Robust Workload and Energy Management for Sustainable Data Centers

Tianyi Chen, Yu Zhang, Student Member, IEEE, Xin Wang, Senior Member, IEEE, and Georgios B. Giannakis, Fellow, IEEE

Abstract—A large number of geo-distributed data centers begin to surge in the era of data deluge and information explosion. To meet the growing demand in massive data processing, the infrastructure of future data centers must be energy-efficient and sustainable. Facing this challenge, a systematic framework is put forth in this paper to integrate renewable energy sources (RES), distributed storage units, cooling facilities, as well as dynamic pricing into the workload and energy management tasks of a data center network. To cope with RES uncertainty, the resource allocation task is formulated as a robust optimization problem minimizing the worst-case net cost. Compared with existing stochastic optimization methods, the proposed approach entails a deterministic uncertainty set where generated RES reside, thus can be readily obtained in practice. It is further shown that the problem can be cast as a convex program, and then solved in a distributed fashion using the dual decomposition method. By exploiting the spatio-temporal diversity of local temperature, workload demand, energy prices and renewable availability, the proposed approach outperforms existing alternatives, as corroborated by extensive numerical tests performed using real data.

Index Terms—Cloud computing, data centers, energy storage, geographical load balancing, renewable energy, robust optimization, smart grids.

NOMENCLATURE

A. Abbreviations

| RES      | Renewable energy sources. |
| DC       | Data center. |
| MN       | Mapping node. |
| QoS      | Quality-of-service. |
| IT       | Information technology. |
| DW       | Delay-tolerant workload. |
| CG       | Conventional generator. |
| RG       | Renewable generator. |
| RWEM     | Robust workload and energy management. |

B. Indices, numbers, and sets

| $T, t, T_i, s$ | Number, index, set and sub-set of time slots. |
| $I, i, I$      | Number, index, and set of DCs. |
| $J, j, J$      | Number, index, and set of MNs. |
| $Q, q$         | Number and index of delay-tolerant workloads. |
| $k, \ell$      | Iteration indexes of the (dual) subgradient ascent. |
| $\mathcal{T}_q$ | Sets of all delay-tolerant workloads and those collected by MN $j$. |
| $\mathcal{X}$  | Feasible set of primal variables. |
| $\mathcal{E}_i$ | Uncertainty set of renewable generation in DC $i$. |

C. Constants

| $A_j^t$         | Demand of interactive workloads at MN $j$ per slot $t$. |
| $B_q^t$         | Demand of $q$th delay-tolerant job. |
| $L_{ji}$        | Bandwidth from MN $j$ to DC $i$ per slot $t$. |
| $S_q, E_q$      | Starting time and deadline for $q$th delay-tolerant job. |
| $\rho$          | Fraction of server peak consumption when in idle state. |
| $\gamma_i, \kappa_i^t$ | Parameters of chilled-water and outside-air cooling. |
| $T_{i,RA}^t, T_{i,OA}^t$ | IT rack and outside air temperatures at DC $i$ per time $t$. |
| $P_{i,a}^t$     | Maximum capacity of outside-air cooling in DC $i$ per slot $t$. |
| $P_{i,TH}^t$    | Cooling source selection threshold in DC $i$ per slot $t$. |
| $C_i^0, C_i^t, \overline{C}_i^t$ | Initial, minimum and maximum energy levels of battery. |
| $\overline{P}_{i,\text{ch}}, \overline{P}_{i,\text{dis}}$ | Maximum charging and discharging rates of battery at DC $i$. |
| $T_{i,g}^t$     | Generation capacity of CG at DC $i$. |
| $R_{i,v}^{\text{up}}, D_{i,v}^{\text{dis}}$ | Maximum ramping-up/down rates of CG at DC $i$. |
| $E_i^t, e_i$    | Renewable generation at DC $i$ per slot $t$, and vector collecting $[E_1^t, \ldots, E_i^t]^\top$. |
| $E_i^t, \overline{E}_i^t$ | Upper and lower bounds of renewable generation at DC $i$ per slot $t$. |
| $E_{T_{i,s}}, \overline{E}_{T_{i,s}}$ | Upper and lower bounds of total renewable generation at DC $i$ over $T_{i,s}$. |
In the new era of big data analytics, cloud computing, and the Internet of Things, data centers are proliferating globally to provide important Internet services such as instant messaging, video distribution, and data backup. For the purposes of reliability and quality-of-service (QoS), a cloud service provider usually has multiple data centers (DCs) geographically distributed in different areas. For instance, Google currently operates seven DCs in the US, and fourteen geographically distributed in different areas. For instance, Google’s DCs in 2010 approximately emitted carbon comparable to that of 280,000 cars [10]. Hence, not only achieving energy efficiency is of essence, but also the sustainability of DCs should be targeted [11], [12]. Exploiting renewable energy sources is clearly a key for sustainable operation of DCs [13]. The main challenge of renewable-integrated energy management is to account for its random and nondispatchable nature, which motivates the use of energy storage units [14]. Energy supply side management with distributed storage units was considered for a homogeneous DC [15], and geo-distributed DCs [16]. Taking advantage of RES, a two-timescale Lyapunov optimization technique was developed to control the energy supply in both ahead-of-time and real-time settings [14]. However, spatial diversity of workloads, prices, and renewables was not utilized therein. Furthermore, in [14]–[16] the server model was simplified while the cooling power consumption was ignored. Yet, a substantial amount of DC energy actually goes to a variety of cooling facilities [17], [18].

Existing approaches of dealing with RES uncertainty include the Lyapunov optimization based on the stochastic approximation technique [7], [14]–[16], and the scenario-based stochastic optimization approaches using RES samples from historical data or a given distribution [19], [20]. To guarantee the convergence and optimality, these methods typically require the condition of independent and identically distributed (i.i.d.) RES samples, which is generally unrealistic [21], [22]. For a large time scale, the distribution of the RES generation changes with the seasons. For a short time scale, hourly RES generation can be highly correlated. How to properly handle the RES uncertainty for DC daily operations is still a challenging problem.

In this paper, we consider the optimal workload and energy management for a cloud network consisting of multiple geodistributed mapping nodes (MNs) and DCs. For each DC, practical models of power supply, cooling, and IT systems are considered. In particular, the power supply system comprises multiple energy sources; the cooling system combines two different cooling sources with time-varying cooling coefficients1; and the IT operating system is able to adaptively activate servers and schedule workloads with guaranteed hard deadlines. Distinct from existing works, a deterministic uncertainty set of the unknown renewable generation, as well as a two-way energy trading mechanism is introduced to account for the stochastic and nondispatchable nature of RES. The proposed uncertainty set of the RES generation only requires easy-to-obtain first-order statistics, and allows the possible sample correlation. Control parameters are further designed to trade off robustness for conservatism of the robust optimization

\[ \mathbf{T}_{i,s} \] Historical sample average of total RES over \( T_{i,s} \) at DC \( i \).
\[ \Delta_{s}^{up}, \Delta_{s}^{low} \] Levels of robustness for upper and lower bounds in RES uncertainty set.
\[ \alpha_{i}, \beta_{i} \] Buying and selling energy prices in DC \( i \) at slot \( t \).
\[ \varphi_{i}, \phi_{i} \] Auxiliary variables of energy prices.

