Aggregating Correlated Wind Power with Full Surplus Extraction

Wenyuan Tang, *Member, IEEE,* and Rahul Jain, *Member, IEEE*

***Abstract*—We study the problem of designing profit-maximizing mechanisms for an aggregator who aggregates wind power from a group of wind power producers (WPPs). The WPPs have more refined forecasts of the wind power generation than the aggregator. Such forecasts are their private information, which also give the reservation utilities of the WPPs. The goal of the aggregator is to elicit the private information truthfully, while paying them as little as possible. Inspiring by the fact that those forecasts are typically correlated due to the geographical proximity of the WPPs, we formally define the full correlation condition, which holds ubiquitously in practice. Under that condition, we construct an optimal mechanism which yields the truthful elicitation, while extracting the full surplus (i.e., with minimum payments equal to the reservation utilities) in expectation. Finally, we conduct a case study based on the real-world data, which empirically validates the results.**

***Index Terms*—Electricity market, mechanism design, wind power aggregation.**

NOMENCLATURE

N	Number of WPPs.
i	WPP index taking values from 1 to N.
W_i	Random variable representing the wind power generation of WPP i.
W	Random vector of all W_i's.
W_{-i}	Random vector of all W_i's except W_i.
$g_i(w_i)$	Probability density function of W_i.
K	Number of components for the Gaussian mixture distribution.
k	Index taking values from 1 to K.
L	Constant equal to K^{N-1}.
l	Index taking values from 1 to L.
$\alpha_{i,k}$	Prior probability of picking the kth component for WPP i.
$\mu_{i,k}$	Mean of the kth component for WPP i.
$\sigma_{i,k}^2$	Variance of the kth component for WPP i.
θ_i	Type of WPP i.
Θ_i	Space of θ_i.
θ	Type of all WPPs.
θ_{-i}	Type of all WPPs except WPP i.
Θ_{-i}	Space of θ_{-i}.
$\phi(\theta)$	Prior distribution of θ.
$\phi_i(\theta_{-i}\|\theta_i)$	Conditional distribution of θ_{-i} given θ_i.
p	Day-ahead contract price.
q	Real-time penalty price.
$F_X(\cdot)$	Cumulative distribution function of a random variable X.
$f_X(\cdot)$	Probability density function of a random variable X.
$V_i(\theta_i)$	Expected profit when WPP i participates in the market individually.
$x_i^*(\theta_i)$	Optimal power contract when WPP i participates in the market individually.
$V(\theta)$	Expected profit when all WPPs participate in the market as an aggregate.
$x^*(\theta)$	Optimal power contract when all WPPs participate in the market as an aggregate.
$V_{-i}(\theta_{-i})$	Expected profit when all WPPs except WPP i participate in the market as an aggregate.
$x_{-i}^*(\theta_{-i})$	Optimal power contract when all WPPs except WPP i participate in the market as an aggregate.
$R_i(\theta_i)$	Reservation utility of WPP i.
$\hat{\theta}_i$	Reported type of WPP i.
$\hat{\theta}$	Reported type of all WPPs.
$\hat{\theta}_{-i}$	Reported type of all WPPs except WPP i.
$t_i(\hat{\theta}, W_i)$	Payment to WPP i.
$\tilde{t}_i(\hat{\theta}, W_i)$	Payment to WPP i in the stochastic VCG mechanism.
A_i	Conditional distribution matrix for WPP i.
$h_i(\theta_i)$, $t_i'(\hat{\theta}_{-i})$	Auxiliary variables for contructing the optimal mechanism.
$t_i^*(\hat{\theta}, W_i)$	Payment to WPP i in the optimal mechanism.

Manuscript received April 8, 2016; revised July 20, 2016, November 2, 2016, and February 10, 2017. A preliminary version of this paper appears in SmartGridComm 2014 [1].

W. Tang is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA, and the Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA (e-mail: tangwenyuan@berkeley.edu).

R. Jain is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA (e-mail: rahul.jain@usc.edu)

I. INTRODUCTION

RENEWABLE energy increasingly constitutes a greater fraction of the energy portfolio. California has set an ambitious goal of requiring 33 percent of retail sales from renewable energy by 2020. On the other hand, due to the variability and unpredictability of renewables such as wind and solar, many system challenges arise in integrating renewable energy into the current grid and the electricity market.

The wholesale electricity markets in the United States are mainly two-settlement markets, each of which consists of a day-ahead forward market and a real-time spot market. Consider a wind power producer (WPP) who participates in the two-settlement market individually. In the day-ahead market, the WPP chooses a power contract which specifies the amount of supply in the real-time market. In the real-time market, the WPP pays a penalty for the deviation from the day-ahead commitment. The trading problem for a WPP has been addressed in the literature [2], [3]. The system operation problem involving uncertain wind power, has also been studied [4], [5], in the framework of stochastic programming or robust optimization. Underlying those trading and operation problems is the task of wind power forecasting, which is challenging and still developing. In [6], a statistical distribution model for the forecast errors is developed. Advanced statistical learning techniques have been employed in various forecasting scenarios [7]–[10].

It is well known that aggregating diverse wind power sources can reduce the overall variability, which benefits both the system and the members of the aggregation. Consequently, the profit under aggregation can be greater than the total profit under individual participation. The question is how to fairly allocate the profit under aggregation among the WPPs with random generation. In [11], they address the question in the framework of cooperative game theory. They show that there always exists a payoff allocation that stabilizes the grand coalition. As related work, [12] introduces risky power contracts in addition to firm power contracts to enable flexible and efficient wind power aggregation.

In this work, we follow a different approach. We assume that there is an aggregator who aggregates wind power from the WPPs. Such an aggregator can be a third-party financial entity, who may not own any physical resources in the grid. It is not new that purely financial entities participate in the electricity market. Some common financial instruments include financial transmission rights and virtual bids. Recently, demand response providers (e.g., OhmConnect) have been entering the market. Those providers assist retail customers to lower electricity bills, and then get paid from the market through load curtailment. They play a similar role to an aggregator.

Clearly, the aggregator can potentially extract the surplus from aggregation. But what are the incentives for a WPP to contract with the aggregator? This can be justified as follows. First, it can be costly for a small WPP to participate in the market directly, due to regulation requirements. Second, a WPP is more risk averse and sensitive to the highly volatile real-time penalties, while the aggregator assumes the risk to make a profit. Third, in the proposed mechanism, each WPP is guaranteed to make at least the profit under individual participation, which we call the reservation utility.

