

Optimization Decomposition of Resistive Power Networks with Energy Storage

Xin Lou, *Student Member, IEEE* and Chee Wei Tan, *Senior Member, IEEE*

Abstract—A fundamental challenge of a smart grid is: to what extent can moving energy through space and time be optimized to benefit the power network with large-scale energy storage integration? With energy storage, there is a possibility to generate more energy when the demand is low and store it for later use. In this paper, we study a dynamic optimal power flow problem with energy storage dynamics in purely resistive power networks. By exploiting a recently-discovered zero duality gap property in the optimal power flow (OPF) problem, we apply optimization decomposition techniques to decouple the coupling energy storage constraints and obtain the global optimal solution using distributed message passing algorithms. The decomposition methods offer new interesting insights on the equilibrium load profile smoothing feature over space and time through the relationship between the optimal dual solution in the OPF and the energy storage dynamics. We evaluate the performance of the distributed algorithms in several IEEE benchmark systems and show that they converge fast to the global optimal solution by numerical simulations.

Index Terms— Optimal power flow, energy storage, decomposition method, distributed optimization, smart grid, message passing algorithm.

I. INTRODUCTION

A key challenge in a smart grid design is the integration of power flow control in the power grid network together with the distributed renewable energy sources and energy storage at the endpoints of the network. Renewable energy is intermittent and difficult to predict and harness, and this in turn makes power flow in the power grid network harder to control and optimize. An interconnected system of hundreds of millions of distributed energy resources introduces rapid, large and random fluctuations in power supply and demand. Such deployments create stronger demands and highly variable (in time and space) operating conditions than those experienced in current networks. The integration of a two-way communication infrastructure to connect the power grid network and different types of demands and supply generation (see Figure 1), while promising in a smart grid, motivates the important question: to what extent can moving energy through space and time be optimized to benefit the power network with large-scale storage integration?

There are clearly benefits in a joint optimization of energy storage and the power flow [1], [2]. Energy production and

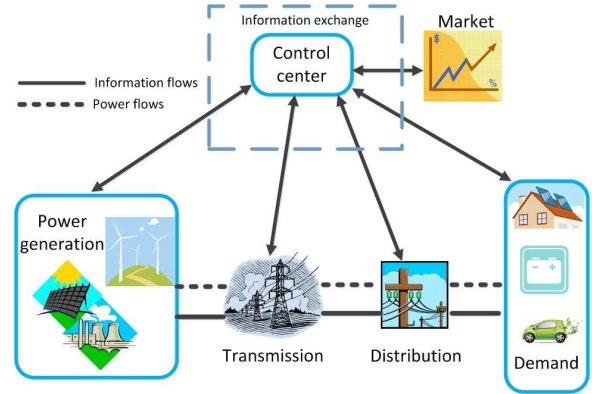


Fig. 1. A general outline of the interconnection in a power grid network with a two-way communication infrastructure for message passing to coordinate the distributed optimization of supply, energy storage and different kinds of workloads with possibly time-varying demand, e.g., plug-in electric vehicles.

consumption can dynamically lead to new equilibrium operating points over time. Energy storage at demand buses can absorb the transients while the power loads are rebalanced. In particular, when the power price is low (due to a lower demand at off-peak hours), energy can be stored in batteries. On the other hand, when the power price is high (due to a higher demand at peak hours), users can first draw energy locally from the batteries before consuming the energy directly from the power grid network. Energy storage can also help in load smoothing [2]. For example, the energy storage battery in plug-in electric vehicles¹ plays an interestingly dual role: they represent a new type of demand load to be satisfied and also as new energy storage resources (when power flows from the vehicle to grid) for load smoothing [3], [4]. With energy storage, how should power flow be optimized to maximize efficiency and minimize power consumption cost?

The optimal power flow (OPF) problem is a classical nonlinear optimization problem, whose solution optimizes a network-wide objective, e.g., generation cost or transmission loss, subject to physical constraints on the power flow, the demand specification and the network connectivity constraints [5], [6]. It is well-known that the OPF is generally hard to solve due to its nonconvexity. In the vast literature on OPF, approximation methodologies, e.g., the direct current (DC) OPF linearization that assumes a constant voltage and small phase angle, have been widely used [7]–[11]. However, recent

Manuscript received Oct 16, 2013; revised Apr 14, 2014. The work in this paper was partially supported by grants from the Research Grants Council of Hong Kong Project No. RGC CityU 122013.

X. Lou and C. W. Tan are with the College of Science and Engineering, City University of Hong Kong, Tat Chee Ave., Hong Kong (email: louxin86@gmail.com and cheewtan@cityu.edu.hk).

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

¹A trend is expected in the growing use of plug-in electric vehicles that use chemical energy stored in rechargeable battery packs and installation of electric vehicle charging infrastructures, and this is one of the driving forces behind a smart power grid network.

developments² have shown that semidefinite programming (SDP) can be used to convexify the OPF and, under certain assumptions, solving this SDP convex relaxation is even exact [14], [15].³ For example, the relaxation is exact for distribution networks that have tree-like network topology [16], [17]. This is important because SDP is a convex optimization problem that can be efficiently solved [18].

These developments have attracted a surge of interests in developing computational algorithms to solve the OPF exactly or to obtain relaxations for the OPF and its special cases [19]–[23]. However, most efforts have been devoted to centralized computation issues (which of course is still useful in the context of a tightly-regulated system) without focusing on energy storage. Indeed, the dynamic OPF with energy storage is a more complex time-dependent power flow optimization problem (also known as the multi-period OPF in the power system literature [24], [25]). The energy storage dynamics have unique time-dependent characteristics that couple with power flow (over time) and the network topology (over space). In addition, the demand profile (over time) and the energy storage dynamics introduce a new degree of freedom to optimize for load smoothing over time that is beneficial to power systems operation.

The authors in [26] first studied such a dynamic OPF with energy storage using the DC OPF linearization and finite horizon optimal control for a simple network topology (a single generator with a single battery connected to a single load). The authors demonstrated a unique feature in the optimal policy in which generation first exceeds demand in order to charge up the battery and subsequently demand exceeds generation with additional energy supplement coming from the battery till the battery is fully depleted by the end of the horizon. The authors in [25] applied the SDP convex relaxation in [15] to a dynamic OPF but did not consider energy storage. Gayme et al [20] solved this dynamic OPF problem with energy storage constraints using the SDP convex relaxation technique in a centralized manner. This however requires a high computational complexity due to solving an SDP of higher dimension (whose complexity increases with the number of periods in the horizon). The authors in [3] proposed valley-filling algorithms (for load smoothing) to tackle the problem in [20] for electric vehicle charging. The authors in [21] proposed message passing algorithms for distributed optimization that can be combined with receding horizon control.

In this paper, we study a dynamic OPF problem for a purely resistive power network (i.e., no phase angle, reactive power variables and reactance parameters) with energy storage dynamics. This allows energy storage dynamics to be optimized jointly with the demand and supply across multiple time periods. Our focus is to study optimization decomposition techniques to enable distributed computation. Decomposition

methods (predominantly Lagrange dual decomposition) have been previously used in the literature to solve the OPF [7], [8], [24], [27], [28], but these prior work use approaches (using the DC OPF linearization to approximate the OPF) different from that in this paper. Decomposition enables fundamental understanding of architectural possibilities, especially in the distributed coordination of functional modules in a large network [29], [30]. We also study the interplay between optimal generation schedule and the time-varying load profile, i.e., how to generate more energy when the demand is low and store it for later use by charging and discharging the energy storage. A resistive power network OPF with energy storage can be practically useful in HVDC (high-voltage direct current) network or microgrid clusters integrated with renewable energy sources, e.g., solar cells and photovoltaic generator, that produce real power flow.

Overall, the contributions of the paper are as follows:

- 1) We formulate a time-dependent dynamic OPF problem for resistive power networks with energy storage dynamics. We leverage a zero duality gap property in the static OPF to solve this dynamic OPF problem.
- 2) We propose various decomposition methodologies (an indirect Lagrange dual-dual and a primal-dual hierarchical decomposition) that lead to different kinds of efficient and distributed algorithms to solve the dynamic OPF.
- 3) We give a physical interpretation to the connection between the dual solution corresponding to the power flow and energy storage as time-varying marginal price and storage price respectively that explains the phenomenon of load profile smoothing over time.
- 4) Through extensive numerical experiments, we show that these algorithms converge fast to the global optimal solution of the dynamic OPF. We also demonstrate the equilibrium load profile smoothing feature of the dynamic OPF with energy storage under various problem settings.

The rest of this paper is organized as follows. In Section II, we introduce the system model and formulate the dynamic OPF with energy storage problem formulation. In Section III, we review some preliminary results on the static OPF problem and its convex relaxation. In Section IV, we propose distributed algorithms to solve the dynamic OPF problem by an indirect dual-dual decomposition technique. In Section V, we propose an alternative primal-dual decomposition. In Section VI, we explain the connection between the Lagrange dual variables and the energy storage dynamics. Numerical evaluations are presented in Section VII. In Section VIII, we conclude the paper.

We refer the readers to Figure 2 for an overview of the overall relationship between the problem formulation, the optimization decomposition, the proposed algorithms and the interpretation of the dynamic OPF with energy storage.

II. OPTIMAL POWER FLOW WITH ENERGY STORAGE

Let us consider a purely resistive power network with buses $\mathcal{N} = \{1, 2, \dots, N\}$ and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ transmission lines. Let

²We refer the readers to [12], [13] for an overview on this development that clarifies the relationships between various models of OPF and its convex relaxation.

³The convex relaxation is exact if the optimal solution of a convex relaxation is feasible for the original OPF problem and hence globally optimal as well. The relaxation gap (as well as the Lagrange duality gap in the OPF considered in this paper) is zero in this case.

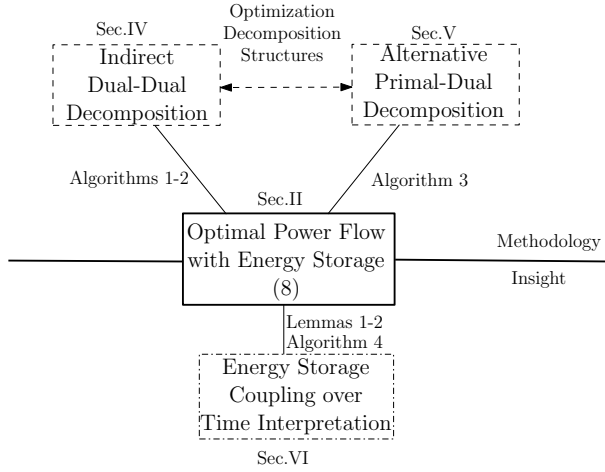


Fig. 2. An overview of the connection on the different optimization decomposition, algorithms and the interpretation of the dynamic OPF with energy storage. The upper half of the line corresponds to the two optimization decomposition techniques, while the lower half corresponds to the insight on how the dual solution of the two decomposition can be interpreted.

us denote the set of generation buses and demand buses by \mathcal{G} and \mathcal{D} respectively. We assume that the bus $i \in \mathcal{D}$ has an energy storage battery attached to it. Let us denote the one-hop neighbors (connected by transmission lines) of bus i by the set Ω_i ($|\Omega_i| \geq 1$). The transmission lines have admittance and this is modeled by the admittance matrix \mathbf{Y} . In particular, \mathbf{Y} is symmetric, i.e., $Y_{ij} = Y_{ji} \in \mathbb{R}_{++}$, if $(i, j) \in \mathcal{E}$ and $Y_{ij} = Y_{ji} = 0$ otherwise [14], [15].

