Value of forecasts in planning under uncertainty

Konstantinos Gatsis, Ufuk Topcu, and George J. Pappas

Abstract— In environments with increasing uncertainty, such as smart grid applications based on renewable energy, planning can benefit from incorporating forecasts about the uncertainty and from systematically evaluating the utility of the forecast information. We consider these issues in a planning framework in which forecasts are interpreted as constraints on the possible probability distributions that the uncertain quantity of interest may have. The planning goal is to robustly maximize the expected value of a given utility function, integrated with respect to the worst-case distribution consistent with the forecasts. Under mild technical assumptions we show that the problem can be reformulated into convex optimization. We exploit this reformulation to evaluate how informative the forecasts are in determining the optimal planning decision, as well as to guide how forecasts can be appropriately refined to obtain higher utility values. A numerical example of wind energy trading in electricity markets illustrates our results.

I. INTRODUCTION

The intrinsically uncertain nature of renewable energy sources, for example their dependence on local weather conditions [1], impedes the reliable operation of the electricity grid [2]. Forecasts about the renewable energy generation provide a means to cope with this uncertainty. Hence the successful incorporation of forecasts into planning grid operation emerges as an important challenge, as well as the problem of obtaining forecasts that provide the most valuable information for planning.

Uncertainty is usually considered in a stochastic framework during the grid planning stage. Generation, for example, is modeled as a random variable, and planning decisions are made a certain length of time before the generation is realized at operation time. Planning in this context can provide probabilistic guarantees, for example, that undesirable events will happen with low probability, or that expected operation costs/utilities are optimized. Examples can be found in, e.g., dispatch problems [3], unit commitment [4], and participation in electricity markets [5], [6].

Solving such stochastic optimization problems relies on the availability of the underlying probability distribution of the uncertain quantities. The renewable generation forecasts however do not provide complete descriptions of the underlying uncertainty, in the sense that they do not completely describe the probability distribution of the generation. Certain questions arise in this context:

1) How can forecasts of the quantities of interest be incorporated in a stochastic planning framework?

2) How informative is a set of given forecasts to a specific planning problem?

The first question is crucial to reliably integrate renewable generation in the grid. The second question becomes practically relevant if obtaining forecasts during the planning stage incurs significant cost. This could be, for example, the computational cost of running complex renewable generation forecasting algorithms, typically combining numerical weather prediction models, historical data, local measurements, and simulations [7]. Alternatively these costs could be monetary, as is the case when grid operators and power producers do not create their own forecasts but instead purchase them in the form of products from companies specializing in forecasting [8], [9]. Understanding how valuable are the forecasts in determining planning decisions could help reduce such associated costs.

In this paper we take a step towards mathematically formalizing and answering the above questions. We consider forecasts in a probabilistic form. This type of forecasts has recently attracted attention as a more sophisticated alternative to the traditional point forecasts for the highly variable wind power [6], [7], [10]. In particular, probabilistic forecasts provide information about the probability distribution of some unknown random quantity of interest, in contrast to the less descriptive point forecast which is a single value thought as the mean value of the random quantity. Motivated by [6], [7], one example of probabilistic forecasts are prediction intervals, i.e., bounds on the probability that the quantity takes values in certain intervals.

We consider a general stochastic planning framework under probabilistic forecasts (Section II). We look for a planning decision that maximizes the expected value of a given utility function, but the probability distribution of the uncertain quantity, with respect to which the expectation is computed, is unknown and only described via the forecasts. We interpret the probabilistic forecasts as constraints defining a set of possible probability distributions for which the decision needs to account. Consequently a robust (max-min) planning problem is formulated to determine the decision that maximizes the worst-case expected utility, where the worst-case expectation is selected from the set of probability distributions consistent with the forecasts.

In Section III we utilize Lagrange duality theory [11] to show that the robust planning problem admits an equivalent reformulation to a single maximization problem with the introduction of auxiliary (dual) variables. This resulting problem is convex under certain relatively mild assumptions on the planning utility function, but it includes an infinite number of constraints. For the special case of forecasts in the form of prediction intervals we show that the constraints can be reduced to a finite number (Section III-A), resulting in

a standard (finite-dimensional) convex optimization problem that can be solved readily [12].

The problem of selecting the worst-case probability distribution given probabilistic descriptions has been discussed previously in the context of optimal uncertainty quantification [13]. The difference in our formulation however is that an additional planning optimization level is considered. Related planning formulations under uncertainty about probability distributions have been considered in, e.g., [14], [15] and references therein, where the optimal planning is solved in a data-driven fashion based on samples from the unknown underlying distribution. Conceptually similar approaches have been considered in the context of power systems, where expected values in stochastic optimization are approximated using samples obtained from wind power forecasting algorithms [4], [16]. In contrast, planning in our work is decoupled from the data collection and does not rely on sample-based approximations, but instead utilizes the output of the forecasting procedures, i.e., the probabilistic forecasts themselves.

A further novelty of our approach is that it allows to characterize how valuable is the forecast information in determining the optimal planning decision. This characterization follows from a sensitivity analysis of the optimal planning objective value with respect to the forecast constraints (Section IV). Moreover, under the hypothesis that forecasts can be refined by, e.g., a forecasting oracle, we develop a sensitivity-driven procedure that sequentially selects forecast refinements to yield a higher robust planning utility value. Numerical examples of the proposed approach for trading wind energy in electricity markets [5], [6] are presented in Section V. Finally, we conclude with a discussion on our contributions and future work in Section VI.

