One common criticism of chance constraints (which are closely related to the VaR_ϵ risk measure) is that they limit only the probability of constraint violation, not the size of the violation. Furthermore, chance constraints are generally non-convex and can be challenging to solve. There exists several convex approximations of chance constraints in the literature, most notably the *CVaR approximation* [19]–[21]. This approximation is a conservative approximation of the chance constraint, but also has the added feature of limiting the size of a violation. The chance constraint can be interpreted as the product between the probability distribution of the constraint function $g_i^S(\mathbf{x}, \mathbf{y}, \xi)$ and a step-function that goes from zero to one at the limit. This step function records the probability of violation, but does not distinguish between small and large violations. In contrast, a CVaR constraint can be interpreted as the product between the probability distribution and a piecewise linear function with slope one above the limit. This linear function measures the size of the violation rather than the probability of violation. We note that it is also possible to leverage more general risk functions, which have been referred to as *weighted chance constraints* [22].

2) Distributionally Robust Chance Constraints: If the probability distribution of ξ is unknown, we may choose to limit risk by enforcing distributionally robust chance constraints. The distributionally robust form of a single chance constraint is given by

$$\min_{P \in A} \; \mathsf{P}\left(h_i^S(\mathbf{x}, \mathbf{y}, \xi) = 0\right) \quad 1 \quad \epsilon_i, \tag{7a}$$

$$\min_{P \in A} \; \mathsf{P}\left(g_j^S(\mathbf{x}, \mathbf{y}, \xi) \quad 0\right) \quad 1 \quad \epsilon_j, \tag{7b}$$

whereas the distributionally robust form of a joint chance constraint can be stated as

$$\min_{P \in A} \; \mathsf{P}_\xi \begin{pmatrix} h^S(\mathbf{x}, \mathbf{y}, \xi) = 0 \\ g^S(\mathbf{x}, \mathbf{y}, \xi) \quad 0 \end{pmatrix} \quad 1 \quad \epsilon. \tag{8}$$

Here, the probability that each individual constraint $h^S(\mathbf{x}, \mathbf{y}, \xi) = 0$ and $g^S(\mathbf{x}, \mathbf{y}, \xi) \quad 0$ hold under the worst-case distribution $\mathsf{P} \in A$ should be greater than or equal to $1 \quad \epsilon$. Note that the worst-case distribution P in the distributionally robust chance constraint (7) (8) is not necessarily identical to the worst-case distribution in the objective function (3a), and that the worst-case distribution can also differ between constraints.

3) Robust Feasibility Satisfaction: A third option to guarantee constraint satisfaction across a range of possible uncertainty realizations is to enforce robust feasibility constraints, which explicitly guarantee feasibility of the constraints for all uncertainty realizations $\in \Xi$, i.e.

$$h^S(\mathbf{x}, \mathbf{y}, \xi) = 0, \qquad \forall \xi \in \Xi, \tag{9a}$$

$$g^S(\mathbf{x}, \mathbf{y}, \xi) \quad 0, \qquad \forall \xi \in \Xi. \tag{9b}$$

Many problems with robust constraints only include first-stage variables $\mathbf{x}$, and the goal is to identify decisions that remain safe regardless of the realization ξ. Robust optimization problems that include recourse actions in response to the uncertainty realization ξ are referred to as *adaptive* robust optimization problems.

IV. Tractable Approximations and Solution Algorithms

Evaluating risk measures such as expected second-stage cost or the probability of constraint violations as a function of the decision variables $\mathbf{x}, \mathbf{y}$ is generally theoretically and numerically challenging. As a result, most optimization problems under uncertainty are difficult to solve. In this section, we discuss some common reformulations and solution algorithms for solving these problems.

A. Obtaining Finite-Dimensional Problem Formulations

In most stochastic programming problems, there tends to be a very large (or possibly infinite) number of decision variables and constraints. For instance, a continuous distribution for ξ implies that the problem is infinite dimensional, i.e., we have an infinite number of possible realizations of ξ, giving rise to an infinite number of second-stage variables $\mathbf{y}$ and second-stage constraints. Before attempting to solve the problem, we therefore have to find a finite-dimensional representation. This is typically achieved either by representing the uncertain parameters through samples (either *i.i.d.* samples or samples representing a discretized distribution) or by parameterizing the second-stage responses as a function of the uncertainty, also referred to as a recourse policy.

Replace ξ with Samples: One of the most common ways to reduce a stochastic programming formulation to a finite-dimensional problem is to replace the random variable ξ by a finite set of sample realizations ξ_s, $s \in S$. This allows us to define second-stage variables $\mathbf{y_s}$ and corresponding second-stage constraints for each realization s, giving rise to a finite-dimensional problem. This is usually a good modeling choice for problems where the decision maker is allowed to optimize their decision $\mathbf{y}$ once ξ is known (e.g., a unit commitment problem where the generation dispatch is optimized to meet realized demand).

[2]For example, if a the AC power flow equations are not satisfied, it may indicate system instability and the steady-state model is no longer valid. In this case, the values of decision variables representing quantities such as voltage variables or power flows no longer have a physical meaning.

These sampled realizations $_{\mathbf{s}}$ are used to obtain estimates of the objective function and constraints that involve the uncertain parameters. Typically, using a larger number of samples allows for a more accurate problem representation, but also increases the complexity of the problem as we introduce a new set of second-stage decision variables $\mathbf{y_s}$ and second-stage constraints for each sample. One way of coping with the large scale of the problem is to carefully select scenarios so as to reduce problem size while concisely capturing the diversity of uncertainty that the system faces. Scenario selection and scenario reduction methods based on the transportation metric have been proposed by [23] and extended by [24], and have found applications in power systems [25], [26]. Although these methods originate from sound theoretical foundations, their generalization to cases of multi-dimensional uncertainty (e.g., multi-area wind production) or composite uncertainty (i.e., the Cartesian product of forecast errors and contingencies) which is inherent to power systems applications is not straightforward. These drawbacks have motivated scenario selection methods inspired by importance sampling [27] as well as methods that explicitly consider the impact of scenarios on costs [26]. The trade-off between the number of scenarios that are input to a model and the accuracy of the solution that can be obtained has also been studied in the literature [28].

Recourse Policies: Another common option is to approximate $\mathbf{y}$ as a function of , where the function parameters are first-stage decision variables. This kind of function is commonly referred to as a *recourse policy*, a control policy, or a decision rule. It is typically a good modeling choice in problems where the decision maker has limited ability to optimize the recourse decisions $\mathbf{y}$, either because they have to react very quickly (e.g., to balance the system during a power imbalance) or cannot communicate decisions in real time (e.g., due to communication bandwidth issues).

Affine Recourse Policy: An affine recourse policy (also referred to as a linear decision rule) [29], [30] expresses the second-stage response as a linear function of the uncertainty realization, i.e.,

$$\mathbf{y}(\xi) = \mathbf{y}_{(1)} \quad + \mathbf{y}_{(0)}. \tag{10}$$

The coefficients $\mathbf{y}_{(1)}$ and $\mathbf{y}_{(0)}$ are first-stage variables to be determined *before* the realization of uncertainty.

More General Recourse Policies: While affine policies are by far the most common (and usually most tractable) policies, they can be generalized. For example, it is possible to consider non-linear dependencies on the uncertain parameter by extending the vector of uncertain parameters to contain additional parameters that represent non-linear dependencies on , e.g., $\sin(\xi_i)$. It is also possible to consider piece-wise linear policies [22], [31] or more general policies including polynomials [32] and truncated distributions [33].

With a recourse policy, we reduce the number of decision variables to a finite number of first-stage decision variables. However, we also need to obtain a tractable reformulation of the second-stage constraints. One possibility is to leverage recourse policies in combination with sampled realizations of $_{\mathbf{s}}$. Compared with a sample-based formulation that introduces one decision variable $\mathbf{y_s}$ for each scenario, the use of an (affine) recourse policy may reduce computational complexity. In some cases, the use of a parameterized recourse policy in combination with samples may even be required by the solution method (e.g., the scenario approach, further discussed below, requires a fixed number of decision variables to determine the number of samples needed). In addition to using samples, there are several other methods to reduce the number of second-stage constraints to a finite number. Such methods will be further discussed below, particularly in the context of chance-constrained, distributionally robust, and robust optimization.

B. Reformulations and Solution Algorithms for Risk-Neutral Two-Stage Stochastic Programs

The main challenge in solving risk-neutral problems is to accurately represent the expected second-stage cost. Although similar challenges arise in other problems that involve expectations, such as in the evaluation of CVaR, we frame the discussion in this section around risk-neutral problems only.

For risk-neutral two-stage stochastic problems in the form of (2), it is common to leverage a sample average approximation to represent the expected second-stage cost, i.e.,

$$\mathsf{E}_P\left[f^{\mathsf{S}}(\mathbf{x}, \mathbf{y} \ , \ )\right] \quad \sum_{\mathbf{s}2S} \pi_{\mathbf{s}} f^{\mathsf{S}}(\mathbf{x}, \mathbf{y}_{\mathbf{s}}, \ _{\mathbf{s}}), \tag{11}$$

where $_{\mathbf{s}}$ represents the realization of the uncertain parameters in scenario $\mathbf{s}$ and $\pi_{\mathbf{s}}$ is the associated probability of occurrence. It is worthwhile to note that the accuracy of the sample average approximation (11) depends on several factors, including the number of samples that is used and the distribution of . However, solution algorithms for two-stage stochastic optimization problems, such as the ones described below, generally assume that the approximation (11) is an accurate representation of the expected cost. An assessment of the true expected cost is often obtained through a posteriori evaluation, as described in Section V.

Although it is possible to solve (2) directly by replacing the second-stage cost by the approximation (11) and introducing copies of the second-stage variables and constraints for each sample $\xi_{\mathbf{s}}$, risk-neutral formulations typically require many samples to accurately estimate the expected cost. As a result, the problem quickly escalates to a size that becomes too large to solve directly, motivating a variety of alternative solution techniques.

1) Decomposition Strategies: We proceed to discuss the general concept of decomposition, as well as two specific methods by which this concept is leveraged in stochastic programs: Lagrange relaxation and cutting plane methods.

Decomposition: Even with a relatively modest number of scenarios, the extensive form of the problem (which includes variables and constraints for each scenario) is challenging to solve directly. As a result, solution methods for stochastic optimization problems typically rely on *decomposition* algorithms which split the problem by scenario to formulate

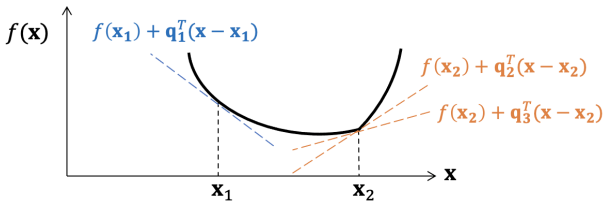

Figure 3: Examples of subgradients of the function $f(\mathbf{x})$. Where $f(\mathbf{x})$ is non-differentiable, there is more than one subgradient.

smaller, easier-to-solve subproblems. Although the number of subproblems may be large, solving them individually typically requires significantly less effort than solving the combined problem, with the added benefit that they can often be solved in parallel. By carefully combining information from the smaller subproblems, these methods iteratively approach the optimal solution to the original problem.

Lagrange relaxation: The idea of Lagrange relaxation is to relax complicating constraints in a large-scale problem and tackle the dual problem for which we have access to geometric information, such as the subgradient of the dual function. A subgradient is a generalization of a gradient for a non-differentiable function. For a function $f(\mathbf{x})$, the vector $\boldsymbol{q}$ is a subgradient of $f(\mathbf{x})$ at $\mathbf{x}_1$ if the following holds true

$$f(\mathbf{x}) \quad f(\mathbf{x}_1) + \boldsymbol{q}^T(\mathbf{x} \quad \mathbf{x}_1) \; . \tag{12}$$

An illustrative example of subgradients is provided in Fig. 3. Note that if $f(\mathbf{x})$ is differentiable at $\mathbf{x}$, the subgradient is simply a gradient. If $f(\mathbf{x})$ is convex, the subgradient is a global underestimator of the function.

The subgradient method is similar to gradient descent in that it seeks for the next best solution in the direction of the subgradient. However, it is not guaranteed that the solution is monotonously improving across iterations, which makes it necessary to keep track of the best known solution and can make it challenging to assess convergence.

Cutting planes: A cutting plane method is an approximation techniques whereby a convex (or concave) function is outer approximated by supporting hyperplanes, referred to as cutting planes, which approximate the function from below (or above, in the case of concave functions). Supporting hyperplanes are closely related to the subgradients of a function; in fact, (12) describes a supporting hyperplane of $f(\mathbf{x})$. The benefit of using cutting planes to approximate a function $f(\mathbf{x})$ is that the approximation can be built iteratively. Cutting plane methods selectively add new cutting planes to improve the approximation at each iteration until we reach the desired solution accuracy. Approximations with cutting planes are an important part of many algorithms for solving stochastic two-stage and multi-stage optimization problems.

2) Decomposition Algorithms: We next review a few common solution methods which leverage one or more of the above concepts. These include Benders decomposition, stochastic dual dynamic programming, dual decomposition, and progressive hedging.

Benders Decomposition: Benders decomposition solves the first-stage problem with an iteratively refined approximation of the expected second-stage cost until an optimal first-stage solution is obtained. The key observation that is leveraged by Benders decomposition is that two-stage stochastic programming problems tend to have (or can be formulated to have) a *complicating-variable structure* [34]. In such problems, if some variables (the complicating ones) of the problem are fixed to given values, the problem decomposes by scenario, making it easier to solve. In the two-stage stochastic programming problem (13) below, the complicating (coupling) variables are the first-stage variables $\mathbf{x}$ that prevent a decomposition by scenario:

$$\min_{\mathbf{x}, \mathbf{y}_s \, 8s2S} \quad f^{\mathrm{F}}(\mathbf{x}) + \sum_{s2S} \pi_s f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \; _s) \tag{13a}$$

$$\text{s. t.} \quad h^{\mathrm{F}}(\mathbf{x}) = \mathbf{0}, \quad g^{\mathrm{F}}(\mathbf{x}) \quad \mathbf{0}, \tag{13b}$$

$$h^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \; _s) = \mathbf{0}, \quad 8 \; s \; 2 \; S, \tag{13c}$$

$$g^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \; _s) \quad \mathbf{0}, \quad 8 \; s \; 2 \; S. \tag{13d}$$

We note that problem (13) is equivalent to

$$\min_{\mathbf{x}} \quad f^{\mathrm{F}}(\mathbf{x}) + \alpha(\mathbf{x}) \tag{14a}$$

$$\text{s. t.} \quad h^{\mathrm{F}}(\mathbf{x}) = \mathbf{0}, \quad g^{\mathrm{F}}(\mathbf{x}) \quad \mathbf{0}, \tag{14b}$$

where the function $\alpha(\boldsymbol{x})$ represents the second-stage cost and is given by

$$\alpha(\mathbf{x}) = \min_{\mathbf{y}_s \, 8s2S} \quad \sum_{s2S} \pi_s f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \; _s) \tag{14c}$$

$$\text{s. t.} \quad h^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \; _s) = \mathbf{0}, \quad 8 \; s \; 2 \; S, \tag{14d}$$

$$g^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \; _s) \quad \mathbf{0}, \quad 8 \; s \; 2 \; S. \tag{14e}$$

The convexity properties of the function $\alpha(\mathbf{x})$ are critical for the derivation and convergence of the Benders decomposition algorithm. Specifically, if $\alpha(\mathbf{x})$ is convex or if it has a convex hull, it can be outer approximated by cutting planes[3]. In this case, Benders' algorithm converges to an optimal solution. Otherwise, it generally does not. Many practical two-stage stochastic programming problems that appear in power system applications give rise to convex functions $\alpha(\mathbf{x})$.

Considering problem (14) and assuming that the function $\alpha(\mathbf{x})$ has the desired convexity properties, Benders decomposition works as follows. An initial feasible solution of first-stage variables, $\mathbf{x}^{(\nu)}$, is obtained by solving the initial master problem:

$$\min_{\mathbf{x}, \alpha} \quad f^{\mathrm{F}}(\mathbf{x}) + \alpha \tag{15a}$$

$$\text{s. t.} \quad h^{\mathrm{F}}(\mathbf{x}) = \mathbf{0}, \quad g^{\mathrm{F}}(\mathbf{x}) \quad \mathbf{0} \tag{15b}$$

$$\alpha \quad \alpha^{\mathrm{down}} \tag{15c}$$

where α^{down} is a lower bound on the expected second-stage cost α that can be easily obtained in most applications. If

[3]This holds true in general for problems where $h^S(\mathbf{x}, \mathbf{y_s}, \boldsymbol{\xi_s}) = A\mathbf{x} + B\mathbf{y_s}$ and $g^S(\mathbf{x}, \mathbf{y_s}, \boldsymbol{\xi_s}) = -b(\mathbf{x}) + c(\mathbf{y_s})$ where $b(\mathbf{x})$ is convex in $\mathbf{x}$.

the original two-stage stochastic programing problem (13) is well-posed, the initial master problem (15) is always feasible.

Given the initial feasible solution $\mathbf{x}^{(\nu)}$, the second-stage subproblem is formulated as:

$$\min_{\mathbf{y}_s, s \in S} \quad \sum_{s \in S} \pi_s f^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \ _s) \tag{16a}$$

$$\text{s. t.} \quad h^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \ _s) = \mathbf{0}, \quad s \in S, \tag{16b}$$

$$g^{\mathrm{S}}(\mathbf{x}, \mathbf{y}_s, \ _s) \quad \mathbf{0}, \quad s \in S, \tag{16c}$$

$$\mathbf{x} = \mathbf{x}^{(\nu)} : \quad ^{(\nu)}. \tag{16d}$$

The solution of subproblem (16) provides $\mathbf{y}_s^{(\nu)}, s \in S$ and sensitivities $^{(\nu)}$. We note that subproblem (16) decomposes by scenario, which makes it drastically easier to solve. For the sake of simplicity, we assume that the subproblem (16) satisfies relatively complete recourse, i.e., is always feasible.