We consider a multi-period power flow system indexed by $t = 1, \dots, T$, where T is the terminating time period. We use $\mathbf{V}(t)$ and $\mathbf{I}(t)$ to denote the column vectors for the voltages $(V_i(t))_{i \in \mathcal{N}}$ and the current $(I_i(t))_{i \in \mathcal{N}}$ for $t = 1, \dots, T$, respectively. By Kirchhoff's Current Law and Ohm's Law, we have $\mathbf{I}(t) = \mathbf{Y}\mathbf{V}(t)$ for $t = 1, \dots, T$, i.e.,

$$\begin{pmatrix} I_1(t) \\ I_2(t) \\ \vdots \\ I_N(t) \end{pmatrix} = \begin{pmatrix} \sum_{j \in \Omega_1} Y_{1j} & \cdots & -Y_{1N} \\ -Y_{21} & \cdots & -Y_{2N} \\ \vdots & \ddots & \vdots \\ -Y_{N1} & \cdots & \sum_{j \in \Omega_N} Y_{Nj} \end{pmatrix} \begin{pmatrix} V_1(t) \\ V_2(t) \\ \vdots \\ V_N(t) \end{pmatrix}. \quad (1)$$

Using $\mathbf{I}(t)$ in (1), the nodal power and voltage constraints are given respectively as follows:

$$\mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t) \leq \bar{p}_i(t) - r_i(t), \quad t = 1, \dots, T, \quad (2)$$

$$\underline{\mathbf{V}} \leq \mathbf{V}(t) \leq \bar{\mathbf{V}}, \quad t = 1, \dots, T, \quad (3)$$

where $(\cdot)^\top$ denotes the transpose, $\mathbf{Y}_i = \frac{1}{2}(\mathbf{E}_i \mathbf{Y} + \mathbf{Y} \mathbf{E}_i)$, $\mathbf{E}_i = \mathbf{e}_i \mathbf{e}_i^\top \in \mathbb{R}^{n \times n}$ with \mathbf{e}_i being the standard basis vector in \mathbb{R}^n (i.e., \mathbf{E}_i is a matrix with a single one in the (i, i) th entry and zero everywhere else), $r_i(t)$ is the amount of power that the energy storage of bus i charges (positive) or discharges (negative) at time t and $\{\underline{\mathbf{V}}, \bar{\mathbf{V}}\}$ are given voltage bound values. For a generation bus $i \in \mathcal{G}$, $\bar{p}_i(t)$ is strictly positive and models the generator capacity. On the other hand, for a demand bus $i \in \mathcal{D}$, $\bar{p}_i(t)$ is strictly negative, and the magnitude $|\bar{p}_i|$ represents the minimum power that has to be provided to satisfy the load at bus i . Similarly to [22], we make the load

over-satisfaction assumption, i.e., a demand bus can absorb the power $|\bar{p}_i|$ to satisfy the basic demand requirement and to charge the attached energy storage battery.

In addition to the basic OPF constraints, each battery at bus $i \in \mathcal{N}$ has to satisfy [20], [26]:⁴

$$b_i(t+1) = \eta_i b_i(t) + r_i(t), \quad t = 1, \dots, T, \quad (4)$$

where storage leakage in a general battery can be modeled by appropriately choosing $0 < \eta_i \leq 1$ for all $i \in \mathcal{N}$ in (4). The initial condition of the battery at $i \in \mathcal{N}$ is given by:

$$b_i(1) = B_i^0, \quad (5)$$

where B_i^0 is a positive constant. At each $i \in \mathcal{N}$, $b_i(t)$ and $r_i(t)$ are constrained respectively by:

$$0 \leq b_i(t) \leq B_i, \quad t = 1, \dots, T+1, \quad (6)$$

$$\underline{r}_i \leq r_i(t) \leq \bar{r}_i, \quad t = 1, \dots, T, \quad (7)$$

where B_i is the battery capacity at bus i . Our objective is to minimize the total transmission loss in the network:

$$\sum_{t=1}^T \mathbf{V}(t)^\top \mathbf{Y} \mathbf{V}(t),$$

and the energy storage cost:

$$h_i(b_i(t), r_i(t)) = h_{1i}(b_i(t)) + h_{2i}(r_i(t)), \quad \forall i \in \mathcal{N}, \quad \forall t,$$

whose first component

$$h_{1i}(b_i(t)) = \alpha_i (B_i - b_i(t)) \text{ for all } i \in \mathcal{N}, t = 1, \dots, T+1,$$

with $\alpha_i \geq 0$ for all $i \in \mathcal{N}$ captures the instantaneous amount of the energy stored in the battery, i.e., the penalty is proportional to the deviation from the storage capacity [26]. The second component $h_{2i}(r_i(t))$ captures the battery dynamics, i.e., the charging/discharging amount of power in the battery. In addition, we assume that $h_{2i}(r_i(t))$ is convex in $r_i(t)$ for all $i \in \mathcal{N}$ and $t = 1, \dots, T$ [31], i.e., the battery cost increases with the charge/discharge rate. In this paper, we let

$$h_{2i}(r_i(t)) = \tilde{\alpha}_i r_i^2(t) \text{ for all } i \in \mathcal{N}, t = 1, \dots, T,$$

with $\tilde{\alpha}_i \geq 0$ for all $i \in \mathcal{N}$.

Consider the following time-dependent dynamic OPF problem:

$$\begin{aligned} & \text{minimize} \sum_{t=1}^T \mathbf{V}(t)^\top \mathbf{Y} \mathbf{V}(t) + \sum_{i \in \mathcal{N}} \sum_{t=1}^{T+1} h_{1i}(b_i(t)) + \sum_{t=1}^T h_{2i}(r_i(t)) \\ & \text{subject to } \mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t) \leq \bar{p}_i(t) - r_i(t), \quad \forall i \in \mathcal{N}, \quad t = 1, \dots, T, \\ & \quad \underline{\mathbf{V}} \leq \mathbf{V}(t) \leq \bar{\mathbf{V}}, \quad t = 1, \dots, T, \\ & \quad b_i(t+1) = \eta_i b_i(t) + r_i(t), \quad i \in \mathcal{N}, \quad t = 1, \dots, T, \\ & \quad \mathbf{b}(1) = \mathbf{B}^0, \\ & \quad \mathbf{0} \leq \mathbf{b}(t) \leq \mathbf{B}, \quad t = 1, \dots, T+1, \\ & \quad \underline{\mathbf{r}} \leq \mathbf{r}(t) \leq \bar{\mathbf{r}}, \quad t = 1, \dots, T, \\ & \text{variables: } \mathbf{V}(t), \mathbf{b}(t), \mathbf{r}(t), \end{aligned} \quad (8)$$

⁴We let the power quantities be in the unit of energy per unit time so that energy and power are used interchangeably [26].

where $\mathbf{b}(t)$ and $\mathbf{r}(t)$ denote the column vectors for $(b_i(t))_{i \in \mathcal{N}}$ and $(r_i(t))_{i \in \mathcal{N}}$, respectively. The problem formulation in (8) is a nonconvex quadratic constrained quadratic programming (QCQP) problem, which is generally hard to solve. In the following, we first review some results in [14], [15], [22] for a static OPF, and then leverage optimization decomposition techniques in Sections IV and V to solve (8).

III. PRELIMINARIES ON CONVEX RELAXATION OF STATIC OPF

Consider a fixed time period in (8) and suppose that there is no battery, then (8) reduces to (omitting the time index t):

$$\begin{aligned} & \text{minimize } \mathbf{V}^\top \mathbf{Y} \mathbf{V} \\ & \text{subject to } \mathbf{V}^\top \mathbf{Y}_i \mathbf{V} \leq \bar{p}_i, \quad \forall i \in \mathcal{N}, \\ & \quad \underline{\mathbf{V}} \leq \mathbf{V} \leq \bar{\mathbf{V}}, \\ & \text{variables: } \mathbf{V}. \end{aligned} \quad (9)$$

Note that (9) is a static OPF problem that minimizes the total power losses in a resistive power network subject to voltage and power constraints. Even though (9) is nonconvex, it has recently been shown in [15] that a convex relaxation using SDP can yield the global optimal solution of (9) under some mild conditions [15]. This remarkable result is a consequence of a special type of quadratic programming problem with nonnegative problem parameters and solution [14], [32]. In particular, the SDP convex relaxation of (9) is given by:

$$\begin{aligned} & \text{minimize } \text{trace}(\mathbf{Y} \mathbf{W}) \\ & \text{subject to } \text{trace}(\mathbf{Y}_i \mathbf{W}) \leq \bar{p}_i, \quad \forall i \in \mathcal{N}, \\ & \quad \underline{V}_i^2 \leq \text{trace}(\mathbf{E}_i \mathbf{W}) \leq \bar{V}_i^2, \quad \forall i \in \mathcal{N}, \\ & \quad \mathbf{W} \succeq 0, \quad W_{ij} \geq 0, \quad \forall i, j \in \mathcal{N}, \\ & \text{variables: } \mathbf{W}, \end{aligned} \quad (10)$$

where $\mathbf{W} \succeq 0$ means that \mathbf{W} is a positive semidefinite matrix and $\text{trace}(\cdot)$ denotes the matrix trace operator. In fact, (10) is equivalent to (9) if the constraint $\text{rank}(\mathbf{W}) = 1$ is included in (10) since $\mathbf{W} = \mathbf{V} \mathbf{V}^\top$. However, by observing that $\mathbf{Y}, \mathbf{Y}_1, \dots, \mathbf{Y}_N$ are matrices with nonpositive off-diagonal elements, the authors in [14], [15] show that (9) can be solved exactly by (10). Furthermore, it has been shown that this tight relaxation result also implies that the Lagrange duality gap of (9) is zero, and (9) can also be solved exactly by a second order cone programming relaxation [12], [13], [22]. A more general result that applies to the AC (alternating current) OPF problem can be found in [12], [13], [15].

In our recent work [22], we leverage this zero duality gap property to design low-complexity distributed algorithms to solve (9) directly instead of solving the SDP convex relaxation in (10) with a centralized SDP solver (e.g., interior-point method). In fact, the optimal solution of (9) can be unique (for different network topologies, e.g., line, radial and mesh networks). In addition, this uniqueness characterization of (9) has implications on how efficient distributed algorithms can be designed to solve (9). The key idea is to exploit the zero duality gap result in [14], [15] and the Poincare-Hopf Index Theorem in [33], [34] to show uniqueness, and then design message passing algorithms to solve (9) using the Lagrange dual decomposition. In another recent work [35], we

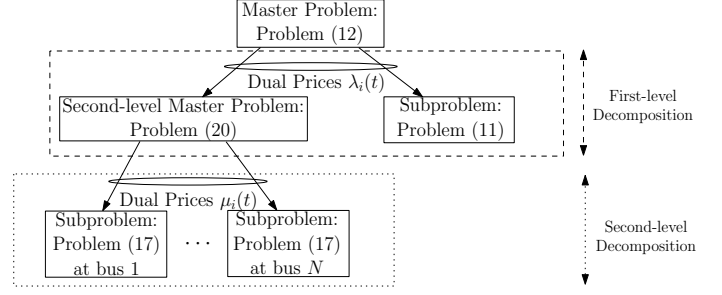


Fig. 3. Indirect dual-dual optimization decomposition to solve the OPF with energy storage problem in (8). The partial dual decomposition is applied at the two levels.

have studied a dynamic OPF problem similar to (8) but the work has some limitations: A simplified inequality constraint $b_i(t+1) \leq b_i(t) + r_i(t)$ for $t = 1, \dots, T$ has been used instead of the equality constraint given in (4) that can model the storage dynamics more accurately, and only the battery amount (i.e., $h_{1i}(t)$) is used to model the battery cost.