Therefore, as a profit-maximizing aggregator, its objective is to extract as much surplus as possible, subject to making some minimum payments to the WPPs. On the other hand, such reservation utilities are private information, which the WPPs, as strategic players, may not have incentives to reveal. The goal of this work is to design mechanisms which elicit truthful information of the WPPs and maximize the profit, on behalf of the aggregator.

Achieving incentive compatibility by itself is not difficult, and the well-known Vickrey-Clarke-Groves (VCG) mechanism [13] provides a widely applicable solution. As an extension, we propose the stochastic VCG mechanism in our earlier work [14] to address the uncertainty of wind power generation. In that setting, the aggregator is a welfare-maximizing social planner. However, for a profit-maximizing aggregator, the stochastic VCG mechanism suffers from the issue of budget deficit.

Since the WPPs are in close proximity to each other, their forecasts of their generation are possibly correlated. Such correlation can be exploited by the aggregator to elicit the forecasts truthfully. Combining the ideas of the existing work on mechanism design with correlated types [15]–[18] and our work on mechanism design with stochastic resources [14], we show that under the full correlation condition, the optimal mechanism extracts the full surplus in expectation. The proposed mechanism naturally handles private reservation utilities, which is an open problem in a general setting [19].

The paper is organized as follows. In Section II, we introduce a Gaussian mixture model of wind power generation, which motivates the mechanism design problem for wind power aggregation. In Section III, we formulate the problem. The main results are presented in Section IV. In Section V, we conduct a case study to validate the theoretical results. Section VI concludes the paper.

II. Statistical Model of Wind Power Generation

Throughout the paper, we use the wind power data from GEFCom2012 [20]. The data contains hourly wind power generation for seven wind farms ("WPPs" hereafter) in the same region of the world from July 1, 2009 to December 31, 2010. Fig. 1 shows the distribution of wind power generation for WPP 1 during hour 1 (the first hour of each day). Note that the generation in the data has been normalized by the respective nominal capacities of the WPPs.

Let W be a random variable denoting the wind power generation. There are many ways to model the distribution of W. We use the Gaussian mixture model, under which the probability density function of W is given by

$$g(w) = \sum_{k=1}^{K} \alpha_k \mathcal{N}(w; \mu_k, \sigma_k^2), \tag{1}$$

where K is the number of mixture components, α_k is the prior probability of picking the kth component, and $\mathcal{N}(\cdot; \mu_k, \sigma_k^2)$ is the Gaussian density of the kth component.

We use K-means clustering to partition the support of W into K clusters. Fix μ_k as the center of each cluster k. For simplicity, μ_k's are sorted such that $\mu_1 < \cdots < \mu_K$. Then we use the expectation-maximization algorithm to obtain α_k and σ_k^2 for each k. When $K = 5$, the result is shown in Fig. 1.

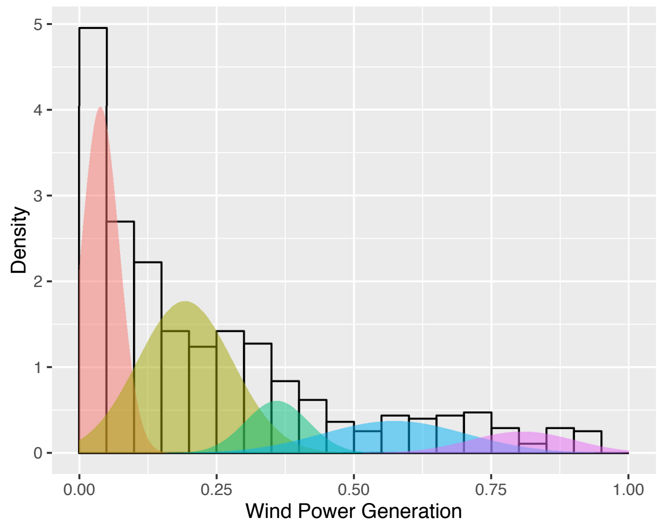

Fig. 1. The distribution of wind power generation for WPP 1 during hour 1, fitted by a Gaussian mixture model.

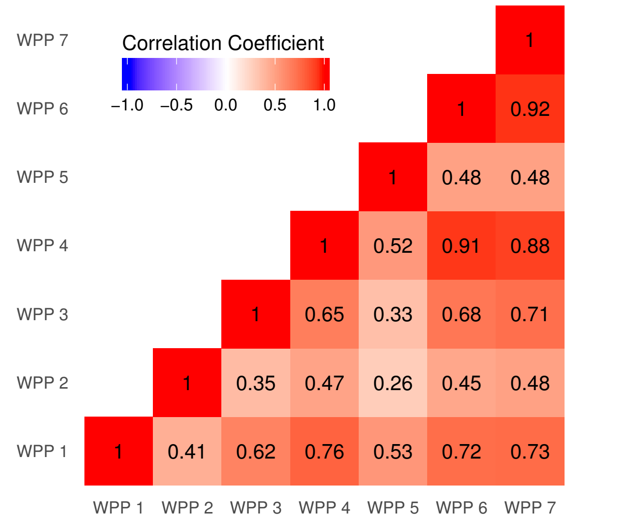

Fig. 2. Correlation matrix heat map of wind power generation for the seven WPPs during hour 1.

Such a mixture model has a meaningful interpretation in the context of wind power aggregation. At the day-ahead stage, the aggregator can only tell that the wind power generation W follows a Gaussian mixture distribution, while the WPP has a more refined forecast. In particular, the WPP knows which of the K Gaussians W is drawn from, i.e., the value of k. Reporting k can be translated into predicting whether the generation of the following day is, say $K = 5$, very low, low, medium, high, or very high. In other words, k gives the posterior distribution of W. We emphasize that we use the mixture model to highlight the layered information structure between the aggregator and the WPPs. The choice of the base distribution is not crucial—it does not have to be Gaussian. The selection of the parameters, including determining the value of K, is a modeling issue, which does not affect our main analysis.

The objective of the profit-maximizing aggregator is to elicit the value of k for each WPP while making minimum payments to them. In general, this task is challenging, as the WPPs may not have incentives to reveal such private information to the aggregator. Fortunately, besides the mixture model, the aggregator can also exploit the correlation structure of the individual forecasts. This is the central idea of this paper.