The solution of problem (16), and particularly the sensitivities $^{(\nu)}$, allow formulating a refined master problem as follows:

$$\min_{\mathbf{x}, \alpha} \quad f^{\mathrm{F}}(\mathbf{x}) + \alpha \tag{17a}$$

$$\text{s. t.} \quad h^{\mathrm{F}}(\mathbf{x}) = \mathbf{0}, \quad g^{\mathrm{F}}(\mathbf{x}) \quad \mathbf{0}, \tag{17b}$$

$$\alpha \quad \alpha^{\mathrm{down}}, \tag{17c}$$

$$\alpha \quad \sum_{s \in S} \pi_s f^{\mathrm{S}}(\mathbf{x}^{(\nu)}, \mathbf{y}_s^{(\nu)}, \ _s) + (\ ^{(\nu)})^{\top}(\mathbf{x} \quad \mathbf{x}^{(\nu)}). \tag{17d}$$

Here, (17d) is referred to as a Benders cut, which approximates the function $\alpha(\mathbf{x})$ from below. Solving the refined problem (17) provides new values for the first-stage variables, $\mathbf{x}^{(\nu+1)}$, and a new value for α, i.e., $\alpha^{(\nu+1)}$. If the original two-stage stochastic programing problem (13) is well-posed, the master problem (17) is always feasible. We note that problem (17) is generally the computational bottleneck of the Benders decomposition algorithm.

The Benders cuts of expression (17d) are referred to as *optimality cuts*. If the problem does not satisfy relative complete recourse, it is possible that the subproblems are not feasible. In this case, we need to generate *feasibility cuts* by using the extreme rays of the subproblem, which an optimization solver can provide as certificates of subproblem infeasibility. Alternatively, we can reformulate the subproblems by introducing slack variables that are penalized in the objective in order to guarantee relatively complete recourse.

Benders decomposition provides an upper and lower bound on the objective function at each iteration (ν), which can be used to assess convergence.

Upper bound (UB): Since subproblem (16) is a further constrained version of the original problem (13) (variables $\mathbf{x}$ are fixed), we can use the optimal objective value of the subproblem (16) to compute an upper bound of the optimal objective value of the original problem (13), i.e.,

$$\mathrm{UB}^{(\nu)} = f^{\mathrm{F}}(\mathbf{x}^{(\nu)}) + \sum_{s \in S} \pi_s f^{\mathrm{S}}(\mathbf{x}^{(\nu)}, \mathbf{y}_s^{(\nu)}). \tag{18}$$

Lower bound (LB): The constraints (17d) approximate the true second-stage cost $\alpha(\mathbf{x})$ from below. Therefore, problem (17) is a relaxation of problem (13), which allows deriving a lower bound of the optimal value of the objective function of problem (13), i.e.,

$$\mathrm{LB}^{(\nu)} = f^{\mathrm{F}}(\mathbf{x}^{(\nu+1)}) + \alpha^{(\nu+1)}. \tag{19}$$

These upper and lower bounds can be used to devise a convergence and termination criterion, giving rise to the following algorithm:

Step 0. Set $\nu = 1$, solve the initial master problem (15) and obtain $\mathbf{x}^{(\nu)}$.

Step 1. Given $\mathbf{x}^{(\nu)}$, solve subproblem (16) to obtain $\mathbf{y}_s^{(\nu)}, s \in S$, and sensitivities $^{(\nu)}$. Compute upper bound (18), $\mathrm{UB}^{(\nu)} = f^{\mathrm{F}}(\mathbf{x}^{(\nu)}) + \sum_{s \in S} \pi_s f^{\mathrm{S}}(\mathbf{x}^{(\nu)}, \mathbf{y}_s^{(\nu)}, \ _s)$.

Step 2. Given $\mathbf{y}_s^{(\nu)}, s \in S$, and sensitivities $^{(\nu)}$, solve master problem (17) and compute $\mathbf{x}^{(\nu+1)}$ and $\alpha^{(\nu+1)}$. Compute the lower bound (19), $\mathrm{LB}^{(\nu)} = f^{\mathrm{F}}(\mathbf{x}^{(\nu+1)}) + \alpha^{(\nu+1)}$.

Step 3. If $\mathrm{UB}^{(\nu+1)} \quad \mathrm{LB}^{(\nu+1)} \quad \epsilon$, where ϵ is a convergence tolerance, stop. Otherwise set $\nu \quad \nu+1$ and continue in Step 1.

Benders decomposition is directly applicable, for example, to the stochastic network-constrained unit commitment problem. If the commitment (binary) variables of a stochastic network-constrained unit commitment problem are fixed to given binary values, the problem decomposes by scenario (note that only commitment variables couple operation scenarios), and each scenario subproblem is continuous and generally easy to solve.

Stochastic dual dynamic programming (SDDP): SDDP is a method for solving multi-stage stochastic linear programs [35], [36], which can be generalized to address stochastic convex programs [37]. It blends cutting plane techniques with Monte Carlo simulation in order to tackle the multi-stage generalization of two-stage stochastic linear programs.

The typical structure of a problem that is amenable to resolution by SDDP is as follows [38]: (i) decisions are made at a given time stage, (ii) there exists a set of linear constraints that couple the decision variables of the current time stage to the decisions of the previous time stage, and (iii) the decisions of the current time stage obey linear constraints that depend on uncertain input parameters (note that a discretization of the uncertain parameters is needed). There are various specific assumptions that enable the more structured resolution of the problem, e.g., right-hand side uncertainty and the serial independence of uncertain parameters [39], although more general uncertainty structures can be handled as well [40]. Generalizations of the method have been proposed, e.g., to stochastic mixed-integer programming [41] as well as risk-averse formulations [42].

The algorithm is typically iterating between *forward passes* and *backward passes*. The backward passes approximate the convex expected cost-to-go functions (i.e., the expected cost of

future problem stages) with an increasing degree of precision using cutting planes. The forward passes try to identify the optimal policy by sampling uncertainty over the horizon of the problem. For more details on how to solve such problems, we refer the reader to [38].

The probabilistic nature of the algorithm poses non-trivial questions of selecting convergence criteria appropriately [43]. Although the algorithm is guaranteed finite convergence in the linear case [44], specific instances can pose serious challenges in terms of narrowing down optimality gaps [42]. Methods for accelerating the basic algorithm include regularization techniques [45] as well as cut selection techniques [46]–[48]. The algorithm is highly amenable to parallel computing [36]. This has inspired numerous parallelization schemes along both forward and backward passes [37], [49]–[52] in order to further accelerate the basic scheme. There exist various open-source packages that implement the algorithm, including in Matlab and Julia [37]. Despite these improvements, the SDDP algorithm remains computationally complex and a topic of active research.

Dual Decomposition: In Benders Decomposition and SDDP, we derive an increasingly accurate description of future cost using cutting planes. In contrast, dual decomposition splits the problem into subproblem by introducing scenario-dependent copies of the first-stage variables and relax the requirement that all copies have the same value. This enables decomposition of the problem by scenario. The key idea behind the solution algorithm is to share information between the subproblems to obtain consensus regarding the optimal value of the first-stage variables.

In dual decomposition, we first create a scenario-dependent copies $\mathbf{x_s}$ of the first-stage variable $\mathbf{x}$ and introduce an explicit constraint $\mathbf{x_s} \quad \mathbf{x} = 0$ for every scenario. This constraint is referred to as a *non-anticipativity constraint*, as it ensures that the scenario-dependent variables $\mathbf{x_s}$ have to be chosen without knowledge (i.e., anticipation) of the scenario-specific uncertainty realization $_{\mathbf{s}}$. Model (2) can then be expressed equivalently as follows:

$$\min_{\mathbf{x},\,\mathbf{x_s},\mathbf{y_s}\ 8\mathbf{s}2S} \quad f^{\mathsf{F}}(\mathbf{x}) + \sum_{\mathbf{s}2S} \ _{\mathbf{s}} f^{\mathsf{S}}(\mathbf{x};\mathbf{y_s};\ _{\mathbf{s}}) \tag{20a}$$

s.t.

$$\boldsymbol{h}^{\mathsf{F}}(\mathbf{x}) = \mathbf{0};\quad \boldsymbol{g}^{\mathsf{F}}(\mathbf{x}) \quad \mathbf{0}; \tag{20b}$$

$$h^{\mathsf{S}}(\mathbf{x_s};\mathbf{y_s};\ _{\mathbf{s}}) = \mathbf{0}; \qquad 8\mathbf{s}\ 2\ S; \tag{20c}$$

$$g^{\mathsf{S}}(\mathbf{x_s};\mathbf{y_s};\ _{\mathbf{s}}) \quad \mathbf{0}; \qquad 8\mathbf{s}\ 2\ S; \tag{20d}$$

$$\mathbf{x} \quad \mathbf{x_s} = \mathbf{0} \quad : \quad _{\mathbf{s}} \qquad 8\mathbf{s}\ 2\ S: \tag{20e}$$

In this problem, the variable vector $_{\mathbf{s}}$ represents the *dual multipliers* associated with the constraints (20e) for scenario $\mathbf{s}$. While the introduction of the variables $\mathbf{x_s}$ is redundant from a modeling perspective, the benefit of the above formulation is that scenario-specific second-stage constraints (20c) and (20d) have been isolated from the first-stage constraints (20b). Thus, (20e) is the only complicating (coupling) constraint that links the first and second stages. This allows us to apply dual decomposition.

Dual decomposition [53] is based on Lagrangian relaxation, a common decomposition method which is applicable in a much broader family of problems than stochastic programming [54]. We define a Lagrangian relaxation of (20), where we remove the constraints (20e) and instead penalize the violation as $^{>}_{\mathbf{s}}(\mathbf{x} \quad \mathbf{x_s})$ in the objective function, giving rise to the following problem:

$$\min_{\mathbf{x},\,\mathbf{x_s},\mathbf{y_s}\ 8\mathbf{s}2S} \quad f^{\mathsf{F}}(\mathbf{x}) + \sum_{\mathbf{s}2S} \ _{\mathbf{s}} f^{\mathsf{S}}(\mathbf{x};\mathbf{y_s};\ _{\mathbf{s}}) + \ ^{>}_{\mathbf{s}}(\mathbf{x} \quad \mathbf{x_s}) \tag{21a}$$

s.t.

$$\boldsymbol{h}^{\mathsf{F}}(\mathbf{x}) = \mathbf{0};\quad \boldsymbol{g}^{\mathsf{F}}(\mathbf{x}) \quad \mathbf{0}; \tag{21b}$$

$$h^{\mathsf{S}}(\mathbf{x_s};\mathbf{y_s};\ _{\mathbf{s}}) = \mathbf{0}; \qquad 8\mathbf{s}\ 2\ S; \tag{21c}$$

$$g^{\mathsf{S}}(\mathbf{x_s};\mathbf{y_s};\ _{\mathbf{s}}) \quad \mathbf{0}; \qquad 8\mathbf{s}\ 2\ S: \tag{21d}$$

For this problem, we define the following dual function,

$$\phi(\ ) = \phi_1(\ ) + \sum_{\mathbf{s}2S} \pi_{\mathbf{s}}\phi_{\mathbf{s}}(\ _{\mathbf{s}}). \tag{22}$$

where represents the collection of $_{\mathbf{s}}$ for all $\mathbf{s}\ 2\ S$. The function $\phi_1(\ )$ can be defined as follows,

$$_1(\ ) = \min_{\mathbf{x}} \ f^{\mathsf{F}}(\mathbf{x}) + \sum_{\mathbf{s}2S} \ _{\mathbf{s}} \ ^{>}_{\mathbf{s}}\mathbf{x} \tag{23a}$$

$$\text{s.t.} \quad \boldsymbol{h}^{\mathsf{F}}(\mathbf{x}) = \mathbf{0};\quad \boldsymbol{g}^{\mathsf{F}}(\mathbf{x}) \quad \mathbf{0}; \tag{23b}$$

while functions $\phi_{\mathbf{s}}(\ _{\mathbf{s}})$ for all $\mathbf{s}\ 2\ S$ are given by

$$_{s}(\ _{s}) = \min_{\mathbf{x_s},\mathbf{y_s}} \sum_{\mathbf{s}2S} (f^{\mathsf{S}}(\mathbf{y_s}) \quad ^{>}_{\mathbf{s}}\mathbf{x_s}) \tag{24a}$$

s.t.

$$h^{\mathsf{S}}(\mathbf{x_s};\mathbf{y_s};\ _{\mathbf{s}}) = \mathbf{0}; \tag{24b}$$

$$g^{\mathsf{S}}(\mathbf{x_s};\mathbf{y_s};\ _{\mathbf{s}}) \quad \mathbf{0}: \tag{24c}$$

For a given set of multipliers , we can now decompose the problem by stage and scenario, i.e., we can compute $\phi(\ )$ by solving the problem $\phi_1(\ )$ and the jSj subproblems $\phi_s(\ _{\mathbf{s}})$ *independently*. We next focus on obtaining a solution by solving the dual problem, which amounts to maximizing the dual function $\phi(\ )$ over . The dual function is concave, giving rise to a convex, but non-differentiable maximization problem. There exist a variety of solution algorithms ranging from basic subgradient algorithms to cutting plane methods, trust region methods, level methods, and several other schemes [55], [56].

Any feasible solution to the dual problem provides a lower bound on the relaxed problem (21), and thus also a lower bound on the original problem (20). As we approach the dual function optimum, the duality gap of the problem, which quantifies the lack of consensus regarding the values of the first-stage variables $\mathbf{x_s}$, starts decreasing. Problem-specific feasibility recovery heuristics can then be used to recover feasible solutions to the original problem (20) at every iteration of the algorithm [57], providing upper bounds on the objective value. Using the upper and lower bounds, we can measure the optimality gap and decide when to terminate. Vanilla subgradient algorithms, which may exhibit unstable behavior, can be further enhanced by replicating scenario-specific constraints in the subproblem which computes $\phi_1(\lambda)$. These techniques

have been used successfully for solving large-scale stochastic unit commitment problems [27], [28], [57].

Progressive Hedging: An alternative approach for decomposing the problem is to consider an augmented Lagrangian of the original problem [53], which amounts to adding a penalty term $(\rho/2)\|\mathbf{x} - \mathbf{x}_s\|^2$ to the objective. With this term, the problem becomes non-separable, i.e., we cannot directly split it up by scenarios. There are however known algorithms, called proximal algorithms [58], [59], which maintain separability while improving the stability of the decomposition. The progressive hedging algorithm is essentially a proximal method applied to two-stage stochastic programs. The algorithm uses only scenario-dependent first-stage variables, and decomposes the problem by scenario while using a quadratic term to penalize a lack of consensus among scenarios. Concretely, the subproblem that is being solved at every iteration for every scenario $s \in S$ is:

$$\min_{\mathbf{x}_s, \mathbf{y}_s} \quad \pi_s f^F(\mathbf{x}_s) + \pi_s f^S(\mathbf{y}_s) + \lambda_s^\top(\mathbf{x}_s - \bar{\mathbf{x}}) + \frac{\rho}{2}\|\mathbf{x}_s - \bar{\mathbf{x}}\|^2 \tag{25a}$$

s.t.

$$\boldsymbol{h}^F(\mathbf{x}_s) = \mathbf{0}, \quad \boldsymbol{g}^F(\mathbf{x}_s) \leq \mathbf{0}, \tag{25b}$$

$$h^S(\mathbf{x}_s, \mathbf{y}_s, \xi_s) = \mathbf{0}, \quad \forall s \in S, \tag{25c}$$

$$g^S(\mathbf{x}_s, \mathbf{y}_s, \xi_s) \leq \mathbf{0}, \quad \forall s \in S, \tag{25d}$$

where $\bar{\mathbf{x}} = \sum_{s \in S} \pi_s \mathbf{x}_s$ is the average first-stage decision. The overall algorithm can be described as follows [60]:

0) Define some non-anticipative $\mathbf{x}^0$, some initial multiplier λ^0, a penalty parameter $\rho > 0$, and a tolerance κ. Set $k = 0$. Go to step 1.
1) Solve problem (25) for each scenario $s \in S$ with λ^k to obtain a solution $(\mathbf{x}_s, \mathbf{y}_s)$.
2) Compute the average first-stage decision $\bar{\mathbf{x}}^{(k+1)}$ and let $\lambda_s^{(k+1)} = \lambda_s^{(k)} + \rho(\mathbf{x}_s - \bar{\mathbf{x}}^{(k+1)})$.
3) If $\|\bar{\mathbf{x}}^{(k+1)} - \bar{\mathbf{x}}^{(k)}\| \leq \kappa$ and $\|\lambda_s^{(k+1)} - \lambda_s^{(k)}\| \leq \kappa$, then stop. Otherwise, let $k = k + 1$ and go to step 1.

The choice of the parameter ρ is especially crucial for the performance of the algorithm [53]. The method has been applied successfully to the stochastic unit commitment problem [61].

C. Solution Methods for Robust Optimization Problems

When solving robust optimization problems of the form (4) (or with robust constraints of the form (9)), the main challenge is that we have to identify the realization of $\mathbf{u}$ within the uncertainty set that gives rise to the worst-case cost (or worst-case constraint violation). Which realization is the worst does, however, depend on our decision variables $\mathbf{x}$ and $\mathbf{y}$ and thus cannot be determined a priori. Instead, solution algorithms for robust optimization problems aim to identify the worst-case realization as part of the solution process. We next explain one one of the most common methods for addressing two-stage robust problems, namely column-and-constraint generation (CCG). We also briefly discuss methods for robust constraint satisfaction.

Column-and-Constraint Generation: CCG is also known as a *primal cut* algorithm and is similar to Benders Decomposition in that we iteratively update a master problem. However, while Benders Decomposition adds constraints that reflect the cost of all possible realizations, the CCG iteratively adds new *scenarios* with corresponding sets of constraints and variables that represent the worst-case conditions. The overall idea is that robust optimization problems (4) may require only a few scenarios to fully reconstruct the optimal solution. Thus, it is not necessary to include all scenarios contained in the uncertainty set (which could be potentially infinite for continuous uncertainty sets).