IV. INDIRECT DUAL-DUAL DECOMPOSITION AND ALGORITHMS

In this section, we study new optimization decomposition methodologies that can offer novel architectural transformation insights, and enable scalable and distributed algorithm design to solve (8). We study an indirect decomposition technique to decompose the dynamic OPF in (8) into simpler subproblems to unravel the coupling over space and time. Figure 3 illustrates the overview of this decomposition. In the first level, Lagrange dual decomposition is first applied. The dual variable corresponding to (2) plays the role of power flow price to control the decomposed subproblems. In the second level, Lagrange dual decomposition is again applied. The dual variable corresponding to (4) plays the role of energy storage price to control the second-level subproblems.

In the following, we explain each of these two levels in details. Let us apply a partial dual decomposition to decompose (8) over time. Then, for the subproblems, they are separable for $t = 1, \dots, T$ and each of them corresponds to one static OPF problem at time period t :

$$\begin{aligned} & \text{minimize } \mathbf{V}(t)^\top \mathbf{Y} \mathbf{V}(t) + \sum_{i \in \mathcal{N}} \lambda_i(t) (\mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t)) \\ & \text{subject to } \underline{\mathbf{V}} \leq \mathbf{V}(t) \leq \bar{\mathbf{V}}, \\ & \text{variables: } \mathbf{V}(t), \end{aligned} \quad (11)$$

where $\lambda_i(t)$ for all i and t are the nonnegative Lagrange dual variables corresponding to the constraint $\mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t) \leq \bar{p}_i(t) - r_i(t)$ in (8). Note that $\lambda_i(t)$ can be interpreted as the power price (by the shadow price interpretation of the optimal dual solution, cf. Chapter 5 in [18]). Due to the zero duality gap, each decomposed subproblem (11) for a fixed t can be solved with the following message passing algorithm (for a fixed $\lambda_i(t)$ for all i and t , and we use ℓ to index the iteration in the algorithm) in [22].

Algorithm 1: Distributed computation of voltages

Compute voltage $\mathbf{V}(t)$:

$$V_i^{\ell+1}(t) = \max \left\{ \underline{V}_i, \min \left\{ \bar{V}_i, \sum_{j \in \Omega_i} A_{ij}(t) V_j^{\ell}(t) \right\} \right\}$$

$\forall i \in \mathcal{N}$, where

$$A_{ij}(t) = \frac{2Y_{ij} + \lambda_i(t)Y_{ij} + \lambda_j(t)Y_{ij}}{2(1 + \lambda_i(t)) \sum_{j \in \Omega_i} Y_{ij}}, \quad \forall (i, j) \in \mathcal{E}.$$

Remark 1: Algorithm 1 converges to the unique optimal solution of (11) for each $t = 1, \dots, T$ [22]. The spectral property of the nonnegative matrix $\mathbf{A}(t)$ in Algorithm 1 can be used to characterize its convergence performance [22]. In addition, the dependency on the network topology for message passing is captured by the matrix $\mathbf{A}(t)$.

The optimal power prices are obtained by solving the following primary master dual problem:

$$\begin{aligned} & \text{maximize} \sum_{t=1}^T g(\lambda(t)) + \lambda(t)^\top (\mathbf{r}(t) - \bar{\mathbf{p}}(t)) \\ & \text{subject to } \lambda(t) \geq \mathbf{0}, \quad t = 1, \dots, T, \\ & \text{variables: } \lambda(t), \end{aligned} \quad (12)$$

where

$$g(\lambda(t)) = \mathbf{V}^*(t)^\top \mathbf{Y} \mathbf{V}^*(t) + \sum_{i \in \mathcal{N}} \lambda_i(t) (\mathbf{V}^*(t)^\top \mathbf{Y}_i \mathbf{V}^*(t)).$$

The dual function is differentiable at each t and can be solved iteratively by the following gradient update:

$$\lambda_i^{k+1}(t) = [\lambda_i^k(t) + \beta (\mathbf{V}^*(t)^\top \mathbf{Y}_i \mathbf{V}^*(t) - \bar{p}_i(t) + r_i(t))]^+, \quad \forall i \in \mathcal{N}, \quad (13)$$

where β is an appropriate stepsize, and $[\cdot]^+$ denotes the projection onto the nonnegative orthant. This update will in turn drive Algorithm 1.

Next, let us consider the decomposed problem corresponding to the energy storage:

$$\begin{aligned} & \text{minimize} \sum_{t=1}^{T+1} \alpha^\top (\mathbf{B} - \mathbf{b}(t)) + \sum_{t=1}^T \lambda(t)^\top \mathbf{r}(t) + \sum_{i \in \mathcal{N}} \tilde{\alpha}_i r_i^2(t) \\ & \text{subject to } b_i(t+1) = \eta_i b_i(t) + r_i(t), \quad i \in \mathcal{N}, \quad t = 1, \dots, T, \\ & \quad \mathbf{b}(1) = \mathbf{B}^0, \\ & \quad \mathbf{0} \leq \mathbf{b}(t) \leq \mathbf{B}, \quad t = 1, \dots, T+1, \\ & \quad \underline{\mathbf{r}} \leq \mathbf{r}(t) \leq \bar{\mathbf{r}}, \quad t = 1, \dots, T, \\ & \text{variables: } \mathbf{b}(t), \mathbf{r}(t). \end{aligned} \quad (14)$$

Since $b_i(t+1) = \eta_i b_i(t) + r_i(t)$ for $t = 1, \dots, T$, and $b_i(1) = B_i^0$, then for each $i \in \mathcal{N}$ we have:

$$b_i(t+1) = \eta_i^t b_i(1) + \sum_{\tau=1}^t \eta_i^{t-\tau} r_i(\tau), \quad t = 1, \dots, T. \quad (15)$$

We thus rewrite (14) as:

$$\begin{aligned} & \text{minimize} \sum_{t=1}^{T+1} \alpha^\top (\mathbf{B} - \eta^t \mathbf{b}(1) - \sum_{\tau=1}^t \eta^{t-\tau} \mathbf{r}(\tau)) \\ & \quad + \sum_{t=1}^T \lambda(t)^\top \mathbf{r}(t) + \sum_{i \in \mathcal{N}} \tilde{\alpha}_i r_i^2(t) \\ & \text{subject to } b_i(t+1) = \eta_i^t b_i(1) + \sum_{\tau=1}^t \eta_i^{t-\tau} r_i(\tau), \quad \forall i \in \mathcal{N}, \\ & \quad \underline{\mathbf{r}} \leq \mathbf{r}(t) \leq \bar{\mathbf{r}}, \\ & \text{variables: } \mathbf{r}(t), \mathbf{b}(t), \quad t = 1, \dots, T. \end{aligned} \quad (16)$$

Now, for a fixed $\mathbf{b}(t)$, if we apply the partial dual decomposition to (16), we get a second-level subproblem at each $i \in \mathcal{N}$:

$$\begin{aligned} & \text{minimize} \sum_{t=1}^{T+1} \alpha_i (B_i - \eta_i^t b_i(1) - \sum_{\tau=1}^t \eta_i^{t-\tau} r_i(\tau)) \\ & \quad + \sum_{t=1}^T \lambda_i(t) r_i(t) - (\sum_{\tau=1}^t \eta_i^{t-\tau} r_i(\tau) \mu_i(t)) \\ & \quad + \mu_i(t) (b_i(t+1) - \eta_i^t b_i(1)) + \tilde{\alpha}_i r_i^2(t) \\ & \text{subject to } \underline{r}_i \leq r_i(t) \leq \bar{r}_i, \quad i \in \mathcal{N}, \quad t = 1, \dots, T, \\ & \text{variables: } r_i(t), \end{aligned} \quad (17)$$

where $\mu_i(t)$ is the Lagrange dual variable corresponding to (4). Then, we have the following update on the charging/discharging amount of power in the energy storage for $t = 1, \dots, T$ and $i \in \mathcal{N}$:

$$\begin{aligned} r_i^{k+1}(t) &= [r_i^k(t) - \theta (\lambda_i^*(t) - \sum_{\tau=t}^T \eta_i^{\tau-t} (\alpha_i + \mu_i^*(\tau)) \\ & \quad + 2\tilde{\alpha}_i r_i^k(t))]_{\underline{r}_i}^{\bar{r}_i}, \end{aligned} \quad (18)$$

$$b_i^k(t+1) = [\eta_i^t b_i(1) + \sum_{\tau=1}^t \eta_i^{t-\tau} r_i^k(\tau)]_0^{B_i}, \quad (19)$$

where θ is an appropriate stepsize, and $[x]_b^a$ denotes the projection $\max\{b, \min\{a, x\}\}$ where $b < a$.

Furthermore, the second-level master dual problem is given by

$$\begin{aligned} & \text{maximize} \sum_{t=1}^T \sum_{i \in \mathcal{N}} g_i(\mu(t)) + \tilde{\alpha}_i r_i^{*2}(t) + \lambda(t)^\top \mathbf{r}^*(t) \\ & \quad + \alpha^\top (\mathbf{B} - \mathbf{b}^*(T+1)) \\ & \text{variables: } \mu(t), \end{aligned} \quad (20)$$

where

$$g_i(\mu(t)) = \alpha_i (B_i - b_i^*(t)) + \mu_i(t) (b_i^*(t+1) - \eta_i b_i^*(t) - r_i^*(t))$$

at each subproblem of (17). In particular, a subgradient update of the dual variable $\mu_i(t)$ for $t = 1, \dots, T$ to solve (20) is given by

$$\mu_i^{k+1}(t) = \mu_i^k(t) + \gamma (b_i^*(t+1) - \eta_i b_i^*(t) - r_i^*(t)), \quad \forall i \in \mathcal{N}, \quad (21)$$

where γ is an appropriate stepsize. We summarize the various solution update of the above dual-dual decomposition in the following (using k to index algorithm iteration).

Algorithm 2: Dual-Dual Decomposition

1) Set the stepsizes $\beta, \theta, \gamma \in (0, 1)$.

