Distributed Stochastic Market Clearing with High-Penetration Wind Power

Yu Zhang, Student Member, IEEE, and Georgios B. Giannakis, Fellow, IEEE

Abstract—Integrating renewable energy into the modern power grid requires risk-cognizant dispatch of resources to account for the stochastic availability of renewables. Toward this goal, day-ahead stochastic market clearing with high-penetration wind energy is pursued in this paper based on the DC optimal power flow (OPF). The objective is to minimize the social cost which consists of conventional generation costs, end-user disutility, as well as a risk measure of the system re-dispatching cost. Capitalizing on the conditional value-at-risk (CVaR), the novel model is able to mitigate the potentially high risk of the recourse actions to compensate wind forecast errors. The resulting convex optimization task is tackled via a distribution-free sample average based approximation to bypass the prohibitively complex high-dimensional integration. Furthermore, to cope with possibly large-scale dispatchable loads, a fast distributed solver is developed based on ADMM. Numerical results tested on a modified benchmark system are reported to corroborate the merits of the novel framework and proposed approaches.

Index Terms—ADMM, conditional value-at-risk, demand response aggregator, market clearing, stochastic optimization, wind power.

NOMENCLATURE

A. Indices, numbers, and sets

- $T, \mathcal{T}$ Number and set of scheduling periods.
- $N_b, N_l$ Number of buses and lines.
- $N_g, N_j$ Number and set of conventional generators.
- $N_a, N_i$ Number and set of aggregators.
- $N_w, N_m$ Number and set of wind farms.
- $N_s, N_r$ Number and set of wind power generation samples.
- $\mathcal{R}_j$ Set of end users served by aggregator $j$.
- $\mathcal{S}_{jr}$ Set of smart appliances of residential user $r$ served by aggregator $j$.
- $\mathcal{P}_{jrs}$ Set of operational constraints of appliance $s$ of residential user $r$ served by aggregator $j$.
- $\mathcal{T}_{jr}^E$ Set of scheduling periods of appliance $s$.
- $k$ ADMM iteration index.

B. Constants

- $P_{\text{min}}^{G_i}$, $P_{\text{max}}^{G_i}$ Minimum and maximum power output of conventional generator $i$.
- $P_{\text{up}}^g$, $P_{\text{down}}^g$ Ramp-up and ramp-down limits of conventional generator $i$.
- $P_{\text{BL}}^t$ Fixed base load power demand in slot $t$.
- $P_{\text{max}}^{\text{DRA}_j}$ Maximum power provided by demand response aggregator $j$.
- $P_{\text{min}}^{\text{DRA}_j}$, $P_{\text{max}}^{\text{DRA}_j}$ Minimum and maximum power consumption of appliance $s$.
- $T_{\text{start}}$, $T_{\text{end}}$, $T_{jr}$ Total energy consumption of appliance $s$.
- $E_{jrs}$ Start and end times of appliance $s$.
- $p_{\text{max}}^t$, $p_{\text{max}}^t$ Minimum and maximum power flow limits.
- $p_{\text{max}}^t$ Maximum committed wind power.
- $A_n$ Branch-node incidence matrix.
- $A_g, A_w, A_a$ Nodal susceptance matrix.
- $B_n$ Matrix relating bus angles to branch power flows.
- $B_f$ Branch susceptance matrix.
- $b_f$ Susceptance of line $f$.
- $s_t$ Vector collecting selling prices in slot $t$.
- $p_t$ Vector of purchase prices in slot $t$.
- $\delta$ Tolerance of the ADMM termination criterion using primal feasibility.
- $\rho$ Weight of augmented Lagrangian.
- $\mu$ Weight of CVaR-based transaction cost.
- $\beta$ CVaR probability level.

C. Decision variables

- $P_{Gi}^t$ Output of conventional generator $i$ in slot $t$.
- $p_{jrs}^t$ Consumption of appliance $s$ in slot $t$.
- $P_{\text{DRA}_j}^t$ Total power consumption of aggregator $j$ in slot $t$.
- $P_{\text{W}_m}^t$ Power committed by wind farm $m$ in slot $t$.
- $\eta$ A variable in the CVaR-based transaction cost.
- $\theta^t$ Vector of nodal voltage phases in slot $t$.
- $P_G^t$ Vector collecting $P_{Gi}^t$ for all $i \in N_g$.
- $P_{\text{DRA}}^t$ Vector collecting $P_{\text{DRA}_j}^t$ for all $j \in N_a$.
- $P_{\text{W}}^t$ Vector collecting $P_{\text{W}_m}^t$ for all $m \in N_w$.
- $p_{jrs}^t$ Vector collecting $p_{jrs}^t$ for all $t \in \mathcal{T}$.
- $p_{0}^t$ Vector collecting $\eta$ and $P_{Gi}^t$, $P_{\text{DRA}_j}^t$, $P_{\text{W}_m}^t$, $\theta^t$ for all $t \in \mathcal{T}$.
- $p_j^t$ Vector collecting $p_{jrs}^t$ for all $r$ and $s$. 

Manuscript received March 12, 2014; revised August 21, 2014, December 20, 2014, and March 29, 2015; accepted April 01, 2015. This work was supported by the Initiative for Renewable Energy & the Environment (IREE) grant RL-0010-13, University of Minnesota, and NSF grants CCF-1423316 and CCF-1442686. Part of this work was presented at the 2014 IEEE 53rd Annual Conference on Decision and Control, Los Angeles, CA, December 15-17, 2014.

The authors are with the Department of Electrical and Computer Engineering and the Digital Technology Center, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA. E-mails: {zhan1220, georgios}@umn.edu.
D. Uncertain quantities

\[
\begin{align*}
    w^t_m & \quad \text{Actual power output of wind farm } m \text{ in slot } t. \\
    \mathbf{w}^t & \quad \text{Vector collecting } w^t_m \text{ for all } m \in \mathcal{N}_w.
\end{align*}
\]

E. Functions

\[
\begin{align*}
    C_i(\cdot) & \quad \text{Cost function of generator } i. \\
    U_{jrs}(\cdot) & \quad \text{Utility function of appliance } s. \\
    F_\beta(\cdot) & \quad \text{CVaR transaction cost.} \\
    \bar{F}_\beta(\cdot) & \quad \text{Sample mean of } F_\beta(\cdot). \\
    L_p(\cdot) & \quad \text{Partial Lagrangian function of the stochastic market clearing problem.} \\
    \Pi_{G_n}, \Pi_{DRA_n}, \Pi_{W_n} & \quad \text{Revenues or payments of the supplier, the aggregator, and the wind farm located at bus } n.
\end{align*}
\]

F. Abbreviations

- **ADMM**: Alternating direction method of multipliers.
- **CVaR**: Conditional value-at-risk.
- **DA**: Day-ahead.
- **DSM**: Demand side management.
- **DR**: Demand response.
- **ED**: Economic dispatch.
- **ISO**: Independent system operator.
- **LMPs**: Locational marginal prices.
- **LOLP**: Loss-of-load probability.
- **MC**: Market clearing.
- **OPF**: Optimal power flow.
- **RES**: Renewable energy sources.
- **RT**: Real-time.
- **SCED**: Security-constrained economic dispatch.
- **SCUC**: Security-constrained unit commitment.
- **SAA**: Sample average approximation.
- **UC**: Unit commitment.
- **VaR**: Value-at-risk.
- **WPPs**: Wind power producers.