To identify the worst-case scenarios, the CCG algorithm exploits a master-adversarial problem framework. Specifically, the CCG problem creates a tractable master problem in order to identify a candidate solution of the first-stage variables $\mathbf{x}^k$. This typically involves relaxing the original problem (4) to only consider a small sample set $S^k \subseteq S$,

$$\min_{\mathbf{x}, \mathbf{y}_s} \quad f^F(\mathbf{x}) + \eta \tag{26a}$$

s.t.

$$\eta \geq f^S(\mathbf{x}, \mathbf{y}_s, \xi_s), \quad \forall s \in S^k, \tag{26b}$$

$$\boldsymbol{h}^F(\mathbf{x}) = \mathbf{0}, \quad \boldsymbol{g}^F(\mathbf{x}) \leq \mathbf{0}, \tag{26c}$$

$$h^S(\mathbf{x}, \mathbf{y}_s, \xi_s) = \mathbf{0}, \quad \forall s \in S^k, \tag{26d}$$

$$g^S(\mathbf{x}, \mathbf{y}_s, \xi_s) \leq \mathbf{0}, \quad \forall s \in S^k. \tag{26e}$$

An *adversarial problem*, also called *oracle subproblem* or just *subproblem*, is designed to check if the candidate first-stage solution $\mathbf{x}^{(k)}$ is inducing a new worst-case (adversarial) scenario of the subproblem. If this is the case, the scenario set used in the master problem for the next iteration $S^{(k+1)}$ is updated to include the new worst-case scenario ξ_s with corresponding second-stage variables $\mathbf{y}_s$ and constraints.

Note that, for linear robust programs with relatively complete recourse, the relaxed master problem (26) always provides a lower bound for the original problem, while the worst-case cost of the subproblem provides an upper bound. Thus, the CCG algorithm provides an optimality gap along the iterative process, and eventually, it provides a global solution with convergence guarantees [62].

Robust Constraint Satisfaction: For problems with robust constraints of the form (9), there exist several classes of problems which can be efficiently represented and solved. In general, the key idea of ensuring robust constraint satisfaction is to find a way to identify the worst-case realizations of the uncertainty set and enforce constraints for those realizations. For some classes of problems, such as linear problems with elliptical or polyhedral uncertainty sets, it is possible to directly enforce constraints that guarantee robust constraint satisfaction. For other classes of problems, it may be necessary to consider a constraint generation procedure (similar to the one above) where we iterate between solving a master problem and identifying and adding constraints for the worst-case realizations. We refer interested readers to [63] for an introduction and overview.

D. Reformulations and Solution Approaches for Chance Constraints

The main challenge in solving chance-constrained (and distributionally robust chance-constrained) problems is efficiently evaluating (or upper bounding) the violation probabilities involved in single (5) and joint (6) chance constraints. In this section, we focus on the different ways in which chance-constrained problems can be reformulated into tractable deterministic optimization problems. We focus less on discussing the solution algorithms themselves since the reformulated problems can often be solved using standard solvers. However, scalability to large problems often requires the development of special purpose algorithms or may leverage solution algorithms related to the ones described above for risk neutral or robust optimization problems.

Approximating joint chance constraints with single chance constraints: Single chance constraints consider the violation probability of each constraint in isolation and thus focus on the scalar probability distribution of a single constraint function. Problems with single chance constraints are typically more straightforward to solve than problems with joint chance constraints, where the correlation and trade-offs of violations across different constraints (with possibly complicated joint probability distributions) become important. As a result, some publications seek to enforce a joint chance-constraint through the use of single chance constraints. A simple way to achieve this, known in the literature as *Bonferroni approximation*, is to split the acceptable joint violation probability ϵ_J equally across all n constraints and enforce single chance constraint with probability $\epsilon_S = \epsilon_J/n$. However, this method typically leads to very conservative results with lower-than-required violation probabilities and a higher cost. This is primarily due to the fact that (*i*) not all constraints will experience violations and thus will not be "using" their allotted violation probability, and (*ii*) this method double counts violations in situations where several constraints are violated simultaneously. There exist several methods in the literature that aim to mitigate these drawbacks, for example through iterative risk allocation [64], estimating the joint violation probability of different constraints [65], and using chance-constraint tuning [66]. However, these methods often remain computationally demanding and may yield sub-optimal results.

Sample Average Approximation: A common method to solve optimization problems with chance constraints is to use a sample average approximation [67], which enforces the desired violation probability empirically based on a set of samples. The key idea is to replace by a set of samples $\mathbf{s} \in S$ and allow for constraint violations in ϵ percent of the samples (e.g., if we have 100 equiprobable samples and desire a violation probability $\epsilon = 0.05$, the sample average approximation would allow violations for 5 samples). This formulation has the appealing properties that it (*i*) can solve joint chance constraints and (*ii*) allows a lot of flexibility in the modeling of the second stage (e.g., it is possible to introduce new decision variables $\mathbf{y_s}$ for each considered realization). However, the classic formulation of the problem [67] uses integer variables to represent whether or not a particular sample is among the violated ones, thus limiting the number of samples that can be considered.

A variety of methods have been proposed to remedy this lack of scalability, which is rooted in the non-convexity of the chance constraint. Methods for obtaining upper bounds on the problem include the conservative CVaR approximation [68] or, more recently, the ALSO-X algorithm [69], [70] which has been shown to be a tighter convex approximation than CVaR. Approximation methods based on non-linear programming, which creates a smoothed, differentiable version of the step function, include [71] and [72].

An important challenge of the sample average approximation is that the number of considered samples may be too small to rigorously guarantee that the chance constraint is satisfied. Specifically, the *empirically observed* violation probability (i.e., the percentage of samples that are violated) may be different from the *true* violation probability since the chosen set of samples may not accurately represent the underlying distribution. It is possible to introduce an a posteriori test to assess whether a solution satisfies a desired violation probability [67]; however, it is challenging to know how to improve the solution if it fails to pass the test [66].

Scenario Approach: The scenario approach [73] is another sample-based method for solving joint chance constraints, which is primarily applied to convex optimization problems. It does not make any assumptions about the underlying distribution, except for the availability of N *i.i.d.* samples of , where the value of N depends on the number of decision variables in the problem. The main difference between the sample average approximation and the scenario approach is that the scenario approach enforces the constraints for *all* the samples (i.e., it is targeting a 0% empirical violation probability). Note that although the number of samples needed can be very large for large problems, it is typically easier to solve a problem with the scenario approach than with sample average approximation. This is because enforcing all constraints does not require consideration of binary variables. Another benefit of the scenario approach is that it provides a priori probabilistic guarantees on solution performance, i.e., this approach guarantees that the *true* violation probability of the joint chance constraint is below ϵ.

However, the scenario approach also tends to give conservative results in power system problems [30], [74]. Specifically, by requiring constraint satisfaction for all samples, it tends to be very sensitive to the worst-case samples $\xi_{\mathbf{s}}$ within the sample set and typically finds solutions that have a much lower violation probability than the acceptable violation probability ϵ. As a result, the scenario approach may fail to find a feasible solution even for problems where one exists. It is also interesting that the scenario approach has a curiously adversarial relationship to data. Where other approaches tend to find better solutions when more data and thus information about the underlying distribution becomes available, the solutions obtained with the scenario approach can only become more

conservative if more samples are added as each sample adds another set of constraints that need to be satisfied.

Another drawback of the scenario approach is that it cannot handle scenario-dependent second-stage variables $\mathbf{y_s}$, but requires the use of a recourse policy to represent second-stage decisions. This is because the required number of scenarios N depends on the number of decision variables. As a result, introducing new variables for each scenario would also require us to update the number of scenarios. Furthermore, the original scenario approach only applies to convex problems. This is partially remedied by [75], which uses the scenario approach to define a robust uncertainty set, and by [76], which provides a posteriori guarantees for non-convex problems.

Moment-based Reformulations: A common method for reformulating single chance constraints into a tractable form is to use a moment-based reformulation. This method typically relies on affine recourse policies in order to express the reaction to uncertainty, and uses the first and second moments (i.e., the mean and standard deviation) of $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ to represent (5) as

$$\mu_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) + \rho(1 \quad \epsilon)\sigma_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) \quad \mathbf{0}, \tag{27}$$

where $\mu_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ and $\sigma_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ indicate the mean and standard deviation of $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$, respectively. In (27), the function $\rho(1 \quad \epsilon)$ represents our knowledge and/or assumptions regarding the distribution of $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$, which is related to, but not the same as, the distribution of . When evaluated for our chosen ϵ, the function $\rho(1 \quad \epsilon)$ is a constant input parameter to our model. Generally, a smaller violation probability ϵ will lead to a larger value of $\rho(1 \quad \epsilon)$. Note that since the standard deviation $\sigma_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ is always positive, a larger value of $\rho(1 \quad \epsilon)$ indicates that the constraint becomes more conservative.

We next discuss several options for defining the function $\rho(1 \quad \epsilon)$. If follows a multivariate normal distribution and $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ is an affine function of , then $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ will be normally distributed with mean (29) and standard deviation (30), respectively. In this case, by choosing $\rho(1 \quad \epsilon)$ as the inverse cumulative distribution function of the standard normal distribution evaluated at $1 \quad \epsilon$, the moment-based reformulation (27) is exact (i.e., it allows the violation probability to reach, but not exceed ϵ and provides the least conservative value of $\rho(1 \quad \epsilon)$). However, the assumption of a normal distribution may be overly restrictive in most practical applications. To address such cases, it is possible to define $\rho(1 \quad \epsilon)$ to guarantee the satisfaction for a family of possible distributions, giving rise to distributionally robust chance constraints. While still maintaining the same tractable form as in (27), we can define $\rho(1 \quad \epsilon)$ to guarantee security for all distributions which share the same moments [77] or add additional assumptions such as symmetry or unimodality [78]. Examples of several such reformulations for an optimal power flow problem are given in [79]. Alternatively, it is possible to use data to tune $\rho(1 \quad \epsilon)$ to achieve the desired performance guarantees [66]. Other methods to handle more general cases of distributionally robust chance constraints are discussed below.

The moment-based reformulation (27) holds for general distributions and constraint functions $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$. In the general case, it can be challenging to evaluate the mean and standard deviation of the constraint functions, though methods such as polynomial chaos expansion can be applied [80]. However, if $g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ is an affine function of , i.e.,

$$g_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) = a(\mathbf{x}, \mathbf{y})\xi + b(\mathbf{x}, \mathbf{y}), \tag{28}$$

we can express the mean and standard deviation as

$$\mu_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) = a(\mathbf{x}, \mathbf{y}) \quad + b(\mathbf{x}, \mathbf{y}), \tag{29}$$

$$\sigma_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) = \sqrt{a(\mathbf{x}, \mathbf{y}) \quad a(\mathbf{x}, \mathbf{y})^{\top}}, \tag{30}$$

where μ and represent the mean vector and the covariance matrix of , respectively. We note that the above equations hold for any distribution of , as long as and are finite. Furthermore, if the coefficients $a(\mathbf{x}, \mathbf{y})$ are linear in the decision variables $\mathbf{x}$ and $\mathbf{y}$, then mean (29) is a linear function and the standard deviation $\sigma_i^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )$ is a second-order cone. In this case, the reformulated constraints (27) are second-order cone constraints, which are convex and give rise to scalable and efficiently solvable optimization problems [81].

E. Reformulations of Problems with Distributionally Robust Objective or Constraints

A common approach for solving distributionally robust optimization problems is to use affine recourse policies, as described above in Section IV-A. Using these policies in the distributionally robust optimization problem (3), the worst-case expected cost $\max_{P \in A} \mathsf{E}_P\left[f^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ )\right]$ in the objective function is reformulated as $\max_{P \in A} \mathsf{E}_P\left[f^{\mathsf{S}}(\mathbf{y_{(1)}} \quad + \mathbf{y_{(0)}})\right]$, where the parameters $\mathbf{y_{(1)}}$, $\mathbf{y_{(0)}}$ are first-stage variables shared among all scenarios. If the ambiguity set A is represented using the moment-based approach, where we assume that the values of the mean and covariance are exactly known, the min-max objective function (3a) can be straightforwardly reformulated as a standard objective function, which minimizes a quadratic term excluding [82]. If the ambiguity set A is represented using an approach based on the Wasserstein metric, the objective function reformulation requires a more complicated mathematical procedure, as thoroughly explained in [83] (Theorem 4.2). For more information on distributionally robust optimization, we refer the reader to [84] and [85].

In the case of a distributionally robust chance constraint with affine policies $\min_{P \in A} \mathsf{P}\left(g^{\mathsf{S}}(\mathbf{x}, \mathbf{y_{(1)}} \quad + \mathbf{y_{(0)}}, \ ) \quad \mathbf{0}\right) \quad 1 \quad \epsilon$, a common approach is to conservatively approximate it by a distributionally robust CVaR constraint, i.e.,

$$\max_{P \in A} \mathrm{CVaR}_P^{\epsilon}\left(g^{\mathsf{S}}(\mathbf{x}, \mathbf{y_{(1)}} \quad + \mathbf{y_{(0)}}, \ )\right) \quad \mathbf{0}, \tag{31a}$$

where

$$\mathrm{CVaR}_P^{\epsilon}\left(g^{\mathsf{S}}(\mathbf{x}, \mathbf{y_{(1)}} \quad + \mathbf{y_{(0)}}, \ )\right) = \min_{\theta} f\theta + \frac{1}{\epsilon}\mathsf{E}_P[g^{\mathsf{S}}(\mathbf{x}, \mathbf{y_{(1)}} \quad + \mathbf{y_{(0)}}, \ ) \quad \theta]^{+}g, \tag{31b}$$

where $[\,]^+ = \max(0, \,)$. This is a conservative approximation since the CVaR accounts for the violation magnitude and may eventually enforce the constraint with a higher probability than the one imposed in the original distributionally robust chance constraint. Nonetheless, the distributionally robust CVaR constraint guarantees the satisfaction of the original distributionally robust chance constraint. In the case where the first two moments (mean and standard deviation) are known exactly, but no other information is known, this reformulation is exact. We refer the reader to [4] for further details on reformulating the resulting max-min problem obtained by substituting the CVaR definition (31b) in (31a).

F. Probabilistic and Robust Equality Constraints

A particular challenge for problems that arise in many power systems settings (e.g., problems with power flow constraints) is to consider probabilistic equality constraints, such as

$$\mathsf{P}_{\xi}\left(h_j^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) = 0\right) \quad 1 \quad \alpha_j. \tag{32}$$

In power system optimization problems, the equality constraints $h_j^{\mathsf{S}}(\mathbf{x}, \mathbf{y}\ ,\ ) = 0$ typically represent the physical laws that underpin the operation of the electric power system, such as the power flow equations. A violation of these constraints thus indicates that the physical laws which govern the entire model are violated and, as a result, all other values that are derived from those equations (e.g., power flows, voltages, and load served) cannot be trusted. A possible way to circumvent this problem is to enforce the equality constraints as robust constraints, i.e., enforce (9a) instead of (5a). In the case of linear constraints, the equality constraints can be used to substitute state variables (e.g., in DC power flow, we can use the nodal power balance constraints to substitute the voltage angle variables [30]) and enforce conditions on the system recourse (e.g., choose the parameters of an affine recourse policy to ensure that the system remains balanced at all times). In the case of non-linear equality constraints (such as the AC power flow constraints), the representation of equality constraints is more involved. Existing approaches include the use of polynomial chaos expansion to represent the propagation of uncertainty through the non-linear system of constraints [86] and the derivation of robust inner approximations [87]. We discuss the challenges of formulating tractable versions of these constraints for linear DC and non-linear AC power flow constraints in Section XI.

V. Evaluation of Solution Quality

In many problems under uncertainty, solution quality may be limited due to assumptions and model approximations (e.g., assumptions regarding the distribution of , limited numbers of scenarios $\mathbf{s}\ 2\ \mathsf{S}$, and the use of linear approximations for non-linear constraints). An ex-post *out-of-sample* simulation is the most common approach to evaluate the quality of a solution, irrespective of the approach that has been used for decision-making under uncertainty, e.g., stochastic programming, robust optimization, chance-constrained optimization, distributionally robust optimization, etc.

For given sources of uncertainty , consider N samples obtained from a probabilistic forecast or empirical observations, with each sample representing a potential realization of . Arbitrarily split these samples into two separate sets of samples, each with M and K samples, such that $M + K = N$ and $K > M$. We may use the former set with M samples to provide scenarios in stochastic programming, build an uncertainty set in robust optimization, or construct a moment- or metric-based ambiguity set in distributionally robust chance-constrained optimization. This set with M samples is widely referred to as the *seen* or *training* set. Building on this set, in the so-called *in-sample* simulation, we solve the underlying problem, e.g., the two-stage stochastic program (2), the distributionally robust problem (3), the robust problem (4), or any of the formulations with probabilistic or robust constraints, in order to obtain the in-sample value for the objective function and a solution for the decision variables. For example, this value in the case of the two-stage stochastic program (2) is $f^{\mathsf{F}}(\mathbf{x}\ ) + \mathsf{E}_P\left[f^{\mathsf{S}}(\mathbf{x}\ , \mathbf{y}\ ,\ )\right]$, where $\mathbf{x}$ and $\mathbf{y}$ are the optimal values obtained for the first- and the second-stage variables, according to the M samples used in the training set. Denote this in-sample value of the objective function as κ^{ins}.