2) Compute the charging/discharging amount of power in storage:

$$r_i^{k+1}(t) = [r_i^k(t) - \theta(\lambda_i^k(t) - \sum_{\tau=t}^T \eta_i^{\tau-t}(\alpha_i + \mu_i^k(\tau)) + 2\tilde{\alpha}_i r_i^k(t))]_{\underline{r}_i}^{\bar{r}_i} \quad \{r_i^k(t), b_i^k(t), \mu_i^k(t)\}$$

$$\forall i \in \mathcal{N} \text{ and } t = 1, \dots, T.$$

3) Compute the battery amount:

$$b_i^k(t+1) = [\eta_i^t b_i(1) + \sum_{\tau=1}^t \eta_i^{t-\tau} r_i^k(\tau)]_0^{B_i}$$

$$\forall i \in \mathcal{N} \text{ and } t = 1, \dots, T.$$

4) Compute the storage dual variable:

$$\mu_i^{k+1}(t) = \mu_i^k(t) + \gamma(b_i^k(t+1) - \eta_i b_i^k(t) - r_i^k(t))$$

$$\forall i \in \mathcal{N} \text{ and } t = 1, \dots, T.$$

5) Run Algorithm 1 to compute the voltage and set Algorithm 1 output as $\mathbf{V}^k(t)$ for $t = 1, \dots, T$.

6) Compute the power price:

$$\lambda_i^{k+1}(t) = [\lambda_i^k(t) + \beta(\mathbf{V}^k(t)^\top \mathbf{Y}_i \mathbf{V}^k(t) - \bar{p}_i(t) + r_i^k(t))]^+$$

$$\forall i \in \mathcal{N} \text{ and } t = 1, \dots, T.$$

Update β, θ and γ until convergence.

Remark 2: Steps 5 and 6 are equivalent to Algorithm 2 in [22] for a fixed $r_i(t)$ at each t , and its convergence proof is given in [22]. Moreover, the subgradient method of the second-level (i.e., Steps 2 and 4) can converge to a closed neighborhood of the optimal solution by appropriately choosing the stepsize (using a diminishing stepsize rule or sufficiently small constant stepsizes) [30], [36].

Remark 3: Algorithm 2 is distributed, because each bus only communicates with its local one-hop neighbors, i.e., at Steps 5 and 6, each bus exchanges nodal voltage and nodal power price information with its one-hop neighbors by message passing. An example is illustrated in Figure 4 for a 5-bus system ([6], Chapter 6, pp.327). Each node $i \in \mathcal{N}$ randomly chooses a feasible $V_i^0(t)$ for $t = 1, \dots, T$ and broadcasts to its one-hop neighbors. Once a node $i \in \mathcal{N}$ receives all $V_j(t)$ where $j \in \Omega_i$, it calculates $\lambda_i(t)$. Next, the updated $\lambda_i(t)$ is then passed to its neighbors. After collecting all $\lambda_j(t)$ ($j \in \Omega_i$), $V_i(t)$ is calculated by Algorithm 1 at bus i . The other variables $r_i(t), b_i(t)$ and $\mu_i(t)$ for $j \in \Omega_i$ and $t = 1, \dots, T$ only require information locally available at bus i . This message passing is repeated until convergence.

V. ALTERNATIVE PRIMAL-DUAL DECOMPOSITION AND ALGORITHMS

We now present an alternative decomposition method (a primal-dual decomposition) that leads to a distributed algorithm (without using Algorithm 1) to solve (8). The difference

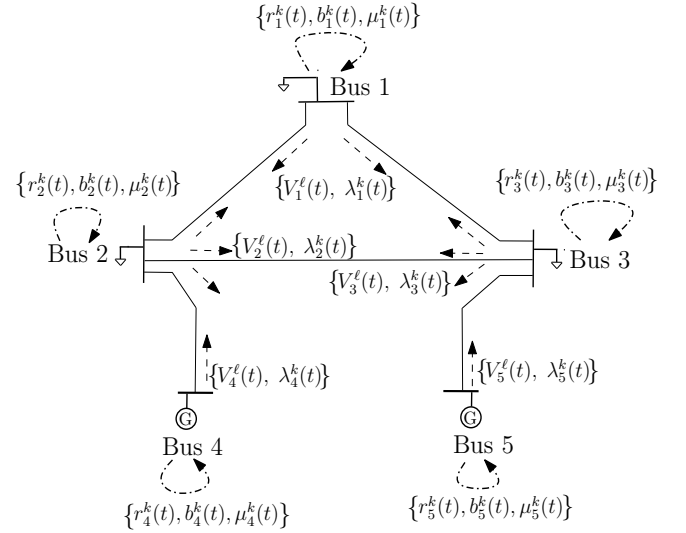


Fig. 4. Message passing in a 5-bus system ([6], Chapter 6, pp.327).

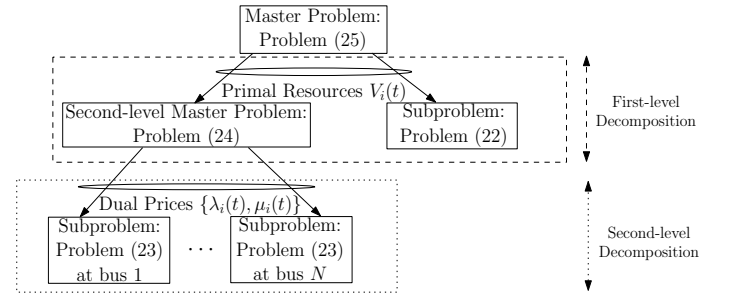


Fig. 5. Alternative primal-dual decomposition to solve the OPF with energy storage problem in (8). The primal and dual decomposition are applied at the first level and the second level respectively.

between these two decompositions is in how power prices drive the voltage computation. Consider a primal decomposition to (8) by fixing $\mathbf{V}(t)$ for $t = 1, \dots, T$. Then, we obtain subproblems of (8) given by:

$$\begin{aligned} & \text{minimize} \quad \sum_{i \in \mathcal{N}} \sum_{t=1}^{T+1} h_{1i}(b_i(t)) + \sum_{t=1}^T h_{2i}(r_i(t)) \\ & \text{subject to} \quad \mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t) \leq \bar{p}_i(t) - r_i(t), \quad \forall i \in \mathcal{N}, \quad t = 1, \dots, T, \\ & \quad b_i(t+1) = \eta_i b_i(t) + r_i(t), \quad i \in \mathcal{N}, \quad t = 1, \dots, T, \\ & \quad \mathbf{b}(1) = \mathbf{B}^0, \\ & \quad \mathbf{0} \leq \mathbf{b}(t) \leq \mathbf{B}, \quad t = 1, \dots, T+1, \\ & \quad \underline{\mathbf{r}} \leq \mathbf{r}(t) \leq \bar{\mathbf{r}}, \quad t = 1, \dots, T, \\ & \text{variables: } \mathbf{b}(t), \mathbf{r}(t). \end{aligned} \tag{22}$$

We use Figure 5 to illustrate this alternative decomposition structure from (22). In the first level, a primal decomposition is used. The primal variable $v_i(t)$ plays the role of resources to control the decomposed subproblems. In the second level, a partial Lagrange dual decomposition is applied to (22). The dual variables $\lambda_i(t)$ and $\mu_i(t)$ play the role of power price and storage dual variable to control these second-level subproblems, respectively.

Now, by applying the partial dual decomposition to (22),

for each $i \in \mathcal{N}$, we have the following subproblem:

$$\begin{aligned}
& \text{minimize} \sum_{t=1}^{T+1} \alpha_i (B_i - b_i(t)) + \sum_{t=1}^T \lambda_i(t) r_i(t) + \tilde{\alpha}_i r_i^2(t) \\
& \quad + \mu_i(t) (b_i(t+1) - \eta_i b_i(t) - r_i(t)) \\
& \text{subject to } b_i(1) = B_i^0, \\
& \quad 0 \leq b_i(t) \leq B_i, \quad t = 1, \dots, T+1, \\
& \quad \underline{r}_i \leq r_i(t) \leq \bar{r}_i, \quad t = 1, \dots, T, \\
& \text{variables: } b_i(t), r_i(t).
\end{aligned} \tag{23}$$

After using the relation $b_i(t+1) = \eta_i b_i(t) + r_i(t)$ for $t = 1, \dots, T$, and fixing $b_i(t)$, we can rewrite the second-level subproblem as (17) at each $i \in \mathcal{N}$. Thus, we have $r_i(t)$ and $b_i(t)$ as given in (18) and (19) for all $i \in \mathcal{N}$ and $t = 1, \dots, T$.

This second-level master dual problem is given by

$$\begin{aligned}
& \text{maximize} \sum_{t=1}^T \sum_{i \in \mathcal{N}} g_i(\mu(t), \lambda(t)) + \tilde{\alpha}_i r_i^{*2}(t) \\
& \quad + \alpha^\top (\mathbf{B} - \mathbf{b}^*(T+1)) \\
& \text{subject to } \lambda(t) \geq \mathbf{0}, \\
& \text{variables: } \mu(t), \lambda(t),
\end{aligned} \tag{24}$$

where

$$\begin{aligned}
g_i(\mu(t), \lambda(t)) = & \lambda_i(t) (\mathbf{V}^*(t)^\top \mathbf{Y}_i \mathbf{V}^*(t) - \bar{p}_i(t) + r_i^*(t)) \\
& + \mu_i(t) (b_i^*(t+1) - \eta_i b_i^*(t) - r_i^*(t)) \\
& + \alpha_i (B_i - b_i^*(t)),
\end{aligned}$$

in each subproblem of (17). The subgradient update of the dual variable $\lambda_i(t)$ and $\mu_i(t)$ for $t = 1, \dots, T$ can be done by (13) and (21) respectively.

Going up one-level, the master primal problem can be written as

$$\begin{aligned}
& \text{minimize } f^*(\mathbf{V}(t)) \\
& \text{subject to } \underline{\mathbf{V}} \leq \mathbf{V}(t) \leq \bar{\mathbf{V}}, \\
& \text{variables: } \mathbf{V}(t),
\end{aligned} \tag{25}$$

where $f^*(\mathbf{V}(t))$ is the optimal value of (8) for a given $\mathbf{V}(t)$ for $t = 1, \dots, T$. Recall that $\lambda(t)$ for $t = 1, \dots, T$ are the set of nodal power prices corresponding to (2). A subgradient method can solve (25) by iteratively updating the voltage $V_i(t)$:

$$V_i^{\ell+1}(t) = \left[V_i^\ell(t) - \delta (2\tilde{\lambda}_i^*(t) \sum_{j \in \Omega_i} Y_{ij} V_j^\ell(t) - \sum_{j \in \Omega_i} \tilde{\lambda}_{ij}^*(t) V_j^\ell(t)) \right]_{\underline{V}_i}^{\bar{V}_i},$$

where

$$\tilde{\lambda}_i^*(t) = (1 + \lambda_i^*(t)),$$

for all $i \in \mathcal{N}$ and

$$\tilde{\lambda}_{ij}^*(t) = (2Y_{ij} + \lambda_i^*(t)Y_{ij} + \lambda_j^*(t)Y_{ij}),$$

for all $(i, j) \in \mathcal{E}$ and δ is an appropriate stepsize.

We summarize the various solution update of the above alternative decomposition in the following (using k to index algorithm iteration).

Algorithm 3: Primal-Dual Decomposition

- 1) Set the stepsizes $\delta, \gamma, \beta, \theta \in (0, 1)$.