Fig. 2 shows the correlation matrix heat map of wind power generation for the seven WPPs. It can be seen that the generation is highly correlated. Furthermore, assuming that each realization is drawn from the nearest component, the component indices can also be shown to be highly correlated. That is, when the generation of one WPP is drawn from the kth component, it is likely that the other's is also drawn from the kth component. In principle, the correlation can be weak or even negative, but our main results do not rely on the sign of correlation.

III. Problem Statement

The conceptual framework of this work is illustrated in Fig. 3. Consider an aggregator who aggregates wind power

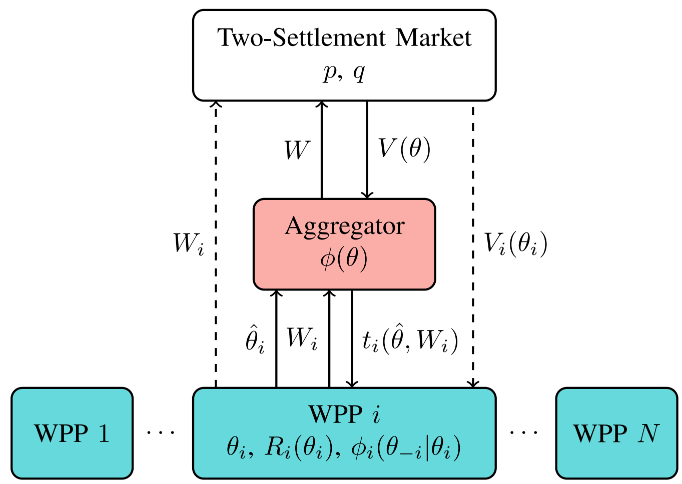

Fig. 3. An aggregator aggregates wind power from N WPPs. Each WPP i's type θ_i, as its private information, gives more refined forecast of its generation W_i, while the aggregator only knows the prior distribution of θ. The type θ_i also gives the reservation utility $R_i(\theta_i)$ and the conditional distribution $\phi_i(\theta_{-i}|\theta_i)$. The goal of the aggregator is to induce each WPP i to report truthfully, i.e., $\hat{\theta}_i = \theta_i$, while minimizing the payment $t_i(\hat{\theta}, W_i)$ to the WPPs. In that case, the expected profit under aggregation is $V(\theta)$, which is greater than or equal to the sum of the expected profits $V_i(\theta_i)$'s under individual participation.

from N WPPs, indexed by $i = 1, \dots, N$. The generation of WPP i is a random variable W_i, following a Gaussian mixture distribution with K components:

$$g_i(w_i) = \sum_{k=1}^{K} \alpha_{i,k} \mathcal{N}(w_i; \mu_{i,k}, \sigma_{i,k}^2), \tag{2}$$

where $\alpha_{i,k}$ is the prior probability of picking the kth component for WPP i, and $\mathcal{N}(\cdot; \mu_{i,k}, \sigma_{i,k}^2)$ is the Gaussian density of the kth component for WPP i. Let $W = (W_1, \dots, W_N)$ and $W_{-i} = (W_1, \dots, W_{i-1}, W_{i+1}, \dots, W_N)$.

While the realization of W_i cannot be known a priori, WPP i learns the value of k, denoted by $\theta_i \in \{1, \dots, K\}$. Then the posterior distribution of W_i given θ_i is $\mathcal{N}(\cdot;\mu_{i,\theta_i},\sigma^2_{i,\theta_i})$. We refer to θ_i as the type of WPP i, and $\Theta_i = \{1,\ldots,K\}$ as the type space (same for all i). Let $\theta = (\theta_1,\ldots,\theta_N)$, $\theta_{-i} = (\theta_1,\ldots,\theta_{i-1},\theta_{i+1},\ldots,\theta_N)$, and $\Theta_{-i} = \{1,\ldots,K\}^{N-1}$. We assume that θ is drawn from a commonly known distribution $\phi(\cdot)$, which is understood as a probability mass function in this paper. Note that $\theta_1,\ldots,\theta_N$ are not necessarily independent. In fact, they are highly correlated in our context. In other words, conditioned on θ_i, each WPP i regards $\phi_i(\theta_{-i}|\theta_i)$ as the distribution of θ_{-i}, whereas the aggregator only knows the prior joint distribution $\phi(\cdot)$.

We consider a two-settlement system of the wholesale electricity market. In the day-ahead market, a market participant offers a power contract (i.e., a constant amount of power) at a price $p > 0$; in the real-time market, when the random generation is realized, the market participant is charged at a price $q > p$ for any shortfall (i.e., negative deviation from the contract). For simplicity, we assume that any surplus (i.e., positive deviation from the contract) is spilled at no cost. Note that the main analysis can be easily generalized when p and q are random variables.

A. Risk Pooling

We first illustrate the advantage of aggregation under uncertainty, which is termed risk pooling in finance.

Define $x_+ = \max\{x, 0\}$. When WPP i participates in the market individually, the expected profit is the optimal value of the following optimization problem:

$$V_i(\theta_i) = \max_{x_i}\ (px_i - \mathbb{E}_{W_i}[q(x_i - W_i)_+|\theta_i]), \tag{3}$$

where x_i is the power contract.

It is easy to derive the optimal power contract.

Proposition 1: The optimal power contract when WPP i participates in the market individually, as the solution to (3), is given by

$$x_i^*(\theta_i) = F^{-1}_{W_i|\theta_i}\left(\frac{p}{q}\right), \tag{4}$$

where $F_X(\cdot)$ the cumulative distribution function of a random variable X.

Proof: The result follows from the first-order condition:

$$\begin{aligned}
&\frac{d}{dx_i}(px_i - \mathbb{E}_{W_i}[q(x_i - W_i)_+|\theta_i])\\
&= p - q\frac{d}{dx_i}\left(\int_0^{x_i}(x_i - w_i)f_{W_i|\theta_i}(w_i)dw_i\right)\\
&= p - q\int_0^{x_i} f_{W_i|\theta_i}(w_i)dw_i\\
&= p - qF_{W_i|\theta_i}(x_i)\\
&= 0,
\end{aligned} \tag{5}$$

where $f_X(\cdot)$ the probability density function of a random variable X. ■

The idea of risk pooling is to aggregate the individual WPPs to form a pool, so that the total variability, or risk, is minimized. When the N WPPs participate in the market as an aggregate, the expected profit is the optimal value of the following optimization problem:

$$V(\theta) = \max_x\ \left(px - \mathbb{E}_W\left[q\left(x - \sum_i W_i\right)_+ \middle| \theta\right]\right). \tag{6}$$

Similarly as before, the optimal power contract under aggregation, as the solution to (6), is given by

$$x^*(\theta) = F^{-1}_{\sum_i W_i|\theta}\left(\frac{p}{q}\right). \tag{7}$$

The advantage of aggregation is formally proved in the following result.