Now, we fix the value of the first-stage variables to that achieved in the in-sample simulation, i.e., $\mathbf{x}$, and *deterministically* solve the second-stage problem K times, each time using a sample that has not been used in the in-sample simulation. Hence, the set of K samples is called the *unseen* or *testing* set. To ensure a valid evaluation, it is typical to use $K \quad M$. For each unseen sample $_i$, we solve the following deterministic problem:

$$\min_{\mathbf{y}} \ f^{\mathsf{S}}(\mathbf{x}\ ; \mathbf{y}; \ _i) \tag{33a}$$

s.t.

$$h^{\mathsf{S}}(\mathbf{x}\ ; \mathbf{y}; \ _i) = \mathbf{0}; \tag{33b}$$

$$g^{\mathsf{S}}(\mathbf{x}\ ; \mathbf{y}; \ _i) \quad \mathbf{0}: \tag{33c}$$

We can now obtain the out-of-sample value of the objective function, denoted as κ^{oos}, by summing up the first-stage value $f^{\mathsf{F}}(\mathbf{x}\ )$ achieved from the in-sample simulation and the average second-stage value, i.e., the average value of the objective function (33a) obtained over the K solutions to the deterministic problem (33).

The difference between κ^{ins} and κ^{oos} is a metric for the solution quality. A comparatively large difference indicates a lower-quality solution, implying that the training set with M samples has not been sufficient to properly represent the underlying uncertainty. Therefore, the training set should be improved either by increasing the number of samples and/or by selecting more representative samples.

One may also exploit the out-of-sample simulation to compare the performance of different approaches discussed for decision-making under uncertainty and to benchmark against deterministic models such as the perfect information model or the certainty equivalent. The difference in solution quality between the stochastic model and the perfect information model is typically referred to as the *Expected Value of Perfect*

Information (EVPI) since it measures the benefit of having access to the true value of . The difference in solution quality between the stochastic model and the certainty equivalent models is referred to as the *Value of Stochastic Solution* (VSS), and measures the benefit of considering as uncertain parameters rather than a fixed forecast. For more information, we refer to the reader to Section 2.6 of [88].

In addition, since the out-of-sample evaluation involves a sequence of simpler optimization problems (i.e., one for each scenario), we may choose to include more detailed constraints (such as an AC power flow approximation instead of a DC approximation). This can allow us to assess whether the approximation of the constraints significantly impacts the solution quality.

Part II: Existing and Emerging Applications of Power System Optimization under Uncertainty

In the following, we present a selection of established and emerging applications of decision making under uncertainty in electric power systems. The list is not meant to be exhaustive, but was selected to provide examples different problem formulations and solution methods. The applications primarily focus on power system operations and planning, taking the perspective of a transmission system operator. We have also included some applications that discuss the perspective of a hydropower generation company (who need to decide on a bidding strategy or generation schedule) or energy consumers (who need to identify a strategy for energy procurement). However, these examples are not as comprehensive as the coverage of the system operator perspective, and we acknowledge that there is a large part of relevant literature on energy trading and load management that is not covered in this paper.

The sections are loosely related to each other, and can be read either from beginning to end or as stand-alone sections that provide an introduction to individual applications. To help guide the reader in identifying the sections that are most relevant to them, we provide a brief overview of the content in this second part of the paper. This content is also summarized in Table I.

Section VI discusses security-constrained optimal power flow, which focuses on securing the system against uncertainty arising from component outages. This problem is traditionally formulated as a robust optimization problem, where constraints are enforced for all component outages in a predefined set (e.g., all $N \quad 1$ situations). Section VII discusses the question of securing the system against variability and uncertainty arising from renewable energy generation, focusing on chance-constrained optimal power flow. This also includes Section VII-B on distributionally robust formulations. Section VIII discusses chance constraints in the context of multi-area reserve scheduling. Section IX extends the discussion of transmission system operations to day-ahead unit commitment, which includes discrete variables to represent the on/off status of generators and focuses on two-stage stochastic programming formulations. Section X discusses the consideration of uncertainty in the context of long-term transmission expansion planning. The sources of uncertainty in long-term planning are different than in short-term planning as the considered time horizon stretches across multiple years or decades, requiring quantification of long-term uncertainty such as load growth, fuel prices, or climate change projections. Section XI is dedicated to a more detailed discussion of common choices and implications of different power flow formulations, which is a common thread in the above applications. In Section XII, we discuss planning for hydro-thermal power systems, where a main source of uncertainty is the inflow of water to hydropower reservoirs. Finally, Section XIII discusses the question of procuring energy for a large consumer, discussing uncertainty associated with price and contract options.

VI. Security-Constrained Optimal Power Flow

Contingencies due to transmission line, transformer, and generator failures can result in large-scale blackouts. Shortly after the optimal power flow problem was first formulated by Carpentier [89], a major blackout in the Northeastern United States and Ontario, Canada in 1965 motivated system operators to explicitly consider contingencies in operational planning. Uncertainty and risk arising from contingencies is traditionally managed by enforcing constraints on the post-contingency operating conditions, resulting in the so-called *Security-Constrained Optimal Power Flow* (SCOPF) problem. The SCOPF problem was first proposed in [90] and continues to receive significant attention due to remaining computational challenges [91]–[93] and the evolving ability to assess the risk and mitigate the impact of contingencies [94]–[97]. Contingencies are discrete events with a certain probability of occurring (individually or in combination). Due to the consideration of potential future contingency events, SCOPF is, at heart, an optimization problem under uncertainty.

The traditional form of the SCOPF problem seeks to identify an operating point which ensures that the system can continue to operate without significant disruption despite experiencing any contingency included in the so-called *contingency list* [98]. Contingency lists typically include the failure of any individual component as specified by the $N \quad 1$ security criterion, but are also frequently expanded to incorporate other events that are considered likely (i.e., have a sufficiently high probability of occurring). Likely contingencies may include, e.g., common mode failures [99] or multiple simultaneous failures in situations where external circumstances such as, e.g., severe weather or terrorist threats increase the probability of such contingencies [100], [101]. The question of which contingencies to include can be understood as the problem of defining an uncertainty set for a robust optimization problem.

While the traditional form of this problem assumes a preventive approach where system set-points are determined to keep the system safe if any of the contingencies occurs [102], the SCOPF problem can also be formulated to allow for post-contingency control actions (often referred to as *corrective* or *remedial* actions) that are taken in response to a specific contingency [103]. Thus, the security-constrained optimal power

TABLE I: Overview of applications and their main characteristics

Application	Problem type (time scale)	Main uncertainty source(s)	Uncertainty representation	Goal of the problem	Optimization formulation	Reformulations/ solution methods
Security-constrained optimal power flow	Power system operations (day-ahead to real-time)	Contingencies (finite set of discrete events)	Robust uncertainty set with selected contingencies	Guarantee post-contingency constraint satisfaction for all contingencies	Robust constraint satisfaction	Algorithms for robust optimization
Chance-constrained optimal power flow	Power system operations (day-ahead to real-time)	Renewable energy (continuous variations in generation)	Probability distribution, moments, or scenarios	Limit probability of constraint violations	Chance-constrained optimization	Various reformulations
Distributionally robust optimal power flow	Power system operations (day-ahead to real-time)	Renewable energy (continuous variations in generation)	Family of probability distributions	Limit the worst-case cost (or violation probability) among family of distributions	Distributionally robust optimization	Moment-based or metric-based reformulations
Multi-area reserve dimensioning	Power system operations (months)	Total power imbalance	Probability distributions, quantiles	Limit probability of running out of reserves	Chance-constrained optimization	Analytic or scenario-based
Stochastic unit commitment	Power system operations (day-ahead)	Renewable energy and contingencies	Scenarios	Minimize expected cost of operations	Two-stage stochastic optimization	Algorithms for two-stage stochastic optimization
Transmission expansion planning	Power system long-term planning (decades)	*Long-term uncertainty* (e.g., climate change, technology trends) and *short-term uncertainty* (renewable generation)	Scenarios	Minimize expected capital cost (CAPEX) and operational cost (OPEX)	Two- or multi-stage stochastic programming	Variety of algorithms
Choice of power flow formulation	Any problem with power flow constraints	Renewable energy (continuous variations in generation)	Focused on continuous uncertainty	Model propagation of uncertainty	Chance-constrained, distributionally robust, robust	Convex restrictions, polynomial chaos expansion, others
Planning for hydro-thermal power systems	Generation planning (yearly)	Reservoir inflows, future energy prices	Scenarios	Optimal generation schedule and "value of water"	Multi-stage stochastic programming	Stochastic dual dynamic programming
Energy procurement	Consumer perspective (months)	Future energy prices	Scenarios	Minimize expected cost of energy across multiple markets	Two-stage stochastic optimization	Algorithms for two-stage stochastic optimization

flow problem is equivalent to an (adaptive) robust optimization problem where the set of considered contingencies represent the uncertainty set. The post-contingency control actions, if included, represent the adaptive recourse actions. It is worth noting that this uncertainty set has finite support, i.e., it consists of a finite number of possible realizations, and the problem can thus be solved by enforcing constraints for each contingency in the specified set. Although the SCOPF problem is often not explicitly characterized as a robust optimization problem, many of the methods that are used to solve the SCOPF, including contingency screening [104], [105] and methods to iteratively build a set of constraints representing worst-case post-contingency operating conditions with load and generation uncertainty [106], [107], are similar to techniques used to solve other robust optimization problems.

Recently, there has been an increasing interest in rethinking whether the traditional SCOPF model is the most appropriate way to mitigate the risk of contingencies. In particular, the definition of a contingency list relies on a vaguely defined notion of "likely contingencies" and the SCOPF problem ensures that no constraint violations will occur due to any contingency on this list. This, on the one hand, can lead to significant resources being spent on mitigating the effects of contingencies that are not particularly dangerous, such as cascading events that do not lead to load loss and only impact a few transmission elements. In such cases, ensuring $N \quad 1$ security may not always be socio-economically desirable. On the other hand, defining the list of $N \quad 1$ events based on the likelihood of occurrence may leave out potentially important low-probability, high-impact events, such as common-mode failures (i.e., multiple failures due to a single cause) that could cause a large-scale blackout. A more comprehensive approach for mitigating the risk of contingencies explicitly considers both the probability and the impact of contingencies. One line of research develops risk functions that describe the risk of post-contingency component overloads [94]. Extensions of this work on risk functions allow the risk to be controlled by FACTS devices [108] and accounted for price formation [109]. Other extensions generalize this work to consider uncertainties in the cost and availability of remedial actions [96]. The inclusion of risk functions make the problem harder to solve, motivating the need for solution algorithms based on relaxation and decomposition [110]. Other researchers take a more ambitious view and aim to simulate actual load shed following a contingency [111]. This is challenging because it might require simulating multiple cascading steps, which has been addressed in [111] by only considering events that significantly contribute to system risk. A related set of works challenge the idea that preventive and corrective control are equally secure by explicitly modeling the probability and impact of corrective action failures [112]–[114].

Although risk-based approaches promise a more comprehensive view of system security, there are several significant challenges that remain. For example, risk-based approaches

can cause the complexity of the problem to explode, making the development of scalable solution algorithms challenging. Furthermore, the additional input data to the problem (i.e., the probability of contingencies and the impact of load shed) is more challenging to obtain, and requires careful consideration as it may significantly impact the results.

VII. Chance-Constrained Optimal Power Flow

Chance-constrained optimal power flow problems consider the problem of how to dispatch generators in day-to-day operations while ensuring that system constraints will be satisfied with a specified probability despite uncertainty in load and renewable generation. Choosing an acceptable violation probability is perceived as an intuitive and transparent way of determining a probabilistic security level by transmission operators [115], and this practice also aligns well with established industry practice for reserve dimensioning [116], [117] and the definition of reliability margins in European market coupling [118].

The problem has been formulated both with or without consideration of contingencies. The simultaneous consideration of chance constraints and (robust) security constraints to secure the system against both renewable variability and contingencies is necessary to ensure system security in practical operations. When security constraints are included, the violation probability does not represent the probability of an actual, physical system overload (such as the overload of a transmission line), but rather the probability of violation of the $N \quad 1$ security constraints. The physical overload would only be realized if the chance-constraint violation happens simultaneously with the contingency. However, many countries regulate transmission system operation to always ensure $N \quad 1$ security. As such, the chance constraint violation probability could be interpreted as the probability of complying with such regulations without further control action.

The uncertainty from renewable energy impacts both generation and transmission constraints. The generators must reserve some capacity to be able to balance the system as renewable generation varies, and transmission capacity has to be limited to accommodate changes in power flow as power injections from both renewable and conventional generators vary. The reduction in available generation and transmission capacity can be understood as an uncertainty margin [119], i.e., a security margin against uncertainty. While early papers on chance-constrained optimal power flow tended to leave system balancing to the slack bus [120], Vrakopoulou et al. [121] proposed modeling these adjustments via an affine control policy which uses participation factors to distribute the power imbalance across several generators in the system. This model is very close to practical operation in the electric grid as it mimics the use of automatic generation control and has since been widely adopted. The participation factors can be treated as known parameters, based, e.g., on the total generation capacity of individual generators [122] or they can be optimized as part of the problem [81]. The affine control policy can be extended to consider that generators may react differently to wind power plants in different locations, thus providing better ability to manage congestion and transmission constraints [123], [124]. A significant benefit of the affine control policy is that it is easy to implement the control policy in real-time operations. Furthermore, the use of an affine control policy allows us to reformulate the chance-constrained optimal power flow problem with a single chance constraint on generators and transmission lines as a second-order cone program using a moment-based reformulation [81], which reduces to a linear program in the case of fixed participation factors [122]. With the integration of unit commitment and $N \quad 1$ constraints [125], the same problem becomes an mixed-integer second-order cone problem that can be solved using a modified Benders decomposition algorithm. Alternatively, using the scenario approach, we can reformulate the chance-constrained OPF with affine control policies and joint chance constraints as a linear program, though the formulation becomes bi-linear if generator outages are considered as part of the constraints [121]. Others have solved the joint chance constrained optimal power flow using a sample average approach, leveraging either non-linear programming [71] or a combination of constraint screening and bound tightening based on tight valid inequalities [126] to obtain a scalable formulation. Many successful solution algorithms [81], [125], [127] take advantage of the fact that only a few of the power flow constraints in the optimal power flow problem are binding (i.e., there are only a few transmission lines that are congested).

A. More Complex Controls and Recourse Policies

Although the affine control policy is an accurate and appropriate model for small power imbalances, it is not realistic if the imbalances are large. In these cases, it is necessary to consider more general generation control policies, which include saturation of reserve capacity (i.e., that generators will stop contributing reserves when they reach maximum capacity [128]), integration of tertiary reserve activation [22], and curtailment of wind power generation above a certain threshold [33]. Furthermore, using generators for balancing and congestion management is only one option. Several other options for mitigating the impact of uncertainty on the system have been proposed in the context of chance-constrained optimal power flow, including demand-side management with thermostatically controlled loads [129], [130] and electric vehicle charging [131]. Other lines of work have proposed changing the settings of phase-shifting transformers or HVDC lines [127], [132], [133] to better control power flows and implementing dynamic line rating [134]. Some methods also consider chance-constrained optimal power flow as part of multi-energy system models, such as natural gas [135] and water distribution systems [136].

B. Distributionally Robust Formulations

The papers on chance-constrained optimal power flow mentioned above consider a range of different reformulation strategies. To highlight the use of distributionally robust formulations, we provide a brief review. Reference [137]

develops a distributionally robust optimal power flow that is robust to ambiguity in the first and second moment and proposes a solution approach based on cutting planes. Reference [138] develops a distributionally robust optimal power flow problem with dynamic line rating, exploring both moment-based and Wasserstein metric-based approaches. The resulting model, after reformulations, is a convex conic program. Reference [139] develops a moment-based distributionally robust optimal power flow problem, accounting for uncertain renewable power generation as well as uncertain reserve provided by flexible loads. Under different schemes and reformulations, the resulting model is either a semidefinite or a second-order cone program. A similar model is developed in [140] for radial distribution systems, where a distribution system operator optimizes grid operation taking into account uncertain power injections from distributed energy resources. A Wasserstein distributionally robust optimal power flow model with uncertain renewable power supply is proposed in [141]. Reference [142] proposes a similar model but considers joint chance constraints. In addition, [143] proposes a similar model to [141] that accounts for an approximate model of the AC power flow equations. Reference [144] develops a multi-stage Wasserstein distributionally robust optimal power flow problem. Finally, [145] proposes a Wasserstein distributionally robust optimal power flow problem with an exact reformulation.

VIII. Multi-Area Reserve Dimensioning

To make up for situations where renewable generators provide less power than expected, generators need to carry sufficient reserve capacity. The dimensioning (also referred to as sizing) of reserve is the problem of determining how much reserve capacity the system needs while accounting for the size of uncertain disturbances (e.g., forecast errors in renewable generation and load demands) as well as line and generator contingencies, i.e., *composite uncertainty* [27], [146].

Bottom-up reserve sizing models trade off the economic costs of committing and operating reserves against the reliability that these reserves afford. Recent European legislation, in particular article 157 of the System Operation Guideline (SOGL) [147], as well as established industry practice [116], [117] motivate reserve sizing methodologies on the basis of probabilistic criteria. In the absence of network constraints, the problem of reserve sizing amounts to estimating quantiles of capacity shortfall [148]–[151].

In the presence of network constraints, the problem can be cast as a chance-constrained optimization [121], [152]. The first-stage decisions amount to the allocation of reserve capacity in different areas and to different generators. In the second stage, the power imbalance is revealed and the reserves are activated with the goal of balancing the system while respecting network constraints. The objective function minimizes the total amount of reserve capacity that the system carries. Probabilistic constraints can be represented using binary variables that indicate whether a scenario corresponds to reliable operation or not, in the spirit of [153]. The formulation bears similarities to formulations for reserve deliverability in US market clearing models [154]–[158] as well as reserve dimensioning in chance-constrained optimal power flow [121], [159], [160].

IX. Stochastic Unit Commitment

The stochastic unit commitment problem aims to determine the optimal day-ahead commitment of generators in order to operate the system at minimum expected cost. This problem considers both the fixed cost of committing generators and the variable cost of dispatching them based on realized uncertainty. The problem is typically formulated as a two-stage stochastic program, with the earliest such formulations proposed by [161] and [162] in the mid-1990s. The problem recently experienced a significant resurgence in interest as a means of quantifying the impact of renewable resource integration on power system operating costs [163], [164] and reserve requirements [165], [166]. Apart from its use as a policy analysis tool, the model has also been applied as an ideal benchmark for market products such as flexiramp [167]. The ambition of using the model as an operational tool in the context of security-constrained unit commitment [168] has been tamed by the complex input and heavy computational requirements and the fact that alternative formulations are better suited to the conservative nature of system operations [169].