- 2) Until convergence (of the index ℓ), compute the voltage:

$$V_i^{\ell+1}(t) = [V_i^\ell(t) - \delta (2\tilde{\lambda}_i^k(t) \sum_{j \in \Omega_i} Y_{ij} V_j^\ell(t) - \sum_{j \in \Omega_i} \tilde{\lambda}_{ij}^k(t) V_j^\ell(t))]_{\underline{V}_i}^{\bar{V}_i},$$

where $\tilde{\lambda}_i^k(t) = (1 + \lambda_i^k(t))$, $\forall i \in \mathcal{N}$ and $\tilde{\lambda}_{ij}^k(t) = (2Y_{ij} + \lambda_i^k(t)Y_{ij} + \lambda_j^k(t)Y_{ij})$, $\forall (i, j) \in \mathcal{E}$ and $t = 1, \dots, T$. Upon convergence, let $\mathbf{V}^k(t)$ be the voltage output.

- 3) Compute the storage dual variable:

$$\mu_i^{k+1}(t) = \mu_i^k(t) + \gamma (b_i^k(t+1) - \eta_i b_i^k(t) - r_i^k(t))$$

$\forall i \in \mathcal{N}$ and $t = 1, \dots, T$.

- 4) Compute the power price:

$$\lambda_i^{k+1}(t) = [\lambda_i^k(t) + \beta (\mathbf{V}^k(t)^\top \mathbf{Y}_i \mathbf{V}^k(t) - \bar{p}_i(t) + r_i^k(t))]^+.$$

- 5) Compute the charging/discharging amount of power in storage:

$$r_i^{k+1}(t) = [r_i^k(t) - \theta (\lambda_i^k(t) - \sum_{\tau=t}^T \eta_i^{\tau-t} (\alpha_i + \mu_i^k(\tau)) + 2\tilde{\alpha}_i r_i^k(t))]_{\underline{r}_i}^{\bar{r}_i}$$

$\forall i \in \mathcal{N}$ and $t = 1, \dots, T$.

- 6) Compute the battery amount:

$$b_i^k(t+1) = [\eta_i^t b_i(1) + \sum_{\tau=1}^t \eta_i^{t-\tau} r_i^k(\tau)]_0^{B_i}$$

$\forall i \in \mathcal{N}$ and $t = 1, \dots, T$.

Update δ, γ, β and θ until convergence.

Remark 4: Unlike Algorithm 2, Algorithm 3 requires more subgradient updates to converge to a closed neighborhood of the optimal solution by appropriately choosing the stepsizes (using a diminishing stepsize rule or sufficiently small constant stepsizes) [30], [36] and also synchronizing the updates at each level of the decomposition [29]. The message passing in Algorithm 3 is similar to that illustrated in Figure 4.

VI. ENERGY STORAGE COUPLING OVER TIME INTERPRETATION VIA LAGRANGE DUALITY

In this section, we discuss the physical interpretation of the optimization structures obtained through the two different decomposition methods in the previous two sections.

A. Physical interpretation of dual solution from indirect dual-dual decomposition

In the indirect dual-dual decomposition, for the subproblem corresponding to the energy storage, i.e., (14), the Lagrangian

at each $i \in \mathcal{N}$ can be expressed as:

$$L_i = \sum_{t=1}^{T+1} \alpha_i (B_i - b_i(t)) + \bar{\mu}_i(t)(b_i(t) - B_i) - \underline{\mu}_i(t)b_i(t) \\ + \sum_{t=1}^T \lambda_i(t)r_i(t) + \mu_i(t)(b_i(t+1) - \eta_i b_i(t) - r_i(t)) \\ + \tilde{\alpha}_i r_i^2(t) + \bar{\nu}_i(t)(r_i(t) - \bar{r}_i) - \underline{\nu}_i(t)(r_i(t) - \underline{r}_i) \\ + \rho_i(1)(b_i(1) - B_i^0), \quad (26)$$

where the dual variable $\mu_i(t)$ corresponds to (4), $\underline{\mu}_i(t)$ and $\bar{\mu}_i(t)$ respectively correspond to the lower and upper bounds of (6), $\rho_i(1)$ corresponds to (5), and $\underline{\nu}_i(t)$ and $\bar{\nu}_i(t)$ respectively correspond to the discharge and charge bounds in (7). Interestingly, the difference between two successive dual variables $\eta_i \mu_i(t) - \mu_i(t-1)$ plays a decoupling role over time:

Lemma 1: For a battery i , if $\mu_i(t-1) - \eta_i \mu_i(t) < \alpha_i$, where $t = 2, \dots, T$ or $\mu_i(T) < \alpha_i$, then it is saturated at t or $T+1$; If the inequality is reversed, then battery i is drained for $t = 2, \dots, T$ or $T+1$. If the battery bound of bus i in (8) is inactive for $t = 1, \dots, T$, then $\mu_i(t) = \sum_{\tau=t}^T \eta_i^{T-\tau} \alpha_i$.

In addition, if $r_i(t) < \frac{1}{2\tilde{\alpha}_i}(\mu_i(t) - \lambda_i(t))$, where $t = 1, 2, \dots, T$, then the charging power amount upper bound is hit at t ; If the inequality is reversed, then battery i reaches its discharging power amount upper bound for $t = 1, 2, \dots, T$. If the charging/discharging power amount bounds of bus i in (8) are inactive for $t = 1, 2, \dots, T$, then $r_i(t) = \frac{1}{2\tilde{\alpha}_i}(\mu_i(t) - \lambda_i(t))$.

Proof: As the dynamic OPF has a zero duality gap, we use the Karush-Kuhn-Tucker (KKT) conditions (cf. [18]) to deduce the following:

$$2\tilde{\alpha}_i r_i(t) = \mu_i(t) - \lambda_i(t) + \underline{\nu}_i(t) - \bar{\nu}_i(t), \quad t = 1, 2, \dots, T, \quad (27)$$

$$\eta_i \mu_i(t) - \mu_i(t-1) = -\alpha_i + \bar{\mu}_i(t) - \underline{\mu}_i(t), \quad t = 2, 3, \dots, T, \quad (28)$$

$$\mu_i(T) = \alpha_i - \bar{\mu}_i(T+1) + \underline{\mu}_i(T+1).$$

Suppose that battery i never drains or saturates, i.e., $0 < b_i(t) < B_i$, then $\eta_i \mu_i(t) - \mu_i(t-1) = -\alpha_i$ for $t = 2, \dots, T$ and $\mu_i(T) = \alpha_i$. Thus, $\mu_i(t) = \sum_{\tau=t}^T \eta_i^{T-\tau} \alpha_i$ for $t = 1, \dots, T$. If a battery i drains, then $\eta_i \mu_i(t) - \mu_i(t-1) \leq -\alpha_i$ for $t = 2, 3, \dots, T$ (drains at t) and $\mu_i(T) \geq \alpha_i$ (drains at $T+1$). If a battery i saturates, then $\eta_i \mu_i(t) - \mu_i(t-1) \geq -\alpha_i$ for $t = 2, 3, \dots, T$ (saturates at t) and $\mu_i(T) \leq \alpha_i$ (saturates at $T+1$).

In addition, suppose the charge/discharge bounds are not tight at the storage of bus i , i.e., $\underline{r}_i < r_i(t) < \bar{r}_i$ for $t = 1, 2, \dots, T$, then the charging/discharging power amount is completely determined by the difference between two dual variables ($\mu_i(t)$ and $\lambda_i(t)$), i.e., $r_i(t) = \frac{1}{2\tilde{\alpha}_i}(\mu_i(t) - \lambda_i(t))$. If a battery hits its charging power amount upper bound for $t = 1, 2, \dots, T$, then $r_i(t) = \frac{1}{2\tilde{\alpha}_i}(\mu_i(t) - \lambda_i(t) - \bar{\nu}_i(t))$. If a battery hits its discharging power amount upper bound for $t = 1, 2, \dots, T$, then $r_i(t) = \frac{1}{2\tilde{\alpha}_i}(\mu_i(t) - \lambda_i(t) + \underline{\nu}_i(t))$. This proves Lemma 1. ■

We now derive the battery amount. Rewriting the battery amount (4) with dual variables for each $i \in \mathcal{N}$ and $t =$

$1, 2, \dots, T$, we have:

$$b_i(t+1) = \eta_i b_i(t) + \frac{1}{2\tilde{\alpha}_i}(\mu_i(t) - \lambda_i(t) - \bar{\nu}_i(t) + \underline{\nu}_i(t)), \\ = \eta_i b_i(t) + \frac{1}{2\tilde{\alpha}_i}(-\lambda_i(t) - \bar{\nu}_i(t) + \underline{\nu}_i(t) + \mu_i(1)\eta_i^{-(t-1)} \\ - \sum_{\tau=1}^{t-1} \eta_i^{-\tau} \alpha_i + \sum_{\tau=2}^t (\bar{\mu}_i(\tau) - \underline{\mu}_i(\tau)) \eta_i^{-(t+1-\tau)}) \quad (29)$$

where the second equality is due to (28). In (29), these two entries corresponding to the lower and upper bound dual variables $\underline{\nu}_i(t) - \bar{\nu}_i(t)$ and $\sum_{\tau=2}^t (\bar{\mu}_i(\tau) - \underline{\mu}_i(\tau))$ can be obtained as follows:

$$\underline{\nu}_i(t) - \bar{\nu}_i(t) = 2\tilde{\alpha}_i r_i(t) - \mu_i(t) + \lambda_i(t), \quad (30)$$

$$\sum_{\tau=2}^t \bar{\mu}_i(\tau) - \underline{\mu}_i(\tau) = \sum_{\tau=2}^t (\eta_i \mu_i(\tau) - \mu_i(\tau-1) + \alpha_i). \quad (31)$$

Therefore, if we rewrite (29) by the iteration updates for $t = 1, 2, \dots, T$ and $i \in \mathcal{N}$ as:

$$b_i^k(t+1) = [\eta_i b_i^k(t) + \frac{1}{2\tilde{\alpha}_i}(-\lambda_i(t) + 2\tilde{\alpha}_i r_i(t) - \mu_i(t) + \lambda_i(t) \\ + \sum_{\tau=2}^t (\eta_i \mu_i(\tau) - \mu_i(\tau-1) + \alpha_i) \eta_i^{-(t+1-\tau)} \\ - \sum_{\tau=1}^{t-1} \eta_i^{-\tau} \alpha_i + \mu_i(1) \eta_i^{-(t-1)})]_{0}^{B_i}, \quad (32)$$

then we have an extension of Algorithm 2 that modifies the battery amount update using the dual price iterates.

Algorithm 4: Extension of Algorithm 2

Replace Step 3 of Algorithm 2 by the update:

$$b_i^k(t+1) = [\eta_i b_i^k(t) + \frac{1}{2\tilde{\alpha}_i}(-\mu_i^k(t) + 2\tilde{\alpha}_i r_i^k(t) \\ + \sum_{\tau=2}^t (\eta_i \mu_i^k(\tau) - \mu_i^k(\tau-1) + \alpha_i) \eta_i^{-(t+1-\tau)} \\ - \sum_{\tau=1}^{t-1} \eta_i^{-\tau} \alpha_i + \mu_i^k(1) \eta_i^{-(t-1)})]_{0}^{B_i}$$

$\forall i \in \mathcal{N}$ and $t = 1, \dots, T$.

Remark 5: Similarly to Algorithm 2, by an appropriate choice of the stepsizes, Algorithm 4 can converge to the global optimal solution, and we compare these modifications in Section VII.