D. Variables

- \( a_{j,i}^{t} \): Interactive workloads distributed from MN \( j \) to DC \( i \).
- \( b_{t,q}^{i}, b_{t,q}^{f} \): Amounts of \( q \)th delay-tolerant job routed to, and being processed in DC \( i \) per slot \( t \).
- \( d_{t}^{i} \): Total IT workload demand in DC \( i \) at slot \( t \).
- \( m_{t}^{i} \): Number of active servers in DC \( i \) at slot \( t \).
- \( P_{t,a}, P_{t,a}^{i} \): IT power consumption allocated for chiller and outside-air cooling.
- \( P_{t,dis}, P_{t,dis, i}^{i} \): (Dis)charging power from battery at DC \( i \) per slot \( t \).
- \( P_{t, in}, P_{t, out}^{i} \): Total energy supply and consumption in DC \( i \) per slot \( t \).
- \( C_{t}^{i} \): Energy level in battery \( i \) at the beginning of slot \( t \).
- \( P_{t}^{q} \): Output generation of CG in DC \( i \) per slot \( t \).
- \( R_{t}^{q} \): Auxiliary variable for market power trading.
- \( x, \bar{x} \): Matrix collecting all primal variables and its running average.
- \( \lambda_{t,q}^{i}, \psi_{t}^{i}, \pi_{t}^{i} \): Lagrange multipliers.
- \( \varpi \): Vectors collecting all Lagrange multipliers.

I. INTRODUCTION

In the new era of big data analytics, cloud computing, and the Internet of Things, data centers are proliferating globally to provide important Internet services such as instant messaging, video distribution, and data backup. For the purposes of reliability and quality-of-service (QoS), a cloud service provider usually has multiple data centers (DCs) geographically distributed in different areas. For instance, Google currently operates seven DCs in the US, and fourteen all over the world [1]. Along with the ever-increasing demand of Internet applications, energy-consuming DCs incur a huge amount of electricity bills. According to [2], DCs among the US consumed about 91 billion kWh electricity in 2013, which is almost twice the amount of power needed by all the households in New York City. Such a consumption is projected to reach 140 billion kWh by 2020.

In order to reduce the electricity cost, considerable efforts from both industry and academia have been made over the last decade [3]. One typical direction is to reduce the energy usage in DCs, through e.g., dynamic thermal management, server speed scaling, and dynamic resizing [4]–[6]. Another line of research aims at information technology (IT) workload management by exploiting the spatio-temporal diversity of energy prices [7], [8]. Specifically, [7] proposed a stochastic optimization based approach to workload routing and server management for geo-distributed DCs, while [8] investigated the means of reducing a single DC’s peak loads by shifting the batch workloads. To deal with the uncertainties of IT workloads and electricity prices, [9] advocated hedging techniques in the financial management to smooth the associated dynamics.

The carbon dioxide emission along with the large energy consumption will also become a major concern. For instance, Google’s DCs in 2010 approximately emitted carbon comparable to that of 280,000 cars [10]. Hence, not only achieving energy efficiency is of essence, but also the sustainability of DCs should be targeted [11], [12]. Exploiting renewable energy sources is clearly a key for sustainable operation of DCs [13]. The main challenge of renewable-integrated energy management is to account for its random and nondispatchable nature, which motivates the use of energy storage units [14]. Energy supply side management with distributed storage units was considered for a homogeneous DC [15], and geo-distributed DCs [16]. Taking advantage of RES, a two-timescale Lyapunov optimization technique was developed to control the energy supply in both ahead-of-time and real-time settings [14]. However, spatial diversity of workloads, prices, and renewables was not utilized therein. Furthermore, in [14]–[16] the server model was simplified while the cooling power consumption was ignored. Yet, a substantial amount of DC energy actually goes to a variety of cooling facilities [17], [18].

Existing approaches of dealing with RES uncertainty include the Lyapunov optimization based on the stochastic approximation technique [7], [14]–[16], and the scenario-based stochastic optimization approaches using RES samples from historical data or a given distribution [19], [20]. To guarantee the convergence and optimality, these methods typically require the condition of independent and identically distributed (i.i.d.) RES samples, which is generally unrealistic [21], [22]. For a large time scale, the distribution of the RES generation changes with the seasons. For a short time scale, hourly RES generation can be highly correlated. How to properly handle the RES uncertainty for DC daily operations is still a challenging problem.

In this paper, we consider the optimal workload and energy management for a cloud network consisting of multiple geo-distributed mapping nodes (MNs) and DCs. For each DC, practical models of power supply, cooling, and IT systems are considered. In particular, the power supply system comprises multiple energy sources; the cooling system combines two different cooling sources with time-varying cooling coefficients1; and the IT operating system is able to adaptively activate servers and schedule workloads with guaranteed hard deadlines. Distinct from existing works, a deterministic uncertainty set of the unknown renewable generation, as well as a two-way energy trading mechanism is introduced to account for the stochastic and nondispatchable nature of RES. The proposed uncertainty set of the RES generation only requires easy-to-obtain first-order statistics, and allows the possible sample correlation. Control parameters are further designed to trade off robustness for conservatism of the robust optimization

1Cooling coefficient is the power consumed for cooling divided by the IT demand, thus representing the cooling efficiency.
formulation.

Built on practical models, the resource allocation task is formulated as a robust optimization problem, which minimizes the system’s worst-case net cost subject to DCs’ operational constraints. Leveraging the problem structure, we show that it can be cast as a convex program. Capitalizing on the dual decomposition approach, an efficient distributed solver is developed. It is shown that the proposed algorithm is guaranteed to obtain the desired robust workload and energy-management strategy, and could also facilitate distributed implementations among the MNs and DCs. Finally, extensive numerical results with real data corroborate the merits of the proposed framework and approaches.

The rest of the paper is organized as follows. The system models are described in Section II. The proposed robust management scheme is developed in Section III. Numerical results are provided in Section IV, followed by conclusions in Section V.

Notation. Boldface lower (upper) case letters represent column vectors (matrices); $(\cdot)^\prime$ stands for vector (matrix) transpose. $\mathbb{R}$ represents the set of real numbers; $r > 0$ denotes entry-wise vector inequality; and $[a]^+ := \max\{a, 0\}$. Finally, the indicator function $\mathbb{1}_{\{A\}}$ takes value 1 when the event $A$ happens and 0 otherwise.

II. SYSTEM MODELS

Consider a network with geographically distributed MNs $\mathcal{J} := \{1, 2, \ldots, J\}$ and DCs $\mathcal{I} := \{1, 2, \ldots, I\}$: see Fig. 1(a). MNs first collect data requests from nearby areas, and then distribute them to different DCs. As shown in Fig. 1(b), each DC has three subsystems: an IT subsystem, a cooling (heat dissipation) subsystem, and a power supply subsystem.

A. Network and workload models

In general, DC workloads are either delay-sensitive (interactive) or delay-tolerant [7]. The interactive workloads entailing real-time user requests must be attended to immediately; e.g., instant messages and voice services. In contrast, delay-tolerant real-time user requests must be attended to immediately; e.g., instant messages and voice services. In contrast, delay-tolerant workloads can be adaptive to the time-varying energy prices and could also facilitate distributed implementations among the MNs and DCs. Finally, extensive numerical results with real data corroborate the merits of the proposed framework and approaches.

The rest of the paper is organized as follows. The system models are described in Section II. The proposed robust management scheme is developed in Section III. Numerical results are provided in Section IV, followed by conclusions in Section V.

Notation. Boldface lower (upper) case letters represent column vectors (matrices); $(\cdot)^\prime$ stands for vector (matrix) transpose. $\mathbb{R}$ represents the set of real numbers; $r > 0$ denotes entry-wise vector inequality; and $[a]^+ := \max\{a, 0\}$. Finally, the indicator function $\mathbb{1}_{\{A\}}$ takes value 1 when the event $A$ happens and 0 otherwise.

II. SYSTEM MODELS

Consider a network with geographically distributed MNs $\mathcal{J} := \{1, 2, \ldots, J\}$ and DCs $\mathcal{I} := \{1, 2, \ldots, I\}$: see Fig. 1(a). MNs first collect data requests from nearby areas, and then distribute them to different DCs. As shown in Fig. 1(b), each DC has three subsystems: an IT subsystem, a cooling (heat dissipation) subsystem, and a power supply subsystem.

A. Network and workload models

In general, DC workloads are either delay-sensitive (interactive) or delay-tolerant [7]. The interactive workloads entailing real-time user requests must be attended to immediately; e.g., instant messages and voice services. In contrast, delay-tolerant workloads such as system updates and data backups are deferrable within a given time interval. This flexibility of delay-tolerant loads enables opportunistic workload management that can be adaptive to the time-varying energy prices and renewables.