The stochastic unit commitment problem is typically formulated as a two-stage decision making problem under uncertainty. In the first stage, units are committed, and then uncertainty is revealed in the form of realized forecast errors and equipment (generator and line) outages. In the second stage, the system is allowed to react by dispatching generators in order to balance the system while respecting network constraints [27], [170]. The model has also been used for generating a policy in rolling simulations [163], [164]. Alternatives that account for uncertainty include robust unit commitment [169], [171], adaptive robust optimization [172] as well as hybrid [173] and chance-constrained [125] formulations.

The stochastic programming formulation amounts to a large-scale mixed-integer stochastic program. A typical approach for solving this problem relaxes the non-anticipativity constraints and then applies Lagrange relaxation combined with feasible recovery heuristics [27]. This process can be parallelized, in either synchronous [28] or asynchronous [57] settings. Alternative dual-based methods have been employed, including augmented Lagrangian methods [162] and progressive hedging [60], [161], [174], which is a specific instance of proximal methods [59]. Benders decomposition [175], and more generally bundle methods [176], have also been employed. The application of stochastic dual dynamic programming to a multi-stage version of the problem has also been proposed [41].

X. Transmission Expansion Planning

The Transmission Expansion Planning (TEP) problem aims to determine the best transmission lines and other equipment to add to a high-voltage power grid in order to support its future operation. The TEP problem accounts for the CAPEX (capital expenditure) of projects related to new transmission lines, the

updates of existing ones, and other equipment upgrades, as well as the OPEX (operational expenditures) related to the daily operation of the power grid for the lifetime of these projects [177]. While the CAPEX is often relatively easy to compute due to the proximity to the actual investments, the OPEX is subject to many stochastic factors such as the price of fuels, the availability of hydrological resources, the cost of storage technology, and investments in new generation capacity. The stochastic nature of the TEP problem is one of the main sources of modeling and computational complexity.

1) Problem structure: The canonical mathematical representation of the TEP problem is usually defined as a centralized decision-making problem under uncertainty [177]. The simplest and most common formulation is a two-stage stochastic program, where the first stage represents the investment decisions, $\mathbf{x}$, while the second stage represents the operational decisions, $\mathbf{y}$, for the final network configuration in a target year. The multi-stage stochastic version of the TEP problem, also known as the dynamic TEP problem [178], has also been extensively addressed in the TEP literature. In this framework, the decisions are optimized for a horizon of N stages, where uncertainty realizations are revealed along each stage.

Conceptually, the TEP problem can be represented as:

$$\min_{\mathbf{x},\mathbf{y}} \ \text{CAPEX}(\mathbf{x}) + \mathsf{E}\big[\text{OPEX}(\mathbf{x},\mathbf{y}\ ,\ )\big] \tag{34a}$$

$$\text{s.t.: Investment restrictions } (\mathbf{x}) \tag{34b}$$

$$\text{Network flow equations } (\mathbf{x},\mathbf{y}\ ,\ ) \tag{34c}$$

$$\text{Network capacity limits } (\mathbf{x},\mathbf{y}\ ,\ ) \tag{34d}$$

$$\text{Generator limits } (\mathbf{y}\ ,\ ) \tag{34e}$$

Here, the main TEP constraints consist of restrictions on investment projects such as budget limitations or available projects and technologies for investment. Power flow equations are used to model the physical operation for the final grid configuration. The most common approach is using a DC approximation of the power flow equations (see Section XI-2 below) to leverage the scalability of linear models. Limits related to transmission lines and generation capacity are also part of the usual TEP constraints.

The transmission grid is the backbone of the infrastructure for delivering power from multiple generation sources, which forms an important part of all modern societies. As a result, the TEP problem reflects not only technical aspects of grid operation, but also energy policy and broader priorities of the society around the grid. Thus, *drivers* for optimal grid updates are not only economic efficiency, i.e., minimal CAPEX and OPEX. The TEP problem typically captures, via constraints or penalization terms in the objective function, issues such as (i) renewable target commitments [179], (ii) operational reliability such as $N \quad 1$ security criteria [179], (iii) climate-aware planning [180], (iv) new generation expansion capacity [181], and (v) incentives and policies [182], among others.

These additional drivers become particularly important in grids where zero or near-zero marginal cost generation takes a larger share of the total demand such that the energy cost alone could be insufficient for providing the right signals for the new network infrastructure updates.

2) Uncertainty modeling: There are many uncertainty sources for the TEP problem. Considering these uncertainties is essential for obtaining consistent TEP solutions. It is crucial to distinguish between two types of uncertainty components [183], [184].

First, the *long-term component* unfolds over many years, e.g., new renewable capacity investments, demand growth, availability of hydro resources, climate variability, etc. This uncertainty is realized only once during the project lifetime (i.e., we will only observe one realization of this uncertainty). The long-term component of uncertainty is a fundamental part of the standard analysis used in existing power systems such WECC, ERCOT, CAISO, and the UK National Grid. It is part of the what-if analysis (also known as scenario analysis) of long-term visions of the future portrayed by stakeholders.

Second, the *short-term component*, e.g., wind production, yearly demand, fuel price, etc., accounts for the variability of stochastic input parameters and is expected to be observed as a large number of realizations that impact OPEX. In many situations, the short-term component can be conditional on the long-term component, such as climate change and renewable resource availability [184]. The short-term component is more commonly considered in the academic environment to represent the infinite or large number of possible stochastic operational states we may find in the future.

The larger the range of scenarios for representing uncertain parameters, the better the approximation of the estimated expected OPEX. However, using a large number of scenarios for representing uncertainty could compromise computational tractability.

3) State-of-the-art: During the last three decades, there have been significant contributions to the modeling, uncertainty characterization, and solution methodologies for TEP problems [185]. Early models based on a transshipment formulation of power flow, i.e., neglecting Kirchhoff's voltage law and using forecast peak demand, were sufficient for conventional power grids with a large mix of dispatchable generators. The decarbonization of the power grid has boosted an increase in renewable generation, the introduction of new business models considering the demand side, and new distributed generation resources, among others. This new panorama makes uncertainty characterization and mathematical frameworks important, but challenging for network planners. We refer to the monographs [177], [186], surveys [185], [187], and other literature [183] for further details on TEP modeling under uncertainty.

While advances in the TEP problem modeling and formulation has brought a deeper understanding and new insights, this complexity has increased the computational burden of TEP models. This increases the complexity of already large- or very-large-scale optimization models under uncertainty, making computational tractability one of the principal challenges for further improvements. This requires the development of advanced solution methodologies using ad hoc algorithms

supported by standard methods such as Benders decomposition, column-and-constraint generation, progressive hedging, and stochastic dual dynamic programming, among others. Aggregation methods [188] are used to reduce the size of TEP mathematical models by grouping similar objects, like network nodes, generating units, or RES profiles, into a single entity. While this leads to a loss of information and reduces modeling accuracy, the computational benefits are significant enough to compensate for the degradation in accuracy.

XI. Choice of Power Flow Formulation

When optimizing system operations in the presence of power injection uncertainty (e.g., stochastic load demands and renewable generation), one must characterize how this uncertainty propagates through the power system to produce uncertainty regarding quantities such as voltages and transmission line flows (which must be kept within bounds). With this information, we can formulate optimization problems that limit the potentially negative impacts from uncertainties.

The propagation of uncertainties in power injections to uncertainties in voltages and power flows is dictated by the power flow equations. The choice of power flow formulation (i.e., whether we use the full non-linear AC power flow equations, a linearized version, a convex relaxation, or a convex restriction) significantly impacts the complexity of the problem. We review next some of the challenges and solution approaches associated with using different power flow formulations in the context of power injection uncertainty.

1) Non-Linear AC Power Flow: The non-linear AC power flow equations relate the active and reactive power injections to the voltages phasors for each bus $i \in N$, where N denotes the set of buses. There are many different ways of representing the AC power flow equations [189]. If we choose polar coordinates for the voltage phasors at bus i, $V_i \angle \theta_i$, the AC power flow equations are

$$P_i = V_i \sum_{k \in N} V_k \left(\mathbf{G}_{ik} \cos(\theta_i - \theta_k) + \mathbf{B}_{ik} \sin(\theta_i - \theta_k) \right), \tag{35a}$$

$$Q_i = V_i \sum_{k \in N} V_k \left(\mathbf{G}_{ik} \sin(\theta_i - \theta_k) - \mathbf{B}_{ik} \cos(\theta_i - \theta_k) \right), \tag{35b}$$

where $\mathbf{Y} = \mathbf{G} + j\mathbf{B}$ is the network admittance matrix and $P_i + jQ_i$ is the complex power injection at bus $i \in N$.

The non-linearity of the AC power flow equations gives rise to non-convex and possibly disconnected feasible regions [190], resulting in many optimization problems being NP-Hard [191], [192] from a theoretical perspective. Note, however, that despite lacking guarantees, recent benchmarking of local solvers such as Ipopt [193] indicate that they tend to often find solutions that are globally optimal [194]. Furthermore, the non-linearity of the power flow equations makes *uncertainty quantification* difficult. Even if the probability distribution of the power injections is known, the non-linearity of the AC power flow equations makes it highly non-trivial to compute probability distributions of voltage magnitudes and line flows. In particular, the non-linear relationships in the AC power flow equations do not preserve the form of the probability distributions even for Gaussian distributions (e.g., Gaussian distributions of power injections generally lead to non-Gaussian distributions of line flows). Computing moments such as the mean and standard deviation of a line flow, which is straightforward with linear equations, is numerically challenging. Moreover, the implicit nature of the AC power flow equations precludes an explicit representation of voltage magnitudes, line flows, etc. in terms of uncertain power injections, thus introducing non-linear equality constraints involving uncertainties into the formulation.

Polynomial Chaos Expansion (PCE) methods provide a promising approach for addressing these challenges. PCE methods propagate uncertainty distributions through non-linear equations by decomposing onto a set of non-linear basis functions, resulting in a hierarchy of increasingly accurate, but more computationally challenging problems. PCE was first applied to power flow problems [86], [195] and then to chance-constrained optimal power flow problems [80], with modeling, computational, and other improvements subsequently developed in [196]–[199]. Despite these improvements, PCE methods remain computationally challenging.

Uncertainty propagation with the non-linear AC power flow equations is further complicated by the fact that these equations may have no solutions [200], [201] or have multiple solutions [202] for a given set of power injections. The lack of an AC power flow solution implies that there is no steady-state operating point for this set of power injections, indicating an imminent threat of instability and voltage collapse [203], [204]. It is difficult to assess which (if any) uncertainty realizations will result in power flow insolvability. Many approaches ignore this issue and assume (implicitly or explicitly) that the AC power flow equations will remain solvable for all uncertainty realizations as long as they hold for a nominal operating point, e.g., [74], [205]–[208]. Other approaches seek to identify worst-case operating conditions (i.e., the realizations that lead to the largest constraint violations) [106], [209] or bound the worst-case impacts of uncertainty [210], [211], but do so under the assumption that the AC power flow equations are solvable for every uncertainty realization. Thus, these methods are unable to identify (unstable) uncertainty realizations for which there exists no AC power flow solution. It is also not known how PCE would perform (i.e., how accurate the results would be) if applied to problems that are close to voltage instability.

Researchers have made progress in addressing these challenges for certain classes of robust optimal power flow problems, including systems with generators or controllable loads at every bus [212] and, using theory from [213], three-phase radial networks representing distribution systems [214]. More recently, progress on this topic in [87] considers general system models by leveraging so-called *convex restrictions*, i.e., convex inner approximations of the AC feasible region [215], [216]. By solving problems which guarantee that all uncertainty realizations are contained within a convex restriction,

these approaches guarantee power flow solvability and can also be formulated to incorporate operational constraints such as power flow limits and bounds on voltage variables [87], [217]. We also note that some sample-based approaches provide a posterori probabilistic guarantees for feasibility with respect to the non-linear AC power flow equations [218], [219].

Although there may be multiple solutions to the power flow equations, there is often a single "high-voltage" solution (i.e., there is a single solution where voltage magnitudes at every bus are within normal operating ranges). However, this is not always the case [220], particularly in systems with high penetrations of distributed energy resources [221]. Furthermore, many algorithms for optimization under uncertainty, e.g., [74], [205]–[207], allow for violations of voltage constraints for a small number of realizations, meaning that the solution algorithms may find undesirable low-voltage solutions. The question of how to handle the multiple solutions that may arise has not yet been solved and is not even frequently considered.

2) Linearized Power Flow Formulations: An alternative approach is to use a linearized version of the power flow equations. For instance, the commonly used DC power flow approximation [222] employs assumptions typical of lightly loaded transmission networks (small angle differences, near-nominal voltages, and a lossless system) to simplify the AC power flow equations to a linear formulation:

$$P_i = \sum_{k2N} \mathbf{B}_{ik} \left(\theta_i \quad \theta_k\right). \tag{36}$$

Many stochastic optimization formulations use the DC power flow approximation to provide rigorous solution methodologies with quality guarantees regarding the approximated problem, e.g., [66], [81], [121], [122], [223], [224]. However, while suitable for many applications, the DC power flow approximation can induce significant errors in the solutions to certain problems [225]–[231]. Quality guarantees for an operating point obtained via the DC power flow approximation do not ensure feasibility or optimality with respect to problems using the AC power flow equations.

To ameliorate these issues, researchers have developed many other power flow linearizations (see [189] for a comprehensive review) and applied them to optimal power flow under uncertainty. Examples include the first-order Taylor approximation [206] around the solution to a deterministic problem where the random variables are replaced by their means [206] as well as linearizations tailored for distribution systems [232]. The approach in [74] uses the full AC power flow equations for the nominal operating point, but models the impact of uncertainty using an iteratively updated first-order Taylor expansion (implicitly assuming that the uncertainty is small).

3) Convex Relaxations: A *convex relaxation* encloses a non-convex feasible region within a larger convex region using carefully formulated constraints that are less restrictive than the non-linear AC power flow equations. There exist many different convex relaxations, usually based on semidefinite programming (SDP) [233]–[235] and second-order cone programming (SOCP) [236], [237], with many variants [189]. Convex relaxations give rigorous guarantees on solution quality (i.e., upper or lower bounds on the objective value and infeasibility certificates) and, if the optimal solution to the convex relaxation happens to be feasible for the original problem, they also provide the globally optimal solution.

To bypass challenges associated with non-convexities from the AC power flow equations, many researchers have leveraged convex relaxations in stochastic optimal power flow problems. Some approaches directly replace the non-linear power flow equations with a relaxation to obtain a convex formulation that is suitable for standard techniques, e.g., methods for chance-constrained optimization using various reformulations [238], [239] and scenario-based methods [240]. Other approaches repeatedly solve relaxations within iterative algorithms for robust optimal power flow problems. Specifically, the approach in [209] uses relaxations to compute candidates for worst-case operating points. Additionally, the approach in [210] uses relaxations to bound the worst-case impacts of uncertainty with respect to the inequality constraints to provide guarantees with respect to inequality constraint satisfaction.

Significant research efforts have focused on assessing when convex relaxations yield globally optimal solutions to optimal power flow problems [189], [241]. However, it is worth noting that these results rely on both the tightness of the relaxation itself (i.e., how closely the relaxation approximates the true non-convex feasible space) and the objective function (which determines in which part of the feasible space we require tightness). In optimal power flow problems under uncertainty, where we are interested in feasibility and optimality not for a single operating point but a range of uncertain operating conditions, it is less clear whether we can expect that the relaxations are tight for all uncertainty realizations. Accordingly, the solution to a relaxed stochastic or robust optimal power flow problem may not be feasible with respect to the AC power flow equations. Likewise, the worst-case uncertainty realizations computed via a convex relaxation may not actually be the worst-case uncertainty realizations for the original problem. Thus, similar to power flow linearizations, guarantees of solution quality obtained using a relaxation may not apply to the original (non-relaxed) problem and care must be taken in their interpretation.

XII. Planning for Hydro-Thermal Power Systems

The key source of uncertainty in hydropower planning is the inflow of water, which determines current and future ability to produce power. Given that the inflow of water carries strong seasonal patterns that are correlated with snow melt and rainy seasons, it is important to consider the use of water on a yearly horizon. Across this horizon, the realized water inflows as well as decisions on how much power to produce can be updated on a regular basis. This naturally gives rise to *multi-stage stochastic programming problems*, with the most common algorithm being stochastic dual dynamic programming (SDDP). Pioneered by Pereira in the late 1980s [35], [36], SDDP has become the go-to methodology for medium- and long-term planning in hydro-thermal systems

for numerous countries [50], [242]. In these hydro-thermal planning models, the key question planners face is whether to "spend" the water in the reservoir to generate power at the current point in time, or whether to save it for later. The dual SDDP solutions provide very useful information on the so-called "value of water" which describes the future value of keeping the water in the reservoir. These dual solutions are as useful (if not more) as the primal solutions, which represent the target level of stored energy in hydro reservoirs. SDDP has recently found applications in other areas, including day-ahead bidding of pumped-hydro plants [48], natural gas storage valuation [243], dairy farm operations [244], short-term operational planning in power systems [245], as well as distribution grid restoration [246].

XIII. Energy Procurement for a Large Consumer

We next consider a large industrial electric energy consumer whose electricity bill amounts for a significant part of its total production cost. Examples of such consumers include an aluminum production company, an air liquefying corporation, and an electrical foundry. This large consumer is concerned with both its electricity cost and the variability of this cost. Seeking minimum expected electricity cost with limited cost variability, the consumer obtains its electricity from three sources, namely, by signing long-term contracts, buying in the spot market, and self-producing.