II. PROBLEM DESCRIPTION

We consider a decision-making framework under uncertainty about some quantity of interest x, for example the unknown renewable power generation. We assume that x takes real values in a subset $\mathcal{X} \subseteq \mathbb{R}$ of the reals and is modeled as a random variable following some probability distribution on $\mathcal{X}$. As we will make clear, this distribution is not known completely. Hence there is uncertainty about the distribution of the random quantity of interest.

Suppose a decision-maker needs to select the value of a controlled parameter b before the random quantity x is realized, i.e., drawn from its probability distribution. We assume b takes values in a convex subset $\mathcal{B} \subseteq \mathbb{R}$ of the reals. To rank the different choices for b we assume that a function $J(x, b)$ models the utility to the user if decision b is taken a priori and the unknown quantity takes value x. Technically we assume the utility function is continuous in both variables x and b, and also concave in b for every value of $x \in \mathcal{X}$, and concave in x for every value of $b \in \mathcal{B}$.

Example 1. Consider the case of a wind power plant participating in a day-ahead electricity market [5], [6]. The producer provides a capacity bid in the market, which we can denote by a decision variable b, representing the estimated renewable power generation that will be supplied to the grid
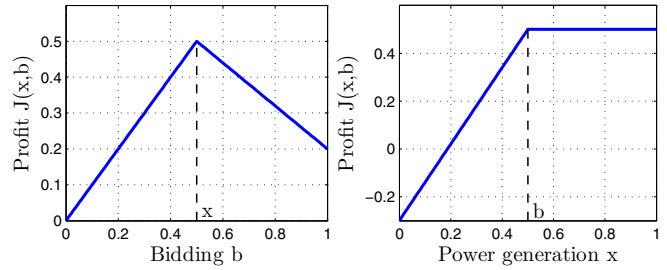

Fig. 1. Profit $J(x, b)$ from bidding in electricity markets for $p = 1$, $q = 1.6$. On the left generation is kept constant $x = 0.5$ while bid b varies. On the right bid is kept constant $b = 0.5$ while generation x varies.

at a future time interval. We let b take values in the unit interval, $b \in [0, 1]$, which can be interpreted as a normalized percentage with respect to the maximum capacity of the generator. If we denote by x the actual realized power, also normalized in $x \in [0, 1]$, a simple model [5] of the producer's monetary profit from a realization x after bidding b is

$$J(x, b) = p\, b - q[b - x]_+, \tag{1}$$

where $[\theta]_+ = \max\{\theta, 0\}$. The constant $p > 0$ rewards high bids, while the constant $q > 0$ penalizes a shortfall of generation compared to the bid, i.e., when $b > x$. It is assumed that $q > p$, implying that the profit decreases for higher shortfalls. The utility function satisfies our general convexity assumptions as can be seen in Fig. 1.

A common assumption in planning problems under uncertainty is that the underlying probability distribution F of the unknown quantity x is known, and one searches for the optimal decision b that maximizes the expected utility $\mathbb{E}_{x \sim F}\, J(x, b)$. In this paper however the underlying distribution F is not completely known but only forecasts about the random quantity x are available. We consider probabilistic forecasts [6], [7], which provide information about the underlying probability distribution F of x. More specifically we consider given (measurable) functions $g_i(x)$, for $i = 1, \ldots, n$, of the random quantity x, and forecasts stating that the expected value of these functions is upper bounded by some parameters ε_i, i.e.,

$$\mathbb{E}_{x \sim F}\, g_i(x) \leq \varepsilon_i, \ i = 1, \ldots, n. \tag{2}$$

Both the functions $g_i(x)$ and the parameters ε_i in these inequalities are given to the decision-maker, e.g., provided by some forecasting algorithm. We interpret the forecasts (2) as a set of constraints on the unknown underlying distribution F, narrowing the uncertainty of the decision-maker regarding the distribution of x. Based on this interpretation, we will interchangeably use the terms forecasts and constraints when referring to (2).

The constraint-based characterization of the uncertainty in (2) matches many common types of forecasts. For example bounds on mean value, second moment, or higher-order moments can be expressed in the form of (2) by appropriately selecting the function $g_i(x)$. In the following example we detail another special case, the prediction intervals. We will revisit this case later.
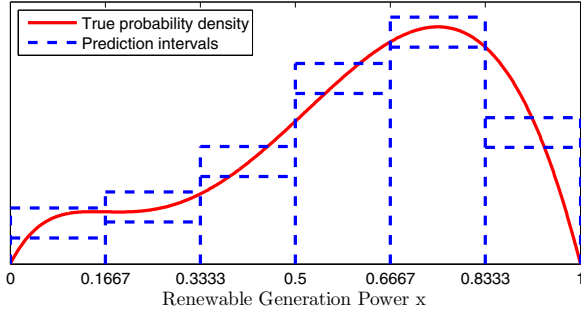
Fig. 2. Pictorial representation of prediction intervals (Example 2). The true but unknown probability density function of the renewable generation is plotted. Forecasts are available in the form of prediction intervals, which are upper and lower bounds on the true probability mass placed on $m = 6$ equal intervals of generation amounts. For illustration purposes the prediction intervals are shown by blocks, with the areas under the blocks equal to the upper and lower bounds, $\overline{\delta}_i$ and $\underline{\delta}_i$ respectively, according to (3).