Consider a discrete-time scheduling horizon $\mathcal{T} := \{1, \ldots, T\}$. For interactive workloads, let $a_{ij}^t$ denote the number of jobs collected by node $j$ to DC $i$ over slot $t$. For delay-tolerant workloads, let $q_{ij}$ denote the jobs collected by node $j$, and $Q := \bigcup_{j=1}^J Q_j$ with $Q_i \cap Q_j = \emptyset, \forall i \neq j$, representing the set of all delay-tolerant jobs. The $q_{ij}$ delay-tolerant job can be specified by its set $S_q$ and active interval $T_q := \{S_q, E_q\}$, with $S_q$ and $E_q$ denoting the start- and end-time slots. Let $b_{ij,q}$ denote the amount of $q_{ij}$ delay-tolerant job routed from $i$ to DC $i$, and $b_{ij,q}$ the amount processed by DC $i$ over slot $t$, respectively; and $L_{ij,q}$ denote the link bandwidth from node $j$ to DC $i$ at time $t$. As shown in Fig. 2, these quantities must satisfy the following constraints:

$$\sum_{i=1}^I a_{ij}^t = A_j^t, \forall j \in \mathcal{J}, t \in \mathcal{T}$$

$$\sum_{t=S_q}^T \sum_{q=1}^Q b_{ij,q}^t = B_q, \forall q \in Q$$

$$a_{ij}^t + \sum_{q=1}^Q b_{ij,q}^t \leq L_{ij,q}, \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}$$

where (1) ensures that interactive workloads are dispatched once arrived; (2) requires routing each delay-tolerant job before its deadline; and (3) captures the bandwidth limitation of data transfer. Clearly, $L_{ij,q} = 0$ if load transfer from node $j$ to DC $i$ is prohibited; e.g., when MN $j$ is not physically linked with DC $i$.

In each DC, interactive workloads are processed immediately while delay-tolerant workloads are deferrable. The unserved portion of delay-tolerant workloads are buffered in separate queues obeying the following dynamic recursions

$$Q_{ij,q}^{t+1} = Q_{ij,q}^t - b_{ij,q}^t + b_{ij,q}^{t+1}, \forall i \in \mathcal{I}, t \in T_q, q \in Q$$

where $Q_{ij,q}^t$ is the queue length of $q_{ij}$ delay-tolerant job in DC $i$ at the beginning of slot $t$. For the deadline completion requirements, queue length $Q_{ij,q}^t$ must satisfy

$$Q_{ij,q}^t \geq 0, \quad Q_{ij,q}^{S_q} = Q_{ij,q}^{E_q+1} = 0.$$
The total IT demand of DC $i$ in slot $t$, is thus given by

$$d_i^t = \sum_{j=1}^J a_{ij}^t + \sum_{q \in Q} b_{i,q}^t, \quad \forall t \in T. \quad (6)$$

The aforementioned IT system mainly deals with the DC network workload balancing tasks by exploiting the heterogeneous server and bandwidth resources in the cloud. The control variables therein are \{\(d_i^t, a_{ij}^t, b_{i,q}^t\), \(b_{i,q}^t\)\}, under constraints (1)–(6). In addition, the IT system also closely connects with the underlying power infrastructure through a power supply and demand relationship, which is instructive to detail in the next two subsections.

### B. Power demand model

The power consumption of a DC generally comes from various sources, but mainly from the running servers and cooling systems [2].

Suppose that each DC $i$ has a set of $M_i$ homogeneous servers, so the number of active servers $m_i^t$ at time $t$ should be in the range $M_i \leq m_i^t \leq M_i$, where $M_i$ stands for the minimum number of servers required for providing basic services. The consumption of each server can be generally modeled as a function of its running speed [23]

$$P_{i,a}(s_i^t) = \mathcal{P}_{i,a}(\varrho(s_i^t)\upsilon + 1 - \varrho), \quad \varrho \in [0, 1]$$

where $\mathcal{P}_{i,a}$ denotes the peak power consumption of a server in DC $i$; $s_i^t \in [0, 1]$ is its actual speed (a.k.a. CPU usage); and parameter $\upsilon$ is typically around 2 for state-of-the-art servers [5]. Clearly, the fraction of peak consumption $1 - \varrho$ denotes the power consumed in the idle state (i.e., $s_i^t = 0$). When the server is in its highest speed $s_i^t = 1$, the actual consumption is $\mathcal{P}_{i,a}$.

Due to the convexity of $P_{i,a}(s_i^t)$, it readily follows that given a total IT demand $d_i^t$, uniform allocation of the workloads to each server is most energy efficient [7]. Accordingly, each server is running at a speed $d_i^t/(m_i^t D_i)$ with $D_i$ denoting the server capacity under the required service level agreement (SLA), and the total power consumption in DC $i$ becomes

$$P_{i,IT}(d_i^t, m_i^t) = \frac{\varrho \mathcal{P}_{i,a} d_i^t}{m_i^t D_i} + (1 - \varrho) \mathcal{P}_{i,a} m_i^t. \quad (7)$$

Since the number of servers is very large, $m_i^t$ can be relaxed to be a positive real number for simplicity [6].

Along with the increasing density of IT equipment in DCs, a considerable amount of electricity is consumed by the cooling system that generally operates in two modes [4], [24]: outside-air and chilled-water cooling.

The energy usage of outside-air cooling is mainly the power consumed by blowers, which can be approximated as a cubic function of the blower speed [25]. As the blower speed under tight control is proportional to the IT consumption $P_{i,IT}$, the outside-air cooling power consumption can be modeled as a convex function of $P_{i,IT}$, namely

$$F_{i,a}(P_{i,IT}) = \kappa_i^3(P_{i,IT})^3, \quad 0 \leq P_{i,IT} \leq \mathcal{T}_{i,a}^3 \quad (8)$$

where $\kappa_i > 0$ depends on the temperature difference between the (hot) exhausting air temperature $T_{i,RDA}^t$ from the IT racks and the outside air temperature $T_{i,OA}^t$ around DC $i$ at time slot $t$. The maximum capacity of outside-air cooling in (8) can be modeled as $\mathcal{T}_{i,a}^3 = C(T_{i,RDA}^t - T_{i,OA}^t)$, with $C > 0$ proportional to the maximal outside air flow rate.

The chilled-water cooling model here is established on the actual measurement of a practical chiller for which the power consumption can be approximated as [26]

$$F_{i,w}(P_{i,IT}) = \gamma P_{i,IT} \quad (9)$$

where $P_{i,IT}$ is again the IT power consumption in (7) and $\gamma > 0$ is a constant depending on the specific chiller characteristics.

Due to different cooling efficiencies and capacities of the two approaches, for a given $P_{i,IT}$, there is an optimal allocation between outside-air cooling and chilled cooling. Let $P_{i,a}^t$ and $P_{i,w}^t$ denote the amounts of IT power consumption allocated for chilled and outside-air cooling, respectively. With (8) and (9), the optimal cooling power consumption is [24]

$$F_i^*(P_{i,IT}) = \min_{0 \leq P_{i,a}^t \leq \mathcal{T}_{i,a}^3, 0 \leq P_{i,w}^t \leq \mathcal{T}_{i,w}^3} F_{i,a}(P_{i,IT}) + F_{i,w}(P_{i,IT})$$

With a temperature-dependent threshold $P_{i,TH}^t := \min\{\mathcal{T}_{i,a}^3, \sqrt{3\mathcal{T}_{i,a}^3}\}$, (10) admits a closed-form solution

$$F_i^*(P_{i,IT}) = \begin{cases} \kappa_i^3(P_{i,IT})^3, & P_{i,IT} \leq P_{i,TH}^t \\ \kappa_i^3(P_{i,TH}^t)^3 + \gamma(P_{i,IT} - P_{i,TH}^t), & \text{otherwise}. \end{cases}$$

For notational convenience, let $P_i^t(\{d_i^t, \{m_i^t\}\})$ include the server and cooling consumptions as $P_i^t(\{d_i^t, \{m_i^t\}\}) := F_i^t(\{d_i^t, \{m_i^t\}\}) + P_{i,IT}(\{d_i^t, \{m_i^t\}\})$, and consider the following lemma.