Proposition 2: The expected profit under aggregation is greater than or equal to the sum of the expected profits under individual participation:

$$V(\theta) - \sum_i V_i(\theta_i) \geq 0. \tag{8}$$

Proof: We have

$$\begin{aligned}
V(\theta) &= \max_x\ \left(px - \mathbb{E}_W\left[q\left(x - \sum_i W_i\right)_+ \middle| \theta\right]\right)\\
&\geq p\sum_i x_i^*(\theta_i) - \mathbb{E}_W\left[q\left(\sum_i x_i^*(\theta_i) - \sum_i W_i\right)_+ \middle| \theta\right]\\
&\geq p\sum_i x_i^*(\theta_i) - \mathbb{E}_W\left[q\sum_i (x_i^*(\theta_i) - W_i)_+ \middle| \theta\right]\\
&= \sum_i (px_i^*(\theta_i) - \mathbb{E}_{W_i}[q(x_i^*(\theta_i) - W_i)_+|\theta_i])\\
&= \sum_i V_i(\theta_i),
\end{aligned} \tag{9}$$

where the second inequality follows from the fact that $(x + y)_+ \leq x_+ + y_+$ for any x and y. ■

When all the WPPs except WPP i participate in the market as an aggregate, the expected profit is the optimal value of the following optimization problem:

$$\begin{aligned}
&V_{-i}(\theta_{-i})\\
&= \max_x\ \left(px - \mathbb{E}_{W_{-i}}\left[q\left(x - \sum_{j\neq i} W_j\right)_+ \middle| \theta_{-i}\right]\right).
\end{aligned} \tag{10}$$

Similarly as before, the optimal power contract under aggregation except WPP i, as the the solution to (10), is given by

$$x^*_{-i}(\theta_{-i}) = F^{-1}_{\sum_{j\neq i} W_i|\theta_{-i}}\left(\frac{p}{q}\right). \tag{11}$$

The next result shows that the individual contribution to the aggregate is at least as much as the individual value.

Proposition 3: The marginal contribution of WPP i to the expected profit under aggregation is greater than or equal to the expected profit under individual participation:

$$V(\theta) - V_{-i}(\theta_{-i}) \geq V_i(\theta_i). \tag{12}$$

Proof: We have

$$V_i(\theta_i) + V_{-i}(\theta_{-i})$$

$$
\begin{aligned}
&= px_i^*(\theta_i) - \mathbb{E}_{W_i}[q(x_i^*(\theta_i) - W_i)_+|\theta_i] + px_{-i}^*(\theta_{-i}) \\
&\quad - \mathbb{E}_{W_{-i}}\left[q\left(x_{-i}^*(\theta_{-i}) - \sum_{j\neq i} W_j\right)_+ \middle| \theta_{-i}\right] \\
&\leq p(x_i^*(\theta_i) + x_{-i}^*(\theta_{-i})) \\
&\quad - \mathbb{E}_W\left[q\left(x_i^*(\theta_i) + x_{-i}^*(\theta_{-i}) - \sum_i W_i\right)_+ \middle| \theta\right] \\
&\leq px^*(\theta) - \mathbb{E}_W\left[q\left(x^*(\theta) - \sum_i W_i\right)_+ \middle| \theta\right] \\
&= V(\theta). \qquad (13)
\end{aligned}
$$

■

Moreover, the total invidual contribution to the aggregate is at least as much as the aggregate value.

Proposition 4: The sum of marginal contributions is greater than or equal to the expected profit under aggregation:

$$\sum_i (V(\theta) - V_{-i}(\theta_{-i})) \geq V(\theta). \qquad (14)$$

Proof: We have

$$
\begin{aligned}
&\sum_i V_{-i}(\theta_{-i}) \\
=& \sum_i px_{-i}^*(\theta_{-i}) \\
&- \sum_i \mathbb{E}_{W_{-i}}\left[q\left(x_{-i}^*(\theta_{-i}) - \sum_{j\neq i} W_j\right)_+ \middle| \theta_{-i}\right] \\
\leq& \sum_i px_{-i}^*(\theta_{-i}) \\
&- \mathbb{E}_W\left[q\left(\sum_i x_{-i}^*(\theta_{-i}) - \sum_i\sum_{j\neq i} W_j\right)_+ \middle| \theta\right] \\
=& \, p\sum_i x_{-i}^*(\theta_{-i}) \\
&- \mathbb{E}_W\left[q\left(\sum_i x_{-i}^*(\theta_{-i}) - (N-1)\sum_i W_i\right)_+ \middle| \theta\right] \\
=& \, (N-1)p\left(\frac{\sum_i x_{-i}^*(\theta_{-i})}{N-1}\right) \\
&- (N-1)\mathbb{E}_W\left[q\left(\frac{\sum_i x_{-i}^*(\theta_{-i})}{N-1} - \sum_i W_i\right)_+ \middle| \theta\right] \\
\leq& \, (N-1)\left(px^*(\theta) - \mathbb{E}_W\left[q\left(x^*(\theta) - \sum_i W_i\right)_+ \middle| \theta\right]\right) \\
=& \, (N-1)V(\theta). \qquad (15)
\end{aligned}
$$

■

B. Reservation Utility and Mechanism Design

By aggregating the WPPs, the aggregator can participate in the market and make an expected profit $V(\theta)$. On the other hand, to provide incentives for the WPPs to contract with the aggregator, the aggregator needs to make a minimum payment to each WPP. Such minimum payment is called the reservation utility, denoted by $R_i(\theta_i)$ for WPP i, which depends on its type.

To motivate the problem formulation, we first consider $R_i(\theta_i) = V_i(\theta_i)$ for all i, where the reservation utility reflects the opportunity cost, while the marginal cost is negligible. If the aggregator knows θ, it can make a payment $R_i(\theta_i)$ to each WPP i, and makes a non-negative expected profit as implied by (8):

$$V(\theta) - \sum_i R_i(\theta_i) \geq 0. \qquad (16)$$

Also, by (12), we have

$$V(\theta) - V_{-i}(\theta_{-i}) \geq R_i(\theta_i). \qquad (17)$$

However, the type θ_i of each WPP i is not known by the aggregator. Our goal is to design mechanisms which elicit the true types of the WPPs and make a maximum profit, on behalf of the aggregator.