I. INTRODUCTION

The future smart grid is an automated electric power grid that capitalizes on modern optimization, monitoring, communication, and control technologies to improve efficiency, sustainability, and reliability of generation, transmission, distribution, and consumption of electric energy. Limited supply and environmental impact of conventional power generation compel industry to aggressively utilize the clean renewable energy sources (RES), such as wind, sunlight, biomass, and geothermal heat, because of their eco-friendly and price-competitive advantages. Growing at an annual rate of 20%, wind power generation already boasted a worldwide installed capacity of 318 GW by the end of 2013, and is widely embraced throughout the world [1]. Recently, both the U.S. Department of Energy (DoE) and the European Union (EU) proposed ambitious blueprints towards a low-carbon economy by meeting 20% of the electricity consumption with renewables by 2030 and 2020, respectively [2], [3].

Towards the goal of boosting the penetration of RES, robust and stochastic planning, operation, and energy management with renewables have been extensively investigated recently. A key challenge of the associated power dispatch tasks is to account for the intrinsically random and non-dispatchable nature of RES so that total power demand can be satisfied by total power supply, while the social cost is minimized. Being resilient to communication outages and malicious cyber-attacks, efficient decentralized algorithms deployed over the interdependent power entities are indispensable as well.

Limiting the loss-of-load probability (LOLP), risk-aware energy management approaches including economic dispatch (ED), unit commitment (UC), and optimal power flow (OPF) were formulated as chance-constrained optimization problems in [4]–[8]. Leveraging scenario sampling, a general non-convex chance-constrained program can be relaxed and solved efficiently as a convex one, which however turns out to be too conservative in certain scenarios [5]. As an alternative, risk-limiting dispatch has been formulated as a multi-stage stochastic control problem [9]; see also [10], where direct coupling of the uncertain energy supply with deferrable demand was accounted for using stochastic dynamic programming.

Additional early works relied on the so-termed committed renewable energy. ED penalizing (under-) over-estimation of wind power was investigated in [11]. Worst-case robust distributed ED with demand side management (DSM) was proposed for grid-connected microgrids [12]. However, the worst-case scenario is unlikely to come up in real-time (RT) operations. Multi-period ED with spatio-temporal wind forecasts was pursued in [13]. The obtained optimal operating point though can be very sensitive to the forecast accuracy.

Turning attention to power system economics, market clearing (MC) is one of the most important routines for a power market, which relies on security-constrained UC or OPF. Independent system operators (ISO) collect generation bids and consumption offers from the day-ahead (DA) electricity market. The MC process is then implemented to determine the market-clearing prices [14]. Deterministic MC without RES has been extensively studied; see e.g., [15]–[17]. Optimal wind power trading or contract offerings have been investigated from the perspective of wind power producers (WPPs) [18]–[21]. MC under uncertain power generation was recently pursued as well. As uncertainty of wind power is revealed on a continuous basis, ISOs are prompted to undertake corrective measures from the very beginning of the scheduling horizon [22]. One approach for an ISO to control the emerging risk is through the deployment of reserves following the contingencies [23]. Electricity pricing and power generation scheduling with uncertainties were accomplished via stochastic programming [24], [25]. In addition, one can co-optimize the competing objectives of generation cost and security indices [26]; see also [27] for a stochastic security-constrained approach. Albeit computationally complex, stochastic bilevel programs are attractive because they can account for the coupling between DA and RT (spot) markets [28], [29].

All existing MC approaches, however, are centralized. Moreover, they are not tailored to address the challenges of emerging large-scale dispatchable loads. Specifically, demand offers come from demand response (DR) aggregators serving large numbers of residential appliances that feature diverse
utility functions and inter-temporal constraints. In this context, the present paper deals with the DC-OPF based MC with high-penetration wind power. Instead of the worst-case or chance-constrained formulations, a novel stochastic optimization approach is proposed to maintain the nodal power balance while minimize (maximize) the grid-wide social cost (welfare). The social cost accounts for the conventional generation costs, the dis-utility of dispatchable loads, as well as a risk measure of the cost incurred by (over-) under-estimating the actual wind generation. This is essentially a cost of re-dispatching the system to compensate wind forecast errors, and is referred as transaction cost throughout this paper. The transaction cost in the spot market is modulated through an efficient solver is developed using the ADMM (Sec. IV). Numerical tests are performed to corroborate the effectiveness of the novel model and proposed approaches using real power market data (Sec. V).

The main contribution of this paper is three-fold: i) a CVaR-based transaction cost is introduced for the day-ahead MC to judiciously control the risk of (over-) under-estimating the wind power generation; ii) a sufficient condition pertinent to transaction prices is established to effect convexity of the CVaR-based cost; and iii) a distributed solver of the resulting stochastic MC task is developed to be run by the market operator and DR aggregators while respecting the privacy of end users.

Notation. Boldface lower (upper) case letters represent column vectors (matrices); calligraphic letters stand for sets. \(\mathbb{R}^{d_1 \times d_2}, \mathbb{R}^d, \mathbb{R}_+\) stand for real spaces of \(d_1 \times d_2\), \(d\), and \(d_-\) stands, respectively; Symbols \(a'\) and \(a \cdot b\) denote the transpose of \(a\), and the inner product of \(a\) and \(b\); \(\lfloor Z \rfloor\) is the lower endpoint of the interval set \(Z\). Operator \([a]^+ := \max\{a, 0\}\) is the projection to the nonnegative reals, while \(\preceq (\succeq)\) indicates the entry-wise inequality. Finally, the expectation is denoted by \(\mathbb{E}[\cdot]\).

II. CVaR revisited: A Convex Risk Measure

Value-at-risk (VaR) and conditional value-at-risk (CVaR) are widely used in various real-world applications, especially in the finance area, as the popular tools to evaluate the credit risk of a portfolio, and reduce the probability of large losses [30–[32]. The following revisit is useful to grasp their role in the present context.

Consider a loss function \(L(x, \xi) : X \times \Xi \mapsto \mathbb{R}\) denoting the real-valued cost associated with the decision variable \(x \in X \subset \mathbb{R}^n\); and the random vector \(\xi\) with probability density function \(p(\xi)\) supported on a set \(\Xi \subset \mathbb{R}^d\). In the context of power grids, \(x\) can represent the power schedules of generators, while \(\xi\) collects the sources of uncertainty due to for instance renewable energy and forecasted load demand.

Clearly, the probability of \(L(x, \xi)\) not exceeding a threshold \(\eta\) is given by the right-continuous cumulative distribution function (CDF)

\[
\Psi(x, \eta) = \int_{L(x, \xi) \leq \eta} p(\xi) \, d\xi. \tag{1}
\]

**Definition 1 (VaR).** Given a prescribed confidence level \(\beta \in (0, 1)\), the \(\beta\)-VaR is the generalized inverse of \(\Psi\) defined as

\[
\eta_{\beta}(x) := \min\{\eta \in \mathbb{R} | \Psi(x, \eta) \geq \beta\}. \tag{2}
\]

\(\beta\)-VaR is essentially the \(\beta\)-quantile of the random \(L(x, \xi)\). Since \(\Psi\) is non-decreasing in \(\eta\), \(\eta_{\beta}(x)\) comes out as the lower endpoint of the solution interval satisfying \(\Psi(x, \eta) = \beta\), and the commonly chosen values of \(\beta\) are, e.g., 0.99, 0.95, and 0.9. Clearly, VaR determines a maximum tolerable loss of an investment, i.e., a threshold the loss will not exceed with a high probability \(\beta\). Hence, given the confidence level \(\beta\), investors are motivated to solve the so-called portfolio optimization problem which yields the optimal investment decisions minimizing the VaR value, \(\eta_{\beta}(x)\) is proportional to the standard deviation if \(\Psi\) is Gaussian. However, for general distributions, \(\beta\)-VaR is non-subadditive which means the VaR of a combined portfolio can be larger than the sum of the VaRs of each component. This violates the common principle “diversification reduces risk”. Moreover, it is generally non-convex rendering the optimization task hard to tackle.