B. Physical interpretation of dual solution from alternative primal-dual decomposition

In the alternative primal-dual decomposition, the following special case result quantifies the effect of load smoothing through the property of the dual solution.

Lemma 2: Suppose we ignore the battery cost $h_i(t)$ (let $\alpha = 0$ and $\tilde{\alpha} = 0$), the battery storage leakage (let $\eta = 1$) and select the battery bounds to be large enough such that the

battery never hits the bounds before $T+1$, then the load profile is strictly smoothed, i.e., $\mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t) = \mathbf{V}(t-1)^\top \mathbf{Y}_i \mathbf{V}(t-1)$ for $t = 2, \dots, T$.

Proof: The Lagrangian that corresponds to the energy storage, i.e., (22) ($\mathbf{V}(t)$ is fixed), at each $i \in \mathcal{N}$ is given by: $\tilde{L}_i = L_i + \sum_{t=1}^T \lambda_i(t) (\mathbf{V}^\top(t) \mathbf{Y}_i \mathbf{V}(t))$. We analyze the KKT conditions of \tilde{L}_i similarly to Lemma 1. Since the battery bound of bus i in (8) is inactive for all $i \in \mathcal{N}$ and $t = 1, \dots, T-1$, $\mu_i(t) = \sum_{\tau=t}^{T-1} \eta_i^{T-\tau-1} \alpha_i + \eta_i^{T-t} \mu_i(T)$. Thus, if $\alpha_i = 0$ and $\eta_i = 1$ for all $i \in \mathcal{N}$, we have $\mu_i(1) = \mu_i(2) = \dots = \mu_i(T)$ for all $i \in \mathcal{N}$. Furthermore, by the stationarity of the Lagrangian in the KKT conditions corresponding to $r_i(t)$, if $\tilde{\alpha}_i = 0$ and battery i does not reach its charge/discharge bound for all $i \in \mathcal{N}$ and $t = 1, \dots, T$, we have $\mu_i(t) = \lambda_i(t)$, which then implies that $\lambda_i(1) = \lambda_i(2) = \dots = \lambda_i(T)$ for all $i \in \mathcal{N}$. The problem in (22) is then no longer coupled over time and hence can be solved individually for each t . Moreover, as the nodal power prices satisfy $\lambda_i(1) = \lambda_i(2) = \dots = \lambda_i(T)$ for all $i \in \mathcal{N}$, this means that the nodal power in (22) at node i for $t = 1, \dots, T$ must be the same, i.e., $\mathbf{V}(t)^\top \mathbf{Y}_i \mathbf{V}(t) = \mathbf{V}(t-1)^\top \mathbf{Y}_i \mathbf{V}(t-1)$ for $t = 2, \dots, T$. ■

In summary, different decomposition methods can lead to different load profile smoothing features. Partial load profile smoothing can still be achieved with a partial decomposition. As the energy storage batteries charge and discharge at different time periods, the dual solution of each OPF in the temporal dimension are coupled to achieve the load profile smoothing (see evaluations in Section VII-A).

VII. NUMERICAL EVALUATIONS

In this section, we evaluate the performance of our proposed message passing algorithms and demonstrate the load profile smoothing features of the dynamic OPF problem in (8) for different problem settings. In the following, unless otherwise indicated, all the power values are normalized to per unit values (pu) like in [20].

A. Illustration of load profile smoothing by energy storage

To illustrate load profile smoothing, we use a 5-bus example ([6], Chapter 6, pp.327) as shown in Figure 4, where a battery is attached to each bus. We divide the time period of a day into six intervals (every four hours in an interval and thus $t = 1, 2, \dots, 6$) and the demand profile starts at 6:00PM (peak hour). After solving this dynamic OPF under the demand variation, we can observe a similar load profile smoothing feature akin to “valley-filling” [3], [4] as shown in Figure 6.

Figure 6(a) shows the ideally smoothed load profile as described in Lemma 2. The peak has been flattened while the valley has been filled to the same level in the whole period. In Figure 6(b), we increase the demand at Bus 2 while keeping other parameters the same as before. Since Bus 2 reaches the battery lower bound at time $t = 3$ (discharge the battery to satisfy the high demand requirement), a strictly smoothed load profile is not shown according to Lemma 1 and Lemma 2. However, for the two time periods $t = 1 \sim 2$ and $t = 3 \sim 6$, as the relationship $\lambda_i(t) = \lambda_i(t-1)$ still holds at each time period, the load profile is also strictly smoothed in each of

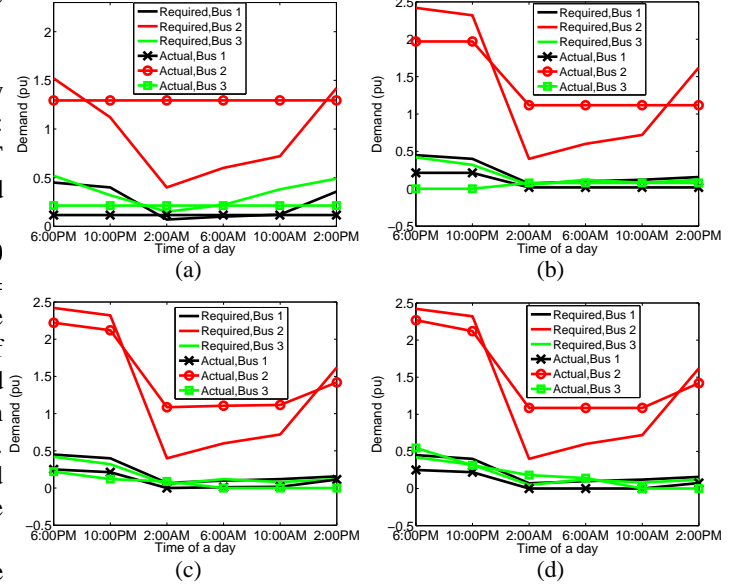


Fig. 6. Illustration of the actual and required demand under different problem settings. (a) Ideal load smoothing (b) Approximately ideal load smoothing (c) A load smoothing case by choosing $\alpha_i = \tilde{\alpha}_i = 0.1$ (d) A load smoothing case by choosing $\alpha_i = 0.2, \tilde{\alpha}_i = 0.3$.

these two periods. In Figures 6(c) and 6(d), we consider the battery dynamic cost $\mathbf{h}(t)$ and set the battery storage leakage parameter $\eta_i = 0.98$. Incidentally, by an appropriate choice of the battery cost parameter $\alpha_i = \tilde{\alpha}_i = 0.1$, the load can also be smoothed. Specifically, at the demand peak hour, e.g., 6:00PM, the demand bus discharges the battery and less power is absorbed from the grid. Conversely, the battery is charged at the demand off-peak hour, e.g., 2:00AM. Thus, the high load usage during the peak hours is shifted to the off-peak hours (valley bottom). However, when α_i increases, the smoothing effect is limited in order to minimize the total battery cost, and thus smoothing holds partially, e.g., the load is not balanced at Bus 3 when $\alpha_i = 0.2, \tilde{\alpha}_i = 0.3$.

Next, in connection to Figure 6(a), we plot the variation of the energy storage at Bus 2 in Figure 7 to illustrate the result in Lemma 2. When the demand is high, the battery is discharged to provide power to the attached bus. During the off-peak hours, the battery absorbs power from the grid to charge the battery. When the next peak hour arrives, the power stored earlier is used to supply the demand for the attached bus (realizing the load profile smoothing over the whole time period).

B. Convergence performance

We evaluate the convergence performance of Algorithm 2 and Algorithm 3 in the 5-bus system (Figure 4) using the time-varying demand of a day (six time intervals) as in Section VII-A. The parameter of the line admittance is $[y_{12}, y_{13}, y_{23}, y_{24}, y_{35}]^\top = [4, 3, 6, 6, 7]^\top$ pu. The initial step-sizes in Algorithm 2 are $\beta = 0.3, \theta = 0.04, \gamma = 0.04$. The initial step-sizes in Algorithm 3 are $\theta = 0.35, \beta = 0.25, \gamma = 0.25, \delta = 0.1$. We use a diminishing stepsize rule (cf. [30], [36]) to update the step-sizes. In Figure 8 and Figure 9, we

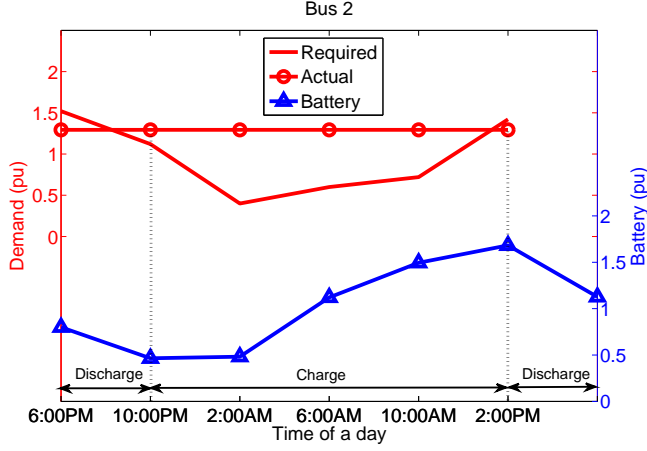


Fig. 7. Illustration of Lemma 2 in the 5-bus system. This figure shows how the actual demand, required demand and the battery storage amount vary in the ideal load smoothing setting of a day as in Figure 6(a) at Bus 2.

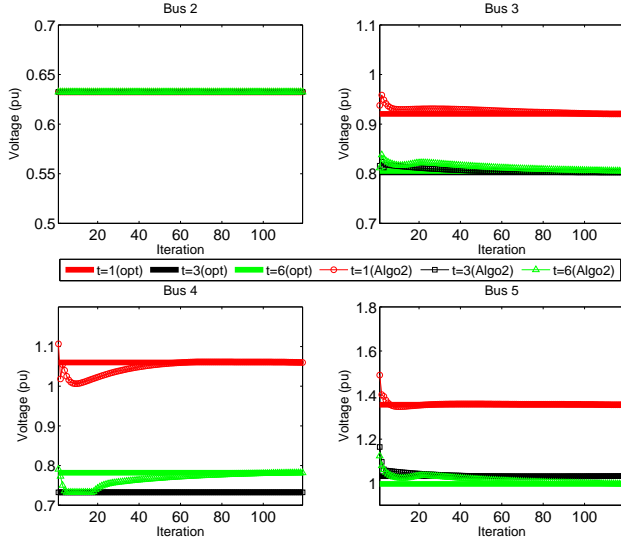


Fig. 8. Illustration of convergence of Algorithm 2 in the 5-bus system with the initial point set close to the optimal solution.

show the convergence of Algorithm 2 for four buses at three time intervals ($t = 1, 3$ and 6) using different initial points. In Figure 8, the initial point is set close to the optimal solution, i.e., in the neighborhood of the optimal solution. Algorithm 2 converges by around 115 iterations. In Figure 9, we let the initial point be set further away from the optimal solution, Algorithm 2 converges in less than 160 iterations. In Figure 10, we choose the initial point the same as that in Figure 9 to show the convergence behavior of Algorithm 3. By comparing Figure 10 and Figure 9, we see that Algorithm 3 converges to the optimal solution slower than that of Algorithm 2.