Now we formalize the problem in the mechanism design framework. In a mechanism, each WPP i reports a type $\hat{\theta}_i$, which may not be the true type θ_i. Similarly as before, let $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_N)$ and $\hat{\theta}_{-i} = (\hat{\theta}_1, \ldots, \hat{\theta}_{i-1}, \hat{\theta}_{i+1}, \ldots, \hat{\theta}_N)$. The key component is the payment $t_i(\hat{\theta}, W_i)$ made by the aggregator to each WPP i, which depends on the reported types of all the WPPs and the realized generation of WPP i.

Definition 1 (Dominant Strategy Incentive Compatibility): A mechanism is dominant strategy incentive compatible if reporting truthfully is a dominant strategy for each WPP i:

$$
\begin{aligned}
&\mathbb{E}_{W_i}[t_i(\theta_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}] \\
&\geq \mathbb{E}_{W_i}[t_i(\hat{\theta}_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}], \ \forall\theta_i, \ \forall\hat{\theta}_i, \ \forall\hat{\theta}_{-i}. \qquad (18)
\end{aligned}
$$

Definition 2 (Interim Individual Rationality): A mechanism is interim individual rational if each WPP i has no incentive to withdraw from the mechanism after it learns its own type but before the others' types are revealed:

$$\mathbb{E}_{\theta_{-i}, W_i}[t_i(\theta_i, \theta_{-i}, W_i)|\theta_i] \geq R_i(\theta_i), \ \forall\theta_i. \qquad (19)$$

Definition 3 (Optimal Mechanism): A mechanism is optimal if it maximizes the expected profit of the aggregator, or equivalently, minimizes the expected total payment to the WPPs, subject to the constraints of dominant strategy incentive compatibility and interim individual rationality:

$$
\begin{aligned}
&\underset{t_i(\cdot), \forall i}{\text{minimize}} \qquad \mathbb{E}_{\theta, W}\left[\sum_i t_i(\theta, W_i)\right] \qquad (20) \\
&\text{subject to} \qquad (18) \text{ and } (19).
\end{aligned}
$$

C. Stochastic VCG Mechanism

The stochastic VCG mechanism [14] characterizes a feasible solution to (20). In that mechanism, the payment function is given by

$$
\begin{aligned}
&\tilde{t}_i(\hat{\theta}, W_i) \\
&= px^*(\hat{\theta}) - V_{-i}(\hat{\theta}_{-i}) \\
&\quad - \mathbb{E}_{W_{-i}}\left[q\left(x^*(\hat{\theta}) - W_i - \sum_{j\neq i} W_j\right)_+ \middle| \hat{\theta}_{-i}\right]. \qquad (21)
\end{aligned}
$$

Proposition 5: The stochastic VCG mechanism is dominant strategy incentive compatible.

Proof: We have

$$\begin{aligned}
&\mathbb{E}_{W_i}[\tilde{t}_i(\hat{\theta}_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}] \\
&= px^*(\hat{\theta}_i, \hat{\theta}_{-i}) - V_{-i}(\hat{\theta}_{-i}) \\
&\quad - \mathbb{E}_W \left[q\left(x^*(\hat{\theta}_i, \hat{\theta}_{-i}) - W_i - \sum_{j \neq i} W_j \right)_+ \middle| \theta_i, \hat{\theta}_{-i} \right] \\
&\leq px^*(\theta_i, \hat{\theta}_{-i}) - V_{-i}(\hat{\theta}_{-i}) \\
&\quad - \mathbb{E} \left[q\left(x^*(\theta_i, \hat{\theta}_{-i}) - W_i - \sum_{j \neq i} W_j \right)_+ \middle| \theta_i, \hat{\theta}_{-i} \right] \\
&= \mathbb{E}_{W_i}[\tilde{t}_i(\theta_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}]. \qquad (22)
\end{aligned}$$

■

Such dominant strategy implementation is very robust, in the sense that each WPP's best response only depends on its own type, regardless of the others' true or reported types. In particular, the best response does not depend on $\phi(\cdot)$, the prior distribution of the true types. Given the constraint of dominant strategy incentive compatibility, we can write $t_i(\hat{\theta}, W_i) = t_i(\theta, W_i)$.

Proposition 6: The stochastic VCG mechanism is interim individual rational.

Proof: We have

$$\begin{aligned}
&\mathbb{E}_{W_i}[\tilde{t}_i(\theta, W_i)|\theta] \\
&= px^*(\theta) - V_{-i}(\theta_{-i}) - \mathbb{E}_W \left[q\left(x^*(\theta) - \sum_i W_i \right)_+ \middle| \theta \right] \\
&= V(\theta) - V_{-i}(\theta_{-i}) \\
&\geq R_i(\theta_i), \qquad (23)
\end{aligned}$$

and therefore

$$\begin{aligned}
&\mathbb{E}_{\theta_{-i}, W_i}[\tilde{t}_i(\theta_i, \theta_{-i}, W_i)|\theta_i] \\
&= \mathbb{E}_{\theta_{-i}}[\mathbb{E}_{W_i}[\tilde{t}_i(\theta_i, \theta_{-i}, W_i)|\theta_i, \theta_{-i}]|\theta_i] \\
&\geq \mathbb{E}_{\theta_{-i}}[R_i(\theta_i)|\theta_i] \\
&= R_i(\theta_i). \qquad (24)
\end{aligned}$$

■

While the VCG mechanism is a feasible solution to (20), it is not optimal. In fact, the total payment made to the WPPs is too high to be recovered by the profit under aggregation:

$$\sum_i \mathbb{E}_{W_i}[\tilde{t}_i(\theta, W_i)|\theta] = \sum_i (V(\theta) - V_{-i}(\theta_{-i})) \geq V(\theta), \quad (25)$$

following from (14).

Such budget deficit is undesirable for the profit-maximizing aggregator. Thus, the aggregator needs to seek for alternative mechanisms which make smaller payments to the WPPs, yet still elicit truthful information. This can be achieved by exploiting the correlation structure of the types, which is not used in the stochastic VCG mechanism.

IV. Optimal Mechanism Design

In the previous section, we considered a particular reservation utility, $R_i(\theta_i) = V_i(\theta_i)$ for all i. From now on, we consider general reservation utilities, and take (16) and (17) as practical constraints.

Assumption 1: The reservation utility $R_i(\theta_i)$ for each i satisfies (16) and (17).