Because of these conceptual and practical drawbacks, CVaR (a.k.a. “tail VaR”, “mean shortfall”, or “mean excess loss”) was proposed as an alternative risk metric that has many superior properties over VaR.

**Definition 2 (CVaR).** The \(\beta\)-CVaR is the mean of the \(\beta\)-tail distribution of \(L(x, \xi)\), which is given as

\[
\Psi_\beta(x, \eta) := \begin{cases} 
0, & \text{if } \eta < \eta_{\beta}(x) \\
\frac{\Psi(x, \eta) - \beta}{1 - \beta}, & \text{if } \eta \geq \eta_{\beta}(x). 
\end{cases} \tag{3}
\]
Truncated and re-scaled from $\Psi$, function $\Psi_\beta$ is non-decreasing, right-continuous, and in fact a distribution function. If $\Psi$ is continuous everywhere (without jumps), $\beta$-CVaR coincides with the lower CVaR $\phi_\beta^L(x) := \mathbb{E}_\xi [L | L \geq \eta_\beta(x)]$, that is the conditional expectation of the loss beyond the $\beta$-

VaR. Hence, roughly speaking, $\beta$-CVaR is the expected loss in the worst 100(1 - $\beta$)% scenarios; i.e., cases of such severe losses occur only 100(1 - $\beta$) percent of the time.

The $\beta$-CVaR can be also defined as the optimal value of the following optimization problem

$$ \phi_\beta(x) := \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\beta} \mathbb{E}_\xi [L(x, \xi) - \eta]^+ \right\}. $$

Let $F_\beta(x, \eta)$ denote the objective function in (4). Key properties of $F_\beta$ and its relationship with $\eta_\beta(x)$ and $\phi_\beta(x)$ are summarized next.

**Theorem 1** ([31], pp. 1454–1457). Function $F_\beta(x, \eta)$ is finite and convex in $\eta$. Values $\eta_\beta(x)$ and $\phi_\beta(x)$ are linked through $F_\beta(x, \eta)$ as

$$ \eta_\beta(x) = \arg\min_{\eta \in \mathbb{R}} F_\beta(x, \eta) \quad \text{(5)} $$

$$ \phi_\beta(x) = F_\beta(x, \eta_\beta(x)) \quad \text{(6)} $$

$$ \min_{x \in X} \phi_\beta(x) = \min_{(x, \eta) \in X \times \mathbb{R}} F_\beta(x, \eta). \quad \text{(7)} $$

Moreover, if $L(x, \xi)$ is convex in $x$, then $F_\beta(x, \eta)$ is jointly convex in $(x, \eta)$, while $\phi_\beta(x)$ is convex in $x$.

From Definition 2, it can be seen that CVaR is an upper bound of VaR, implying that portfolios with small CVaR also have small VaR. As a consequence of Theorem 1, minimizing the convex $\phi_\beta(x)$ amounts to minimizing $F_\beta(x, \eta)$, which is not only convex, but also easier to approximate. A readily implementable approximation of the expectation function $F_\beta$ is its empirical estimate using $N_s$ Monte Carlo samples $\{\xi_s\}_{s=1}^{N_s}$, namely

$$ \hat{F}_\beta(x, \eta) = \eta + \frac{1}{N_s(1-\beta)} \sum_{s=1}^{N_s} (L(x, \xi_s) - \eta)^+. \quad \text{(8)} $$

Clearly, the sample average approximation method is distribution free, and the law of large numbers ensures $\hat{F}_\beta$ approximates well $F_\beta$ for $N_s$ large enough. Furthermore, $\hat{F}_\beta(x, \eta)$ is convex with respect to $(x, \eta)$ if $L(x, \xi_s)$ is convex in $x$. The non-differentiability due to the projection operator can be readily overcome by leveraging the epigraph form of $F$, which will be shown explicitly in Section III-C.

With the function $F_\beta(x, \eta)$, it is now possible to develop the CVaR-based stochastic market clearing, as detailed in the next section.

### III. Stochastic Market Clearing

In a day-ahead electricity market, participants including power generation companies and load service entities (LSEs) first submit their hourly supply bids and demand offers to market operators for the next operating day. Then, the ISO or regional transmission organization (RTO) clear the forward markets yielding least-cost unit commitment decisions, power dispatch outputs, and the corresponding DA clearing prices. The MC procedure proceeds in two stages. A security-constrained unit commitment (SCUC) is performed first by solving a large-scale mixed integer program to commit generation resources after simplifying or omitting transmission constraints. The second stage involves security-constrained economic dispatch (SCED) obtaining the economical power generation outputs and the locational marginal prices (LMPs) as a byproduct. With unit commitment decisions fixed, SCED is usually in the form of DC-OPF, including the transmission network constraints [33].

The MC process is implemented with a goal of minimizing the system net cost, or equivalently maximizing the social welfare. With the trend of increasing penetration of renewables, WPPs are able to directly bid in the forward market [34]. Under uncertainty of wind generation, it now becomes challenging but imperative for the ISOS/RTOs and market participants to extract forecast information and make efficient decisions, including reserve requirements, day-ahead scheduling, market clearing, reliability commitments, as well as the real-time dispatch [35]. In this section, a stochastic MC approach using the CVaR-based transaction cost will be developed as follows.

#### A. CVaR-based Energy Transaction Cost

Consider a power system comprising $N_b$ buses, $N_l$ lines, $N_g$ conventional generators, $N_w$ wind farms and $N_a$ aggregators, each serving a large number of residential end-users with controllable smart appliances. Let $T := \{1, 2, \ldots, T\}$ denote the scheduling horizon of interest, e.g., one day ahead. If a wind farm is located at bus $m$, two quantities will be associated with it: the actual wind power generation $w_m$, and the power scheduled to be injected $p_{W,m}$. Note that the former is random, whereas the latter is a decision variable. For notational simplicity, define also two $N_w$-dimensional vectors $w^t := [w_{w,m}^t, \ldots, w_{w,N_w}^t]^t$, and $p_{W,t} := [p_{W,1}^t, \ldots, p_{W,N_w}^t]^t$.

Since $w^t$ varies randomly, either energy surplus or shortage should be included to satisfy the nodal balance with the committed quantity $p_{W,t}$. When surplus occurs, the wind farms can sell the excess wind energy back to the spot market, or simply curtail it. For the case of shortage, in order to accomplish the promised bid in the DA contract, farms can buy the energy shortfall from the RT market in the form of ancillary services.

Let $b^t := [b_{1}^t, \ldots, b_{N_b}^t]$ and $s^t := [s_{1}^t, \ldots, s_{N_b}^t]$ collect the purchase and selling prices at time $t$, respectively. Clearly, with the power shortfall and surplus being $[p_{W,t} - w^t]^+$ and $[w^t - p_{W,t}]^+$ at time $t$, the grid-wide net transaction cost is

$$ T(p_w, w) = \sum_{t=1}^{T} \left( b^t \cdot [p_{W,t} - w^t]^+ - s^t \cdot [w^t - p_{W,t}]^+ \right) $$

$$ = \sum_{t=1}^{T} \left( w^t - p_{W,t} \right) + \vartheta^t \cdot (p_{W,t} - w^t) \quad \text{(9)} $$

where $\vartheta^t := \frac{b^t - s^t}{2}$ and $\vartheta^t := \frac{b^t + s^t}{2}$; $p_{W,t}$ and $w$ collect $p_{W,t}^t$ and $w^t$ for all $t \in T$, respectively.
Replacing $L(\cdot)$ in (4) with $T(\cdot, \cdot)$, function $F_\beta$ can be expressed through the conditional expected transaction cost as

$$F_\beta(p_W, \eta) = \eta + \frac{1}{1 - \beta} \mathbb{E}_w \left[ \sum_{t=1}^{T} \left( \varepsilon_t \cdot |p_W^t - w^t| + \vartheta_t \cdot (p_W^t - w^t) \right) \right].$$

(10)

A condition guaranteeing convexity of $F_\beta(p_W, \eta)$ is established next.

