Two-Timescale Stochastic Dispatch of Smart Distribution Grids

Luis M. Lopez-Ramos, *Member, IEEE,* Vassilis Kekatos, *Senior Member, IEEE,*
Antonio G. Marques, *Senior Member, IEEE,* and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Smart grids should efficiently integrate stochastic renewable resources while effecting voltage regulation. Energy management is challenging since it is a multistage problem where decisions are not all made at the same timescale and must account for the variability during real-time operation. The joint dispatch of slow- and fast-timescale controls in a smart distribution grid is considered here. The substation voltage, the energy exchanged with a main grid, and the generation schedules for small diesel generators have to be decided on a slow timescale; whereas optimal photovoltaic inverter setpoints are found on a more frequent basis. While inverter and looser voltage regulation limits are imposed at all times, tighter bus voltage constraints are enforced on the average or in probability, thus enabling more efficient renewable integration. Upon reformulating the two-stage grid dispatch as a stochastic convex-concave problem, two distribution-free schemes are put forth. An average dispatch algorithm converges provably to the optimal two-stage decisions via a sequence of convex quadratic programs. Its non-convex probabilistic alternative entails solving two slightly different convex problems and is numerically shown to converge. Numerical tests on real-world distribution feeders verify that both schemes yield lower costs over competing alternatives.**

***Index Terms*—Multistage economic dispatch, voltage regulation, stochastic approximation, convex-concave problem.**

I. Introduction

With increasing renewable generation, energy management of power distribution grids is becoming a computationally challenging task. Solar energy from photovoltaic (PV) units can change significantly over one-minute intervals. The power inverters found in PV units can be commanded to curtail active power generation or adjust their power factor within seconds [1], [2]. At a slower timescale, distribution grid operators exchange energy with the main grid hourly or on a 10-minute basis, and may experience cost penalties upon deviating from energy market schedules [3]. Moreover, voltage regulation equipment and small diesel generators potentially installed in microgrids respond at the same slower timescale. As a result, comprehensive designs to optimize such diverse tasks call for *multistage* smart grid dispatch solutions.

Spurred by demand-response programs and the use of PV inverters to accomplish various grid tasks [4], *single-stage* dispatch schemes for distribution grids have been an active area of research. Power inverters can be controlled using localized rules for voltage regulation, see e.g., [5], [6], [7], [8]. Assuming two-way communication between buses and the utility operator, dispatching a distribution system can be posed as an optimal power flow (OPF) problem. Centralized schemes use nonlinear program solvers [9]; or rely on convex relaxations of the full AC model of balanced [10], [11], or unbalanced grids [12]. Distributed solvers with reduced computational complexity have been devised in [13], [14], [15].

Nevertheless, the efficient and secure operation of distribution grids involves decisions at different timescales. A dynamic programming approach for a two-stage dispatch is suggested in [10]: The taps of voltage regulators are set on a slow timescale and remain fixed for consecutive shorter time slots over which elastic loads are dispatched; yet the flexibility of loads is assumed known *a priori*. Alternatively, centrally computed OPF decisions can be communicated to buses at a slow timescale, while on a faster timescale, PV power electronics are adjusted to optimally track variations in renewable generation and demand [16], [17]. Relying on approximate grid models and ignoring the effect of uncertainty on the dispatch of slow-responding units, the latter schemes yield a partially decentralized real-time allocation of the power flows across fast-responding units.

Multistage dispatching under uncertainty is routinely used in transmission systems and microgrids [18]. Robust approaches find optimal slow-timescale decisions for the worst-case fast-timescale outcome; see [19] and references therein. To avoid the conservativeness of robust schemes, probabilistic approaches postulate a probability density function (pdf) for demand, wind generation, and system contingencies to find day-ahead grid schedules [20], [21]. The risk-limiting dispatch framework adjusts multistage decisions as the variance of the random variables involved decreases while approaching actual time [3]. Decisions can be efficiently calculated only for convenient pdfs for a network-constrained risk-limiting dispatch and under congestion assumptions [22]. As a third alternative, sample average approximation (SAA) approaches yield optimal slow-timescale decisions using samples drawn from the postulated pdf; see e.g., [23], [19]. Recent works

Manuscript submitted August 3, 2016; revised December 2, 2016; accepted January 11, 2017. Date of publication DATE; date of current version DATE. Paper no. TSG.01019.2016.