Next, we evaluate the algorithm performance on energy storage dynamics, i.e., the charging/discharging power amount using the same 5-bus system example. We choose the initial points randomly and let $\mathbf{r}^0(t) = 0$. In Figure 11 and Figure 12, we illustrate the convergence of Algorithm 2 and Algorithm 4 on $\mathbf{r}(t)$ for two buses at all the six intervals with differently

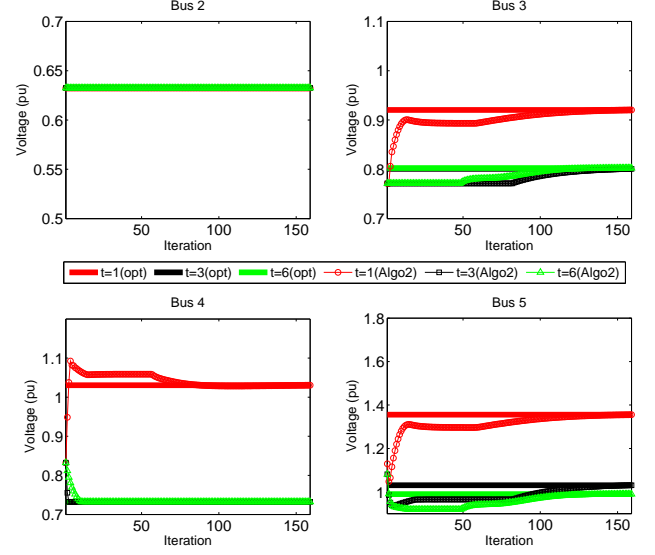


Fig. 9. Illustration of convergence of Algorithm 2 in the 5-bus system with the initial point set further away from the optimal solution.

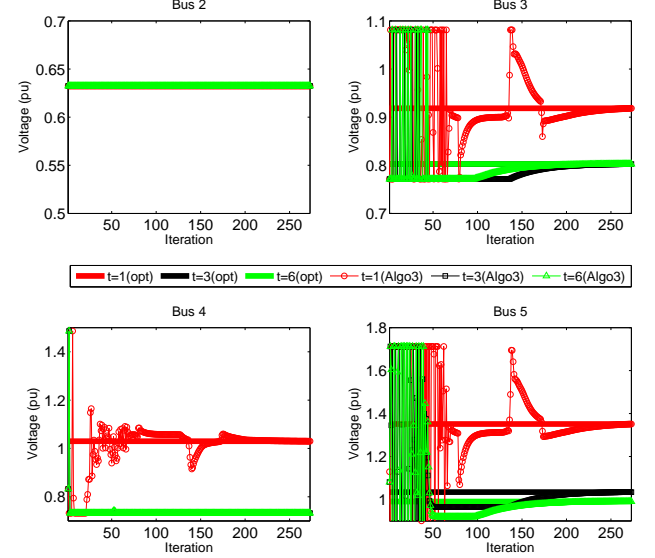


Fig. 10. Illustration of convergence of Algorithm 3 in the 5-bus system with the initial point set further away from the optimal solution.

random initial points, respectively. In Figure 11, the battery amount is computed only by the primal variable updates (Step 3 in Algorithm 2). We see that it converges in around 150 iterations. In Figure 12, the battery amount is computed by using the dual variable update (Step 3 in Algorithm 4), which also converges fast to the optimal solution like in Algorithm 2.

C. Algorithm performance in IEEE 14-bus and 30-bus systems

In this section, we evaluate the convergence performance of Algorithm 2 and Algorithm 3 for medium-sized systems using the IEEE 14-bus system and 30-bus system that have richer connectivity. The IEEE 14-bus system and the IEEE 30-bus system correspond to a portion of the Midwestern U.S. Electric Power System as of February 1962 and December 1961 respectively [37]. As originally there is no energy storage, we

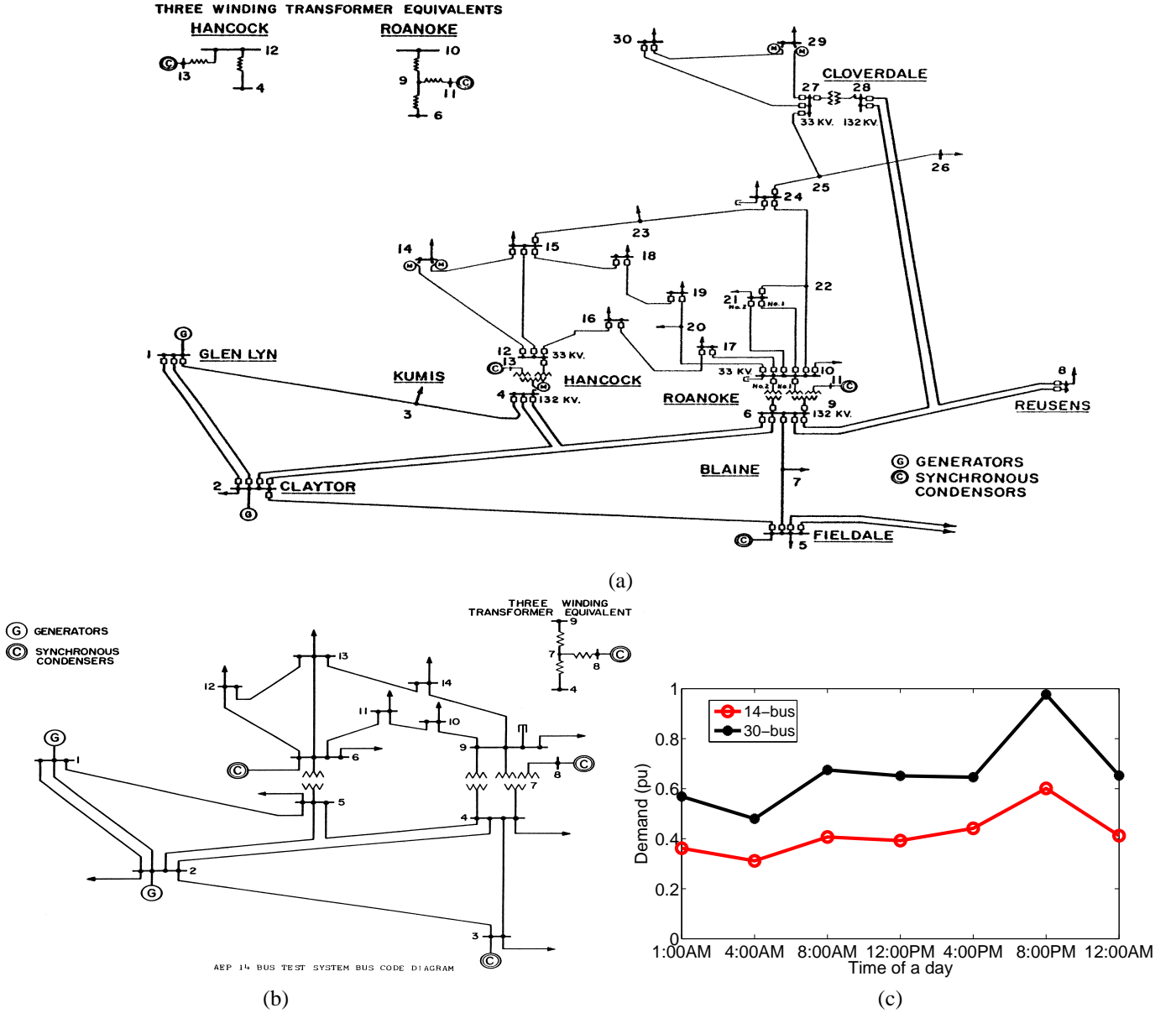


Fig. 13. The IEEE test systems and the time-varying demand profiles used for numerical evaluations. (a) The topology of the IEEE 30-bus system. (b) The topology of the IEEE 14-bus system. (c) The total demand profiles created by using typical hourly demands averaged over 14 and 30 different days in January 2009 of the domestic customers in USA [38] for the IEEE 14-bus and 30-bus system respectively.

have used the network topology and select appropriate values for the other system parameters (e.g., line admittance) and battery parameters to model energy storage. In both systems, Buses 1 and 2 are the generation buses. Moreover, similarly to [20], the demand profiles for each bus are created by using typical hourly demands averaged over 14 and 30 different days in January 2009 of the domestic customers in USA for the IEEE 14-bus and 30-bus systems respectively (available in [38]). Figure 13(a) and Figure 13(b) show the topology of the IEEE 30-bus system and the IEEE 14-bus system respectively. Figure 13(c) plots the scaled aggregated demand profiles used in these two test systems, and the time period of a day is divided into seven intervals (i.e., $t = 1, 2, \dots, 7$) and the demand profile starts at 1:00AM.

By appropriately choosing the stepsizes, we plot the conver-

gence behavior of Algorithms 2 and 3. Figure 14 and Figure 15 show the convergence of both algorithms in the IEEE 14-bus system and the IEEE 30-bus system respectively. In the top two sub-figures of Figure 14, we show the convergence of Algorithm 2 at three buses (Buses 1, 4 and 14) for two time intervals ($t = 1$ and 4). In the bottom two sub-figures of Figure 14, we show the convergence of Algorithm 3 at three buses (Buses 2, 6 and 12) for two time intervals ($t = 3$ and 6). Similarly, in the top two sub-figures of Figure 15, we show the convergence of Algorithm 2 at three buses (Buses 1, 15 and 30) for the time intervals $t = 1$ and 4. In the bottom two sub-figures of Figure 15, we show the convergence of Algorithm 3 at three buses (Buses 9, 14 and 23) for the time intervals $t = 1$ and 4). From Figures 14 and Figure 15, we see that the algorithms can converge to the optimal solution typically in

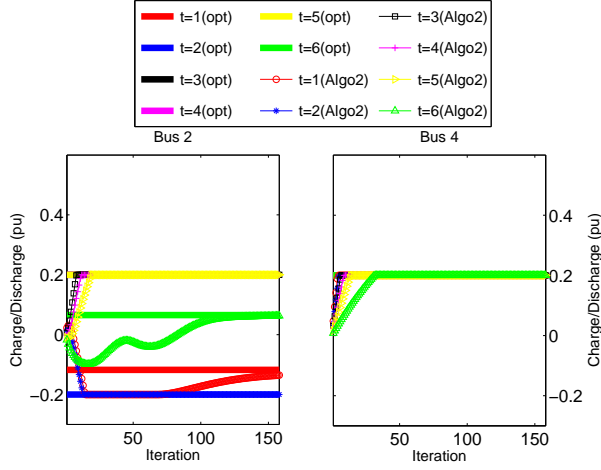


Fig. 11. Illustration of convergence for the storage dynamics in Algorithm 2 in the 5-bus system.

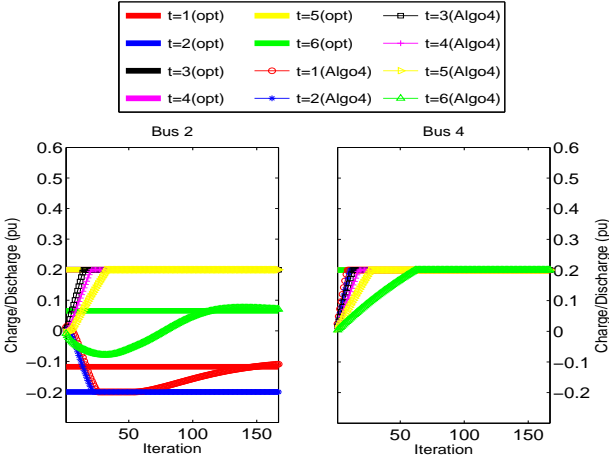


Fig. 12. Illustration of convergence for the storage dynamics in Algorithm 4 in the 5-bus system.

hundreds of iterations.