The long-term contracts that can be signed by the consumer to procure its electricity requirements include both forward contracts and options. A forward contract allows the consumer to buy electricity at a fixed price, thus eliminating price volatility. An option allows, for a fee, the consumer to decide at a later time (with reduced uncertainty) whether or not to use a forward contract to buy electricity. The spot market (day-ahead and real-time markets) allows the consumer to buy electricity, at the risk of facing potentially volatile prices. Self-producing is possible when the consumer owns a generation facility that can cover a portion of its electricity consumption. This facility may also be used for trading, i.e., to sell its generation in the spot market if the price is high enough (provided that the consumer's demand is satisfied).

We consider an electricity procurement problem which involves three sources of uncertainty, (i) the electricity price in the spot market (both day ahead and real time), (ii) the consumer demand, and (iii) the fuel cost of the self-production facility. The spot price uncertainty is typically significantly larger than the uncertainties pertaining to either consumer demand or fuel cost. The problem faced by the consumer in a specific period of time (for example, on an hourly basis for several months into the future) consists of determining which forward contracts or options to sign and the subsequent energy procurement strategy for any possible realization of the uncertainty. The objective is to minimize the expected cost of electricity throughout the procurement horizon while controlling the variability of such cost using a risk metric.

Describing the uncertainty via scenarios $\mathbf{s} \in S$, the procurement problem can be formulated as the two-stage stochastic programming problem below (note that multi-stage versions can also be easily formulated):

$$\min_{\mathbf{x};\mathbf{y_s},\mathbf{z_s},\forall\omega} \quad f^{\mathrm{F}}(\mathbf{x}) + \beta \mathsf{E}_P\left[f^{\mathrm{S}}(\mathbf{y},\mathbf{z},\ )\right] + (1-\beta) R_{cost}\left[f^{\mathrm{S}}(\mathbf{y},\mathbf{z},\ )\right] \tag{37a}$$

$$\text{s.t.} \quad h^{\mathrm{F}}(\mathbf{x}) = \mathbf{0}, \quad g^{\mathrm{F}}(\mathbf{x}) \leq \mathbf{0}, \tag{37b}$$

$$h^{\mathrm{S}}(\mathbf{y_s},\mathbf{z_s},\ _{\mathbf{s}}) = \mathbf{0}, \quad \forall \mathbf{s} \in S, \tag{37c}$$

$$g^{\mathrm{S}}(\mathbf{y_s},\mathbf{z_s},\ _{\mathbf{s}}) \leq \mathbf{0}, \quad \forall \mathbf{s} \in S. \tag{37d}$$

We note that the time (every hour of the procurement horizon spanning, e.g., the following three months) is implicitly represented in the above formulation. The variable vector $\mathbf{x}$ of first-stage variables represents contracting decisions spanning the procurement horizon, while variable vectors $\mathbf{y_s}$ and $\mathbf{z_s}$ are spot trading decisions and self-production decisions, respectively, per scenario $_{\mathbf{s}}$ and throughout the procurement horizon. In this problem, the first-stage cost $f^{\mathrm{F}}(\mathbf{x})$ is the contracting cost (not affected by uncertainty), while the second-stage cost $f^{\mathrm{S}}(\mathbf{y_s},\mathbf{z_s})$ is the spot market and self-production cost (a random variable). The parameter β is used to manage the tradeoff between expected cost and the risk of losses ($0 \leq \beta \leq 1$). Eqs. (37b) are contracting constraints, while (37c), (37d) are demand-supply (and risk-related) constraints. For a given β, weighting expected cost and cost variability, the solution of the two-stage stochastic programming problem (37) provides the optimal contracting strategy, $\mathbf{x}$, the optimal spot market involvement, $\mathbf{y_s}, \forall \mathbf{s} \in S$, and the optimal self-production, $\mathbf{z_s}, \forall \mathbf{s} \in S$.

For further reading, a seminal work on electricity procurement for large consumers with risk control is [247]. Among a significant number of relevant contributions by different research communities, [248] explores similar models and analyses to those in [247], but using a broader energy context and an analytical focus. Finally, [249] carries out a real-world insightful case study using a multi-stage stochastic programming model.

Part III: Summary, Conclusions, and Outlook

XIV. Summary and Conclusions

In this paper, we have provided an overview of methods for modeling and solving optimization problems under uncertainty arising in the context of electric power systems, with the goal of providing a summary and suggestions for further reading to researchers who are interested in applying optimization under uncertainty in their own work. The first part of the paper describes some of the most commonly used modeling techniques and algorithms. The second part of the paper reviews several applications. In this section, we offer overarching insights and an outlook towards future research directions.

An important aspect of power system optimization under uncertainty is that modeling and solution algorithms go hand in hand. How we choose to model and represent risk impacts which solution algorithm is going to be most successful at obtaining a solution with reasonable time and computational complexity. Conversely, our modeling choices (in particular,

the approximations we choose to apply) are often impacted by the need to obtain a computationally tractable problem. As a result, when modeling and solving optimization problems under uncertainty, we often encounter inherent trade-offs between solution quality (e.g., how optimal a solution is in terms of first-stage and expected second-stage cost?), providing probabilistic guarantees (e.g., how accurately are we able to model risk and probability of violations?) and computational tractability (e.g., how quickly can we solve the problem? Do we need a supercomputer or is a laptop enough?). How we choose to manage the trade-offs among these different aspects will vary by application and will also depend on how we choose to model and solve the problem. Solution evaluation is vitally important in these applications, as it can help us understand which modeling assumptions are reasonable (i.e., can be applied without deteriorating the quality of the decisions made by the model) and where we should invest more time and effort to improve our results. As the saying goes: "All models are wrong, but some are useful".

XV. Outlook and Future Directions

To conclude, we discuss several common observations and provide an outlook to challenges that we believe constitute important directions for future research.

1) Time scales of renewable energy uncertainty: Across all the applications, one of the primary drivers of uncertainty is the availability and variability in power generation from renewable energy sources. Uncertainty from wind and solar power is frequently considered at a shorter time-scale, whereas hydro-power uncertainty typically considers seasonal variability in precipitation. Given the rapid adoption of wind and solar power and the decline in other sources of generation capacity, it is becoming increasingly important to consider the seasonal and yearly variations in wind and solar power availability and how they correlate with load. Considering larger time horizons increases the complexity of solving the problems and also raises several important modeling questions. For example, how do we represent the risk and uncertainty associated with prolonged periods with lower wind and solar output such as the wind droughts in Europe in October 2021? How does wind and solar correlate with extreme weather to exaggerate or mitigate the impacts of corresponding changes in load, such as the correlation between very cold weather with low wind predicted for both Europe and the United States? Do we need the same or different risk metrics in situations with oversupply of renewable energy (California in Spring) and potential scarcity (California in Summer)?

These questions affect models ranging from short-term operations, which deal with the current scarcity or oversupply of power, to long-term planning, which need to consider how these correlations may be impacted by climate change.

2) High-impact, low-probability events: There are several emerging drivers of uncertainty in electric power systems. There is significant uncertainty associated with the impacts of climate change on the grid. Accordingly, there is increasing interest in modeling and mitigating the impacts of extreme weather, such as heatwaves, cold spells, hurricanes, and increased wildfire risk, both in operations and long-term planning. At the same time, increasing reliance on electricity for transportation, heating, and cooling also implies that the impacts of power outages and thereby the risk associated with them is changing, motivating the development of models that integrate technical aspects with the societal context around grid operation. Emerging examples include models that integrate information about hurricane evacuation orders or data on wildfire risk into emergency dispatching models. Accurately modeling the impacts of extreme weather events also requires considering new source of common mode outages (e.g., a cold wave that impacts both electricity and natural gas systems) and how to model (and weigh) the impacts of high-impact, low-probability events.

3) Interdependent systems: In some cases, risk can propagate between the electric grid and adjacent systems, including interactions with the natural gas system, water supply systems, cyber-physical layers of system operations, large-scale electric vehicle charging, and large-scale computing infrastructure. The need to characterize this uncertainty and mitigate associated risks while maintaining practically plausible models of reasonable complexity is a significant challenge.

In some situations, optimization under uncertainty, which inherently accounts for the fact that some information is unknown, can provide important tools for managing coupled system operations with limited exchange of information.

From a market design perspective, some assets might be able to provide services for improving the stability of the entire system against the systemic risk and, in the extreme case, against potential cascading failures. It is of interest to explore whether the current market products, e.g., ancillary services, sufficiently remunerate those assets for their services and, if not, whether new market products are necessary.

4) Scalability: Across all applications and methods, one of the principal challenges is computational scalability. Scalability limits the fidelity of our models, including how many scenarios we are able to include, how we choose to represent our recourse actions, what type of power flow model we choose (which impacts the detail in which we are able to represent the system operating state), which emerging technologies we are able to represent in the model, how many decision variables we can include, and the length of time-horizons that we are able to resolve. Significant ongoing advances in computer hardware and optimization algorithms have offered many new possibilities. Although these general advances continually increase the range of tractable problems, the structure specific to power system problems (such as the sparsity of the power flow equations or the fact that only a few transmission lines tend to be congested) provides opportunities to develop special-purpose algorithms that exploit this structure, often in combination with standard methods such as cutting plane algorithms, column-and-constraint generation, and stochastic dual dynamic programming, among others. Developing such algorithms is an important avenue for future

research that requires an interdisciplinary understanding of both power systems and optimization.

5) Formulating and solving uncertain distributed optimization problems: The rapid growth of distributed energy resources motivates the application of distributed optimization algorithms where multiple computing agents representing different portions of the a power system cooperatively solve optimization problems. In distributed optimization, agents iterate between solving local subproblems and exchanging information regarding the values of shared variables at the agents' interconnections. With some limited exceptions such as [196], the existing literature on distributed optimization for power systems primarily focuses on deterministic problems [250]. However, many practical applications of distributed optimization algorithms will likely require consideration of uncertainty. There are a number of related open questions, including what information regarding the uncertainties should be shared with neighboring agents (e.g., samples of uncertainty realizations, probability distributions, or uncertainty sets), how to efficiently use this information within distributed optimization algorithms, and how to design incentives so that the agents accurately report this information.

6) Pricing of risk and uncertainty: The consideration of risk and uncertainty in system operations tend to increase the nominal cost of operation. Furthermore, adverse realizations of the uncertain parameters may lead to very high operational cost. The question of how to share this increased cost among market participants remains an important question in the design of energy markets. Some methods for decision-making under uncertainty, e.g., robust and distributionally robust optimization, may require complicated formulations, leading to non-linear problems or the addition of auxiliary integer variables. This may complicate deriving efficient and equilibrium-supporting market-clearing prices using dual variables and is an important avenue for further research.

This list of research directions for power systems under uncertainty is intended to illustrate some potential future directions and is inherently non-exhaustive. We nevertheless hope it will serve as an inspiration for our readers and that they will use the material in this paper on the modeling, formulation, and solution of optimization problems under uncertainty to identify their own problems and research directions.

ACKNOWLEDGEMENT

We would like to acknowledge the very helpful comments provided by our anonymous reviewers. We would also like to thank the PhD students and postdocs who provided feedback in the process of writing this paper.

REFERENCES

[1] P. Kall, S. W. Wallace, and P. Kall, *Stochastic Programming*. Springer, 1994, vol. 6.

[2] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming, 2nd Edition*. Springer Science & Business Media, 2011.

[3] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Operations Research*, vol. 62, no. 6, pp. 1358–1376, 2014.

[4] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Mathematical Programming*, vol. 137, no. 1-2, pp. 167–198, 2013.

[5] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.

[6] B. P. Van Parys, P. J. Goulart, and D. Kuhn, "Generalized Gauss inequalities via semidefinite programming," *Mathematical Programming*, vol. 156, no. 1, 2016.

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

[8] B. Li, R. Jiang, and J. L. Mathieu, "The value of including unimodality information in distributionally robust optimal power flow," 2018. [Online]. Available: https://arxiv.org/abs/1811.10217

[9] L. V. Kantorovich and G. S. Rubinshtein, "On a space of totally additive functions," *Vestnik Leningradskogo Universiteta*, vol. 13, pp. 52–59, 1958.

[10] D. Kuhn, P. Mohajerin Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, "Wasserstein distributionally robust optimization: Theory and applications in machine learning," *INFORMS TutORials in Operations Research*, pp. 130–166, 2019.

[11] A. Ben-Tal and A. Nemirovski, "Robust solutions of uncertain linear programs," *Operations Research Letters*, vol. 25, no. 1, pp. 1–13, 1999.

[12] D. Bertsimas and M. Sim, "The price of robustness," *Operations Research*, vol. 52, no. 1, pp. 35–53, 2004.

[13] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Mathematical Programming*, vol. 167, no. 2, pp. 235–292, 2018.

[14] R. B. Nelsen, *An Introduction to Copulas*. Springer Science & Business Media, 2007.

[15] J. Nair, A. Wierman, and B. Zwart, "The fundamentals of heavy-tails: Properties, emergence, and identification," in *Proceedings of the ACM SIGMETRICS/international conference on Measurement and modeling of computer systems*, 2013, pp. 387–388.

[16] H. M. Markowitz, "Portfolio selection," in *Portfolio Selection*. Yale University Press, 1968.

[17] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[18] R. T. Rockafellar, S. Uryasev *et al.*, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, pp. 21–42, 2000.

[19] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 969–996, 2007.

[20] T. Summers, J. Warrington, M. Morari, and J. Lygeros, "Stochastic optimal power flow based on conditional value at risk and distributional robustness," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 116–125, 2015.

[21] Y. Zhang and G. B. Giannakis, "Robust optimal power flow with wind integration using conditional value-at-risk," in *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2013, pp. 654–659.

[22] L. Roald, S. Misra, M. Chertkov, and G. Andersson, "Optimal power flow with weighted chance constraints and general policies for generation control," in *IEEE Conference on Decision and Control (CDC) 2015*, Osaka, Japan, 2015.

[23] J. Dupacova, N. Gröwe-Kuska, and W. Römisch, "Scenario reduction in stochastic programming: An approach using probability metrics," *Mathematical Programming*, vol. 95, no. 3, pp. 493–511, 2003.

[24] H. Heitsch and W. Römisch, "Scenario reduction algorithms in stochastic programming," *Computational Optimization and Applications*, vol. 24, no. 2-3, pp. 187–206, 2003.

[25] N. Gröwe-Kuska, K. C. Kiwiel, M. P. Nowak, W. Romisch, and I. Wegner, *Decision Making Under Uncertainty: Energy and Power*, ser. IMA Volumes in Mathematics and Its Applications. New York: Springer-Verlag, 2002, vol. 128, ch. Power Management in a Hydro-Thermal System Under Uncertainty by Lagrangian Relaxation, pp. 39–70.

[26] J. M. Morales, S. Pineda, A. J. Conejo, and M. Carrion, "Scenario reduction for futures trading in electricity markets," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 878–888, May 2009.

[27] A. Papavasiliou and S. S. Oren, "Multi-area stochastic unit commitment for high wind penetration in a transmission constrained network," *Operations Research*, vol. 61, no. 3, pp. 578–592, 2013.

[28] A. Papavasiliou, S. S. Oren, and B. Rountree, "Applying high performance computing to transmission-constrained stochastic unit commitment for renewable penetration," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1690–1701, 2015.

[29] D. Kuhn, W. Wiesemann, and A. Georghiou, "Primal and dual linear decision rules in stochastic and robust optimization," *Mathematical Programming*, vol. 130, no. 1, pp. 177–209, 2011.

[30] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "Probabilistic guarantees for the N-1 security of systems with wind power generation," in *Probabilistic Methods Applied to Power Systems (PMAPS) 2012*, Istanbul, Turkey, 2012.

[31] M. Vrakopoulou, "Optimal decision making for secure and economic operation of power systems under uncertainty," Ph.D. dissertation, ETH Zurich, 2013.

[32] T. Mühlpfordt, T. Faulwasser, V. Hagenmeyer, L. Roald, and S. Misra, "On polynomial real-time control policies in stochastic AC optimal power flow," *Electric Power Systems Research*, vol. 189, p. 106792, 2020, presented at *21st Power Systems Computation Conference (PSCC 2020)*.

[33] L. Roald, S. Misra, M. Chertkov, S. Backhaus, and G. Andersson, "Optimal power flow with wind power control and limited expected risk of overloads," in *19th Power System Computation Conference (PSCC 2016)*, Genova, Italy, 2016.

[34] A. J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition Techniques in Mathematical Programming: Engineering and Science Applications*. Springer Science & Business Media, 2006.

[35] M. V. Pereira, "Optimal stochastic operations scheduling of large hydroelectric systems," *International Journal of Electrical Power and Energy Systems*, vol. 11, no. 3, pp. 161–169, 1989.

[36] M. V. F. Pereira and L. M. V. G. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, vol. 52, pp. 359–375, 1991.

[37] O. Dowson and L. Kapelevich, "SDDP.jl: A Julia package for stochastic dual dynamic programming," *INFORMS Journal on Computing*, vol. 33, no. 1, pp. 27–33, 2021.

[38] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, ser. Springer Series in Operations Research and Financial Engineering. Springer, 2010.

[39] S. Rebennack, "Combining sampling-based and scenario-based nested benders decomposition methods: Application to stochastic dual dynamic programming," *Mathematical Programming*, vol. 156, no. 1-2, pp. 343–389, 2016.

[40] N. Löhndorf and A. Shapiro, "Modeling time-dependent randomness in stochastic dual dynamic programming," *European Journal of Operational Research*, vol. 273, no. 2, pp. 650–661, 2019.

[41] J. Zou, S. Ahmed, and X. A. Sun, "Stochastic dual dynamic integer programming," *Mathematical Programming*, vol. 175, no. 1, pp. 461–502, 2019.

[42] A. Shapiro, W. Tekaya, J. P. da Costa, and M. P. Soares, "Risk neutral and risk averse stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 224, pp. 375–391, 2013.

[43] A. Shapiro, "Analysis of stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 209, pp. 63–72, 2011.

[44] A. B. Philpott and Z. Guan, "On the convergence of stochastic dual dynamic programming and related methods," *Operations Research Letters*, vol. 36, no. 4, pp. 450–455, 2008.