Example 2. An example motivated by recent forecast methods for renewable generation [6] are prediction intervals. Suppose the renewable generation x takes values in a normalized interval $\mathcal{X} = [0, 1]$, partitioned into m sub-intervals $[x_{i-1}, x_i)$ for $i = 1, \ldots, m$, where $x_0 = 0$, $x_m = 1$, and the forecasts take the form

$$\underline{\delta}_i \leq \mathbb{P}(x_{i-1} \leq x < x_i) \leq \overline{\delta}_i, \quad i = 1, \ldots, m. \quad (3)$$

In other words, this forecast provides upper and lower bounds on the probability that x takes value in each of the intervals. These forecasts can be reformulated into the form of (2) by defining the indicator functions

$$g_i(x) = \mathbb{1}\{[x_{i-1}, x_i)\} = \begin{cases} 1 & \text{if } x_{i-1} \leq x < x_i, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

for $i = 1, \ldots, m$. Then the constraints (3) are equivalently written in the form of (2) as

$$\mathbb{E}_{x \sim F}\, g_i(x) \leq \overline{\delta}_i, \text{ and } \mathbb{E}_{x \sim F}\, -g_i(x) \leq -\underline{\delta}_i. \quad (5)$$

A pictorial representation is given in Fig. 2.

We make the following technical assumption about the forecasts provided in (2).

Assumption 1. There exists a probability distribution F on $\mathcal{X}$ such that the forecasts (2) are satisfied with strict inequality, i.e., $\mathbb{E}_{x \sim F}\, g_i(x) < \varepsilon_i$ for all $i = 1, \ldots, n$.

This assumption states that the constraints (2) are feasible, and furthermore that strict feasibility holds. To motivate the feasibility, we can think of the actual underlying distribution of the random quantity, which is not known to the decision-maker, as satisfying the constraints (2). The strict feasibility is assumed for certain technical reasons in order to establish the results in the sequel of this paper, but is not practically restrictive. The parameters ε_i can always be perturbed to make the strict inequality assumption hold.

Given the constraint-based interpretation of the forecasts in (2), we can alternatively describe the planner's uncertainty concerning the true distribution F of the unknown quantity x by a set $\mathcal{F}$ containing the (possibly uncountably many) distributions that satisfy the forecasts (2). The set of all distributions consistent with forecasts (2) is

$$\mathcal{F}(\varepsilon) = \{F \in \mathcal{D}_{\mathcal{X}} : \mathbb{E}_{x \sim F}\, g_i(x) \leq \varepsilon_i, \; i = 1, \ldots, n\}, \quad (6)$$

where we denote the set of all probability distributions on $\mathcal{X}$ by $\mathcal{D}_{\mathcal{X}}$. Here we parametrize the set with the forecast bounds ε_i, $i = 1, \ldots, n$, grouped in a vector ε. For the current problem development, ε is fixed, but we introduce this parametrization for later use in Section IV. By Assumption 1 the set $\mathcal{F}(\varepsilon)$ has a nontrivial interior, e.g., the true underlying distribution belongs in the set.

Having defined the utility $J(x, b)$ of different choices b to the decision-maker, as well as the possible distributions $F \in \mathcal{F}(\varepsilon)$ of the random quantity x that the decision-maker can anticipate according to the probabilistic forecasts, we now pose a planning problem. The decision-maker needs to determine the value b that robustly maximizes the worst-case expected utility anticipated from the forecasts, mathematically captured as a robust optimization problem

$$P^*(\varepsilon) = \underset{b \in \mathcal{B}}{\text{maximize}} \; \underset{F \in \mathcal{F}(\varepsilon)}{\inf} \; \mathbb{E}_{x \sim F}\, J(x, b). \quad (7)$$

In other words, the planner looks for a decision that leads to the most favorable objective value assuming the worst-case distribution among the ones consistent with the forecast $\mathcal{F}(\varepsilon)$. We denote the robustly optimal objective value and decision by $P^*(\varepsilon)$, $b^*(\varepsilon)$ respectively, assuming they exist, again parametrized by the forecast bounds ε for later use.

The difficulty in solving the robust planning problem (7) lies in the max-min formulation. For every choice b, one needs to solve a minimization problem with respect to the distribution F to evaluate how good the choice is, and then infer how to improve on b to maximize the objective function. To overcome this complexity, in Section III, we examine how the robust planning problem (7) can be equivalently written as a single-layer optimization problem. For the special case of forecasts given by prediction intervals (Example 2), we obtain an equivalent convex optimization problem from which $b^*(\varepsilon)$ can be determined efficiently.

Then in Section IV we characterize how informative the n forecasts given by (2) are in determining the optimal planning decisions by examining the sensitivity of the robust planning value $P^*(\varepsilon)$ to changes in the forecast parameters ε. Based on this analysis, we also develop a methodology for improving the planning utility by appropriately refining forecasts.