VIII. CONCLUSION

In this paper, we studied a dynamic OPF problem in a purely resistive power network with energy storage. By exploiting a recently-discovered zero duality gap result for the static OPF, we decomposed the dynamic OPF into simpler subproblems using an indirect dual-dual decomposition method and an alternative primal-dual decomposition method for distributed optimization by message passing algorithms. The optimization decomposition technique also revealed an interesting coupling over space and time of the Lagrange dual solution (to be interpreted as power prices) and the energy storage dynamics. We conducted extensive numerical evaluations on the performance of the proposed algorithms and to demonstrate the equilibrium load profile smoothing feature in the dynamic OPF problem.

An important direction for future research is the extension for the general alternating current (AC) network. Further studies that unify the various approximation methodologies for solving the AC OPF (e.g., the DC OPF linearization) and special cases (purely resistive network as given here or

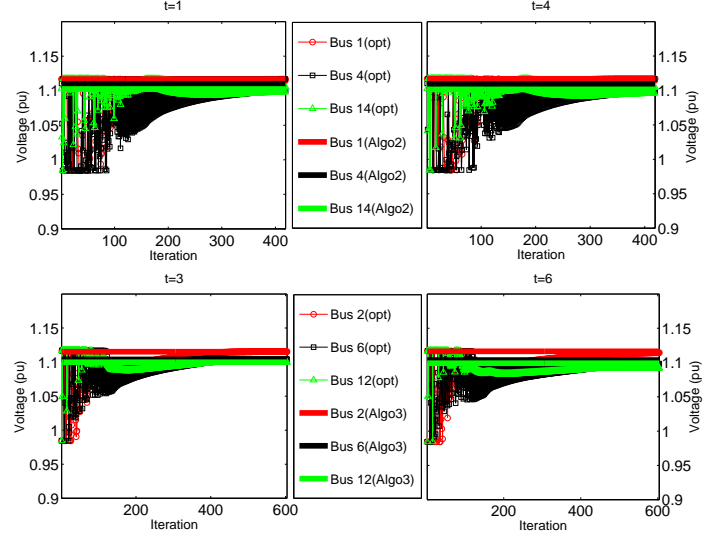


Fig. 14. Illustration of convergence of Algorithms 2 and 3 in the IEEE 14-bus system. The top two subfigures show the convergence of Algorithms 2 for buses 1, 4, and 14 over two time intervals. The bottom two subfigures show the convergence of Algorithms 3 for buses 2, 6, and 12 over two time intervals.

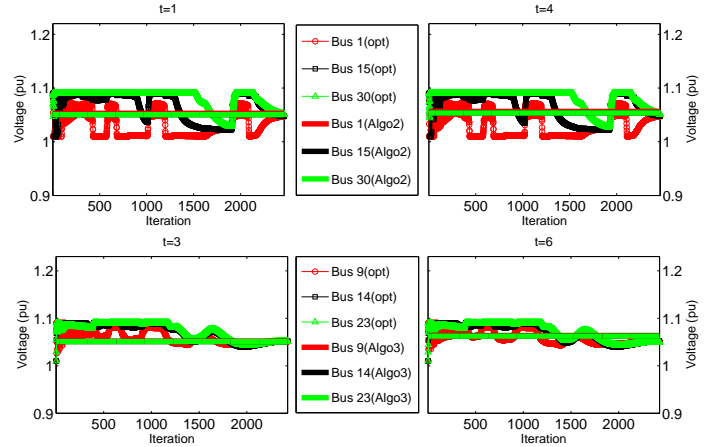


Fig. 15. Illustration of convergence of Algorithms 2 and 3 in the IEEE 30-bus system. The top two subfigures show the convergence of Algorithms 2 for buses 1, 15, and 30 over two time intervals. The bottom two subfigures show the convergence of Algorithms 3 for buses 9, 14, and 23 over two time intervals.

AC with special network topology) can allow us to better understand which configuration of data, parameters and design variables are important to load profile smoothing, and to shed further insights on the interaction between the physics of power flow and energy storage dynamics.

Another promising direction for future research is to examine other important constraints in power system such as security constraints, stability constraints and chance constraints. These coupling constraints can lead to new decomposability structures and algorithms for a smart grid. Incorporating the causality constraint (whenever the demand cannot be predicted and resources can only be optimized using only currently available knowledge and not those in the future time) is even harder to analyze, but would be more practically useful. A possibility is to study the extent to which online algorithm design can be used for distributed computation and to benchmark optimality

and performance with the offline setting.

REFERENCES

- [1] J. P. Barton and D. G. Infield. Energy storage and its use with intermittent renewable energy. *IEEE Trans. on Energy Conversion*, 19(2):441–448, 2004.
- [2] P. Denholm and M. Hand. Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. *Energy Policy*, 39(3):1817–1830, 2011.
- [3] N. Chen, C. W. Tan, and T. Quek. Electric vehicle charging in smart grid: Optimality and valley-filling algorithms. *IEEE Journal of Selected Topics in Signal Processing*, submitted, 2014.
- [4] L. Gan, U. Topcu, and S. H. Low. Optimal decentralized protocol for electric vehicle charging. *IEEE Trans. on Power Systems*, 28(2):940–951, 2013.
- [5] J. Carpentier. Contribution to the economic dispatch problem. *Bulletin de la Societe Francoise des Electriciens*, 3(8):431–447, 1962. In French.
- [6] J. D. Glover, M. S. Sarma, and T. J. Overbye. *Power System Analysis and Design*. Cengage Learning, 5th edition, 2011.
- [7] A. J. Conejo and J. A. Aguado. Multi-area coordinated decentralized DC optimal power flow. *IEEE Trans. on Power Systems*, 13(4):1272–1278, 1998.
- [8] B. H. Kim and R. Baldick. A comparison of distributed optimal power flow algorithms. *IEEE Trans. on Power Systems*, 15(2):599–604, 2000.
- [9] H. Wei, H. Sasaki, J. Kubokawa, and J. Yokoyama. An interior point nonlinear programming for optimal power flow problems with a novel data structure. *IEEE Trans. on Power Systems*, 13(3):870–877, 1998.
- [10] B. Stott, J. Jardim, and O. Alsac. DC power flow revisited. *IEEE Trans. on Power Systems*, 24(3):1290–1300, 2009.
- [11] F. Dorfler and F. Bullo. Novel insights into lossless AC and DC power flow. *Proc. of IEEE PES General Meeting*, 2013.
- [12] S. H. Low. Convex relaxation of optimal power flow part I: Formulations and equivalence. *IEEE Trans. on Control of Network Systems*, 1(1):15–27, 2014.
- [13] S. H. Low. Convex relaxation of optimal power flow part II: Exactness. *IEEE Trans. on Control of Network Systems*, to appear, 2014.
- [14] J. Lavaei, A. Rantzer, and S. H. Low. Power flow optimization using positive quadratic programming. *Proc. of IFAC World Congress*, 2011.
- [15] J. Lavaei and S. H. Low. Zero duality gap in optimal power flow problem. *IEEE Trans. on Power Systems*, 27(1):92–107, 2012.
- [16] S. Bose, D. F. Gayme, S. H. Low, and K. M. Chandy. Quadratically constrained quadratic programs on acyclic graphs with application to power flow. *arXiv:1203.5599*, 2012.
- [17] J. Lavaei, D. Tse, and B. Zhang. Geometry of power flows and optimization in distribution networks. *IEEE Trans. on Power Systems*, 29(2):572–583, 2014.
- [18] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [19] A. Y. S. Lam, B. Zhang, and D. Tse. Distributed algorithms for optimal power flow problem. *Proc. of IEEE Conf. on Decision and Control (CDC)*, 2012.
- [20] D. F. Gayme and U. Topcu. Optimal power flow with large-scale storage integration. *IEEE Trans. on Power Systems*, 28(2):709–717, 2013.
- [21] M. Kraning, E. Chu, J. Lavaei, and S. Boyd. Dynamic network energy management via proximal message passing. *Foundations and Trends in Optimization*, 1(2):70–122, 2014.
- [22] C. W. Tan, D. W. H. Cai, and X. Lou. Resistive network optimal power flow: uniqueness and algorithms. *IEEE Trans. on Power Systems*, to appear, 2014.
- [23] D. K. Molzahn, J. T. Holzer, B. C. Lesieutre, and C. L. DeMarco. Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Trans. on Power Systems*, 28(4):3987–3998, 2014.
- [24] N. Alguacil and A. J. Conejo. Multiperiod optimal power flow using Benders decomposition. *IEEE Trans. on Power Systems*, 15(1):196–201, 2000.
- [25] S. Ghosh, D. A. Iancu, D. Katz-Rogozhnikov, D. T. Phan, and M. S. Squillante. Power generation management under time-varying power and demand conditions. *Proc. of IEEE PES General Meeting*, 2011.
- [26] K. M. Chandy, S. H. Low, U. Topcu, and H. Xu. A simple optimal power flow model with energy storage. *Proc. of IEEE Conf. on Decision and Control (CDC)*, 2010.
- [27] R. C. Burchett, H. H. Happ, and K. A. Wirgau. Large scale optimal power flow. *IEEE Trans. on Power Apparatus and Systems*, 10(10):3722–3730, 1982.
- [28] R. Romero and A. Monticelli. A hierarchical decomposition approach for transmission network expansion planning. *IEEE Trans. on Power Systems*, 9(1):373–380, 2000.
- [29] L. S. Lasdon. *Optimization Theory for Large Systems*. Macmillan, New York, 1970.
- [30] D. P. Palomar and M. Chiang. Alternative distributed algorithms for network utility maximization: framework and applications. *IEEE Trans. on Automatic Control*, 52(12):2254–2269, 2007.
- [31] N. Li, L. Chen, and S. H. Low. Optimal demand response based on utility maximization in power networks. *Proc. of IEEE PES General Meeting*, 2011.
- [32] S. Kim and M. Kojima. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. *Computational Optimization and Applications*, 26:143–154, 2003.
- [33] A. Simsek, A. E. Ozdaglar, and D. Acemoglu. Uniqueness of generalized equilibrium for box constrained problems and applications. *Proc. of Allerton*, 2005.
- [34] A. Simsek, A. E. Ozdaglar, and D. Acemoglu. Generalized Poincare-Hopf theorem for compact nonsmooth regions. *Mathematics of Operations Research*, 32(1):193–214, 2007.
- [35] X. Lou and C. W. Tan. Convex relaxation and decomposition in large resistive power networks with energy storage. *Proc. of IEEE SmartGridComm*, 2013.
- [36] N. Z. Shor. *Minimization Methods for Non-Differentiable Functions*. Springer-Verlag, Berlin, 1985.
- [37] Power System Test Case Archive. University of Washington, Seattle, WA. Available: <http://www.ee.washington.edu/research/pstca/>.
- [38] SCE website. Available: <https://www.sce.com/wps/portal/home/regulatory/load-profiles/>.