This work was supported by the Spanish Ministry of Education FPU Grant AP2010-1050; CAM Grant S2013/ICE-2933; MINECO Grant TEC2013-41604-R; and NSF grants 1423316, 1442686, 1508993, and 1509040.

L.-M. Lopez-Ramos and Antonio G. Marques are with the Dept. of Signal Theory and Communications, King Juan Carlos Univ., Fuenlabrada, Madrid 28943, Spain. V. Kekatos is with the ECE Dept., Virginia Tech, Blacksburg, VA 24061, USA. G. B. Giannakis is with the Digital Technology Center and the ECE Dept., University of Minnesota, Minneapolis, MN 55455, USA. Emails: luismiguel.lopez@urjc.es, kekatos@vt.edu, antonio.garcia.marques@urjc.es, georgios@umn.edu.

Color versions of one or more of the figures is this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier XXXXXX

impose limits on the probability of undesirable events, either relying on convex approximation of chance constraints [24], or via the (sample-based) scenario approximation approach [25] to reduce computations; e.g., [26].

Focusing to distribution grids, PV inverters could be overloaded sporadically in time and across buses to accommodate solar fluctuations and prevent overvoltages [27]. The spatiotemporal overloading of power system components (such as inverters, bus voltages, line flows) could thus constitute an additional means for integrating renewables in smart grids. Nonetheless, ensuring that overloading occurs sparingly couples decisions across time. The single-stage scheme of [28] finds optimal PV setpoints while limiting time averages of overloaded quantities. The latter approach has been also adopted in [29] for dispatching a transmission system in a day-ahead/real-time market setup under load shedding.

Jointly dispatching slow- and fast-timescale grid resources under *average* or *probabilistic* constraints over fast-timescale decisions is considered here. Our contribution is three-fold. First, Section III formulates a two-stage grid dispatch as a convex-concave problem: The expected cost over a slow control period is minimized, while looser voltage limits are satisfied at all times and tighter voltage limits are enforced on the average or in probability. Second, upon adapting the stochastic saddle-point approximation scheme from [30], the provably convergent algorithm in Sec. IV provides optimal slow-timescale decisions for the average-constrained formulation. Different from SAA approaches, this stochastic approximation (SA) scheme processes random samples one at a time to improve computational efficiency. Third, in the case of non-convex *probabilistic* constraints, an algorithm solving two similar convex problems for each second stage is put forth in Sec. V. Although the expected cost enjoys zero-duality gap [31], the overall two-stage dispatch is not convex-concave, which explains why the algorithm's performance is validated numerically. Both schemes require only samples of loads and solar generation (rather than their joint pdfs), and can rely either on an approximate, or a convexified grid model. Numerical tests using the linearized distribution flow model on 56- and 123-bus feeders corroborate the validity of our findings in Sec. VI.

Regarding *notation*, lower-(upper-)case boldface letters denote column vectors (matrices), with the only exception of the power flow vectors, which are uppercase. Calligraphic letters are used to denote sets. Symbol $^\top$ denotes transposition, while $\mathbf{0}$ and $\mathbf{1}$ are the all-zeros and all-ones vectors of appropriate dimensions. The indicator function $\mathbb{1}\{\cdot\}$ equals 1 when its argument is true, and 0 otherwise. A diagonal matrix with the entries of vector $\mathbf{x}$ on its main diagonal is denoted by $\mathrm{dg}(\mathbf{x})$. The operator $[\cdot]_+$ projects its argument onto the positive orthant; $\mathbb{E}[\cdot]$ denotes expectation and $\Pr\{\cdot\}$ probability.

II. Problem Formulation

Consider a distribution grid whose energy needs are procured by distributed renewable generation, distributed conventional (small diesel) generators, and the main grid. The distribution grid operator aims at serving load at the minimum cost while respecting voltage regulation and network constraints. Energy is exchanged with the main grid at wholesale electricity prices through the feeder bus. To effectively integrate stochastic renewable generation, the focus here is on short-term grid dispatch. To that end, the distribution grid is operated at two timescales: a slower timescale corresponds to 5- or 10-min real-time energy market intervals, while the inverters found in PVs are controlled at a faster timescale of say 10-sec intervals. One period of the slower timescale is comprised by T faster time slots indexed by $t = 1, \ldots, T$.