III. ROBUSTLY OPTIMAL PLANNING

The robust planning problem described in (7) involves a max-min structure that makes it hard to determine the robustly optimal decision $b^*(\varepsilon)$. In this section we follow an alternative path based on Lagrange duality theory [11]. Under certain technical conditions the inner minimization in (7) over probability distributions F of the unknown quantity x is equivalent to its Lagrange dual (maximization) problem. Replacing then the inner minimization in (7) with the equivalent maximization yields a maximization problem (with a single optimization layer) over the planning decision

variable b and additional (dual) variables. This result is stated in the following theorem, and the proof relies on [11].[1]

Theorem 1. *If Assumption 1 holds, the robust planning problem defined in (7) is equivalent to the following optimization problem*

$$\underset{b,\lambda,\eta}{\text{maximize}} \quad -\sum_{i=1}^{n} \lambda_i\, \varepsilon_i - \eta \tag{8}$$

$$\text{subject to} \quad J(x,b) + \sum_{i=1}^{n} \lambda_i\, g_i(x) + \eta \geq 0,\ \forall x \in \mathcal{X}, \tag{9}$$

$$b \in \mathcal{B},\ \lambda \in \mathbb{R}^n_+,\ \eta \in \mathbb{R}. \tag{10}$$

According to the theorem, the optimal choice $b^*(\varepsilon)$ of the robust planning problem (7) can be equivalently found by the optimization problem (8)-(10), and the optimal values of the two problems are equal. The advantage of the representation in (8)-(10) is that it bypasses the max-min structure of (7). There is a finite number of optimization variables in (8)-(10), i.e., the decision b, as well as the dual variables $\eta \in \mathbb{R}$ and the vector $\lambda \in \mathbb{R}^n_+$. The caveat, although, is that the number of constraints is infinite and uncountable. Such optimization problems are called semi-infinite – see [18] for an overview. Nevertheless it is a convex optimization problem since the objective is linear and the constraint (9) for each value of x defines a convex set due to the concavity of function $J(x,b)$ in variable b for any value x.

General approaches for solving semi-infinite programs can be found in [18]. We examine however next the special case in which the forecasts (2) are given in the form of prediction intervals, described in Example 2. In that case it turns out that the number of constraints in the equivalent problem (8)-(10) can be reduced to a finite number. Hence we can pose the robust planning decision under prediction interval forecasts as a standard finite-dimensional convex optimization problem, for which efficient algorithms exist.

After this special case, we show in the following section that the optimal values of the additional variables λ in the optimization problem (8)-(10) can be interpreted as indicators of how valuable are the given forecasts in determining the optimal planning. Based on this interpretation, we also examine in the following section how updates on the forecasts can yield a higher planning objective value.

A. Robust planning under prediction intervals

Consider forecasts given in the form of prediction intervals described in (3). By their reformulation into the generic form of forecasts presented in (5), we can pose the robust planning decision problem (7) into its equivalent form provided by Theorem 1 in (8)-(10). In particular, introducing dual variables $\bar{\lambda}_1, \ldots, \bar{\lambda}_m$ for the upper bound constraints in (5), and $\underline{\lambda}_1, \ldots, \underline{\lambda}_m$ similarly for the lower bounds, we have that

---

[1] The proofs of the results in this paper are omitted due to space limitations, but can be found in [17].

robust planning follows from the solution to the problem

$$\underset{b\in\mathcal{B},\, \lambda\geq 0,\, \eta}{\text{maximize}} \quad \sum_{i=1}^{m} \left(\underline{\lambda}_i\, \underline{\delta}_i - \bar{\lambda}_i\, \bar{\delta}_i \right) - \eta \tag{11}$$

$$\text{subject to} \quad J(x,b) + \sum_{i=1}^{m} (\bar{\lambda}_i - \underline{\lambda}_i)\, \mathbb{1}\{[x_{i-1}, x_i]\}$$
$$+ \eta \geq 0, \quad \text{for all } x \in [0,1]. \tag{12}$$

We note that the continuum of constraints (12) can be equivalently separated into the given intervals

$$J(x,b) + (\bar{\lambda}_i - \underline{\lambda}_i) + \eta \geq 0, \quad \text{for all } x \in [x_{i-1}, x_i), \tag{13}$$

for all $i = 1, \ldots, m$. Note also that by the assumption that the function $J(x,b)$ is concave in x for all b, it follows that $J(x,b) \geq \min\{J(x_{i-1},b), J(x_i,b)\}$, where the inequality is tight by continuity of function $J(x,b)$. Hence instead of checking (13) for a continuum of values $x \in [x_{i-1}, x_i)$ for each $i = 1, \ldots, m$, we only need to check at the two endpoints x_{i-1} and x_i. In other words, the problem (11)-(12) can be equivalently written as

$$\underset{b\in\mathcal{B},\, \lambda\geq 0,\, \eta}{\text{maximize}} \quad \sum_{i=1}^{m} \left(\underline{\lambda}_i\, \underline{\delta}_i - \bar{\lambda}_i\, \bar{\delta}_i \right) - \eta \tag{14}$$

$$\text{subject to} \quad J(x_{i-1},b) + \bar{\lambda}_i - \underline{\lambda}_i + \eta \geq 0, \tag{15}$$
$$J(x_i,b) + \bar{\lambda}_i - \underline{\lambda}_i + \eta \geq 0, \quad i = 1, \ldots, m.$$

This optimization problem is convex and has a finite number of constraints. The number of constraints is twice the number of initial forecast intervals because essentially each forecast of the robust optimization is converted into two constraints, one for each of the two interval endpoints. The solution of (14) can be determined efficiently by standard convex optimization algorithms.