**Proposition 1.** If the selling price $s^t_m$ does not exceed the purchase price $b^t_m$ for any $m \in N_w$ and $t \in T$, function $F_\beta(p_W, \eta)$ is jointly convex with respect to $(p_W, \eta)$.

**Proof:** Thanks to Theorem 1, it suffices to show that $T(p_W, w) = \sum_{t=1}^{T} \varepsilon_t \cdot |p_W^t - w^t| + \vartheta_t \cdot (p_W^t - w^t)$ is convex in $p_W$ under the proposition’s condition. Clearly, the stated condition is equivalent to $\varepsilon^t \geq 0$ for all $t \in T$. Thus, by the convexity of the absolute value function, and the convexity-preserving operators of summation and expectation [36, Sec. 3.2], the claim follows readily.

In this paper, a perfectly competitive market is assumed such that all participants act as price takers. That is, every competitor is atomistic to have small enough market share so that there is no market power affecting the price [37]. For American electricity markets, a single pricing mechanism is used such that $s^t = b^t$ holds in most of the scenarios. This is a special case of the pricing condition in Prop. 1, which facilitates calculating the function (10) since the absolute value functions vanish. Note that it is possible that different WPPs may buy (sell) wind energy from (to) different sellers (purchasers) in a competitive electricity pool as an ancillary service, which can yield different purchase and selling prices.

For most of the European markets including UK, France, Italy, and Netherlands, the imbalance prices \{\text{b}^t, \text{s}^t\}_t are commonly set in an ex-post way that is known as dual imbalance pricing [38]. Specifically, if the system RT imbalance is negative, i.e., the overall market is short, then $s^t = \chi^t \leq b^t$ holds, where $\chi^t := [\chi_1^t, \ldots, \chi_{N_w}^t]^T$ collects the DA prices at the buses attached with all $N_w$ wind farms. In this case, the RT purchase price is typically higher than the DA price, reflecting the cost of acquiring the balancing energy [39]. Wind farms with excess energy can sell this part to reduce the system imbalance but only be paid the DA prices. On the other hand, we have $s^t \leq \chi^t = b^t$ if the market is long. Hence, market participants selling excess energy receive a balancing price which is lower than the DA one, while those running negative imbalance pay the DA price. Note that the relationship $s^t \leq \chi^t \leq b^t$ always holds even when the market imbalance outcome is unknown at the time of the DA bids. Such a pricing mechanism drives bidders to match their forward offers with the true forecasts of generation or consumption.

Leveraging the CVaR-based transaction cost, a stochastic MC problem based on the DC-OPF will be formulated next.

**B. CVaR-based Market Clearing**

Let $p_G^t :=[P_{G_1}^t, \ldots, P_{G_{N_g}}^t]$ and $p_{DRA}^t := [P_{DRA_1}^t, \ldots, P_{DRA_{N_j}}^t]$ denote the power outputs of the thermal generators, and the power consumption of the aggregators at slot $t$, respectively. Define further the sets $N_a := \{1, 2, \ldots, N_a\}$ and $N_g := \{1, 2, \ldots, N_g\}$. Each aggregator $j \in N_a$ serves a set $R_j$ of residential users, and each user $r \in R_j$ has a set $S_{rj}$ of controllable appliances. Let $p_j$ be the power consumption of appliance $s$ with user $r$ corresponding to aggregator $j$ across the slots. The operational constraints of $p_j$ are captured by a set $P_{j}$, while the end user satisfaction is modeled by a concave utility function $U_j(s)$. Furthermore, let convex functions $C_i(.)$ denote the generation costs, and $p_{BL}$ the base load demand. For brevity, let vector $p_i$ collect variables $\eta$ and $\{p_G^t, p_{DRA}^t, p_{BL}, \theta^t\}_{t \in T}$, and vector $\{p_j\}_{j \in N_a}$ the power consumption of all appliances with the aggregator $j$.

Hinging on three assumptions: a1) lossless lines, a2) small voltage phase differences, and a3) approximated one p.u. voltage magnitudes, the DC-OPF based stochastic MC stands with the goal of minimizing the social cost:

$$\min_{t=1}^{T} \sum_{i=1}^{N_g} C_i(P_{G_i}^t) - \sum_{j=1}^{N_a} \sum_{r \in R_j} U_j(s) + \mu F_\beta(p_W, \eta)$$

subject to:

$$A_p p_G^t + A_w p_W^t - A_a p_{DRA}^t - p_{BL}^t = B_n \theta^t, \ t \in T$$

(11a)

$$P_{gmin} \leq P_{G_i}^t \leq P_{gmax}, \ i \in N_g, \ t \in T$$

(11b)

$$- R_{l}^{down} \leq P_{G_i}^t - P_{G_i}^{t-1} \leq R_{l}^{up}, \ i \in N_g, \ t \in T$$

(11c)

$$f_{min} \leq B_j \theta^t \leq f_{max}, \ t \in T$$

(11d)

$$\theta^t_1 = 0, \ t \in T$$

(11e)

$$0 \leq p_W \leq P_{wmax}$$

(11f)

$$0 \leq P_{DRA_j} \leq F_{DRA_{j}}, \ j \in N_a, \ t \in T$$

(11g)

$$\sum_{r \in R_j, s \in S_{rj}} p_j^t, \ j \in N_a, \ t \in T$$

(11h)

$$p_j \in P_j, \ s \in S_{rj}, \ r \in R_j, \ j \in N_a$$

(11i)

where the nodal susceptance matrix $B_n := -A_n^T B_n A_n \in \mathbb{R}_{N_a \times N_a}$ and the angle-to-flow matrix $B_j := -B_j A_n \in \mathbb{R}_{N_a \times N_a}$. The $t$th row of the branch-node incidence matrix $A_n \in \mathbb{R}_{N_b \times N_a}$ has 1 and -1 in its entry corresponding to the from and to nodes of branch $\ell$, and 0 elsewhere; and the square diagonal matrix $B_s := \text{diag}(b_1, \ldots, b_{N_b})$ is the branch susceptibility matrix collecting the primitive susceptance across all branches.