[45] T. Asamov and W. B. Powell, "Regularized decomposition of high-dimensional multistage stochastic programs with markov uncertainty," *SIAM Journal on Optimization*, vol. 28, no. 1, pp. 575–595, 2018.

[46] V. L. D. Matos, A. B. Philpott, and E. C. Finardi, "Improving the performance of stochastic dual dynamic programming," *Journal of Computational and Applied Mathematics*, vol. 290, pp. 196–208, 2015.

[47] V. Guigues, "Dual dynamic programing with cut selection: Convergence proof and numerical experiments," *European Journal of Operational Research*, vol. 258, no. 1, pp. 47–57, 2017.

[48] N. Löhndorf, D. Wozabal, and S. Minner, "Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming," *Operations Research*, vol. 61, no. 4, pp. 810–823, 2013.

[49] E. L. da Silva and E. C. Finardi, "Parallel processing applied to the planning of hydrothermal systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 14, no. 8, pp. 721–729, 2003.

[50] R. J. Pinto, C. Borges, and M. E. P. Maceira, "An efficient parallel algorithm for large scale hydrothermal system operation planning," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4888–4896, 2013.

[51] F. D. Machado, A. L. Diniz, C. L. Borges, and L. C. Brandao, "Asynchronous parallel stochastic dual dynamic programming applied to hydrothermal generation planning," *Electric Power Systems Research*, vol. 191, p. 106907, 2021.

[52] D. Ávila, A. Papavasiliou, and N. Löhndorf, "Parallel and distributed computing for stochastic dual dynamic programming," *Computational Management Science*, 2021.

[53] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[54] G. L. Nemhauser and L. A. Wolsey, *Integer Programming and Combinatorial Optimization*. Springer, 1988, vol. 191.

[55] C. Lemaréchal, "Lagrangian relaxation," in *Computational Combinatorial Optimization, Lecture Notes in Computer Science*. Springer, 2001, vol. 2241, pp. 112–156.

[56] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. Springer Science & Business Media, 2003, vol. 87.

[57] I. Aravena and A. Papavasiliou, "Asynchronous lagrange scenario decomposition," *Mathematical Programming Computation*, vol. 13, no. 1, pp. 1–50, 2021.

[58] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[59] R. T. Rockafellar and R. J.-B. Wets, "Scenarios and policy aggregation in optimization under uncertainty," *Mathematics of Operations Research*, vol. 16, no. 1, pp. 119–147, 1991.

[60] J.-P. Watson and D. L. Woodruff, "Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems," *Computational Management Science*, vol. 8, pp. 355–370, 2011.

[61] K. Cheung, D. Gade, C. Silva-Monroy, S. M. Ryan, J.-P. Watson, R. J.-B. Wets, and D. L. Woodruff, "Toward scalable stochastic unit commitment," *Energy Systems*, vol. 6, no. 3, pp. 417–438, 2015.

[62] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.

[63] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton University Press, 2009, vol. 28.

[64] M. Ono and B. C. Williams, "Iterative risk allocation: A new approach to robust model predictive control with a joint chance constraint," in *2008 47th IEEE Conference on Decision and Control*, 2008, pp. 3427–3432.

[65] K. Baker and B. Toomey, "Efficient relaxations for joint chance constrained AC optimal power flow," *Electric Power Systems Research*, vol. 148, pp. 230–236, 2017, presented at the *21st Power Systems Computation Conference (PSCC 2020)*.

[66] A. M. Hou and L. A. Roald, "Data-driven tuning for chance-constrained optimization: Two steps towards probabilistic performance guarantees," *IEEE Control Systems Letters*, vol. 6, pp. 1400–1405, 2022.

[67] J. Luedtke and S. Ahmed, "A sample approximation approach for optimization with probabilistic constraints," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 674–699, 2008.

[68] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking and Finance*, vol. 26, pp. 1443–1471, 2002.

[69] S. Ahmed, J. Luedtke, Y. Song, and W. Xie, "Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs," *Mathematical Programming*, vol. 162, no. 1, pp. 51–81, 2017.

[70] N. Jiang and W. Xie, "ALSO-X and ALSO-X+: Better convex approximations for chance constrained programs," to appear in *Operations Research*, 2022.

[71] A. Peña-Ordieres, J. R. Luedtke, and A. Wächter, "Solving chance-constrained problems via a smooth sample-based nonlinear approximation," *SIAM Journal on Optimization*, vol. 30, no. 3, pp. 2221–2250, 2020.

[72] Y. Cao and V. M. Zavala, "A sigmoidal approximation for chance-constrained nonlinear programs," *arXiv preprint arXiv:2004.02402*, 2020.

[73] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Transactions on Automatic Control*, vol. 51, no. 5, pp. 742–753, 2006.

[74] L. A. Roald and G. Andersson, "Chance-constrained AC optimal power flow: Reformulations and efficient algorithms," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2906–2918, 2018.

[75] K. Margellos, P. Goulart, and J. Lygeros, "On the road between robust optimization and the scenario approach for chance constrained optimization problems," *IEEE Transactions on Automatic Control*, vol. 59, no. 8, pp. 2258–2263, 2014.

[76] M. C. Campi, S. Garatti, and F. A. Ramponi, "A general scenario theory for nonconvex optimization and decision making," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4067–4078, 2018.

[77] F. Cantelli, *Intorno ad un teorema fondamentale della teoria del rischio*. Tip. degli operai, 1910.

[78] B. Van Parys, P. Goulart, and D. Kuhn, "Generalized Gauss inequalities via semidefinite programming," *Mathematical Programming*, pp. 1–32, 2015.

[79] L. Roald, F. Oldewurtel, B. Van Parys, and G. Andersson, "Security constrained optimal power flow with distributionally robust chance constraints," Tech. Rep., 2015. [Online]. Available: http://arxiv.org/abs/1508.06061

[80] T. Mühlpfordt, L. Roald, V. Hagenmeyer, T. Faulwasser, and S. Misra, "Chance-constrained AC optimal power flow: A polynomial chaos approach," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4806–4816, 2019.

[81] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.

[82] W. Xie and S. Ahmed, "Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1860–1867, 2018.

[83] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.

[84] A. Shapiro, "Distributionally robust stochastic programming," *SIAM Journal on Optimization*, vol. 27, no. 4, pp. 2258–2275, 2017.

[85] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: A review," 2019. [Online]. Available: www.optimization-online.org/DB_HTML/2019/08/7332.html

[86] T. Mühlpfordt, T. Faulwasser, and V. Hagenmeyer, "Solving stochastic AC power flow via polynomial chaos expansion," in *IEEE Conference on Control Applications (CCA)*, 2016, pp. 70–76.

[87] D. Lee, K. Turitsyn, D. K. Molzahn, and L. Roald, "Robust AC optimal power flow with convex restriction," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 4953–4966, November 2021.

[88] A. J. Conejo, M. Carrión, J. M. Morales *et al.*, *Decision Making Under Uncertainty in Electricity Markets*. Springer, 2010, vol. 1.

[89] J. L. Carpentier, "Contribution a l'etude du dispatching économic," in *Bulletin de la Societé Française des Électriciens*, ser. 8, August 1962, no. vol. III, pp. 431–434.

[90] T. E. Dy Liacco, "The adaptive reliability control system," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, no. 5, pp. 517–531, 1967.

[91] F. Capitanescu, J. M. Ramos, P. Panciatici, D. Kirschen, A. M. Marcolini, L. Platbrood, and L. Wehenkel, "State-of-the-art, challenges, and future trends in security constrained optimal power flow," *Electric Power Systems Research*, vol. 81, no. 8, pp. 1731–1741, 2011.

[92] F. Capitanescu, "Critical review of recent advances and further developments needed in AC optimal power flow," *Electric Power Systems Research*, vol. 136, pp. 57 – 68, 2016.

[93] I. Aravena, D. K. Molzahn, S. Zhang, C. G. Petra, F. E. Curtis, S. Tu, A. Wächter, E. Wei, E. Wong, A. Gholami, K. Sun, X. A. Sun, S. T. Elbert, J. T. Holzer, and A. Veeramany, "Recent developments in security-constrained AC optimal power flow: Overview of Challenge 1 in the ARPA-E Grid Optimization Competition," *arXiv:2206.07843*, 2022.

[94] W. Fu and J. D. McCalley, "Risk-based optimal power flow," in *IEEE PowerTech Conference*, Porto, Portugal, 2001.

[95] Q. Wang, J. D. McCalley, T. Zheng, and E. Litvinov, "A computational strategy to solve preventive risk-based security-constrained OPF," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1666–1675, 2013.

[96] L. Roald, M. Vrakopoulou, F. Oldewurtel, and G. Andersson, "Risk-based optimal power flow with probabilistic guarantees," *International Journal of Electrical Power and Energy Systems*, vol. 72, pp. 66 – 74, 2015.

[97] D. Shchetinin and G. Hug, "Risk-constrained AC OPF with risk limits on individual system states," in *2015 IEEE Eindhoven PowerTech*, 2015, pp. 1–6.

[98] B. Stott, O. Alsaç, and A. J. Monticelli, "Security analysis and optimization," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1623–1644, December 1987.

[99] PACME, The RRPA Subcommittee Working Group, M. Papic, S. Agarwal, R. N. Allan, R. Billinton, C. J. Dent, S. Ekisheva, D. Gent, K. Jiang, W. Li, J. Mitra, A. Pitto, A. Schneider, C. Singh, V. V. Vadlamudi, and M. Varghese, "Research on common-mode and dependent (CMD) outage events in power systems: A review," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1528–1536, 2017.

[100] J. Salmeron, K. Wood, and R. Baldick, "Analysis of electric grid security under terrorist threat," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 905–912, 2004.

[101] M. Panteli and P. Mancarella, "Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies," *Electric Power Systems Research*, vol. 127, pp. 259–270, 2015.

[102] O. Alsac and B. Stott, "Optimal load flow with steady-state security," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 745–751, 1974.

[103] A. Monticelli, M. V. F. Pereira, and S. Granville, "Security-constrained optimal power flow with post-contingency corrective rescheduling," *IEEE Transactions on Power Systems*, vol. 2, no. 1, pp. 175–180, 1987.

[104] F. Capitanescu, M. Glavic, D. Ernst, and L. Wehenkel, "Contingency filtering techniques for preventive security-constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1690–1697, 2007.

[105] A. J. Ardakani and F. Bouffard, "Identification of umbrella constraints in DC-based security-constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3924–3934, 2013.

[106] F. Capitanescu, S. Fliscounakis, P. Panciatici, and L. Wehenkel, "Cautious operation planning under uncertainties," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1859–1869, Nov 2012.

[107] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, "Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive, and corrective actions," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4909–4917, Nov 2013.

[108] G. Hug, "Generation cost and system risk trade-off with corrective power flow control," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012, pp. 1324–1331.

[109] Q. Wang, G. Zhang, J. McCalley, T. Zheng, and E. Litvinov, "Risk-based locational marginal pricing and congestion management," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2518–2528, 2014.

[110] Q. Wang, J. D. McCalley, T. Zheng, and E. Litvinov, "Solving corrective risk-based security-constrained optimal power flow with Lagrangian relaxation and Benders decomposition," *International Journal of Electrical Power & Energy Systems*, vol. 75, pp. 255–264, 2016.

[111] E. Karangelos and L. Wehenkel, "Probabilistic reliability management approach and criteria for power system real-time operation," in *2016 Power Systems Computation Conference (PSCC)*, 2016, pp. 1–9.

[112] E. Karangelos, P. Panciatici, and L. Wehenkel, "Whither probabilistic security management for real-time operation of power systems?" in *2013 IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid*, 2013, pp. 1–17.

[113] E. Karangelos and L. Wehenkel, "Post-contingency corrective control failure: a risk to neglect or a risk to control?" in *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2018, pp. 1–6.

[114] ——, "An iterative AC-SCOPF approach managing the contingency and corrective control failure uncertainties with a probabilistic guarantee," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3780–3790, 2019.

[115] UMBRELLA FP7 Project, "Toolbox for Common Forecasting, Risk Assessment, and Operational Optimisation in Grid Security Coopera-

tions of Transmission System Operators (TSOs)," www.e-umbrella.eu, 2012–2015.

[116] "Operational reserve ad hoc team report," ENTSO-E, Tech. Rep., May 2012. [Online]. Available: https://www.entsoe.eu/fileadmin/user_upload/_library/resources/LCFR/2012-06-14_SOC-AhT-OR_Report_final_V9-3.pdf

[117] "Supporting Document for the Network Code on Load-Frequency Control and Reserves," ENTSO-E, Tech. Rep., June 2013. [Online]. Available: https://www.entsoe.eu/fileadmin/user_upload/_library/resources/LCFR/130628-NC_LFCR-Issue1.pdf

[118] "CWE Enhanced Flow-Based MC Feasibility Report," Amprion, APX, BelPEX, Creos, Elia, EnBW, EPEX, RTE, Tennet, Tech. Rep., October 2011. [Online]. Available: https://www.epexspot.com/document/12597/CWE_FB-MC_feasibility_report.pdf

[119] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical reformulation of security constrained optimal power flow with probabilistic constraints," in *IEEE PowerTech Conference*, Grenoble, France, 2013.

[120] E. Sjodin, D. F. Gayme, and U. Topcu, "Risk-mitigated optimal power flow for wind powered grids," in *American Control Conference (ACC)*, Montreal, Canada, 2012.

[121] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A probabilistic framework for reserve scheduling and N-1 security assessment of systems with high wind power penetration," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3885–3896, 2013.

[122] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical reformulation of security constrained optimal power flow with probabilistic constraints," in *IEEE Grenoble Conference*, 2013.

[123] L. Roald, T. Krause, and G. Andersson, "Integrated balancing and congestion management under forecast uncertainty," in *IEEE EnergyCon*, Leuven, Belgium, 2016.

[124] D. Bienstock and A. Shukla, "Variance-aware optimal power flow: Addressing the tradeoff between cost, security, and variability," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1185–1196, 2019.

[125] K. Sundar, H. Nagarajan, M. Lubin, L. Roald, S. Misra, R. Bent, and D. Bienstock, "Unit commitment with N-1 security and wind uncertainty," in *19th Power System Computation Conference (PSCC 2016)*, Genova, Italy, 2016.

[126] Á. Porras, C. Domínguez, J. M. Morales, and S. Pineda, "Tight and compact sample average approximation for joint chance constrained optimal power flow," *arXiv preprint arXiv:2205.03370*, 2022.

[127] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective control to handle forecast uncertainty: A chance constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1626–1637, 2017.

[128] R. Kannan, J. R. Luedtke, and L. A. Roald, "Stochastic DC optimal power flow with reserve saturation," *Electric Power Systems Research*, vol. 189, p. 106566, 2020.

[129] M. Vrakopoulou, B. Li, and J. L. Mathieu, "Chance constrained reserve scheduling using uncertain controllable loads – Part I: Formulation and scenario-based analysis," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1608–1617, 2017.

[130] B. Li, M. Vrakopoulou, and J. L. Mathieu, "Chance constrained reserve scheduling using uncertain controllable loads – Part II: Analytical reformulation," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1618–1625, 2017.

[131] M. G. Vayá and G. Andersson, "Integrating renewable energy forecast uncertainty in smart-charging approaches for plug-in electric vehicles," in *2013 IEEE Grenoble Conference*. IEEE, 2013, pp. 1–6.

[132] R. Wiget, M. Vrakopoulou, and G. Andersson, "Probabilistic security constrained optimal power flow for a mixed HVAC and HVDC grid with stochastic infeed," in *Power Systems Computation Conference (PSCC)*, Wroclaw, Poland, 2014.

[133] A. Venzke and S. Chatzivasileiadis, "Convex relaxations of probabilistic AC optimal power flow for interconnected AC and HVDC grids," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2706–2718, 2019.

[134] M. A. Bucher, M. Vrakopoulou, and G. Andersson, "Probabilistic N-1 security assessment incorporating dynamic line ratings," in *2013 IEEE Power & Energy Society General Meeting*. IEEE, 2013, pp. 1–5.

[135] L. A. Roald, K. Sundar, A. Zlotnik, S. Misra, and G. Andersson, "An uncertainty management framework for integrated gas-electric energy systems," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1518–1540, 2020.

[136] A. Stuhlmacher and J. L. Mathieu, "Chance-constrained water pumping to manage water and power demand uncertainty in distribution networks," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1640–1655, 2020.

[137] M. Lubin, Y. Dvorkin, and S. Backhaus, "A robust approach to chance constrained optimal power flow with renewable generation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3840–3849, 2015.

[138] C. Wang, R. Gao, F. Qiu, J. Wang, and L. Xin, "Risk-based distributionally robust optimal power flow with dynamic line rating," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6074–6086, 2018.

[139] Y. Zhang, S. Shen, and J. L. Mathieu, "Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1378–1388, 2017.

[140] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 363–368, 2018.

[141] B. K. Poolla, A. R. Hota, S. Bolognani, D. S. Callaway, and A. Cherukuri, "Wasserstein distributionally robust look-ahead economic dispatch," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 2010–2022, 2021.

[142] C. Ordoudis, V. A. Nguyen, D. Kuhn, and P. Pinson, "Energy and reserve dispatch with distributionally robust joint chance constraints," *Operations Research Letters*, vol. 49, no. 3, pp. 291–299, 2021.

[143] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu, "Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4924–4936, 2018.

[144] Y. Guo, K. Baker, E. Dall'Anese, Z. Hu, and T. H. Summers, "Data-based distributionally robust stochastic optimal power flow — Part I: Methodologies," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1483–1492, 2019.

[145] A. Arrigo, C. Ordoudis, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, "Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation," *European Journal on Operational Research*, vol. 296, no. 1, pp. 304–322, 2022.

[146] M. V. F. Pereira and N. J. Balu, "Composite generation/transmission reliability evaluation," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 470–491, 1992.

[147] European Commission, "Commission regulation (EU) 2017/1485 of 2 August 2017 establishing a guideline on electricity transmission system operation," *Official Journal of the European Union*, 2017.