**Lemma 1.** Function $P_i^t(\{d_i^t, \{m_i^t\}\})$ is jointly convex in $\{d_i^t, \{m_i^t\}\}$.

**Proof:** See Appendix A. 

### C. Power supply model

A rapidly increasing use of microgrids characterizes the transformative change from our aging power grid to a smart grid over the last decade [27]. While the traditional geographical workload balancing operates separately from the local power balancing, recently advocated dynamic pricing and demand response programs motivate the interactions between them. In this context, we consider each DC to be supplied by a RES-integrated microgrid consisting of a conventional generator (CG) (e.g., a fuel generator), an on-site renewable
Solar generation (4 PM)  

where bounded by the following so-termed ramping constraints:

\[
0 \leq P_{i,g}^t - P_{i,g}^{t-1} \leq \Delta_{i,g}^{up}, \quad P_{i,g}^{t-1} - P_{i,g}^t \leq \Delta_{i,g}^{low}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}.
\]  

The change of CG energy output in two consecutive slots is bounded by the following so-termed ramping constraints:

\[
P_{i,g}^t - P_{i,g}^{t-1} \leq R_{i,g}^{up}, \quad P_{i,g}^{t-1} - P_{i,g}^t \leq R_{i,g}^{low}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]  

where \( R_{i,g}^{up} \) and \( R_{i,g}^{low} \) are the ramping-up and ramping-down rates of CG at DC \( i \).

Consider now the RES vector \( \mathbf{e}_i := [E_i^1, \ldots, E_i^T]' \) generated at DC \( i \) across all slots. Due to the unpredictable and intermittent nature of RES, \( \mathbf{e}_i \) is unknown a priori. In general, uncertain quantities can be modeled by postulating either an underlying probability distribution or an uncertainty region. Probability distributions (possibly mixed discrete/continuous) of the RES generation are seldom available in practice. Although (non-)parametric approaches can be used to learn these distributions, the processes can be very complicated due to the spatio-temporal correlations incurred by various meteorological factors [28]. On the contrary, the proposed method of postulating an uncertainty region provides the decision maker with ranges instead of point forecasts, which is essentially a distribution-free deterministic set and robust to prediction errors.

The actual RES generation \( \mathbf{e}_i \) is assumed to lie in an uncertainty set \( \mathcal{E}_i \), which can be obtained via forecasting or inference using historical measurements. In particular, the following polyhedral uncertainty set is considered (see also [29], [30])

\[
\mathcal{E}_i := \left\{ \mathbf{e}_i \mid E_i^t \leq E_i^t \leq \overline{E}_i^t, \; \mathcal{T} = \bigcup_{s=1}^{S} \mathcal{T}_{i,s}, \; \Delta_{i,s}^{low} E_{\mathcal{T}_{i,s}} + (1 - \Delta_{i,s}^{low}) E_{\mathcal{T}_{i,s}}^{\text{avg}} \leq \sum_{t \in \mathcal{T}_{i,s}} E_t^i \leq (1 - \Delta_{i,s}^{up}) E_{\mathcal{T}_{i,s}}^{\text{avg}} + \Delta_{i,s}^{up} E_{\mathcal{T}_{i,s}} \right\}
\]  

where \( E_i^t \) (\( \overline{E}_i^t \)) denotes the lower (upper) bound on the actual \( E_i^t \); \( \mathcal{T} \) is partitioned into consecutive but non-overlapping sub-horizons \( \mathcal{T}_{i,s}, \; s = 1, \ldots, S; E_{\mathcal{T}_{i,s}}^{\text{avg}} \) is the sample average of total renewables \( \sum_{t \in \mathcal{T}_{i,s}} E_t^i \); the total renewables over \( \mathcal{T}_{i,s} \) are bounded by \( E_{\mathcal{T}_{i,s}}^i \) and \( E_{\mathcal{T}_{i,s}}^{\text{avg}} \); and the parameter \( \Delta_{i,s}^{low}(\Delta_{i,s}^{up}) \in [0,1] \) represents the level of robustness for the lower (upper) bound of the \( s \)th sub-horizon. Note that all the aforementioned statistics can be directly obtained using real RES generations from public sources.

Clearly, parameters \( \Delta_{i,s}^{low}, \; \Delta_{i,s}^{up} \) trade off robustness for conservatism of the resultant solutions. Based on the CAISO solar generations during Mar. 1 – Oct. 30, 2012 [31], examples are given in Fig. 3 for \( |\mathcal{T}_{i,s}| = 2 \); e.g., each sub-horizon consists of two time slots. When the robustness levels \( \Delta_{i,s}^{low} \) and \( \Delta_{i,s}^{up} \) are high, the uncertainty set is large, which includes most of historical samples thus usually leading to conservative solutions when the degree of uncertainty is high; if \( \Delta_{i,s}^{low} \) and \( \Delta_{i,s}^{up} \) are low, some of samples are excluded from the uncertainty set, which reduces the robustness of the resultant solutions.

Here it is also instructive to point out that by further capturing the maximum variation of RES over two consecutive slots \( \partial E_{\mathcal{T}_{i,s}}^i \), a more accurate polyhedral set can be written as (see Fig. 3(d))

\[
\mathcal{E}_i := \left\{ \mathbf{e}_i \mid E_i^t \leq E_i^t \leq \overline{E}_i^t, \; |E_i^t - E_i^{t-1}| \leq \partial E_{\mathcal{T}_{i,s}}^i, \; \mathcal{T} = \bigcup_{s=1}^{S} \mathcal{T}_{i,s}, \; \Delta_{i,s}^{low} E_{\mathcal{T}_{i,s}} + (1 - \Delta_{i,s}^{low}) E_{\mathcal{T}_{i,s}}^{\text{avg}} \leq \sum_{t \in \mathcal{T}_{i,s}} E_t^i \leq (1 - \Delta_{i,s}^{up}) E_{\mathcal{T}_{i,s}}^{\text{avg}} + \Delta_{i,s}^{up} E_{\mathcal{T}_{i,s}} \right\}.
\]  

This modification allows a decision maker to reduce the uncertainty region, without missing many potential samples. It makes sense intuitively since in a short time scale (e.g., 15 min, or an hour), the RES are highly correlated over successive slots. Although the aforementioned practical models only capture RES uncertainty across the scheduling horizons per DC, our proposed approach could be easily extended to include joint spatio-temporal uncertainty models.

To mitigate the variability of RES, energy storage devices are recently considered so as to store the surplus renewables for later shortage [32]. We consider a storage unit with finite capacity \( \bar{C}_u \), and let \( C_{0}^u \) and \( C_{1}^u \) denote the initial energy level of the storage unit in DC \( i \) at the beginning of time slot \( t \). Since storage devices become unreliable with high depth-of-
discharge (DoD), a nonzero minimum level $C_i$ can avoid high DoD. Such a level could also support the DC operation in the event of a grid outage. Let $P_{i,\text{ch}}^t$ and $P_{i,\text{dis}}^t$ denote the amounts of power charging and discharging the storage unit (battery) in DC $i$ at slot $t$. Due to AC/DC power conversion during the (dis-)charging process, the power conversion losses need to be accounted for by the (dis-)charging efficiency $\delta \in (0,1]$. In addition, dissipation losses due to energy leakages are captured by the efficiency coefficient $\eta \in (0,1]$, which renders a decreasing energy level even if there is no (dis-)charging operation. In short, the energy storage unit can be compactly described as
\[
C_i^{t+1} = C_i^t + \eta P_{i,\text{ch}}^t - \delta P_{i,\text{dis}}^t, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
\[
C_i \leq C_i^t \leq \overline{C}_i, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
\[
0 \leq P_{i,\text{dis}}^t \leq P_{i,\text{dis}}^t; 0 \leq P_{i,\text{ch}}^t \leq P_{i,\text{ch}}^t, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
where the bounds $P_{i,\text{dis}}^t$ and $P_{i,\text{ch}}^t$ on (dis-)charging amounts are dictated by physical limits.