The grid is operated as a radial network with $N+1$ buses rooted at the substation bus indexed by $n = 0$. The distribution line feeding bus n is also indexed by n for $n = 1, \ldots, N$. Let $p_{n,t}$ and $q_{n,t}$ denote respectively the net active and reactive power injections at bus n and slot t; the N-dimensional vectors $\mathbf{p}_t$ and $\mathbf{q}_t$ collect the net injections at all buses except for the substation. Diesel generators are dispatched at the slower timescale to generate $\mathbf{p}^d$ throughout the subsequent T slots at unit power factor. During slot t, PVs can contribute solar generation up to $\overline{\mathbf{p}}_t^r$ that is modeled as a random process. Smart inverters perform active power curtailment and reactive power compensation by following the setpoints $\mathbf{p}_t^r$ and $\mathbf{q}_t^r$ commanded by the utility operator. Load demands $\mathbf{p}_t^l$ and $\mathbf{q}_t^l$ are also modeled as random processes. To simplify the exposition, $(\mathbf{p}_t^l, \mathbf{q}_t^l)$ are assumed inelastic and known at the beginning of slot t; although elastic loads can be incorporated without any essential differences. The operator buys a power block p_0^a from the main grid at the slow timescale, which can be adjusted to $p_{0,t} := p_0^a + p_{0,t}^\delta$ in actual time.

Voltage regulation is effected by controlling (re)active power injections at slot t. Let $v_{n,t}$ denote the squared voltage magnitude at bus n and slot t, and $\mathbf{v}_t$ the vector collecting $\{v_{n,t}\}_{n=1}^N$. The substation voltage v_0^a is controlled at the slower timescale [10], while voltage magnitudes at all buses must adhere to voltage regulation standards, e.g., ANSI C84.1 and EN50160 in [32], [33]. These standards differentiate between a narrower voltage regulation range denoted here by $\mathcal{V}_A$ in which voltages should lie most of the time; and a wider range $\mathcal{V}_B$ (with $\mathcal{V}_A \subset \mathcal{V}_B$) whom voltages should not exceed at any time. One of the goals of this work is to leverage this flexibility to design dispatch schemes that: i) guarantee that voltages lie in $\mathcal{V}_B$ at all times, while ii) they belong to $\mathcal{V}_A$ in a stochastic fashion. To this end, two alternative schemes are presented, the difference between them being how constraint ii) is formulated. The first scheme guarantees that the *average* voltage lies in $\mathcal{V}_A$, whereas the second one maintains the *probability* of under-/over-voltage at a specified low value.

A. Grid modeling

To account for voltage and network limitations, the distribution grid is captured by the approximate linear distribution flow (LDF) model [34]. To briefly review this model, let $\mathbf{r}$ and $\mathbf{x}$ be accordingly the vectors of line resistances and reactances across lines. Define also the branch-bus incidence matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{N \times (N+1)}$ whose (i,j)-th entry is

$$\tilde{A}_{ij} = \begin{cases} +1 & , \text{ if } j-1 \text{ is the source bus of line } i \\ -1 & , \text{ if } j-1 \text{ is the destination bus of line } i \\ 0 & , \text{ otherwise.} \end{cases} \quad (1)$$

Partition $\tilde{\mathbf{A}}$ into its first column and the *reduced* branch-bus incidence matrix $\mathbf{A}$ as $\tilde{\mathbf{A}} = [\mathbf{a}_0 \ \mathbf{A}]$. Ignoring line losses, the LDF model asserts that the vectors of active and reactive line power flows at time t can be approximated by

$$\mathbf{P}_t = \mathbf{F}^\top \mathbf{p}_t \text{ and } \mathbf{Q}_t = \mathbf{F}^\top \mathbf{q}_t \tag{2}$$

where $\mathbf{F} := \mathbf{A}^{-1}$. Moreover, the squared voltage magnitudes can be expressed as [34], [6]