Matrices $A_g \in \mathbb{R}_{N_g \times N_s}$, $A_w \in \mathbb{R}_{N_b \times N_w}$ and $A_a \in \mathbb{R}_{N_a \times N_a}$ in (11b) are the incidence matrices of the conventional generators, the wind farms, and the aggregators, respectively. Take $A_g$ as an example, $(A_g)_{mn} = 1$ if the $n$th generator is injected to the $m$th bus, and $(A_g)_{mn} = 0$, otherwise. Matrices $A_w$ and $A_a$ can be constructed likewise. Consider the power network in Fig. 2 adapted from the Western Electricity Coordinating Council (WECC) system [40].
With $N_b = 6$, $N_l = 6$, $N_g = 3$, and $N_a = 4$, matrices $A_g$, $A_w$, and $A_a$ take the following form:

\[
A_g = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix},
A_w = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix},
A_a = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

A smart appliance example is charging a plug-in hybrid electric vehicle (PHEV), which typically amounts to consumer $E_{jrs}$ over a specific horizon from a start time $T^s_{jrs}$ to a termination time $T^r_{jrs}$. The consumption must remain within a range between $p^\text{min}_{jrs}$ and $p^\text{max}_{jrs}$ per period. With $T^E_{jrs} := \{T^s_{jrs}, \ldots, T^r_{jrs}\}$, set $P_{jrs}$ takes the form

\[
P_{jrs} = \left\{ p_{jrs} \left| \sum_{t \in T^E_{jrs}} p^t_{jrs} = E_{jrs}, p^t_{jrs} \in [p^\text{min}_{jrs}, p^\text{max}_{jrs}], \right. \right\}.
\]

Further examples of $P_{jrs}$ and $U_{jrs}(P_{jrs})$ can be found in [17], where it is argued that $P_{jrs}$ is a convex set for several appliance types of interest.

Linear equality (11b) is the nodal balance constraint; i.e., the load balance at bus levels dictated by the law of conservation of power. Limits of generator outputs and ramping rates are specified in constraints (11c) and (11d). Network power flow constraints are accounted for in (11e). Without loss of generality, the first bus can be set as the reference bus with zero phase in (11f). Constraints (11h) and (11i) capture the lower and upper limits of the energy consumed by the aggregators and the committed wind power, respectively. Equality (11i) amounts to the aggregator-user power balance equation; and constraints (11j) define the feasible set of appliances. Finally, the pre-determined risk-aversion parameter $\mu > 0$ controls the trade off between the transaction cost and the generation cost as well as the end-user utility.

**Remark 1.** (Availability of real-time prices). In this paper, the real-time prices $\{b^t, s^t\}_{t \in T}$ are assumed to be perfectly known to the ISO for the DA market clearing. However, such an assumption can be readily extended to a more practical setup by taking the price stochasticity into account. Specifically, imperfect price information can be modeled by appropriately designing the function $F_{\beta}(p_W, \eta) [\text{cf. (9)}]$. For example, the expectation can be also taken over the random RT prices in (10) as $F_{\beta}(p_W, \eta) = \eta + \frac{1}{|\Delta|} \mathbb{E}_{\mathbf{w}, \mathbf{s}^t, \mathbf{b}^t} [T(p_W, \mathbf{w}) - \eta]^+$. The dependence between $\{b^t, s^t\}$ and $\mathbf{w}$ can be further investigated. In addition, worst-case analysis is available upon postulating an uncertainty set $\Delta$ for $\{b^t, s^t\}$. This results in a novel risk measure given as $F_{\beta}(p_W, \eta) = \eta + \frac{1}{|\Delta|} \mathbb{E}_{\mathbf{w}, \mathbf{s}^t, \mathbf{b}^t} [\operatorname{sup}_{\{b^t, s^t\}} T(p_W, \mathbf{w}) - \eta]^+$.

It is worth mentioning that SCED and SCUC yield two different market pricing systems: locational marginal pricing and convex hull pricing (a.k.a. extended LMP). The ED formulation produces the LMPs given by the dual variables associated with the supply-demand balance constraint. Prices supporting the equilibrium solution are found at the intersection of the supply marginal cost curve with the demand bids. However, if discrete operations of UC are involved, there is no exact price that supports such an economic equilibrium. This issue prompted the introduction of the convex hull pricing to reduce the uplift payments [41]. In the present paper, the core ED model is considered to deal with the high penetration of renewables and large-scale DR programs. Therefore, the formulation (11) relies on re-solving the dispatch problem with fixed UC decisions.

**Remark 2.** (Reliability assessment commitment). The proposed dispatch model can be cast as a two-stage program. The first stage is the DA MC, and the second is simply the balancing operation (recourse action) dealing with differences between the pre-dispatch amount and the actual wind power generation. Between the DA and RT markets, ISOs implement the reliability assessment commitment (RAC) as a reliability backstop tool to ensure sufficient resources are available and cover the adjusted forecast load online. One principle of the RAC process is to commit the capacity deemed necessary to reliably operate the grid at the least commitment cost. In this step, based on the updated information of the wind power forecast, WPPs have an opportunity to feedback to the ISO if they are able to commit the scheduled wind power decided by the DA MC. Then, the ISO is able to adjust UC decisions as necessary to ensure reliability.

To this end, reformulation of problem (11) as a smooth convex minimization is useful for developing distributed solvers, as detailed next.

**C. Smooth Convex Minimization Reformulation**

It is clear that under the condition of Proposition 1, the objective and the constraints of (11) are convex, which renders it not hard to solve in principle. Nevertheless, due to the high-dimensional integration present in $F_{\beta}(p_W, \eta) [\text{cf. (10)}]$, an analytical solution is typically impossible. To this end, it is necessary to re-write the resulting problem in a form suitable for off-the-shelf solvers.
First, as shown in (8), an efficient approximation of $F_\beta(p_W, \eta)$ is offered by the empirical expectation using i.i.d. samples $\{w_s\}_{s=1}^{N_\eta}$; that is,
\[
\hat{F}_\beta(p_W, \eta) = \eta + \frac{1}{N_\eta(1-\beta)} \sum_{s=1}^{N_\eta} \left[ \sum_{t=1}^{T} (\varphi^t \cdot |p_W^t - w_s^t|) + \vartheta^t \cdot (p_W^t - w_s^t) \right] - \eta^+.
\] (13)

Next, by introducing auxiliary variables $\{u_s\}_{s=1}^{N_\eta}$, the non-smooth convex program (11) can be equivalently re-written as the following smooth convex minimization:
\[
\begin{align*}
\min_{p_p} & \quad \sum_{t=1}^{T} \sum_{i=1}^{N_d} C_i(p_{G_i}^t) - \sum_{j=1}^{N_a} \sum_{s \in S_{jrs}} U_{jrs}(p_{jrs}) \\
\text{subject to} & \quad (\varphi^t \cdot |p_{W}^t - w_s^t| + \vartheta^t \cdot (p_{W}^t - w_s^t)) \leq u_s + \eta, \\
& \quad s \in S_{jrs}, \quad jrs \in N_{jrs}, \quad s \in S_{jrs}.
\end{align*}
\] (14a)

variables: $\{p_p\}_{jrs}^{N_{jrs}}$, $\{u_s \in \mathbb{R}_+\}_{s=1}^{N_\eta}$.

Under mild conditions, the optimal solution set of (14) converges exponentially fast to its counterpart of (11), as the sample size $N_\eta$ increases. The proof is based on the theory of large deviations [42], but is omitted here due to space limitations.

Problem (14) can be solved centrally at the ISO in principle. However, with large-scale DR, distributed solvers are well motivated not only for computational efficiency but also for privacy reasons. Specifically, functions $U_{jrs}(p_{jrs})$ and sets $\{P_{jrs}\}$ are private, and are not revealed to the ISO; and (ii) the operational sets $\{P_{jrs}\}_{jrs}$ of very large numbers of heterogeneous appliances may become prohibitively complicated; e.g., mix-integer constraints can even be involved to model the ON/OFF status and un-interruptible operating time of end-user appliances [43], [44]. This renders the overall problem intractable for the ISO. To this end, the DR aggregators can play a critical role to split the resulting optimization task as detailed next.

IV. DISTRIBUTED MARKET CLEARING VIA ADMM

Selecting how to decompose the optimization task as well as updating the associated multipliers are crucial for the distributed design. Fewer updates simply imply lower communication overhead between the ISO and the aggregators. One splitting approach is the dual decomposition with which the dual subgradient ascent algorithm is typically very slow. Instead, a fast-convergent solver via the ADMM [45] is adapted in this section for the distributed MC.