Let $P_{i,\text{out}}^t$ denote the total energy consumption of DC $i$ per slot $t$ including the IT operating consumption, cooling power consumption, and battery charged power; that is, [cf. Lemma 1]
\[
P_{i,\text{out}}^t = P_{i}^t + P_{i,\text{ch}}^t.
\]
Likewise, the total energy supply $P_{i,\text{in}}^t$ in DC $i$ per slot $t$ is given by
\[
P_{i,\text{in}}^t = P_{i,g}^t + E_t^i + P_{i,\text{dis}}^t.
\]
Under constraints (11)-(19), the power supply optimization variables are CG and battery power amounts \{$P_{i,g}^t$, $P_{i,\text{in}}^t$, $P_{i,\text{dis}}^t$, $C_i^t$\}.

D. Cost-revenue model

In addition to the internal energy resources (namely, CG, RG, storage unit), DCs can resort to the main grid market in an on-demand manner. With a two-way energy trading facility, each DC can buy energy from external energy markets in the case of a deficit ($P_{i,\text{out}}^t > P_{i,\text{in}}^t$), or, sell energy to the markets in the case of a surplus ($P_{i,\text{out}}^t < P_{i,\text{in}}^t$). Clearly, the shortage energy that needs to be purchased by the DC is $[P_{i,\text{out}}^t - P_{i,\text{in}}^t]^+$; while the surplus energy that can be sold is $[P_{i,\text{in}}^t - P_{i,\text{out}}^t]^+$. Note that both the shortage and surplus energies are non-negative, and at most one of them is positive at any slot $t$.

Suppose that the energy can be purchased from the wholesale electricity market around DC $i$ in period $t$ at price $\alpha_i$, while the energy is sold at price $\beta_i$.

Notwithstanding, we shall always set $\alpha_i \geq \beta_i$ to avoid less relevant buy-and-sell activities of the DC for profit. For DC $i$, the worst-case transaction cost for the whole scheduling horizon is defined as
\[
G_i(\{P_{i,\text{out}}^t\}, \{P_{i,\text{in}}^t\}) := \max_{E_t \in E} \sum_{t=1}^{T} \alpha_i [P_{i,\text{out}}^t - P_{i,\text{in}}^t]^+ - \beta_i [P_{i,\text{in}}^t - P_{i,\text{out}}^t]^+.
\]

which is the point-wise maximum over any realization of the random RES generation in the uncertainty set.

In addition, let function $G_{C_i}(P_{i,g}^t)$ denote the cost of CG at DC $i$ in slot $t$, which is convex piecewise linear or smooth quadratic. The revenue considered here comes from processing delay-tolerant workloads. Specifically, for the $q$th job, the revenue earned per slot $t$ can be generally modeled as a concave function $U_q^t(\beta_{i,q}^t)$, which reflects the diminishing marginal sensitivity of end users to the increasing gains.

III. ROBUST WORKLOAD AND ENERGY MANAGEMENT

Based on the practical models in Section II, we pursue in this section a robust workload and energy management approach for the considered DC network. Over the scheduling horizon $\mathcal{T}$, the system operator per MN performs an (e.g. hour-) ahead-of-time schedule to optimize workloads routing \{$a_{ji}^t, b_{i,q}^t$\}, while the system operator in each DC optimizes servers and workloads scheduling \{$m_i^t, b_{i,q}^t$\}, CG generation \{$P_{i,g}^t$\}, and battery (dis-)charging energy \{$P_{i,\text{ch}}^t, P_{i,\text{dis}}^t$\}. The optimal management strategy minimizes the worst-case net cost $\tilde{\Psi}$, which includes the worst-case transaction cost, the CG cost and the revenue of delay-tolerant workloads, subject to DC operating and power supply constraints. Note that the worst-case net cost here is the maximum net cost for any realization of the random RES generation in the uncertainty set.

With $\tilde{x}$ collecting all the optimization variables \{$a_{ji}^t, b_{i,q}^t, b_{i,q}^t, d_i^t, m_i^t, b_{i,q}^t, P_{i,g}^t, P_{i,\text{ch}}^t, P_{i,\text{dis}}^t, C_i^t$\}, we wish to solve
\[
\tilde{\Psi} := \min_{\tilde{x}} \sum_{i=1}^{I} \sum_{t=1}^{T} \left( G_i(\{P_{i,\text{out}}^t\}, \{P_{i,\text{in}}^t\}) + \sum_{t=1}^{T} \sum_{q=1}^{Q} \left( G_{C_i}(P_{i,g}^t) - \sum_{q=1}^{Q} U_q^t(\beta_{i,q}^t) \right) \right)
\]
subject to:
\[
C_i^{t+1} = C_i^t + \eta P_{i,\text{ch}}^t - \delta P_{i,\text{dis}}^t, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
\[
C_i \leq C_i^t \leq \overline{C}_i, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
\[
0 \leq P_{i,\text{dis}}^t \leq P_{i,\text{dis}}^t; 0 \leq P_{i,\text{ch}}^t \leq P_{i,\text{ch}}^t, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
\[
0 \leq P_{i,\text{g}}^t \leq P_{i,\text{g}}^t, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}
\]
\[
0 \leq d_i^t \leq \sum_{j \in J} \sum_{i \in I} \sum_{t=1}^{T} a_{ji}^t = A_j^t, \quad \forall j \in J, i \in I, t \in \mathcal{T}
\]
\[
\sum_{i=1}^{I} \sum_{t=1}^{T} \tilde{b}_{i,q}^t = B_q, \quad \forall q \in \mathcal{Q}
\]
variables

A. Convex reformulation

Next, a reformulation relying on epigraph-based relaxation is pursued. Since (21) is nonsmooth and nonconvex, which is hard to be handled; thus, (21) contains only deterministic variables. CHEN

\[ Q_{i,q}^{t+1} = Q_{i,q}^t - b_{i,q}^t + \hat{b}_{i,q}^t, \forall i \in \mathcal{I}, t \in \mathcal{T}, q \in \mathcal{Q} \quad (21n) \]

\[ P_{i,\text{out}}^t = P_{i,\text{in}}^t(d_{i,\text{out}}^t, m_{i,\text{out}}^t) + P_{i,\text{ch}}^t, \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (21o) \]

\[ P_{i,\text{in}}^t = P_{i,\text{in}}^t + E_{i}^t + P_{i,\text{dis}}^t, \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (21p) \]

\[ Q_{i,q}^t \in \mathcal{Q}_{i,q} = 0, \forall i \in \mathcal{I}, q \in \mathcal{Q}, t \in \mathcal{T}_q \quad (21q) \]

\[ d_i^t = \sum_{j=1}^J a_{i,j}^t + \sum_{q \in \mathcal{Q}} b_{i,q}^t, \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (21r) \]

\[ \leq B_i, \forall t \in \mathcal{T}, b_{i,q}^t = 0, \forall i \in \mathcal{I}, q \in \mathcal{Q}, t \notin \mathcal{T}_q \quad (21s) \]

\[ \leq \tau_{i,q}, \forall t \in \mathcal{T}, \hat{b}_{i,q}^t = 0, \forall i \in \mathcal{I}, q \in \mathcal{Q}, t \notin \mathcal{T}_q \quad (21u) \]

Convexity of the worst-case net cost \( \Psi \) is established in the following proposition.

**Proposition 1.** If \( \alpha_i^t \geq \beta_i^t \) holds for all \( i \in \mathcal{I} \) and \( t \in \mathcal{T} \), then RWEM problem (23) is convex and strong duality holds.