A. Full Surplus Extraction

As shown in (16), the maximum expected profit of the aggregator is achieved when the aggregator has full information about θ, and makes a payment equal to the reservation utility to each WPP i. In that case, we have

$$\sum_i t_i(\theta, W_i) = \sum_i R_i(\theta_i). \quad (26)$$

In general, (26) is too strong for a practical mechanism to hold, since the aggregator does not have full information. Instead, we consider a relaxed property which states that (26) holds in expectation, with respect to both W and θ.

Definition 4 (Full Surplus Extraction): A mechanism is said to extract the full surplus, if

$$\mathbb{E}_{\theta, W}\left[\sum_i t_i(\theta, W_i)\right] = \mathbb{E}_\theta\left[\sum_i R_i(\theta_i)\right]. \quad (27)$$

By Definition 3, if a mechanism satisfying (18) and (19) also satisfies (27), then it is an optimal mechanism.

B. Full Correlation

Since the WPPs are geographically close together, their types can be highly correlated. We introduce a definition that characterizes the correlation structure of the types.

Recall that for each i, we have $\Theta_i = \{1, \ldots, K\}$, $|\Theta_i| = K$, $\Theta_{-i} = \{1, \ldots, K\}^{N-1}$, and $|\Theta_{-i}| = K^{N-1}$. Let $(\theta_i^1, \ldots, \theta_i^K)$ be a permutation of Θ_i, and $(\theta_{-i}^1, \ldots, \theta_{-i}^L)$ be a permutation of Θ_{-i}, where we define $L = K^{N-1}$.

Definition 5 (Conditional Distribution Matrix): The conditional distribution matrix A_i for each i is defined as a K by L matrix:

$$A_i = \begin{pmatrix} \phi_i(\theta_{-i}^1|\theta_i^1) & \cdots & \phi_i(\theta_{-i}^L|\theta_i^1) \\ \vdots & \ddots & \vdots \\ \phi_i(\theta_{-i}^1|\theta_i^K) & \cdots & \phi_i(\theta_{-i}^L|\theta_i^K) \end{pmatrix}, \quad (28)$$

where the rows are indexed by the elements in Θ_i, the columns are indexed by the elements in Θ_{-i}, and $\phi_i(\theta_{-i}^l|\theta_i^k)$ is the probability of $\theta_{-i} = \theta_{-i}^l$ conditioned on $\theta_i = \theta_i^k$.

The concept of full correlation is based on the conditional distribution matrix.

Definition 6 (Full Correlation): When A_i has rank K for all i, we say that the types are fully correlated, or that the joint distribution $\phi(\cdot)$ satisfies the full correlation condition.

One special case of full correlation is perfect correlation, i.e., for any i, the value of θ_i corresponds to a unique value of θ_{-i}. In that case, each A_i has exactly one entry 1 in each row and 0s elsewhere. Another special case is established in the following.

Proposition 7: When $N = 2$, define the joint distribution matrix A as

$$A = \begin{pmatrix} \phi(\theta_1^1, \theta_2^1) & \cdots & \phi(\theta_1^1, \theta_2^K) \\ \vdots & \ddots & \vdots \\ \phi(\theta_1^K, \theta_2^1) & \cdots & \phi(\theta_1^K, \theta_2^K) \end{pmatrix}. \tag{29}$$

If A is invertible, then $\phi(\cdot)$ satisfies the full correlation condition.

Proof: Since A is invertible, $\det A \neq 0$. Note that A_1 is obtained by dividing each row k of A by the marginal distribution $\phi(\theta_1^k, \cdot)$. Thus, we have

$$\det A_1 = \frac{\det A}{\prod_{k=1}^{K} \phi(\theta_1^k, \cdot)} \neq 0. \tag{30}$$

So A_1 is invertible. Similarly, A_2 is also invertible. The result follows immediately. ■

We note that the full correlation condition is fairly mild, which only requires that knowing the type of one WPP gives more refined information about the types of other WPPs.

C. Optimal Mechanism under Full Correlation

Under the full correlation condition, we can show that the optimal mechanism extracts the full surplus. This is proved by construction based on the stochastic VCG mechanism.

Define

$$h_i(\theta_i) = \sum_{\theta_{-i}} \phi_i(\theta_{-i}|\theta_i)(V(\theta) - V_{-i}(\theta_{-i})) - R_i(\theta_i). \tag{31}$$

For each i, since A_i has rank K, there exists a $t_i' = (t_i'(\theta_{-i}^1), \dots, t_i'(\theta_{-i}^L))^T$ such that

$$A_i \begin{pmatrix} t_i'(\theta_{-i}^1) \\ \vdots \\ t_i'(\theta_{-i}^L) \end{pmatrix} = \begin{pmatrix} h_i(\theta_i^1) \\ \vdots \\ h_i(\theta_i^K) \end{pmatrix}. \tag{32}$$

Subtracting $t_i'(\hat{\theta}_{-i})$ from the payment function of the stochastic VCG mechanism, we obtain the payment function of the proposed mechanism:

$$t_i^*(\hat{\theta}, W_i) = \tilde{t}_i(\hat{\theta}, W_i) - t_i'(\hat{\theta}_{-i}). \tag{33}$$

Proposition 8: When the joint distribution $\phi(\cdot)$ satisfies the full correlation condition, the proposed mechanism specified by (33) is an optimal mechanism.

Proof: First, the mechanism is dominant strategy incentive compatible, since $t_i'(\hat{\theta}_{-i})$ does not depend on $\hat{\theta}_i$:

$$\begin{aligned} &\mathbb{E}_{W_i}[t_i^*(\hat{\theta}_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}] \\ &= \mathbb{E}_{W_i}[\tilde{t}_i(\hat{\theta}_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}] - t_i'(\hat{\theta}_{-i}) \\ &\leq \mathbb{E}_{W_i}[\tilde{t}_i(\theta_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}] - t_i'(\hat{\theta}_{-i}) \\ &= \mathbb{E}_{W_i}[t_i^*(\theta_i, \hat{\theta}_{-i}, W_i)|\theta_i, \hat{\theta}_{-i}]. \end{aligned} \tag{34}$$