$$\mathbf{v}_t = 2\mathbf{R}\mathbf{p}_t + 2\mathbf{X}\mathbf{q}_t + v_0^d \mathbf{1} \tag{3}$$

where $\mathbf{R} := \mathbf{F}\mathrm{dg}(\mathbf{r})\mathbf{F}^\top$ and $\mathbf{X} := \mathbf{F}\mathrm{dg}(\mathbf{x})\mathbf{F}^\top$. The LDF model applies to both radial and meshed networks and, different from the so termed DC power grid model, it does not ignore line resistances [35]. It can be derived by assuming that voltage magnitudes are close to unity and voltage angle differences across neighboring buses are small. Alternatively, it can be obtained upon linearizing power injections at the flat voltage profile [36].

Let us define the voltage regulation regions

$$\mathcal{V}_A := \{\mathbf{v} : \ \underline{v}_A \mathbf{1} \leq \mathbf{v} \leq \overline{v}_A \mathbf{1}\} \tag{4a}$$

$$\mathcal{V}_B := \{\mathbf{v} : \ \underline{v}_B \mathbf{1} \leq \mathbf{v} \leq \overline{v}_B \mathbf{1}\} \tag{4b}$$

with $\overline{v}_B \geq \overline{v}_A$ and $\underline{v}_B \leq \underline{v}_A$. Compliance with $\mathcal{V}_A$ can be imposed either on the average as $\mathbb{E}_t[\mathbf{v}_t] \in \mathcal{V}_A$, or in probability as $\Pr\{\mathbf{v}_t \in \mathcal{V}_A\} \geq 1-\alpha$ for some small α. Either way, safe grid operation requires that $\mathbf{v}_t \in \mathcal{V}_B$ at all times t. Within the optimization horizon, the random processes involved (demand and renewable generation) can be assumed ergodic, i.e., their time averages converge to their ensemble averages. For this reason, voltage constraints pertaining to $\mathcal{V}_A$ will be referred to as *ergodic*.

According to (2), if $\mathbf{f}_n$ is the n-th column of $\mathbf{F}$, the squared power flow on line n can be written as $P_{n,t}^2 = \mathbf{p}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{p}_t$ and $Q_{n,t}^2 = \mathbf{q}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{q}_t$. Imposing the upper limit $\overline{S}_n$ on the apparent flow on line n is thus expressed as the convex quadratic constraint

$$\mathbf{p}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{q}_t \leq \overline{S}_n^2. \tag{5}$$

Although losses have been dropped in (2), upon assuming that voltage magnitudes are close to unity, active power losses can be approximated as [37]

$$\sum_{n=1}^{N} r_n (P_{n,t}^2 + Q_{n,t}^2) = \mathbf{P}_t^\top \mathrm{dg}(\mathbf{r}) \mathbf{P}_t + \mathbf{Q}_t^\top \mathrm{dg}(\mathbf{r}) \mathbf{Q}_t.$$

Using (2), the latter can be equivalently expressed as $\mathbf{p}_t^\top \mathbf{R} \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{R} \mathbf{q}_t$, so the active power injection at the substation is approximately

$$p_{0,t} = -\mathbf{1}^\top \mathbf{p}_t + \mathbf{p}_t^\top \mathbf{R} \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{R} \mathbf{q}_t \tag{6}$$

Regarding smart inverters, the tuple $(p_{n,t}^r, q_{n,t}^r)$, which denotes the power injection from the inverter located on bus n at slot t, should belong to the feasible set

$$\Omega_{n,t} := \big\{(p_{n,t}^r, q_{n,t}^r) : 0 \leq p_{n,t}^r \leq \overline{p}_{n,t}^r, \tag{7a}$$

$$|q_{n,t}^r| \leq \phi_n p_{n,t}^r, \tag{7b}$$

$$(p_{n,t}^r)^2 + (q_{n,t}^r)^2 \leq \overline{s}_n^2 \big\} \tag{7c}$$

that is random and time-variant due to the variability of $\overline{p}_{n,t}^r$. Constraint (7a) limits the active power generation according to the available solar power; constraint (7b) enforces the lower limit $\cos(\arctan(\phi_n))$ on the power factor (lagging or leading); and (7c) limits the inverter apparent power.