**Proof:** See Appendix B.

With \( \mathbf{x}^* \) and \( \mathbf{x}^* \) denoting the optimal solutions for (21) and (23), we arrive at the following claim.

**Proposition 2.** Problem (21) is equivalent to problem (23) in the sense that \( \Psi^* = \hat{\Psi}^* \) and \( \mathbf{x}^* = \mathbf{x}^* \).

**Proof:** See Appendix C.

B. Lagrange dual approach

Notice that constraints (23c)–(23e) couple variables across MNs, DCs, workloads, and the RES, so a system operator over the entire network is essential to collect all the information and solve the problem in a centralized way, which may not be feasible in an Internet-scale network [34]. However, since (23) is a convex problem [cf. Proposition 1], a Lagrange dual approach can be developed to efficiently find its optimal dual solution with zero duality gap in a decentralized manner [35]. Let \( \{\pi_i^t\}, \{\lambda_i^t\} \) and \( \{v_i^t\} \) denote the Lagrange multipliers associated with the constraints (23c)–(23e). For notational convenience, let \( \lambda_i^t = 0, \forall i, t \in \mathcal{I}, t \notin \mathcal{T}_q \). And with \( \varpi \) collecting all the Lagrange multipliers, the partial Lagrangian function of (23) is

\[ L(\mathbf{x}, \varpi) := \sum_{i=1}^l \left[ G_i(\{R_i^t\}) + \sum_{t=1}^T G_{C_i}(P_{i,g}^t) - \sum_{q \in \mathcal{Q}} U_i^t(b_{i,q}^t) \right] \]

\[ + \sum_{i=1}^l \sum_{t=1}^T \left[ \sum_{j=1}^J d_{i,j}^t - \sum_{q \in \mathcal{Q}} b_{i,q}^t \right] \]

\[ + \sum_{i=1}^l \sum_{q \in \mathcal{Q}} \sum_{t=1}^T \left[ \sum_{\tau=1}^q v_i^{t,\tau} - \sum_{\tau=1}^q \hat{b}_{i,q}^t \right] \]

\[ + \sum_{i=1}^l \sum_{t=1}^T \left[ \sum_{q \in \mathcal{Q}} \sum_{\tau=1}^q \lambda_i^{t,\tau} \right] \]

\[ \leq \sum_{i=1}^l \sum_{t=1}^T \left[ \sum_{q \in \mathcal{Q}} \sum_{\tau=1}^q \pi_i^{t,\tau} \right] \]

If \( \mathcal{X} \) denotes the set of \( \mathbf{x} \) satisfying constraints (21b)–(21m), and (21s)–(21u), the Lagrange dual function is given by

\[ D(\varpi) := \min_{\mathbf{x} \in \mathcal{X}} L(\mathbf{x}, \varpi) \quad (26) \]

and the dual problem of (23) is

\[ \max \quad D(\{\pi_i^t\}, \{\lambda_i^t\}, \{v_i^t\}) \]

s.t. \( \pi_i^t \geq 0, v_i^t \in \mathbb{R}, \forall i \in \mathcal{I}, t \in \mathcal{T} \)

\[ \lambda_i^{t,\tau} \geq 0, \forall i \in \mathcal{I}, q \in \mathcal{Q}, t \in [S_q, E_q - 1] \]

For the dual problem (27), standard subgradient iterations can be employed to obtain the optimal \( \varpi^* \), namely

\[ \pi_i^t(k+1) = \pi_i^t(k) + \mu \lambda_i^{t,\tau}(k)^+, \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (28a) \]

\[ \lambda_i^{t,\tau}(k+1) = \lambda_i^{t,\tau}(k) + \mu \lambda_i^{t,\tau}(k)^+, \forall i \in \mathcal{I}, \tau \in [S_q, E_q - 1] \]

\[ \nu_i^t(k+1) = \nu_i^t(k) + \mu \lambda_i^{t,\tau}(k)^+, \forall i \in \mathcal{I}, t \in \mathcal{T} \quad (28b) \]
Note that because of the exact relaxation [cf. Proposition 2], subproblem (33) is either an LP or a quadratic program. Hence, the optimal solution can be obtained by existing solvers. And for nonsmooth subproblems (34), a standard subgradient iteration can be employed to obtain the optimal solution as

$$R_i^e(\ell + 1) = R_i^e(\ell) - \mu(\ell)g_{R_i^e}(\ell), \quad \forall \ell \in T$$

where $\ell$ denotes iteration index, and $\{\mu(\ell)\}$ is a non-summable but square-summable stepsize sequence; while the partial subgradient of $G_i(\{R_i^e\})$ with respect to $R_i^e$ is obtained as

$$g_{R_i^e}(\ell) := \frac{\partial G_i(\{R_i^e\})}{\partial R_i^e} = \begin{cases} \alpha_{\ell}^i - \pi_{\ell}^i, & \text{if } R_i^e(\ell) \geq E_i^e(\ell) \\ \beta_{\ell}^i - \pi_{\ell}^i, & \text{if } R_i^e(\ell) < E_i^e(\ell) \end{cases}$$

where $e_i^* := [E_i^1(\ell), \ldots, E_i^{T_e}(\ell)]'$ for the given $\{R_i^e(\ell)\}$ is found using

$$e_i^* \in \arg \max_{e_i \in E_i} \sum_{t=1}^T (\psi_t^i[R_i^e(\ell) - E_i^e] + \phi_t^i[R_i^e(\ell) - E_i^e])$$

It can be seen that the objective function in (37) is convex in $e_i$ under the condition $\alpha_i^\ell \geq \beta_i^\ell, \forall \ell \in T$. However, computing where the maximum of a convex function is attained can be NP-hard, in general. Fortunately, the globally optimal solution is attainable at the extreme points of $E_i$ for convex maximization [36, Sec. 2.4]. Leveraging the polyhedral structure of $E_i$, we utilize an efficient vertex enumerating algorithm to evaluate the objective in (37), and obtain $e_i^*$ directly; see Algorithm 1. Although the number of vertices may increase exponentially with the number of variables and constraints, all the vertices of $E_i$ need be generated only once before running Algorithm 1, which means that Algorithm 1 is computationally affordable. In fact, our simulations in Section IV will corroborate that the vertex generating procedure can be completed within several seconds.

### C. Optimality and distributed implementation

For the subgradient iterations (35), if a diminishing stepsize satisfying (i) $\sum_{\ell=0}^{\infty} \mu(\ell) = \infty$, and (ii) $\sum_{\ell=0}^{\infty} \mu(\ell)^2 < \infty$ is adopted, the sequence (35) converges as $\ell \to \infty$ to the optimal $\{R_i^e(\ell)^*\}$ [36]. As a constant stepsize $\mu$ is used in (28), the
Algorithm 1: Subgradient iteration for solving (34)
1: **Initialize:** Generate all the vertices of the polyhedral uncertainty set $\mathcal{E}_i$; choose a proper $\{R_i^*(0)\}$ and stepsize sequence $\mu(\ell)$
2: **repeat** $\ell = 0, 1, 2, \ldots$
3: Evaluate all the vertices in $\mathcal{E}_i$ and find $e_i^*(\ell)$ in (37)
4: Calculate subgradients via (36)
5: Update $\{R_i^*(\ell)\}$ via (35)
6: **until** Convergence

Algorithm 2: Distributed workload and energy management
1: **Initialize:** Choose a proper $\mathcal{X}(0)$ and stepsize $\mu$
2: **repeat** $k = 0, 1, 2, \ldots$
3: Each DC obtains $\{b_{i,q}^t(k), m_i^t(k), P_{i,ch}(k), P_{i,dis}(k), C_i^t(k), P_{i,ch}(k), d_i^t(k)\}$ by solving (31) and (33)–(34) separately
4: Each MN solves (30) and sends $\{a_{j,i}^t(k), \tilde{b}_{i,q}^t(k)\}$ to each DC
5: DCs update $\mathcal{X}(k)$ via (28) and send them to MNs
6: Run averages to recover primal variables via (38)
7: **until** Convergence

Subgradient iterations will converge to a neighborhood of the optimal solution $\mathcal{X}^*$ [36]. The size of the neighborhood is proportional to the stepsize $\mu$. Since the objective (23a) is not strictly convex, running averages of the primal sequence $\{x(k)\}$ can be used to recover the optimal primal solutions, which are given by

$$\mathcal{X}(k) = \frac{1}{k}x(k-1) + \frac{k-1}{k}\mathcal{X}(k-1), \quad \forall k$$  (38)

where $\mathcal{X}(k)$ is the average of all primal solutions up to iteration $k-1$. Since set $\mathcal{X}$ is convex and $x(i) \in \mathcal{X}, 0 < i < k-1$, it turns out that $\mathcal{X}(k)$ is a feasible point in $\mathcal{X}$. In addition, it can be shown that $\mathcal{X}(k)$ is also asymptotically feasible for primal problem (23) [37].