A. The ADMM Method

Consider the following separable convex minimization problem with linear equality constraints:
\[
\begin{align}
\min_{x \in \mathcal{X}, y \in \mathcal{Y}} & \quad f(x) + g(y) \\
\text{subject to} & \quad Ax + By = c.
\end{align}
\] (15a)

For the stochastic MC problem (14), the primal variable $x$ comprises the group $\{u_s\}_{s \in N_{jrs}}$ and $p_0$, while $y$ collects $\{p_{jrs}\}_{jrs}$. Hence, set $\mathcal{X}$ captures constraints (11b)–(11h) and (14b) while $\mathcal{Y}$ represents (11i). The linear equality constraint (15b) corresponds to (11i).

Let $\lambda := [\lambda_1^T, \ldots, \lambda_{N_\eta}^T]^T \in \mathbb{R}^{TN_\eta}$ denote the Lagrange multiplier vector associated with the constraint (11i). The partially augmented Lagrangian of (14) is thus given by (16), where the weight $\rho > 0$ is a penalty parameter controlling the violation of primal feasibility, which turns out to be the step size of the dual update. As the iterative solver of (16) proceeds, the primal residual converges to zero that ensures optimality. Judiciously selecting $\rho$ thus strikes a desirable tradeoff between the size of primal vis-à-vis dual residuals. Note also that by varying $\rho$ over a finite number of iterations may improve convergence [45]. In a nutshell, finding the “optimal” value of $\rho$ is generally application-dependent that requires a trial-and-error tuning.

Different from [46] where the power balance and phase consistency constraints are relaxed, in this work only the aggregator-user power balance equation (11i) is dualized so that the nodal balance equation (11b) is kept in the subproblem of the ISO. Decomposing the problem (14) in such a way can reduce the heavy computational burden at the ISO while respect the privacy of end users within each aggregator. The ADMM iteration cycles between primal variable updates using block coordinate descent (a.k.a. Gauss-Seidel), and dual variable updates via gradient ascent. The resulting distributed MC is tabulated as Algorithm 1, where $k$ is the iteration index. The last step is a reasonable termination criterion based on the primal residual (see [45, Sec. 3.3.1])
\[
\xi := \left[ \sum_{t=1}^{T} \sum_{j=1}^{N_a} (P_{DRA_j}^t - \sum_{r,s} p_{jrs}^t)^2 \right]^{1/2}.
\] (17)

\[
\begin{align*}
L_\rho(x, y, \lambda) := & \sum_{t=1}^{T} \sum_{i=1}^{N_d} C_i(p_{G_i}^t) - \sum_{j=1}^{N_a} \sum_{s \in S_{jrs}} U_{jrs}(p_{jrs}) + \frac{\eta}{N_\eta(1-\beta)} \sum_{s=1}^{N_\eta} u_s \\
& + \sum_{t=1}^{T} \sum_{j=1}^{N_a} \lambda_j^t \left( P_{DRA_j}^t - \sum_{r,s} p_{jrs}^t \right) + \frac{\rho}{2} \sum_{t=1}^{T} \sum_{j=1}^{N_a} \left( p_{DRA_j}^t - \sum_{r,s} p_{jrs}^t \right)^2.
\end{align*}
\] (16)
Algorithm 1 ADMM-based Distributed Market Clearing

1: Initialize $\lambda(0) = 0$
2: repeat for $k = 0, 1, 2, \ldots$
3: update primal variables:
   \[
   x(k + 1) = \arg\min_{x \in X} \sum_{i} C_i(P_{G_i}^t) + \mu \left( \eta + \sum_{s=1}^{N_{s}} u_s \right) \\
   y(k + 1) = \arg\min_{y \in Y} \sum_{i} \lambda_i^j(k) P_{G_i}^t + \frac{\rho}{2} \sum_{i} \sum_{s=1}^{N_{s}} (P_{G_i}^t - P_{G_i}^r)^2 
   \]
4: update dual variables: for all $j \in \mathcal{N}_{a}$ and $t \in \mathcal{T}$
   \[
   \lambda_i^j(k + 1) = \lambda_i^j(k) + \rho (P_{G_i}^t(k + 1) - \sum_{r,s} p_{jrs}^t(k + 1))
   \]
5: until $\xi \leq \epsilon_{pr}$

Specifically, given the Lagrangian multipliers $\lambda(k)$ and the power consumption $\{p_{jrs}(k)\}_{jrs}$ of the end-user appliances, the ISO solves the convex subproblem (18) given as follows:

\[
\begin{align*}
\mathbf{p}_0(k + 1) &= \arg\min_{\mathbf{p}_0, \{u_s\}} \sum_{i \in \mathcal{T}_r} C_i(P_{G_i}^t) + \mu \left( \eta + \sum_{s=1}^{N_{s}} u_s \right) \\
&+ \sum_{i \in \mathcal{T}_r} \lambda_i^j(k) P_{G_i}^t + \frac{\rho}{2} \sum_{i \in \mathcal{T}_r} \sum_{s=1}^{N_{s}} (P_{G_i}^t - P_{G_i}^r)^2
\end{align*}
\]

subject to:

\[
\mathbf{A}_d \mathbf{p}_G^t + \mathbf{A}_u \mathbf{p}_W^t - \mathbf{A}_0 \mathbf{p}^t_{\text{DRA}} - \mathbf{p}^t_{\text{BL}} = \mathbf{B}_n \mathbf{\theta}^t, \quad t \in \mathcal{T}
\]

\[
P_{G_i}^t + P_{G_i}^r \leq P_{\text{max}} G_i, \quad i \in \mathcal{N}_{g}, \quad t \in \mathcal{T}
\]

\[
R_{\text{down}}^i \leq P_{G_i}^t - P_{G_i}^r \leq R_{\text{up}}^i, \quad i \in \mathcal{N}_{g}, \quad t \in \mathcal{T}
\]

\[
\mathbf{f}^\text{min} \leq \mathbf{B}_j \mathbf{\theta}^t \leq \mathbf{f}^\text{max}, \quad t \in \mathcal{T}
\]

\[
\mathbf{\theta}_i^t = 0, \quad t \in \mathcal{T}
\]

\[
0 \leq \mathbf{p}_W^t \leq \mathbf{p}^\text{max}_W, \quad t \in \mathcal{T}
\]

\[
0 \leq P_{\text{DRA}}^j \leq P_{\text{max}} DRA_j, \quad j \in \mathcal{N}_a, \quad t \in \mathcal{T}
\]

\[
\sum_{i=1}^{T} \left( \mathbf{w}_i \cdot (\mathbf{p}_i W^t - \mathbf{w}_s^t) + \mathbf{\theta}^t \cdot (\mathbf{p}_i W^t - \mathbf{w}_s^t) \right) \leq u_s + \eta, \quad s \in \mathcal{N}_s,
\]

and $u_s \geq 0, \quad s \in \mathcal{N}_s$.

Interestingly, (19) is decomposable so that $\{p_{jrs}(k)\}_{jrs}$ can be separately solved by each aggregator:

\[
\begin{align*}
\{p_{jrs}(k + 1)\}_{jrs} &= \arg\min_{\{p_{jrs}\}_{jrs}} \sum_{t=1}^{T} \lambda_i^j(k) \sum_{r,s} p_{jrs}^t \\
&- \sum_{r \in \mathcal{R}_j} U_{jrs}(p_{jrs}) + \frac{\rho}{2} \sum_{t=1}^{T} \sum_{r,s} (p_{jrs}^t - P_{\text{DRA}}^t(k + 1))^2
\end{align*}
\]

subject to: $\{p_{jrs} \in \mathcal{P}_{jrs}\}_{jrs}$.