B. Operation costs

If PV owners are compensated at price $\boldsymbol{\pi}$ for the active power surplus they inject into the distribution grid, the related utility cost at slot t is $C_{\mathrm{PV}}(\mathbf{p}_t^r) := \boldsymbol{\pi}^\top [\mathbf{p}_t^r - \mathbf{p}_t^l]_+$ with $[\cdot]_+ := \max\{0, \cdot\}$ applied entrywise on vector $\mathbf{p}_t^r - \mathbf{p}_t^l$. The diesel generation cost is represented by $C_{\mathrm{D}}(\mathbf{p}^d)$. Regarding energy transactions with the main grid, the power block p_0^a bought in advance is charged at a fixed and known price β. Deviating from p_0^a by $p_{0,t}^\delta$ at slot t is charged at

$$C^t(p_{0,t}^\delta) := \gamma_b [p_{0,t}^\delta]_+ - \gamma_s [-p_{0,t}^\delta]_+ \tag{8}$$

for known prices (γ_b, γ_s). To avoid arbitrage, it is assumed that $0 < \gamma_s < \beta < \gamma_b$; see e.g., [3], [22]. Then, the deviation charge can also be expressed as $C^t(p_{0,t}^\delta) = \max\{\gamma_b p_{0,t}^\delta, \gamma_s p_{0,t}^\delta\}$, which is certainly convex [19].

C. Optimal grid dispatch

Depending on the way compliance with voltage regulation region $\mathcal{V}_A$ is enforced, two grid dispatch formulations are developed next. Commencing with the *average dispatch*, the optimal grid operation is posed as

$$\mathsf{P}_a^* := \min \ C_{\mathrm{D}}(\mathbf{p}^d) + \beta p_0^a + \mathbb{E}_t\left[C^t(p_{0,t}^\delta) + C_{\mathrm{PV}}(\mathbf{p}_t^r)\right] \tag{9a}$$

$$\text{s.to:} \ \mathbf{p}_t = \mathbf{p}_t^r - \mathbf{p}_t^l + \mathbf{p}^d \tag{9b}$$

$$\mathbf{q}_t = \mathbf{q}_t^r - \mathbf{q}_t^l \tag{9c}$$

$$p_{0,t} = p_0^a + p_{0,t}^\delta \tag{9d}$$

$$p_{0,t} \geq -\mathbf{1}^\top \mathbf{p}_t + \mathbf{p}_t^\top \mathbf{R} \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{R} \mathbf{q}_t \tag{9e}$$

$$\mathbf{p}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{p}_t + \mathbf{q}_t^\top \mathbf{f}_n \mathbf{f}_n^\top \mathbf{q}_t \leq \overline{S}_n, \ \forall n \in \mathcal{N} \tag{9f}$$

$$\underline{\mathbf{p}}^d \leq \mathbf{p}^d \leq \overline{\mathbf{p}}^d \tag{9g}$$

$$(p_{n,t}^r, q_{n,t}^r) \in \Omega_{n,t}, \ \forall n \in \mathcal{N} \tag{9h}$$

$$\underline{v}_0 \leq v_0^a \leq \overline{v}_0 \tag{9i}$$

$$\mathbf{v}_t = 2\mathbf{R}\mathbf{p}_t + 2\mathbf{X}\mathbf{q}_t + v_0^a \mathbf{1} \tag{9j}$$

$$\mathbf{v}_t \in \mathcal{V}_B \tag{9k}$$

$$\mathbb{E}_t[\mathbf{v}_t] \in \mathcal{V}_A \tag{9l}$$

$$\text{over} \ \ v_0^a, p_0^a, \mathbf{p}^d, \{\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta\}_{t=1}^T.$$

The slow-timescale variables $\{v_0^a, p_0^a, \mathbf{p}^d\}$ are set in advance, and remain fixed throughout the T subsequent control slots over which the fast-timescale variables $\{\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta\}_{t=1}^T$ are implemented. The latter variables depend on the randomness of slot t as well as slow-timescale decisions.