It is also worth noting that RWEM can afford a distributed implementation, where optimization tasks are distributed among MNs and individual DCs; see Algorithm 2. In RWEM, dual variable updates (28) are all implemented at each DC locally. Subproblem (30) is solved by each MN operator, while each DC operator solves subproblems (31) and (33)–(34). To make these distributed implementations possible, a bidirectional message passing between MNs and DCs is necessary. At every iteration, workload routing variables $\{a_{j,i}^t(k), \tilde{b}_{i,q}^t(k)\}$ are sent from each node to each DC, while the dual variables $\{\lambda_{i,q}^t\}$ and $\{\nu_i^t\}$ are fed back to each MN in turn to solve (30). Note that instead of real-time power and workload schedules, only Lagrange multipliers are sent back to MNs, which could be further leveraged to enhance privacy-preserving operations.

In addition, the worst-case complexity of solving a general convex program is on the order $O(\max\{N_c, N_c\}^4\sqrt{N_c} \log(1/\epsilon))$, where $N_c$, $N_c$ are the total number of variables and constraints, and $\epsilon > 0$ is the given accuracy [38]. Hence, solving subproblems (30)–(31), (33)–(34) in a distributed fashion incurs a markedly lower complexity than directly tackling (23) in a centralized fashion. Faster implementations are possible if we further decentralize (33), and let operators of conventional generation, storage units, delay-tolerant workloads, solve subproblems separately.

### IV. Numerical Evaluation

In this section, results of simulated tests are presented to demonstrate the merits of the proposed approach.

#### A. Experiment setup

The Matlab-based modeling package CVX 2.1 [39] is used to solve the optimization problems involved. The DC network includes 4 DCs and 4 MNs uniformly located in the eastern, central, mountain and western parts of the US. Each DC is connected to a microgrid, of which the power supply parameters are listed in Table I. A polyhedral uncertainty set (13) with a single sub-horizon (no partition, $T_{is} = T$) is considered for the RES. The upper/lower limits $\{\bar{E}_i, \underline{E}_i, \bar{T}_{E_i}, \underline{T}_{E_i}\}$ and the average RES over scheduling horizon $E_{avg}$ were rescaled from the CAISO solar generations during Mar. 1-Oct. 30, 2012; see [31] for detailed description. The upper and lower levels of robustness are set by default to $\Delta_{up} = \Delta_{low} = 1$.

Table II lists the energy purchase prices $\alpha_{ij}$, which are obtained by scaling the hourly electricity prices of the New York City [40]. The selling price is set to $\beta_{ij} = \xi \alpha_{ij}$ with $\xi = 0.6$, while the CG cost is considered as $G_{ci}(P_{i,ch}) = \omega_i P_{i,ch}^2$ with $\omega_i = (1/T) \sum_{t=1}^{T} \alpha_{ij}$.

The total number of servers $M_i$ is set to 80, with common $\varrho = 0.4$, $\bar{P}_{i,ch} = 500$W and $M_i = 5$ for all DC $i$. The cooling parameters are set to $\gamma = 0.2, \kappa_{i} = 2 \times 10^{-9}, \bar{T}_{i,ch} = 30$K, $\forall i \in I, t \in T$. For simplicity, the unit of workloads in this setting is normalized by the server is per-slot capacity $D_{s}$, with common $D_i$ for all DCs. The interactive workload arrival rates at MNs are depicted in Fig. 4, rescaled from the real traffic of Wikipedia [41]. Eight different delay-tolerant
workloads (DWs) are specified in Table III. For job $q$, the revenue function is considered $R_q(b_{ij,q}) = -0.01 u_i^q (b_{ij,q})^2 + u_i^q b_{ij,q}$ $$/unit, where $u_i^q$ is linearly decreasing from $\pi_q$ to $u_q$ across its active interval $[S_q, E_q]$. The bandwidth limits $L_{ji}$ are assumed to be time-invariant, given by

$$L_i = \begin{bmatrix} 90 & 60 & 40 & 40 \\ 40 & 80 & 40 & 50 \\ 50 & 40 & 100 & 30 \\ 40 & 40 & 50 & 90 \end{bmatrix}, \forall t \in T.$$  

Note that the homogenous settings of data centers are considered here in order to exemplify the impact of other factors (e.g., prices and RES).

The time horizon spans $T = 12$ hours, corresponding to the interval 1PM–12AM in Eastern Time Zone. Here we use the Eastern Time Zone for time-keeping, and the real data have been shifted to show the effect of time zone differences. As a result, the peaks of workload demands, RES and prices are different in the four areas, which provide an opportunity for spatio-temporal workload and energy management. Finally, two benchmarks are compared in this setting: A robust local policy allocating all workloads from each MN to its nearest DC, and a non-robust geographical load balancing policy, which predicts the RES at each slot via its sample mean from the historical dataset. Interestingly, our proposed approach can be reduced to the local policy if the bandwidth matrix $L_i$ only has positive diagonal entries, and to the non-robust policy if we set $\Delta_{i}^{up} = 0, \Delta_{i}^{low} = 0, \text{and } |T_{i,s}| = 1$ in (13).

### B. Numerical results

Fig. 5 depicts the evolution of the worst-case net cost for the proposed algorithm, as well as the two alternatives. Within 300 iterations, the proposed algorithm converges to a worst-case net cost 19% lower than that of the non-robust approach, and 51% lower than that of the local policy. Recall that the non-robust approach is sensitive to the RES prediction error, while the local policy can not perform geographical load balancing. In contrast, the proposed RWEM takes advantage of both factors, and purchases less amount of expensive energy from the spot market that results in a smaller worst-case net cost.

To better illustrate this point, sensitivity analysis to the level of robustness $\Delta_{i}^{low}$ is first studied in Fig. 6. As expected, the proposed RWEM outperforms the non-robust approach in...
all cases. Meanwhile, the worst-case net costs of both robust and non-robust scheme grow up as the $\Delta^\text{low}_i$ increases. This makes sense intuitively because a larger $\Delta^\text{low}_i$ implies a bigger uncertainty set [cf. (13)], which will eventually increase the worst-case net cost. Hence, the selection of $\Delta^\text{low}_i$ is critical for various scenarios. While a large $\Delta^\text{low}_i$ guarantees robustness of the resultent solution, a small one can moderately reduce its conservatism. We waive the analysis of $\Delta^\text{up}_i$ here, because it plays a less important role in cost minimization (21), due to the monotonicity of the objective (21a) with respect to $E^t_{i,t}$.

The optimal workload schedules $d^t_{i,t}$ of the proposed RWEM and the local policy are compared in Figs. 7 and 8, respectively. The worst-case renewable generations $\{E^t_{i,t}\}$ are shown to illustrate the principle of geographical workload distribution. Compared with the local policy, the proposed RWEM can intelligently route workloads to a remote DC where the system IT demand is lower, RES availability is higher, or, the local energy price is more affordable. To see this, both interactive and flexible workloads are uniformly routed to each DC in Fig. 7, thus the entire DC network can process more workloads when the RES generations are ample (1PM-6PM). An interesting observation is that even if DC 4 in the Western US enjoys a relatively higher RES, due to its high energy purchase price $\alpha^t_{1}$ [cf. Table II], the MNs are more likely to route workloads to the areas having lower prices, when there is no renewable surplus in DC 4.