Having found $\mathbf{p}_0(k)$ and $\{\mathbf{p}_{jrs}(k)\}_{jrs}$, the multipliers $\{\mu_i^j\}_{j,t}$ are updated using gradient ascent as in (20). To solve the convex problem (22), each aggregator must collect the corresponding users’ information including $U_{jrs}$ and $P_{jrs}$. This is implementable via the advanced metering infrastructure [47].

**Remark 3.** (Distributed demand response). It must be further pointed out that the quadratic penalty $(P_{\text{DRA}}^t - \sum_{r,s} p_{jrs}^t)^2$ in (16) couples load consumptions $\{p_{jrs}^t\}$ over different residential users. Hence, the ADMM-based distributed solver may not be applicable whenever $p_{jrs}^t$ must be updated per end user rather than the aggregator. This may arise either to strictly protect the privacy of end users from DR aggregators, or, to accommodate large-scale DR programs where each aggregator cannot even afford solving the subproblem (22). In this case, leveraging the plain Lagrangian function (no coupling term), the dual decomposition based schemes can be utilized by end users to separately update $\{p_{jrs}^t\}$ in parallel; see e.g., [17] and [48].

The convergence of the ADMM solver and its implications for the market price are discussed next.

**B. Pricing Impacts**

Suppose two additional conditions hold for the convex problem (14): c1) functions $\{C_i(\cdot)\}_i$ and $\{-U_j \} j$ are closed and proper convex; and c2) the plain Lagrangian $L_0$ has a saddle point. Then, the ADMM iterates of the objective (14a) and the dual variables $\lambda^j_p$ are guaranteed to converge to the optimum [45]. In addition, if the objective is strongly convex, then the primal variable iterates including $\mathbf{p}_G^t$, $\mathbf{p}_{\text{DRA}}^t$, $\mathbf{p}_W^t$ and $\{p_{jrs}^t\}_{jrs}$ converge to the globally optimal solutions.

The guaranteed convergence of the dual variables also facilitates the calculation of LMPs. Let $\mathbf{\lambda}^t := [\lambda_1^t, \ldots, \lambda_{N_a}^t]$, $\mathbf{\bar{\lambda}}^t := [\bar{\lambda}_1^t, \ldots, \bar{\lambda}_{N_a}^t]$ denote the optimal Lagrange multipliers associated with the aggregator-user balance constraint (11i), and the nodal balance constraint (11b), respectively. Note that with the optimal solutions $\mathbf{\lambda}^t$ and $\{p_{jrs}^t\}_{jrs}$ obtained by the ADMM solver, the LMPs $\{\bar{\lambda}^t\}_t$ can be found by solving the subproblem (21) with primal-dual algorithms. In addition, if $0 < P_{\text{DRA}}^t < P_{\text{max}} DRA_j, \forall j, t$ holds at the optimal solution $\bar{\mathbf{P}}_{\text{DRA}}^t$, then $\mathbf{\lambda}^t = \mathbf{A}_0^t \mathbf{\bar{\lambda}}^t$; i.e., $\lambda_i^j = \bar{\lambda}_i^j$ for all aggregators $j$ attached with bus $n$ (see also [17]). To this end, payments of the market participants can be calculated with the obtained LMPs and optimal DA dispatches. In the RT market of a twoslot system, if the supplier at bus $n$ delivers $\bar{P}_{G_n}^t$ with the real-time price $\tau_n^t$, then the supplier gets paid

\[
\Pi_{G_n} = \sum_{t=1}^{T} \bar{\lambda}_n^t \bar{P}_{G_n}^t + \bar{\lambda}_n^t (\bar{P}_{G_n}^t - \bar{P}_{G_n}^t).
\]

Likewise, the aggregator at bus $n$ needs to pay

\[
\Pi_{\text{DRA}} = \sum_{t=1}^{T} \tau_n^t P_{\text{DRA}}^t + \tau_n^t (\bar{P}_{\text{DRA}}^t - \bar{P}_{\text{DRA}}^t).
\]

The revenue of the wind farm at bus $n$ is

\[
\Pi_{W_n} = \sum_{t=1}^{T} \tau_n^t P_{W_n}^t + s_n^t (w_n^t - \bar{P}_{W_n}^t)^+ + b_n^t (P_{W_n}^t - w_n^t)^+.
\]
Remark 4. (Pricing consistence). In a perfectly competitive market, any arbitrage opportunities between the DA and RT markets are exploited by market participants. Hence, the DA nodal prices are consistent with the DT nodal prices meaning the expectations of the latter converge to the former. The concepts of price distortions and revenue adequacy have been recently proposed for the stochastic MC in [49]. In the setup of a single snapshot therein, it has been proved that the medians and expectations of RT prices converge to the DA counterparts for the $\ell_1$ and $\ell_2$ penalties between the RT and DA power schedules, respectively. Building upon this solid result, it is possible to establish bounded price distortions for the proposed model, while its consistent pricing property can also be analyzed in a similar fashion. The involved important analysis is however beyond the scope of this paper, and is left for future work.

V. Numerical Tests

In this section, simulated tests are presented to verify the merits of the proposed CVaR-based MC. The tested power system is modified from the WECC system as illustrated in Fig. 2. Each of the 4 DR aggregators serves 200 residential customers. The scheduling horizon starts from 12am until 23pm, a total of 24 hours.

Time-invariant generation cost functions were chosen quadratic as $C_i(P_{Gi}^t) = a_i(P_{Gi}^t)^2 + b_i P_{Gi}^t$, for all $i$ and $t$. For simplicity, each end user has one PHEV to charge from midnight. All detailed parameters of the conventional generators and loads are listed in Tables I and II. The upper bound of each aggregator’s consumption is $P_{\text{max, DRA}} = 50$ MW. At a base of 100 MVA, the values of the network reactances are $\{X_{16}, X_{62}, X_{25}, X_{53}, X_{34}, X_{41}\} = \{0.2, 0.3, 0.25, 0.1, 0.3, 0.4\}$ p.u. Finally, no flow limits were imposed, while the utility functions $\{U_{jrs}(\cdot)\}$ were set to zero. The resulting convex program (21) and (22) were modeled using the Matlab-based package CVX [50], and solved by SeDuMi [51].

Variable characteristics of the daily power market are captured via two groups of parameters shown in Fig. 3: the fixed base load demand $\{p_{BL}\}$, and the purchase prices $\{b^t\}$ at the buses attached with three wind farms. The prices were obtained by scaling the real data from the Midcontinent ISO (MISO) [52]. Two peaks of $\{b^t\}$ appear during the morning 7am to 12pm, and early night 6pm to 9pm. The selling prices $\{s^t\}$ were set to $s^t = 0.9b^t$ satisfying the convexity condition in Proposition 1. The rated capacity of each wind farm was set to 20 MW, yielding a 23% wind power penetration of the total power generation capacity.

Wind power output samples $\{w^t\}_{s,t}$ are needed as inputs of (21). These samples can be obtained either from forecasts of wind power generation, or, by using the distributions of wind speed together with the wind-speed-to-wind-power mappings [cf. [5]]. In this paper, the needed samples were obtained from the model $w^t = w^t + n^t_s$, $\forall t \in T$. The DA wind power forecasts $\{w^t\}$ were taken from the MISO market on March 8, 2014. The forecast error $n^t_s$ was assumed zero-mean white Gaussian. Possible negative-valued elements of the generated samples $\{w^t\}_{s,t}^{N_s}$ were truncated to zero. Finally, the sample size $N_s = 200$, the probability level $\beta = 0.95$, the trade-off weight $\mu = 1$, and the primal-residual tolerance $\xi_{pri} = 10^{-4}$ were set for all simulations, unless otherwise stated.