Alternatively to (9), optimal grid operation can be posed as a *probabilistic dispatch* that is identical to (9) with the exception that (9l) is replaced by the probabilistic constraint

$$\Pr\{\mathbf{v}_t \notin \mathcal{V}_A\} \leq \alpha \tag{10}$$

for some small parameter $\alpha > 0$, say $\alpha = 0.05$. The optimal cost for the probabilistic dispatch will be denoted by P_p^*.

The objective function in (9a) involves the cost of energy dispatched at the slow timescale plus the average fast-timescale energy management cost. Nodal (re)active power balance is ensured via (9b)–(9c). Constraint (9e) accounts for active power losses. Since the cost in (9a) is non-decreasing with respect to $(p_{0,t}^{\delta}, p_0^a)$, relaxing (6) to the convex inequality in (9e) does not incur loss of optimality. Constraint (9f) limits line apparent power flows based on (5). Constraints (9i)–(9l) are voltage regulation constraints: In detail, (9j) relates squared voltage magnitudes to power injections [cf. (3)]; (9i) constraints the substation bus voltage; and (9k) constraints voltages in $\mathcal{V}_B$. While (9l) maintains the average voltage magnitudes in $\mathcal{V}_A$, its alternative in (10) limits the probability of voltage magnitudes being outside $\mathcal{V}_A$.

A pertinent question is which of the two proposed dispatch formulations is to be preferred. The probabilistic formulation is more sophisticated and aligned with voltage regulation standards, emerging as the default option. However, as it will be explained in Section V, enforcing even the single grid-level probabilistic constraint in (10) gives rise to a non-convex problem, which comes with computational challenges. The average dispatch does not suffer from these problems, which can be critical in scenarios where the duration of the slow period is short and the optimization has to be frequently re-run. Furthermore, when renewable generation and loads vary only slightly during a slow period and/or local control loops are in place, enforcing probabilistic guarantees may not be justified and the simpler average constraints suffice.

D. Convexified AC grid model

Although (9) relies on the approximate LDF model, it can be readily customized to the exact AC power flow model [34]. Upon introducing the optimization variable $\boldsymbol{\ell}_t := [\ell_{1,t} \ \ldots \ \ell_{N,t}]^\top$ with the squared line current magnitudes, constraints (9e)–(9f) should be substituted respectively by

$$p_{0,t} \geq -\mathbf{1}^\top \mathbf{p}_t + \mathbf{1}^\top \boldsymbol{\ell}_t \tag{11a}$$

$$P_{n,t}^2 + Q_{n,t}^2 \leq \bar{S}_n, \ \forall n \in \mathcal{N}. \tag{11b}$$

Constraint (9j) defining $\mathbf{v}_t$ should be replaced by

$$\mathbf{v}_t = 2\mathbf{F}\mathrm{dg}(\mathbf{r})\mathbf{P}_t + 2\mathbf{F}\mathrm{dg}(\mathbf{x})\mathbf{Q}_t + v_0^a\mathbf{1} \tag{12}$$

and variable $\boldsymbol{\ell}_t$ is linked to power flows and voltages through the additional constraints:

$$\mathbf{P}_t = \mathbf{F}^\top \mathbf{p}_t + \mathbf{F}^\top \mathrm{dg}(\mathbf{r})\boldsymbol{\ell}_t \tag{13a}$$

$$\mathbf{Q}_t = \mathbf{F}^\top \mathbf{q}_t + \mathbf{F}^\top \mathrm{dg}(\mathbf{x})\boldsymbol{\ell}_t \tag{13b}$$

$$P_{n,t}^2 + Q_{n,t}^2 \leq v_{\pi_n,t}\ell_{n,t}, \qquad \forall n \in \mathcal{N} \tag{13c}$$

where π_n is the parent bus of bus n. In fact, constraint (13c) constitutes a relaxation, since in the actual grid model it is satisfied with equality [10]. Nevertheless, the relaxation has been shown to be exact in radial grids and under different conditions; see [38] for details. Critical for the ensuing sections is that the differences between the formulation in (11)–(13) and that for the LDF model pertain to the fast-timescale operation, whereas the slow-timescale formulation and the constraints coupling slow with fast timescale variables remain intact.