IV. RELATIVE VALUE OF FORECASTS

In this section we aim to quantify the value of the given forecasts (2) to the planner who needs to solve the robust planning problem (7). In particular, (2) defines a set of n given forecasts and we are interested in determining how valuable information does each one of them provide to the planner. To this end, we examine how sensitive the robustly optimal objective $\mathcal{P}^*(\varepsilon)$, defined in (7), is to the given forecast values ε, i.e., what would the objective value become for deviations from the given parameters ε.

The forecast parameters ε determine the objective $P^*(\varepsilon)$ defined in (7) via the set of distributions $\mathcal{F}(\varepsilon)$ appearing in the constraint of the inner minimization. To examine the sensitivity of $P^*(\varepsilon)$ when the given forecast parameters ε change to some new values $\varepsilon - \Delta\varepsilon$ for some $\Delta\varepsilon \in \mathbb{R}^n$, with the negative sign chosen for convention, we leverage the problem reformulation provided by Theorem 1. We exploit the well-known convex optimization fact that in general the optimal values of the Lagrange dual variables λ express the sensitivity of the objective of a convex optimization problem with respect to changes in the constraints - see [12, Ch. 5.6]. In our case, the robust planning (7) involves a two-layer (max-min) optimization objective, but a similar sensitivity result can be obtained, as we state next.

Theorem 2. *If Assumption 1 holds, then*

$$P^*(\varepsilon - \Delta\varepsilon) \geq P^*(\varepsilon) + \sum_{i=1}^{n} \lambda_i^*(\varepsilon) \, \Delta\varepsilon_i, \qquad (16)$$

holds for any $\Delta\varepsilon \in \mathbb{R}^n$, where $\lambda_i^*(\varepsilon)$, for $i = 1, \ldots, n$, is the optimal solution to problem (8).

The theorem provides a lower bound on the optimal planning objective value of (7) after changes in the forecast bounds ε provided in (2). The change in the lower bound compared to $P^*(\varepsilon)$ is proportional to the change from ε to $\varepsilon - \Delta\varepsilon$. The rate of change in each direction $i = 1, \ldots, n$ is given by the optimal value of the (dual) variable $\lambda_i^*(\varepsilon)$. We can thus interpret the values $\lambda_i^*(\varepsilon)$ as indicators of how *valuable* each one of the forecasts (2) is to the planner. A forecast i with a high value $\lambda_i^*(\varepsilon)$ is more informative compared to other forecasts $j \neq i$ because a small change in the given forecast bound ε_i gives a significant change in the (lower bound of the) objective. Moreover the values $\lambda_i^*(\varepsilon)$ are already obtained during solving problem (8), hence no further computation is required to determine them.

Remark 1. The sensitivity analysis provided by Theorem 2 is performed with respect to the given forecast parameters ε, i.e., they are obtained locally for ε. At some other forecast parameter ε', the sensitivities, i.e., the values $\lambda_i^*(\varepsilon')$ will differ. This difference matches the intuition that different forecasts provide different information. On the other hand the theorem provides a lower bound on the objective value at any deviation $\varepsilon - \Delta\varepsilon \in \mathbb{R}^n$, not just locally around ε. It is also worth noting that the theorem does not provide any guarantee on how far the lower bound is with respect to the actual objective for such deviations.

Given the sensitivity-based characterization of how informative some given forecasts are, we next propose a method for refining/updating the forecasts in a way that increases the utility of the decision-making problem.

A. Forecast refinements

Suppose the decision-maker has access to a forecasting oracle, which upon request can refine one of the n given forecasts in (2) – see Remark 2. Inspecting the form of the forecasts in (2) written as inequalities, a refined forecast of some inequality i can be represented by decreasing the bound ε_i to a lower value $\varepsilon_i - \Delta\varepsilon_i$. Equivalently, such decrease has the effect of shrinking the set $\mathcal{F}(\varepsilon)$ of possible probability distributions that the decision-maker has to consider, as defined in (6), to a subset $\mathcal{F}(\varepsilon - \Delta\varepsilon) \subseteq \mathcal{F}(\varepsilon)$.

A question that arises in this context is how the decision-maker should select which of the n forecasts of (7) to refine. This question is particularly important if the cost of refining forecasts is high, e.g., the computational cost of running extra simulations of a forecasting algorithm. The characterization of Theorem 2 suggests that the forecast that provides the most valuable information should be selected, i.e., the forecast i with the highest value $\lambda_i^*(\varepsilon)$. Based on this intuition we propose the iterative forecast refinement procedure shown in Algorithm 1. On every iteration sensitivity values are obtained by solving the planning problem, and a

---

Algorithm 1 Sensitivity-driven forecast refinement

Input: Utility function $J(x, b)$, forecast functions $g_i(x)$ for $i = 1, \ldots, n$, initial forecast parameters $\varepsilon(0) \in \mathbb{R}^n$
Output: Robustly optimal planning decision b
1: $k \leftarrow 0$
2: **repeat**
3:     Solve robust planning (8)-(10) with forecast parameters $\varepsilon(k)$, and determine optimal planning $b^*(\varepsilon(k))$ and sensitivity values $\lambda^*(\varepsilon(k)) \in \mathbb{R}^n$
4:     Select $j = \operatorname{argmax}_{1 \leq i \leq n} \lambda_i^*(\varepsilon(k))$
5:     Request refinement in jth forecast $\varepsilon_j(k+1) \leq \varepsilon_j(k)$
6:     Keep $\varepsilon_i(k+1) = \varepsilon_i(k)$ for all $i \neq j$
7:     $k \leftarrow k + 1$
8: **until** Termination condition
9: **return** Final planning decision $b^*(\varepsilon(k))$

---

refinement for the forecast bound with the highest sensitivity is requested (e.g. from the oracle) while all other bounds are kept constant. The procedure terminates, for example, if no further refinements are possible, or if the objective value improvements become insignificant.