Second, the mechanism is interim individual rational:

$$\begin{aligned} &\mathbb{E}_{\theta_{-i}, W_i}[t_i^*(\theta_i, \theta_{-i}, W_i)|\theta_i] \\ &= \mathbb{E}_{\theta_{-i}}[\mathbb{E}_{W_i}[t_i^*(\theta_i, \theta_{-i}, W_i)|\theta_i, \theta_{-i}]|\theta_i] \\ &= \mathbb{E}_{\theta_{-i}}[\mathbb{E}_{W_i}[\tilde{t}_i(\theta_i, \theta_{-i}, W_i)|\theta_i, \theta_{-i}] - t_i'(\theta_{-i})|\theta_i] \\ &= \mathbb{E}_{\theta_{-i}}[V(\theta) - V_{-i}(\theta_{-i}) - t_i'(\theta_{-i})|\theta_i] \\ &= \sum_{\theta_{-i}} \phi_i(\theta_{-i}|\theta_i)(V(\theta) - V_{-i}(\theta_{-i}) - t_i'(\theta_{-i})) \\ &= \sum_{\theta_{-i}} \phi_i(\theta_{-i}|\theta_i)(V(\theta) - V_{-i}(\theta_{-i})) - h_i(\theta_i) \\ &= R_i(\theta_i). \end{aligned} \tag{35}$$

It follows that the mechanism extracts the full surplus:

$$\begin{aligned} \mathbb{E}_{\theta, W}\left[\sum_i t_i^*(\theta, W_i)\right] &= \sum_i \mathbb{E}_{\theta, W}[t_i^*(\theta, W_i)] \\ &= \sum_i \mathbb{E}_{\theta_i}[\mathbb{E}_{\theta_{-i}, W_i}[t_i^*(\theta, W_i)|\theta_i]] \\ &= \sum_i \mathbb{E}_{\theta_i}[R_i(\theta_i)] \\ &= \mathbb{E}_{\theta}\left[\sum_i R_i(\theta_i)\right]. \end{aligned} \tag{36}$$

■

TABLE I
PARTITION OF WPP NOMINAL CAPACITY

WPP	μ_1	μ_2	μ_3	μ_4	μ_5
1	0.03	0.16	0.31	0.53	0.79
2	0.03	0.17	0.35	0.59	0.85
3	0.04	0.18	0.37	0.62	0.90
4	0.03	0.20	0.39	0.61	0.86
5	0.03	0.16	0.34	0.58	0.88
6	0.05	0.24	0.47	0.67	0.87
7	0.04	0.27	0.50	0.72	0.90

V. CASE STUDY

In this section, we present a case study based on the GEFCom2012 data to validate the main results. Recall that the seven WPPs in the data are in close proximity to each other, and the wind power generation has been normalized. For simplicity, we focus on the data for hour 1.

As introduced in Section II, for each WPP, we use K-means clustering to partition the nominal capacity into K clusters. We choose $K = 5$ in the study. The centers of the clusters are listed in Table I.

As an approximation, we assume that each realization is drawn from the nearest component in the mixture model. For example, when the realized generation for WPP 1 is 0.15, it is considered as drawn from the component centering at 0.16, and therefore $\theta_1 = 2$, which is the private information of WPP 1 at the day-ahead stage.

Let $p = \$50$/MWh and $q = \$100$/MWh in the study. We calculate the profit under individual participation when various information is available: (i) (ex ante) only the mixture model is known; (ii) (type) the component index based on the realization is known; and (iii) (ex post) the realization is known, which is clairvoyant. The results are shown in Fig. 4. The profit under ex post information is the highest, and that

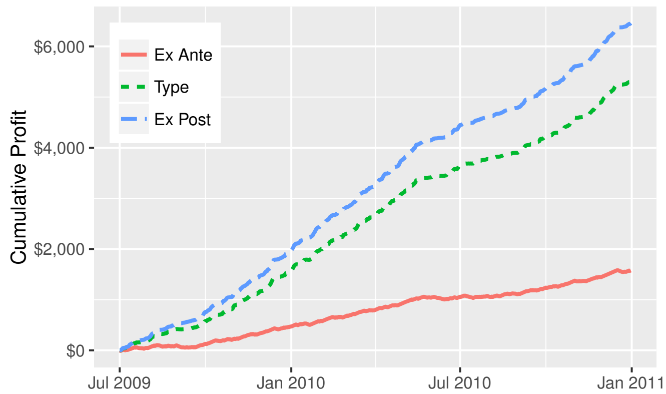

Fig. 4. The cumulative profit for WPP 1 when various information is available.

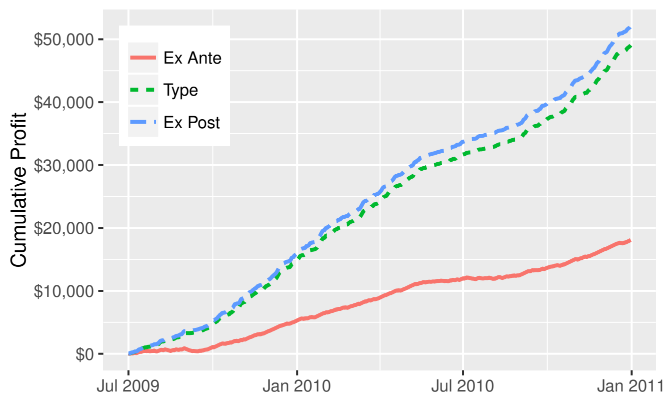

Fig. 5. The cumulative profit under aggregation when various information is available.

under ex ante information is the lowest. When the type is known, the profit is close to the ex post scenario.

Similarly, the profit under the aggregation of the seven WPPs is shown in Fig. 5. It exhibits a similar pattern as the profit under individual participation.

Consider the surplus as defined in (8), which is shown in Fig. 6. There is no surplus under ex post information, since there is no uncertainty. Different from the previous plots, the surplus based on ex ante information is higher than that based on types. That is, aggregation adds more value when less information is available.

Next, we examine the empirical conditional distribution matrices. The matrix A_i for each i is a 5 by 5^6 matrix. It can be verified that all the A_i's have rank 5, i.e., full rank. In fact, the full correlation condition is fairly mild and holds in most circumstances.

To demonstrate the issue of budget deficit for the stochastic VCG mechanism, we plot the profit under aggregation and the total payment made to the WPPs in Fig. 7. It can easily seen that the total payment exceeds the profit under aggregation, which is undesirable for the aggregator.