III. Problem Analysis

To facilitate algorithmic developments, the problem in (9) is expressed in a compact form next. Collect the slow-timescale variables in vector $\mathbf{z}^\top := [v_0^a, p_0^a, \mathbf{p}^d]$; the fast-timescale variables at slot t in $\mathbf{y}_t^\top := [\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^{\delta}]$; and the random variables involved at slot t in $\boldsymbol{\xi}_t^\top := [\overline{\mathbf{p}}_t^r, \mathbf{p}_t^l, \mathbf{q}_t^l]$.

The constraints in (9) can be classified into four groups:
(i) Constraints involving fast-timescale variables only, such as (9c), (9f), (9h), and (9k), that will be abstracted as $\mathbf{y}_t \in \mathcal{Y}_t$.
(ii) Constraints (9g) and (9i) that involve slow-timescale variables only, and they will be denoted as $\mathbf{z} \in \mathcal{Z}$.
(iii) The *linear* constraints (9b), (9d), and (9j), coupling slow- and fast-timescale variables as well as random variables. These constraints are collectively expressed as $\mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t$ for appropriate matrices $\mathbf{K}$, $\mathbf{B}$, and $\mathbf{H}$.
(iv) The ergodic constraints (9l) and (10) depend on the voltage sequence $\{\mathbf{v}_t\}_{t=1}^T$, hence coupling decisions across time. A substantial difference between (9l) and (10) is that the latter is a non-convex constraint.

If the exact grid model of Section II-D is used, the additional variables $\mathbf{P}_t$, $\mathbf{Q}_t$, and $\boldsymbol{\ell}_t$ are added, and set $\mathcal{Y}_t$ in *(i)* is modified to incorporate (11)–(13). Under these considerations, the two dispatch problems can be compactly rewritten as

$$\mathsf{P}_{(a,p)}^* := \min_{\mathbf{z}, \{\mathbf{y}_t\}_{t=1}^T} \ f(\mathbf{z}) + \mathbb{E}_t\left[g_t(\mathbf{y}_t)\right] \tag{14a}$$

$$\text{s.to:} \ \ \mathbf{z} \in \mathcal{Z} \tag{14b}$$

$$\mathbf{y}_t \in \mathcal{Y}_t \qquad \forall t \tag{14c}$$

$$\mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \qquad \forall t \tag{14d}$$

$$\mathbb{E}_t\left[\mathbf{h}(\mathbf{y}_t)\right] \leq \mathbf{0} \tag{14e}$$

where $f(\mathbf{z}) := C_{\mathrm{D}}(\mathbf{p}^d) + \beta p_0^a$ and $g_t(\mathbf{y}_t) := C^t(p_{0,t}^{\delta}) + C_{\mathrm{PV}}(\mathbf{p}_t^r)$. For the average dispatch, the optimal cost in (14) is P_a^* and the function in (14e) is $\mathbf{h}(\mathbf{y}_t) = [\mathbf{v}_t - \overline{v}_A\mathbf{1}, \underline{v}_A\mathbf{1} - \mathbf{v}_t]$. For the probabilistic dispatch, the optimal cost is P_p^* and the function in (14e) is $\mathbf{h}(\mathbf{y}_t) = \mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} - \alpha$.

The optimal values for the slow-timescale variables $\mathbf{z}$ must be decided in advance. Once the optimal $\mathbf{z}$ is found, it remains fixed over the slow-timescale interval. The fast-timescale decisions $\mathbf{y}_t(\mathbf{z})$ for slot t depend on $\mathbf{z}$, while the subscript t indicates their dependence on the realization $\boldsymbol{\xi}_t$. Both the average and the probabilistic dispatch are stochastic programming problems with recourse [3]. Their costs can be decomposed as $\mathsf{P}_{(a,p)}^* = \min_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z}) + G_{(a,p)}(\mathbf{z})$, where the so termed *expected recourse function* is defined as

$$G_{(a,p)}(\mathbf{z}) := \min_{\{\mathbf{y}_t \in \mathcal{Y}_t\}} \ \mathbb{E}_t\left[g_t(\mathbf{y}_t)\right] \tag{15a}$$

$$\text{s.to:} \ \ \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \qquad \forall t \tag{15b}$$

$$\mathbb{E}_t\left[\mathbf{h}(\mathbf{y}_t)\right] \leq \mathbf{0}. \tag{15c}$$

Since problem (15) depends on $\mathbf{z}$, its minimizer can be written as $\{\mathbf{y}_t^*(\mathbf{z})\}_{t=1}^T$ and the recourse function as $G_{(a,p)}(\mathbf{z}) = \mathbb{E}_t[g_t(\mathbf{y}_t^*(\mathbf{z}))]$. The ensuing two sections solve the average and the probabilistic dispatches.