Even though this procedure is motivated mainly by intuition, without theoretical claims in terms of optimality, convergence, etc., we note the following fact. On each iteration, according to (16) of Theorem 2, the robustly optimal utility value increases by at least an amount equal to $\lambda_j^*(\varepsilon(k)) \, (\varepsilon_j(k+1) - \varepsilon_j(k))$. This minimum increase is proportional to the amount of change in the selected forecast parameter. If the selected forecast j cannot be refined, i.e., $\varepsilon_j(k+1) = \varepsilon_j(k)$, then no improvement in decision-making is achieved. In general, however, the result matches the intuition that with a better forecast, i.e., imposing a more constraining set $\mathcal{F}(\varepsilon - \Delta\varepsilon)$ in the minimization of (7), the objective value of the planning can only improve, i.e., the optimal value of (7) cannot decrease.

In the following section we present a numerical example for a wind power plant participating in an electricity market (Example 1) under prediction interval forecasts (Example 2). We demonstrate the methodology derived by Theorem 1 for determining the robustly optimal bidding decision. Additionally we perform the sensitivity analysis developed in this section and we implement the sensitivity-driven procedure for refining forecasts.

Remark 2. The methodology for refining forecasts in this section is developed without considering a specific forecasting procedure. For example, if an oracle obtains the forecasts by sampling from the true distribution of the unknown quantity through simulations of an underlying stochastic model, forecasts can be refined by drawing extra samples. However, a large number of extra samples might be required to improve upon the estimates of all the integrals in (2), e.g., in a Monte Carlo fashion. By focusing on only one of the n integrals during refinement, as the oracle in our hypothesis does, the number of required samples can be reduced. This can be performed by variance reduction techniques, such as importance sampling [19], which aim at sampling more often
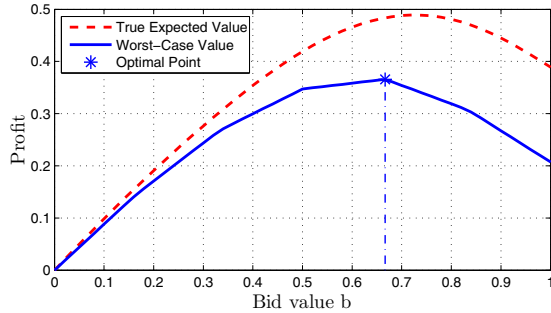
Fig. 3. Expected profit and worst-case expected profit as a function of the chosen bidding for the case considered in the numerical example. The worst-case value is always an under-approximation of the true one. The robustly optimal bidding balances between the shortfall penalties and gains from a high bid.

	$\bar{\lambda}_1$	$\bar{\lambda}_2$	$\bar{\lambda}_3$	$\bar{\lambda}_4$	$\bar{\lambda}_5$	$\bar{\lambda}_6$
Iter 1	0.66	0.39	0.12	0.00	0.00	0.00
Iter 2	0.80	0.53	0.27	0.00	0.00	0.00
Iter 3	0.80	0.53	0.27	0.00	0.00	0.00
Iter 4	1.07	0.80	0.53	0.27	0.00	0.00

TABLE I

SENSITIVITIES TO UPPER FORECAST BOUNDS

	$\underline{\lambda}_1$	$\underline{\lambda}_2$	$\underline{\lambda}_3$	$\underline{\lambda}_4$	$\underline{\lambda}_5$	$\underline{\lambda}_6$
Iter 1	0.00	0.00	0.00	0.14	0.41	0.41
Iter 2	0.00	0.00	0.00	0.00	0.27	0.27
Iter 3	0.00	0.00	0.00	0.00	0.27	0.27
Iter 4	0.00	0.00	0.00	0.00	0.00	0.00

TABLE II

SENSITIVITIES TO LOWER FORECAST BOUNDS

from values important in estimating the selected integral.

V. APPLICATION: BIDDING IN ELECTRICITY MARKETS UNDER PREDICTION INTERVALS

In this section we apply our theoretical results in a numerical example of bidding renewable energy in electricity markets, introduced in Example 1, with the utility function in (1) representing the profit from a bid $b \in [0,1]$ and a generation $x \in [0,1]$. Suppose the generator obtains forecasts about the future generation value x in the form of prediction intervals, introduced in Example 2. We consider the problem of determining the value of bidding that robustly maximizes the expected profit to the generator subject to the given forecasts about the renewable generation. This problem is mathematically described by (7). We adopt the convex optimization reformulation presented in Section III-A, in particular the form (14)-(15).