In contrast to Fig. 7, the workload schedules in Fig. 8 are more isolated and thus inefficient. Without coordinating all DCs, the local policy cannot “smooth” the IT demand with the additional degree of freedom in space. Specifically, when the system demand is low in DC 1 as well as DC 3 (10-12 PM), no flexible workloads can be scheduled. Likewise, the RES is not fully utilized in DC 2 (1-3PM). This high fluctuation of system workloads will also cause switching on/off servers frequently, thus incurring an implicitly higher wear-and-tear cost.

The optimal power consumption and supply schedules of DC 1 are depicted in Figs. 9 and 10, respectively. The scaled fluctuation of energy purchase price $\alpha^t_{1}$ is also plotted to gain intuition on the optimal power schedules. Clearly, less power is consumed when $\alpha^t_{1}$ is higher (6PM). Using combined cooling sources, the average cooling coefficient of the pro-
posed algorithm is around 0.17, which is more efficient than the simple chilled-water cooling with a constant coefficient $\gamma = 0.2$. Furthermore, with the goal of mitigating the high variability of RES, batteries are encouraged to charge when the worst-case renewable generations are high and the energy prices are low (1PM-4PM). Thus, batteries can be discharged when less renewables are available at night (e.g., 5-7PM in Fig. 10). Likewise, Fig. 11 shows that all the batteries exhibit a similar trend in response to RES and price fluctuations. From the power supply perspective, the lower purchase price $\alpha_1^d$ encourages purchasing more energy from the external grid market, while the peak of $\alpha_1^d$ results in a higher power usage from the CG.

V. CONCLUSIONS

Robust ahead-of-time workload and energy management for green DCs was considered in this paper. Practical models of IT, cooling, and power supply subsystems were first introduced. Taking into account the spatio-temporal variations of workloads, renewables and electricity market prices, a resource allocation problem was formulated to minimize the system net cost including the network operational cost and the worst-case energy transaction cost. Relying on the strong duality of the convex reformulation, a Lagrange dual based distributed solver was developed to yield the optimal solution. Extensive numerical tests further corroborated the effectiveness and merits of the proposed scheme. Without specific assumption on the underlying distribution of the renewable generation process, we offer a new perspective in dealing with the uncertainties involved in operating sustainable DCs. The novel robust optimization framework of the present paper opens up some interesting research directions, such as integrating additional uncertainties from the interactive workloads and real-time energy prices.

APPENDIX

A. Proof of Lemma 1

Function $P_{i,IT}(d_{i,t}^0, m_{i,t}^0)$ is jointly convex in $\{d_{i,t}^0, m_{i,t}^0\}$ because of its quadratic-over-linear form [cf. (7)]. Hence, the composite function $F_1^d(P_{i,IT}(d_{i,t}^0, m_{i,t}^0))$ is jointly convex in $\{d_{i,t}^0, m_{i,t}^0\}$ since $F_1^d(P_{i,IT})$ is convex and nondecreasing [33, Sec. 3.2]; and, so is $P_{i,t}^d(d_{i,t}^0, m_{i,t}^0)$.

B. Proof of Proposition 1

Since the absolute value function is convex, and the operations of nonnegative weighted summation and pointwise maximum preserve convexity, it is easy to see that $G_1(\{R_{i,t}^0\})$ is convex in $\{R_{i,t}^0\}$. In addition, $G_C(P_{i,g}^0)$ is convex in $P_{i,g}^0$ and $U_{i,q}^t(b_{i,q}^t)$ is concave in $b_{i,q}^t$; hence, the objective function (23a) is jointly convex in $\{R_{i,t}^0, P_{i,g}^0, P_{i,ch}^t, P_{i,dis}^t\}$. The constraints except (23c) are linear, while (23c) defines a feasible set which is actually the epigraph of a convex function [cf. Lemma 1]. Since the epigraph of a convex function is a convex set [33, Chap. 3.1], it follows that (23) is a convex problem, and strong duality holds.

C. Proof of Proposition 2

Compared to (21), (21n)–(21q) are replaced by (23d), and (21o)–(21p) are substituted by (23c) to convexify the problem, and facilitate distributed implementation in (23). By summing (21n) from $S_q$ to $t \in \{S_q, \ldots, E_q - 1\}$, we find $Q_{i,q}^t = Q_{i,q}^n - \sum_{t = S_q}^{t} b_{i,q}^t + \sum_{t = S_q}^{t} b_{i,q}^t$. Then, due to (21q), we have that $\sum_{t = S_q}^{t} b_{i,q}^t \geq \sum_{t = S_q}^{t} b_{i,q}^t$. Likewise, we obtain $\sum_{t = E_q}^{t} b_{i,q}^t \geq \sum_{t = S_q}^{t} b_{i,q}^t Q_{i,q}^t$. This establishes the equivalence of (21n)–(21q) with (23d). In addition, since the objective (23a) is monotonically increasing with $R_{i,t}^0$, it is easy to see that (23c) is always binding at the optimal solution $x^*$, which implies that the optimal solution $x^*$ is also an optimal solution (21); and thus, $\Psi^* = \Psi^*$.

REFERENCES


Tianyi Chen received the B. Eng. degree (with highest honors) in Communication Science and Engineering from Fudan University, China, in 2014. Since August 2014, he has been with SPiNCOM, working toward his Ph.D. degree in the Dept. of ECE at the University of Minnesota. His research interests lie in network optimization with applications to green communications, and sustainable data centers. He received the Student Travel Grant from the IEEE Communications Society in 2013, a National Scholarship from China in 2013, and the UMN ECE Department Fellowship in 2014.

Yu Zhang (S’11) received his B.Eng. and M.Sc. degrees (both with highest honors) from Wuhan University of Technology, Wuhan, China, and from Shanghai Jiao Tong University, Shanghai, China, in 2006 and 2010, respectively, and the Ph.D. degree from the University of Minnesota, Minneapolis, MN, USA, in 2015, all in electrical engineering.

He is currently a Postdoctoral Associate with the Department of Electrical and Computer Engineering, University of Minnesota. His research interests span the areas of smart power grids, optimization theory, machine learning, and wireless communications. Current research focuses on distributed energy management with renewables, and energy data analytics. He received the Huawei Scholarship and the Infineon Scholarship from the Shanghai Jiao Tong University (2009), the ECE Dept. Fellowship from the University of Minnesota (2010), and the Student Travel Awards from the SIAM and the IEEE Signal Processing Society (2014).

Xin Wang (SM’09) received the B.Sc. and M.Sc. degrees from Fudan University, Shanghai, China, in 1997 and 2000, respectively, and the Ph.D. degree from Auburn University, Auburn, AL, USA, in 2004, all in electrical engineering.

From September 2004 to August 2006, he was a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. In August 2006, he joined the Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA, as an Assistant Professor and, in August 2010, was promoted and became an Associate Professor. He is currently a Professor with the Department of Communication Science and Engineering, Fudan University. His research interests include stochastic network optimization, energy-efficient communications, cross-layer design, and signal processing for communications. He served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and as an Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.
Georgios B. Giannakis (F’97) received his Diploma in Electrical Engr. from the Ntl. Tech. Univ. of Athens, Greece, 1981. From 1982 to 1986 he was with the Univ. of Southern California (USC), where he received his MSc. in Electrical Engineering, 1983, MSc. 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 385 journal papers, 650 conference papers, 22 book chapters, two edited books and two research monographs (h-index 115). Current research focuses on learning from Big Data, wireless cognitive radios, and network science with applications to social, brain, and power networks with renewables. He is the (co-) inventor of 25 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 Fourier Technical Field Award (2015). 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.