Now we study the optimal mechanism. We assume that the reservation utility of each WPP is the profit under individual participation. Since A_i is not square, we consider the Moore-Penrose pseudoinverse of A_i when computing t_i'. The results are presented in Fig. 8. While the payment in the stochastic

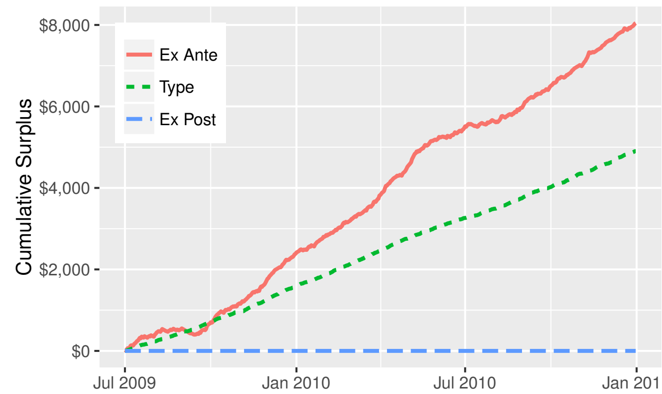

Fig. 6. The cumulative surplus when various information is available.

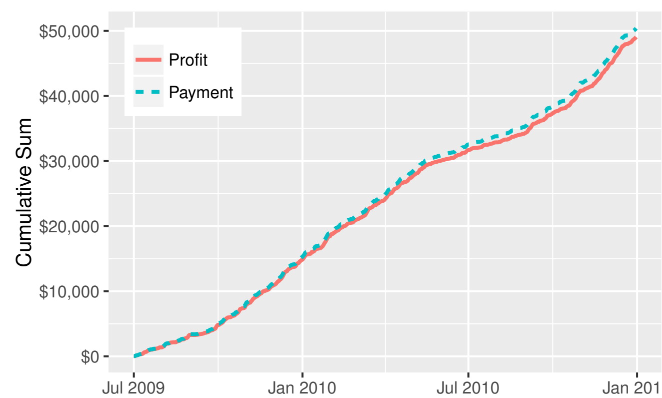

Fig. 7. The cumulative profit under aggregation and the total payment in the stochastic VCG mechanism when various information is available.

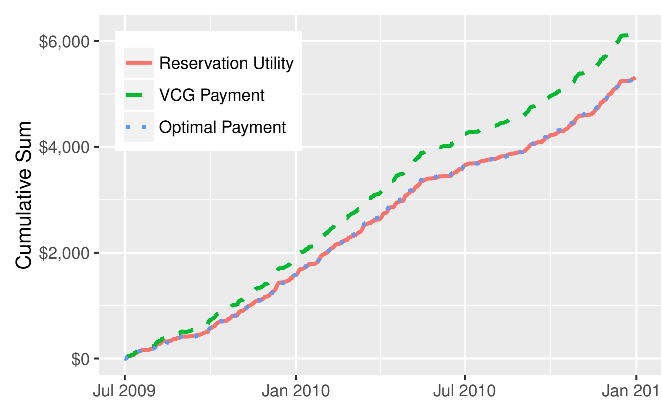

Fig. 8. The cumulative reservation utility for WPP 1, and the payments made to WPP 1 in the stochastic VCG mechanism and in the optimal mechanism.

VCG mechanism exceeds the reservation utility, that in the optimal mechanism is approximately equal to the reservation utility.

VI. Conclusion

Aggregating diverse wind power sources can reduce the overall variability. The surplus provides an incentive for a third-party aggregator to enter the market, by assuming the risk due to the uncertainty of wind power generation. To extract the surplus, the aggregator needs to elicit the private information of the WPPs. While the objective of truthful elicitation can be achieved by the stochastic VCG mechanism, the total payment made to the WPPs cannot be recovered by the profit under aggregation. On the other hand, the aggregator can exploit the correlation structure of the types. Under the full correlation condition, the proposed mechanism extracts the full surplus, subject to the constraints of dominant strategy incentive compatibility and interim individual rationality. It can be shown that there does not exist a mechanism with stronger properties. The results reveal an interesting trade-off in wind power aggregation: aggregation achieves the maximum profit when the types are independent, while the surplus can be fully extracted when the types are fully correlated.

REFERENCES

[1] W. Tang and R. Jain, "Buying random yet correlated wind power," in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, 2014, pp. 145–150.
[2] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Short-term trading for a wind power producer," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 554–564, 2010.
[3] E. Y. Bitar, R. Rajagopal, P. P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *Power Systems, IEEE Transactions on*, vol. 27, no. 3, pp. 1225–1235, 2012.
[4] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 206–215, 2012.
[5] C. Zhao, J. Wang, J.-P. Watson, and Y. Guan, "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2708–2717, 2013.
[6] S. Tewari, C. J. Geyer, and N. Mohan, "A statistical model for wind power forecast error and its application to the estimation of penalties in liberalized markets," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2031–2039, 2011.
[7] C. Wan, Z. Xu, and P. Pinson, "Direct interval forecasting of wind power," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4877–4878, 2013.
[8] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1033–1044, 2014.
[9] ——, "Optimal prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1166–1174, 2014.
[10] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Transactions on Power Systems*, 2016.
[11] E. Baeyens, E. Y. Bitar, P. P. Khargonekar, and K. Poolla, "Coalitional aggregation of wind power," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3774–3784, 2013.
[12] Y. Zhao, J. Qin, R. Rajagopal, A. Goldsmith, and H. V. Poor, "Wind aggregation via risky power markets," *Power Systems, IEEE Transactions on*, vol. 30, no. 3, pp. 1571–1581, 2015.
[13] A. Mas-Colell, M. Whinston, and J. Green, *Microeconomic theory*. Oxford, 1995.
[14] W. Tang and R. Jain, "Market mechanisms for buying random wind," *Sustainable Energy, IEEE Transactions on*, vol. 6, no. 4, pp. 1615–1623, 2015.
[15] J. Crémer and R. P. McLean, "Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent," *Econometrica*, vol. 53, no. 2, pp. 345–361, 1985.
[16] ——, "Full extraction of the surplus in bayesian and dominant strategy auctions," *Econometrica*, vol. 56, no. 6, pp. 1247–1257, 1988.
[17] M. H. Riordan and D. E. Sappington, "Optimal contracts with public ex post information," *Journal of Economic Theory*, vol. 45, no. 1, pp. 189–199, 1988.
[18] R. P. McAfee and P. J. Reny, "Correlated information and mechanism design," *Econometrica*, vol. 60, no. 2, pp. 395–421, 1992.
[19] D. Fudenberg and J. Tirole, *Game theory*. MIT Press, 1991.
[20] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *International Journal of Forecasting*, vol. 30, no. 2, pp. 357–363, 2014.