Figure 4 demonstrates the fast convergence of the proposed ADMM-based solver. The pertinent parameters were set to $\rho = 35$ and $\lambda_f^t(0) = p_{jrs}^t(0) = 0$. Clearly, both the cost and

**Table I**

<table>
<thead>
<tr>
<th>Unit</th>
<th>$a_i$</th>
<th>$b_i$</th>
<th>$P_{Gi}^{\text{max}}$</th>
<th>$P_{Gi}^{\text{min}}$</th>
<th>$R_{Gi}^{\text{up}}$</th>
<th>$R_{Gi}^{\text{down}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3</td>
<td>50</td>
<td>90</td>
<td>10</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>30</td>
<td>50</td>
<td>5</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>40</td>
<td>60</td>
<td>8</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

**Table II**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{\text{PHEV}}$ (kWh)</td>
<td>Max energy consumption</td>
<td>Uniform on ${10, 11, 12}$</td>
</tr>
<tr>
<td>$p_{\text{PHEV}}$ (kWh)</td>
<td>Min energy consumption</td>
<td>Uniform on ${2.1, 2.3, 2.5}$</td>
</tr>
<tr>
<td>$T^{\text{pr}}_{jrs}$</td>
<td>Time of purchase</td>
<td>1am</td>
</tr>
<tr>
<td>$T^{\text{rs}}_{jrs}$</td>
<td>Time of sale</td>
<td>6am w.p. 70%, 7am w.p. 30%</td>
</tr>
</tbody>
</table>

Fig. 3. Fixed base load demand $\{p_{BL}\}$ and energy purchase prices $\{b^t\}$.

Fig. 4. Convergence of the objective value (14a) and the primal residual (17).
are no CV aR-pertinent terms in the objective and constraint s
for the last two alternatives. For all three approaches, the gen-
wind power generation
limiting MC; (ii) the no risk-limiting MC with the expected
the optimal dispatch and cost: (i) the novel CV aR-based risk-
to the latter.

Three methods were tested to show the performance of
the optimal dispatch and cost: (i) the novel CVaR-based risk-
limiting MC; (ii) the no risk-limiting MC with the expected
wind power generation \( \{w^i\} \); and (iii) the MC without wind
power integration. Specifically, \( p_{W}^i = w^i \) was simply used in
the nodal balance (21b) for (ii), while \( p_{W}^i = 0 \) for (iii). There
are no CVaR-pertinent terms in the objective and constraints
for the last two alternatives. For all three approaches, the

generation cost \( \sum_{t=1}^{T} \sum_{i=1}^{N_{g}} C_i(P_{Gi}^t) \) is fixed after solving (14).
Hence, randomness of the optimal total cost stems from the
transaction cost due to the stochasticity of the actual wind
power generation \( \{w^i\} \) [cf. (9)]. In Fig. 5, the cumulative
distribution functions (CDFs) of the optimal total costs were
plotted using 100,000 i.i.d. wind samples with mean \( \{w^i\} \).
Clearly, the two competing alternatives always incur higher
costs than the novel CVaR-based approach. The values of the
mean and standard deviation (std) of the optimal total cost
are listed in Table III. It can be seen that, compared with
the other two methods, the proposed scheme has a markedly
reduced expected total cost and small changes in the std.

Figures 6, 7, and 8 compare the optimal power dispatches
\( \{P_G^t, P_W^t, p_{DRA}\} \in \mathcal{T} \) of the proposed scheme with those of
the scheme (ii). In Fig. 6, it can be clearly seen that over a
single day the CVaR-based MC dispatches lower and smoother
\( p_{G} \) than the one with (ii). Furthermore, for the novel method,
generators 1 and 3 are dispatched to output their minimum
generation \( P_{Gmi} \), while the output of the generator 2 changes
within its generation limits across time. Such a dispatch results
from the economic incentive since the unit 2 has the lowest
generation cost among all three generators [cf. Table I]. On the
contrary, both generators 2 and 3 fluctuate within a relatively
large range in (ii), mainly to meet the variation of base load
demand \( p_{BL} \); see Fig. 3.

As shown in Fig. 7, the novel CVaR-based approach also
dispatches more \( p_{W} \) than that of (ii). This is because the
energy purchase prices \( b^i \) are smaller than the conventional
generation costs [cf. Table I and Fig. 3]. In addition, \( p_{W1}^t \) and
\( p_{W2}^t \) contribute most of the committed wind power at 1pm and
2pm due to the cheaper buying prices during the corresponding
slots [cf. Fig. 3]. Interestingly, Fig. 8 shows that the PHEVs
are scheduled to start charging earlier for the CVaR-based MC,
where \( p_{DRA} \) is jointly optimized with \( p_{C} \) and \( p_{W} \).

Finally, Fig. 9 shows the effect of the weight parameter
\( \mu \) on the optimal costs of the conventional generation and
the CVaR-based transaction. As expected, the CVaR-based
transaction cost decreases with the increase of \( \mu \). For a larger
\( \mu \), less \( p_{W} \) is scheduled so that more wind power is likely to
be sold in the RT market that yields selling revenues rather
than purchase costs. Consequently, to keep the supply-demand
balance, higher conventional generation cost is incurred by the
increase of \( p_{G} \).
The authors are grateful to the anonymous reviewers and the Editor for the insightful comments and valuable suggestions that led to substantial improvement of the manuscript. The authors also would like to thank Dr. Nikolaos Gatsis (Department of Electrical and Computer Engineering, The University of Texas at San Antonio) for the fruitful discussions on the decomposition algorithms.

**References**


Yu Zhang (S’11) received his B.Eng. and M.Sc. degrees (both with highest honors) in Electrical Engineering from Wuhan University of Technology, Wuhan, China, in 2006 and 2010, respectively. Since September 2010, he has been working towards the Ph.D. degree with the Dept. of Electrical and Computer Engineering (ECE) and the Digital Technology Center (DTC) at the University of Minnesota (UMN). During the summer of 2014, he was a research intern with ABB US Corporate Research Center, Raleigh, NC. His research interests span the areas of smart grids, cyber-physical systems, optimization theory, and machine learning. Mr. Zhang received the Huawei Scholarship and the Infinion Scholarship from the Shanghai Jiao Tong University (2009), the ECE Dept. Fellowship from University of Minnesota (2010), and the Student Travel Awards from the SIAM and the IEEE Signal Processing Society (2014).

Georgios B. Giannakis (F’97) received his Diploma in Electrical Engr. from the NTU, Tech. Univ. of Athens, Greece, 1983, M.Sc. in Mathematics, 1986, and Ph.D. in Electrical Engr., 1986. Since 1999 he has been a professor with the Univ. of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the ECE Department, and serves as director of the Digital Technology Center. His general interests span the areas of communications, networking and statistical signal processing - subjects on which he has published more than 370 journal papers, 630 conference papers, 21 book chapters, two edited books and two research monographs (h-index 112). Current research focuses on sparsity and big data analytics, wireless cognitive radios, mobile ad hoc networks, renewable energy, power grid, gene-regulatory, and social networks. He is the (co-)inventor of 23 patents issued, and the (co-)recipient of 8 best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, the G. W. Taylor Award for Distinguished Research from the University of Minnesota, and the IEEE/SP Technical Field Award (2014). He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.