We consider forecasts for the probability that the random generation x takes values in each of m equal intervals that partition the total range of values $[0,1]$, i.e., $x_i = i/m$ for $i = 0, \ldots, m$ – see also Example 2. To derive the lower and upper bounds, $\underline{\delta}_i$ and $\bar{\delta}_i$ respectively according to (3), we adopt the following procedure. We fix some true underlying probability distribution of the generation x and we produce bounds $\underline{\delta}_i, \bar{\delta}_i$ by perturbing the true value $\mathbb{P}(x_{i-1} \leq x < x_i)$ computed with the underlying distribution. Specifically we consider perturbation by a constant amount below and above $\mathbb{P}(x_{i-1} \leq x < x_i)$ for each $i = 1, \ldots, m$. Then only the resulting lower and upper bounds, $\underline{\delta}_i$ and $\bar{\delta}_i$ respectively, are available to the generator, not the underlying probability distribution used to create them.

Given $m = 6$ prediction intervals shown in Fig. 2, we first solve the convex optimization (14)-(15) for reward and penalty values $p = 1, q = 1.6$ in (1) and get an optimal bidding value $b^* \approx 0.67$. In Fig. 3 we illustrate the value of the worst-case expected profit, i.e., the objective in the planning problem (7), for all bid values b as well as the optimal point. For comparison we also illustrate the expected profit computed using the underlying true probability distribution of x for all bid values b. The worst-case expected profit is an under-approximation of the true expected value, since the forecasts provide incomplete information about the underlying distribution. The optimal choice b^* balances between the two extremes 0 and 1, as expected by intuition,

with $b = 0$ being the most "secure" bid but yielding a zero expected profit, and $b = 1$ being the most "risky" bid since it is certain that a shortfall $x < b = 1$ will occur.

Moreover, we examine how the information provided by the given forecasts of Fig. 2 is valued by the decision maker. According to Theorem 2 of Section IV the sensitivity of the robust bidding profit to the set of given forecasts is captured by the optimal values of the dual variables. In particular the sensitivity to upper and lower bounds of prediction intervals, $\bar{\delta}_i$ and $\underline{\delta}_i$ respectively, for $i = 1, \ldots, m$, is captured by the dual variables $\bar{\lambda}_1, \ldots, \bar{\lambda}_m$ and $\underline{\lambda}_1, \ldots, \underline{\lambda}_m$ respectively, as introduced in Section III-A. The sensitivity values in our example are shown in the first row of Tables I, II.

We observe that at each interval $i = 1, \ldots, m$, at most one of the upper or lower bound sensitivities $\bar{\lambda}_i, \underline{\lambda}_i$ is non-zero. The reason is that at most one of two bounds (cf. (3)) is tight for the worst-case distribution at the robustly optimal point of problem (7). Moreover, we see that the optimal profit is sensitive to the probability of both having a low generation (captured by $\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3$) as well as having a high generation (captured by $\underline{\lambda}_4, \underline{\lambda}_5, \underline{\lambda}_6$). In other words information about these events is the most informative. The sensitivity is higher in the low generation case because the profit is significantly affected by shortfalls in this case.

Next we adopt the sensitivity-driven methodology of Section IV to refine the given forecasts so that they become more informative for robust bidding. Following Algorithm 1 we iteratively select to refine the forecast with the highest sensitivity value, i.e., one of the upper or lower bounds $\bar{\delta}_i$ and $\underline{\delta}_i$ for an interval i in (3). If an upper bound $\bar{\delta}_i$ is selected it gets decreased, and respectively if a lower bound $\underline{\delta}_i$ is selected it gets increased. In our example, the resulting new bound needs to be consistent with the true underlying distribution. For simplicity here we model a forecasting oracle who can upon request refine each upper/lower bound by a constant decrease/increase, and the new bound is close enough to the true value so that the oracle cannot refine it any further. In general, hovever, Algorithm 1 can be applied regardless of how the forecasting oracle operates.

The sensitivity values after each iteration of Algorithm 1 are shown in Tables I, II. As already mentioned, $\bar{\lambda}_1$ has the higher value initially, so the upper bound on the first interval is refined (i.e., reduced). On the second iteration, after solving again the optimal planning (14)-(15), the new
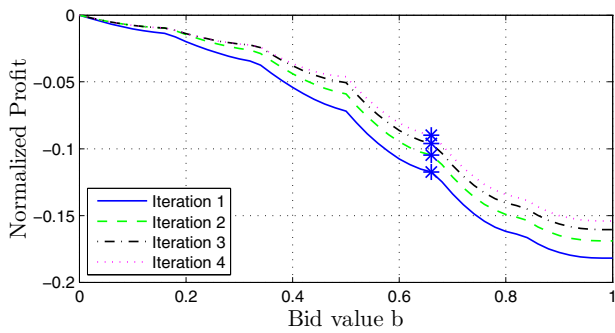
Fig. 4. Normalized worst-case profit as a function of bid value after iterations of the forecast refinement algorithm. The plotted values are normalized by subtracting the true underlying expected profit. The refinements increase the worst-case profit and bring it closer to the true value. The robustly optimal biding values are also indicated.

sensitivity values are obtained and again $\bar{\lambda}_1$ is the largest. Since the forecasting oracle described in our example cannot reduce the corresponding bound any further, the bound with the second largest sensitivity is selected to be refined, which is $\bar{\lambda}_2$. On the third iteration, again $\bar{\lambda}_1, \bar{\lambda}_2$ are the largest, but due to the oracle, the third largest $\bar{\lambda}_3$ is selected, and so on. We note that after some iterations the sensitivities of all lower bounds $\underline{\lambda}_i$ become zero, meaning that they do not provide any valuable information, and refining/increasing them does not necessarily offer an improvement in profit.