[148] M. Bucksteeg, L. Niesen, and C. Weber, "Impacts of dynamic probabilistic reserve sizing techniques on reserve requirements and system costs," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, 2016.

[149] N. Menemenlis, M. Huneault, and A. Robitaille, "Computation of dynamic operating balancing reserve for wind power integration for the time-horizon 1–48 hours," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 692–702, 2012.

[150] K. De Vos, J. Morbee, J. Driesen, and R. Belmans, "Impact of wind power on sizing and allocation of reserve requirements," *IET Renewable Power Generation*, vol. 7, no. 1, pp. 1–9, 2013.

[151] K. De-Vos, N. Stevens, O. Devolder, A. Papavasiliou, B. Hebb, and J. Matthys-Donnadieu, "Dynamic dimensioning approach for operating reserves: Proof of concept in Belgium," *Energy Policy*, vol. 124, pp. 272–285, 2019.

[152] A. Papavasiliou, A. Bouso, S. Apelfröd, E. Wik, T. Gueuning, and Y. Langer, "Multi-area reserve dimensioning using chance-constrained optimization," to appear in *IEEE Transactions on Power Systems*, 2022.

[153] Y. Zhang, J. Wang, B. Zeng, and Z. Hu, "Chance-constrained two-stage unit commitment under uncertain load and wind power output using bilinear Benders decomposition," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3637–3647, 2017.

[154] T. Zheng and E. Litvinov, "Contingency-based zonal reserve modeling and pricing in a co-optimized energy and reserve market," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 277–286, 2008.

[155] J. D. Lyon, K. W. Hedman, and M. Zhang, "Reserve requirements to efficiently manage intra-zonal congestion," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 251–258, 2013.

[156] F. Wang and Y. Chen, "Market implications of short-term reserve deliverability enhancement," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1504–1514, 2021.

[157] Y. Chen, P. Gribik, and J. Gardner, "Incorporating post zonal reserve deployment transmission constraints into energy and ancillary service co-optimization," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 537–549, 2013.

[158] B. Park, Z. Zhou, A. Botterud, and P. Thimmapuram, "Probabilistic zonal reserve requirements for improved energy management and deliverability with wind power uncertainty," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4324–4334, 2020.

[159] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective control to handle forecast uncertainty: A chance constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1626–1637, 2016.

[160] M. A. Bucher, M. A. Ortega-Vazquez, D. S. Kirschen, and G. Andersson, "Robust allocation of reserves considering different reserve types and the flexibility from HVDC," *IET Generation, Transmission & Distribution*, vol. 11, no. 6, pp. 1472–1478, 2017.

[161] S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1497–1508, August 1996.

[162] P. Carpentier, G. Cohen, J.-C. Culioli, and A. Renaud, "Stochastic optimization of unit commitment: A new decomposition framework," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 1067–1073, May 1996.

[163] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit commitment for systems with high wind penetration," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 592–601, May 2009.

[164] P. Meibom, R. Barth, B. Hasche, H. Brand, C. Weber, and M. O'Malley, "Stochastic optimization model to study the operational impacts of high wind penetrations in ireland," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1367–1379, 2010.

[165] J. M. Morales, A. J. Conejo, and J. Perez-Ruiz, "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 900–910, May 2009.

[166] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2197–2206, 2011.

[167] B. Wang and B. F. Hobbs, "Real-time markets for flexiramp: A stochastic unit commitment-based analysis," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 846–860, 2016.

[168] L. Wu, M. Shahidehpour, and T. Li, "Stochastic security-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 800–811, May 2007.

[169] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, February 2013.

[170] P. A. Ruiz, R. C. Philbrick, E. Zack, K. W. Cheung, and P. W. Sauer, "Uncertainty management in the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 642–1651, May 2009.

[171] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 800–810, May 2012.

[172] A. Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32–51, 2016.

[173] Y. Dvorkin, H. Pandžić, M. A. Ortega-Vazquez, and D. S. Kirschen, "A hybrid stochastic/interval approach to transmission-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 621–631, 2014.

[174] D. Gade, G. Hackebeil, S. M. Ryan, J.-P. Watson, R. J.-B. Wets, and D. L. Woodruff, "Obtaining lower bounds from the progressive hedging algorithm for stochastic mixed-integer programs," *Mathematical Programming*, vol. 157, pp. 47–67, 2016.

[175] Y. Wen, C. Guo, H. Pandzic, and D. S. Kirschen, "Enhanced security-constrained unit commitment with emerging utility-scale energy storage," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 652–662, January 2016.

[176] W. van Ackooij and J. Malick, "Decomposition algorithm for large-scale two-stage unit-commitment," *Annals of Operations Research*, vol. 238, no. 1-2, pp. 587–613, 2016.

[177] A. J. Conejo, L. Baringo, S. J. Kazempour, and A. S. Siddiqui, *Investment in Electricity Generation and Transmission*. Springer, 2016.

[178] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, 2007, vol. 703.

[179] A. Moreira, D. Pozo, A. Street, and E. Sauma, "Reliable renewable generation and transmission expansion planning: Co-optimizing system's resources for meeting renewable targets," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3246–3257, 2016.

[180] A. Moreira, D. Pozo, A. Street, E. Sauma, and G. Strbac, "Climate-aware generation and transmission expansion planning: A three-stage robust optimization approach," *European Journal of Operational Research*, 2021.

[181] D. Pozo, E. E. Sauma, and J. Contreras, "A three-level static MILP model for generation and transmission expansion planning," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 202–210, 2012.

[182] D. Quiroga, E. Sauma, and D. Pozo, "Power system expansion planning under global and local emission mitigation policies," *Applied Energy*, vol. 239, pp. 1250–1264, 2019.

[183] B. F. Hobbs, Q. Xu, J. Ho, P. Donohoo, S. Kasina, J. Ouyang, S. W. Park, J. Eto, and V. Satyal, "Adaptive transmission planning: Implementing a new paradigm for managing economic risks in grid expansion," *IEEE Power and Energy Magazine*, vol. 14, no. 4, pp. 30–40, 2016.

[184] A. Velloso, D. Pozo, and A. Street, "Distributionally robust transmission expansion planning: A multi-scale uncertainty approach," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3353–3365, 2020.

[185] S. Lumbreras and A. Ramos, "The new challenges to transmission expansion planning. Survey of recent practice and literature review," *Electric Power Systems Research*, vol. 134, pp. 19–29, 2016.

[186] S. Lumbreras, H. Abdi, A. Ramos *et al.*, *Transmission Expansion Planning: The Network Challenges of the Energy Transition*. Springer, 2021.

[187] V. Krishnan, J. Ho, B. F. Hobbs, A. L. Liu, J. D. McCalley, M. Shahidehpour, and Q. P. Zheng, "Co-optimization of electricity transmission and generation resources for planning and policy analysis: review of concepts and modeling approaches," *Energy Systems*, vol. 7, no. 2, pp. 297–332, 2016.

[188] N. Helistö, J. Kiviluoma, H. Holttinen, J. D. Lara, and B.-M. Hodge, "Including operational aspects in the planning of power systems with large amounts of variable generation: A review of modeling approaches," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 8, no. 5, p. e341, 2019.

[189] D. K. Molzahn and I. A. Hiskens, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.

[190] D. K. Molzahn, "Computing the feasible spaces of optimal power flow problems," *IEEE Transactionson Power Systems*, vol. 32, no. 6, pp. 4752–4763, November 2017.

[191] D. Bienstock and A. Verma, "Strong NP-hardness of AC power flows feasibility," *Operations Research Letters*, vol. 47, no. 6, pp. 494–501, 2019.

[192] K. Lehmann, A. Grastien, and P. Van Hentenryck, "AC-feasibility on tree networks is NP-hard," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 798–801, 2016.

[193] A. Wächter and L. T. Biegler, "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 1, no. 106, pp. 25–57, 2006.

[194] S. Gopinath and H. L. Hijazi, "Benchmarking large-scale ACOPF solutions and optimality bounds," in *IEEE Power & Energy Society General Meeting*, 2022.

[195] H. Wu, Y. Zhou, S. Dong, and Y. Song, "Probabilistic load flow based on generalized polynomial chaos," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 820–821, 2017.

[196] A. Engelmann, T. Mühlpfordt, Y. Jiang, B. Houska, and T. Faulwasser, "Distributed stochastic AC optimal power flow based on polynomial chaos expansion," in *Annual American Control Conference (ACC)*, 2018, pp. 6188–6193.

[197] X. Sun, Q. Tu, J. Chen, C. Zhang, and X. Duan, "Probabilistic load flow calculation based on sparse polynomial chaos expansion," *IET*

Generation, Transmission & Distribution, vol. 12, no. 11, pp. 2735–2744, 2018.

[198] T. Mühlpfordt, T. Faulwasser, V. Hagenmeyer, L. Roald, and S. Misra, "On polynomial real-time control policies in stochastic AC optimal power flow," *Electric Power Systems Research*, vol. 189, p. 106792, 2020, presented at *21st Power Systems Computation Conference (PSCC 2020)*.

[199] D. Métivier, M. Vuffray, and S. Misra, "Efficient polynomial chaos expansion for uncertainty quantification in power systems," *Electric Power Systems Research*, vol. 189, p. 106791, 2020, presented at *21st Power Systems Computation Conference (PSCC 2020)*.

[200] F. D. Galiana, "Analytic properties of the load flow problem," in *International Symposium on Circuits and Systems, Special Session on Power System*, Philadelphia, PA, USA, 1977, pp. 5–7.

[201] ——, "Analytical investigation of the power flow equations," in *American Control Conference*, 1983, pp. 411–415.

[202] A. Klos, "The non-uniqueness of load-flow solution," in *5th Power Systems Comptuation Conference (PSCC)*, 1975.

[203] V. A. Venikov, V. A. Stroev, V. I. Idelchick, and V. I. Tarasov, "Estimation of electrical power system steady-state stability in load flow calculations," *IEEE Transactions on Power Apparatus and Systems*, vol. 94, no. 3, pp. 1034–1041, 1975.

[204] P. W. Sauer and M. A. Pai, *Power System Dynamic Equilibrium, Power Flow, and Steady-State Stability*. Springer International Publishing, 2014, pp. 1–26.

[205] D. Phan and S. Ghosh, "Two-stage stochastic optimization for optimal power flow under renewable generation uncertainty," *ACM Transactions on Modeling and Computer Simulation*, vol. 24, no. 1, pp. 2:1–2:22, Jan. 2014. [Online]. Available: http://doi.acm.org/10.1145/2553084

[206] M. Lubin, Y. Dvorkin, and L. Roald, "Chance constraints for improving the security of AC optimal power flow," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1908–1917, 2019.

[207] M. Chamanbaz, F. Dabbene, and C. M. Lagoa, "Probabilistically robust AC optimal power flow," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1135–1147, 2019.

[208] O. Kuryatnikova, B. Ghaddar, and D. K. Molzahn, "Adjustable robust two-stage polynomial optimization with application to AC optimal power flow," *arXiv:2104.03107*, 2021.

[209] A. Lorca and X. A. Sun, "The adaptive robust multi-period alternating current optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1993–2003, 2018.

[210] D. K. Molzahn and L. A. Roald, "Towards an AC optimal power flow algorithm with robust feasibility guarantees," *20th Power Systems Computation Conference (PSCC)*, June 2018.

[211] ——, "Grid-aware versus grid-agnostic distribution system control: A method for certifying engineering constraint satisfaction," in *52nd Hawaii International Conference on System Sciences (HICSS)*, January 2019.

[212] R. Louca and E. Bitar, "Robust AC optimal power flow," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1669–1681, May 2019.

[213] C. Wang, A. Bernstein, J.-Y. Le Boudec, and M. Paolone, "Explicit conditions on existence and uniqueness of load-flow solutions in distribution networks," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 953–962, March 2018.

[214] C. Wang, J.-Y. Le Boudec, and M. Paolone, "Controlling the electrical state via uncertain power injections in three-phase distribution networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1349–1362, 2017.

[215] D. Lee, H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Convex restriction of power flow feasibility sets," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1235–1245, 2019.

[216] B. Cui and X. A. Sun, "Solvability of power flow equations through existence and uniqueness of complex fixed point," *arXiv preprint arXiv:1904.08855*, 2019.

[217] N. Nazir and M. Almassalkhi, "Convex inner approximation of the feeder hosting capacity limits on dispatchable demand," in *IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 4858–4864.

[218] J. F. Marley, M. Vrakopoulou, and I. A. Hiskens, "An AC-QP optimal power flow algorithm considering wind forecast uncertainty," in *IEEE Innovative Smart Grid Technologies–Asia (ISGT-Asia)*, 2016, pp. 317–323.

[219] ——, "Towards the maximization of renewable energy integration using a stochastic AC-QP optimal power flow algorithm," in *10th IREP Symposium on Bulk Power System Dynamics and Control*, Espinho, Portugal, August 2017.

[220] A. J. Korsak, "On the question of uniqueness of stable load-flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-91, no. 3, pp. 1093–1100, 1972.

[221] H. D. Nguyen and K. S. Turitsyn, "Appearance of multiple stable load flow solutions under power flow reversal conditions," in *IEEE PES General Meeting*, 2014.

[222] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, August 2009.

[223] X. Geng and L. Xie, "Data-driven decision making with probabilistic guarantees (Part 2): Applications of chance-constrained optimization in power systems," *arXiv:1904.06755*, 2019.

[224] A. Peña Ordieres, D. K. Molzahn, L. A. Roald, and A. Wächter, "DC optimal power flow with joint chance constraints," *IEEE Transactions on Power Systems*, vol. 36, no. 1, pp. 147–158, January 2021.

[225] T. J. Overbye, X. Cheng, and Y. Sun, "A comparison of the AC and DC power flow models for LMP calculations," in *37th Annual Hawaii International Conference on System Sciences*, January 2004.

[226] C. Coffrin, P. Van Hentenryck, and R. Bent, "Accurate load and generation scheduling for linearized DC models with contingencies," in *IEEE Power and Energy Society General Meeting*, 2012.

[227] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *IEEE Power Engineering Society General Meeting*, 2005, pp. 454–459.

[228] C. Barrows, S. Blumsack, and P. Hines, "Correcting optimal transmission switching for AC power flows," in *47th Hawaii International Conference on System Sciences*, January 2014, pp. 2374–2379.

[229] K. Dvijotham and D. K. Molzahn, "Error bounds on the DC power flow approximations: A convex relaxation approach," *IEEE 55th Conference on Decision and Control (CDC)*, December 2016.

[230] H. Cetinay, S. Soltan, F. A. Kuipers, G. Zussman, and P. Van Mieghem, "Comparing the effects of failures in power grids under the AC and DC power flow models," *IEEE Transactions on Network Science and Engineering*, vol. 5, no. 4, pp. 301–312, 2018.

[231] K. Baker, "Solutions of DC OPF are never AC feasible," in *Twelfth ACM International Conference on Future Energy Systems (ACM e-Energy)*, June 2021, pp. 264–268.

[232] E. Dall'Anese, K. Baker, and T. Summers, "Chance-constrained AC optimal power flow for distribution systems with renewables," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3427–3438, 2017.

[233] N. Z. Shor, "Quadratic optimization problems," *Soviet Journal of Computer and Systems Sciences*, vol. 25, pp. 1–11, 1987.

[234] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 67, pp. 383 – 392, 2008.

[235] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.

[236] R. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, August 2006.

[237] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, "The QC relaxation: A theoretical and computational study on optimal power flow," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3008–3018, 2016.

[238] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, "Convex relaxations of chance constrained AC optimal power flow," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2829–2841, May 2018.

[239] A. Venzke and S. Chatzivasileiadis, "Convex relaxations of probabilistic AC optimal power flow for interconnected AC and HVDC grids," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2706–2718, 2019.

[240] M. Vrakopoulou, M. Katsampani, K. Margellos, J. Lygeros, and G. Andersson, "Probabilistic security-constrained AC optimal power flow," in *IEEE Grenoble PowerTech*, June 2013.

[241] S. H. Low, "Convex relaxation of optimal power flow–Part II: Exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, June 2014.

[242] B. Flach, L. A. Barroso, and M. V. F. Pereira, "Long-term optimal allocation of hydro generation for a price-maker company in a competitive market: Latest developments and a stochastic dual dynamic

programming approach," *IET Generation, Transmission and Distribution*, vol. 4,, no. 2, pp. 299–314, 2010.
[243] N. Löhndorf and D. Wozabal, "Gas storage valuation in incomplete markets," *European Journal of Operational Research*, vol. 288, no. 1, pp. 318–330, 2021.
[244] O. Dowson, A. Philpott, A. Mason, and A. Downward, "A multi-stage stochastic optimization model of a pastoral dairy farm," *European Journal of Operational Research*, vol. 274, no. 3, pp. 1077–1089, 2019.
[245] A. Papavasiliou, Y. Mou, L. Cambier, and D. Scieur, "Application of stochastic dual dynamic programming to the real-time dispatch of storage under renewable supply uncertainty," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 547–558, 2018.
[246] H. Yang and H. Nagarajan, "Optimal power flow in distribution networks under N-1 disruptions: A multistage stochastic programming approach," *INFORMS Journal on Computing*, vol. 34, no. 2, pp. 690–709, 2022.
[247] M. Carrión, A. B. Philpott, A. J. Conejo, and J. M. Arroyo, "A stochastic programming approach to electric energy procurement for large consumers," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 744–754, 2007.
[248] N. Secomandi and S. Kekre, "Optimal energy procurement in spot and forward markets," *Manufacturing & Service Operations Management*, vol. 16, no. 2, pp. 270–282, 2014.
[249] Q. Zhang, A. M. Bremen, I. E. Grossmann, and J. M. Pinto, "Long-term electricity procurement for large industrial consumers under uncertainty," *Industrial & Engineering Chemistry Research*, vol. 57, no. 9, pp. 3333–3347, 2018.
[250] D. Molzahn, F. Dörfler, H. Sandberg, S. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.