Finally, to examine the improvements in the bidding profit after these iterations of the refinement algorithm, we plot in Fig. 4 the worst-case expected profit, i.e., the objective in the planning problem (7), for all bid values b. For visualization reasons the values are normalized by subtracting the true expected profit (the latter is larger, as shown in Fig. 3, so the plotted values are negative). As the forecasts are refined, the worst-case expected profit increases and gets closer to the actual expected profit (the baseline zero). In other words, the refined forecasts provide a more accurate model of the true objective. The largest discrepancy between worst-case and true expected value happens at the high bid region, since the worst-case scenario assumes that generation takes the lowest possible values with the highest possible probability, making the associated shortfall costs large. In this example even though forecast refinements increase the optimal profit value, the optimal bidding decision in Fig. 4 is the same. This is a consequence of the piecewise linear utility function in (1). For more general utility functions the robust decision may change as well during refinements, along with the optimal objective values.

VI. CONCLUDING REMARKS AND FUTURE WORK

We examine how planning problems under uncertainty can utilize forecasts about uncertain quantities. By equivalently expressing forecasts as constraints on the possible probability distributions that the uncertain quantity can follow, the problem is cast as a robust optimization. The optimal planning seeks to maximize the expected value of a given utility function, integrated with respect to the worst-case distribution consistent with the forecasts. A reformulation into a convex optimization problem is presented, from which we can extract information about how valuable are the forecasts in determining the optimal planning decision.

Supposing a forecasting oracle can refine the given forecasts upon request, a sensitivity-driven approach to iterative forecast refinement is presented. Coupling the theoretical approach with practical forecasting algorithms requires further exploration to account for, e.g., the ability of such algorithms to provide new forecasts instead of strict refinements, or the limited ability for further refining in other cases. Future work also includes generalizations of the approach to uncertain quantities correlated over time, as in the scenario-based forecasts of [20], or over space, e.g., renewable generation at different physical locations in the grid.

REFERENCES

[1] A. Mills, M. Ahlstrom *et al.*, "Dark shadows," *IEEE Power and Energy Magazine*, vol. 9, no. 3, pp. 33–41, 2011.
[2] Y. V. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational impacts of wind generation on California power systems," *IEEE Trans. on Power Systems*, vol. 24, no. 2, pp. 1039–1050, 2009.
[3] P. P. Varaiya, F. F. Wu, and J. W. Bialek, "Smart operation of smart grid: Risk-limiting dispatch," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 40–57, 2011.
[4] E. M. Constantinescu, V. M. Zavala, M. Rocklin, S. Lee, and M. Anitescu, "A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation," *IEEE Trans. on Power Systems*, vol. 26, no. 1, pp. 431–441, 2011.
[5] E. Bitar, R. Rajagopal, P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *IEEE Trans. on Power Systems*, vol. 27, no. 3, pp. 1225–1235, 2012.
[6] P. Pinson, C. Chevallier, and G. N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," *IEEE Trans. on Power Systems*, vol. 22, no. 3, pp. 1148–1156, 2007.
[7] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann *et al.*, "Wind power forecasting: state-of-the-art 2009." *Online Technical Report*, 2009.
[8] K. Parks, Y.-H. Wan, G. Wiener, and Y. Liu, "Wind energy forecasting: A collaboration of the National Center for Atmospheric Research (NCAR) and Xcel Energy," *Online Technical Report*, October 2011.
[9] *Innovations Across The Grid: Partnerships Transforming The Power Sector*. The Edison Foundation, December 2013, Available online.
[10] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.
[11] A. Shapiro, "On duality theory of conic linear problems," in *Semi-infinite programming*. Springer, 2001, pp. 135–165.
[12] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.
[13] S. Han, U. Topcu, M. Tao, H. Owhadi, and R. M. Murray, "Convex optimal uncertainty quantification: Algorithms and a case study in energy storage placement for power grids," in *Proc. of the American Control Conference*, 2013, pp. 1130–1137.
[14] 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.
[15] E. Erdoğan and G. Iyengar, "Ambiguous chance constrained problems and robust optimization," *Mathematical Programming*, vol. 107, no. 1-2, pp. 37–61, 2006.
[16] Z. Zhou, A. Botterud, J. Wang, R. Bessa, H. Keko, J. Sumaili, and V. Miranda, "Application of probabilistic wind power forecasting in electricity markets," *Wind Energy*, vol. 16, no. 3, pp. 321–338, 2013.
[17] K. Gatsis, U. Topcu, and G. J. Pappas, "Value of forecasts in planning under uncertainty: Extended version," *Technical report available at arXiv.org*, 2014.
[18] M. López and G. Still, "Semi-infinite programming," *European Journal of Operational Research*, vol. 180, no. 2, pp. 491–518, 2007.
[19] B. D. Ripley, *Stochastic Simulation*. John Wiley & Sons, 2009.
[20] P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klöckl, "From probabilistic forecasts to statistical scenarios of short-term wind power production," *Wind Energy*, vol. 12, no. 1, pp. 51–62, 2009.