IV. Average Dispatch Algorithm

This section tackles problem (14) with the ergodic constraint in (14e), for which $\mathbf{h}(\mathbf{y}_t) = [\underline{v}_A\mathbf{1} - \mathbf{v}_t, \mathbf{v}_t - \overline{v}_A\mathbf{1}]$. Although convex, problem (14) is challenging due to the coupling across $\{\mathbf{y}_t\}_{t=1}^T$ and between fast- and slow-timescale variables. Dual decomposition is adopted to resolve the coupling across $\{\mathbf{y}_t\}_{t=1}^T$. The partial Lagrangian function for (15) is

$$L_a(\{\mathbf{y}_t\}, \boldsymbol{\nu}) := \mathbb{E}_t\left[g_t(\mathbf{y}_t) + \boldsymbol{\nu}^\top \mathbf{h}(\mathbf{y}_t)\right] \tag{16}$$

with the entries of $\boldsymbol{\nu}$ being the multipliers associated with the upper and lower per-bus constraints in (14e). The corresponding dual function is

$$D_a(\boldsymbol{\nu}; \mathbf{z}) := \min_{\{\mathbf{y}_t \in \mathcal{Y}_t\}} L_a(\{\mathbf{y}_t\}, \boldsymbol{\nu}) \quad \text{s.to: } \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \quad \forall t. \tag{17}$$

Observe that after dualizing, the minimization in (17) is separable over the realizations $\{\boldsymbol{\xi}_t\}$. Precisely, the optimal fast-timescale variable for fixed $(\boldsymbol{\nu}, \mathbf{z})$ and for a specific realization $\boldsymbol{\xi}_t$ can be found by solving:

$$\mathbf{y}_t^*(\boldsymbol{\nu}, \mathbf{z}) \in \arg\min_{\mathbf{y}_t \in \mathcal{Y}_t} \; g_t(\mathbf{y}_t) + \boldsymbol{\nu}^\top \mathbf{h}(\mathbf{y}_t) \tag{18a}$$

$$\text{s.to: } \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t. \tag{18b}$$

For future reference, let us also define $\boldsymbol{\lambda}_t^*(\boldsymbol{\nu}, \mathbf{z})$ as the optimal Lagrange multiplier associated with (18b). If $\boldsymbol{\nu}$ is partitioned as $\boldsymbol{\nu}^\top = [\underline{\boldsymbol{\nu}}^\top, \overline{\boldsymbol{\nu}}^\top]$ with $\underline{\boldsymbol{\nu}}$ corresponding to constraint $\mathbb{E}_t[\mathbf{v}_t] \geq \underline{v}_A\mathbf{1}$ and $\overline{\boldsymbol{\nu}}$ to $\mathbb{E}_t[\mathbf{v}_t] \leq \overline{v}_A\mathbf{1}$, then (18) simplifies to

$$\mathbf{y}_t^*(\boldsymbol{\nu}, \mathbf{z}) \in \operatorname{argmin} \; C^t(p_{0,t}^\delta) + C_{\text{PV}}(\mathbf{p}_t^r) + (\overline{\boldsymbol{\nu}} - \underline{\boldsymbol{\nu}})^\top \mathbf{v}_t \tag{19}$$
$$\text{s.to: (9b)} - \text{(9f)}, \text{(9h)}, \text{(9j)}, \text{(9k)}$$
$$\text{over } \{\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{p}_t^r, \mathbf{q}_t^r, p_{0,t}, p_{0,t}^\delta\}$$

and can be solved as a convex quadratic program. If the relaxed AC grid model of Section II-D is used, then (19) becomes a second-order cone program (SOCP) which is also convex. Given the optimal pair $(\boldsymbol{\nu}^*, \mathbf{z}^*)$, the optimal fast-timescale variables $\mathbf{y}_t$ can be thus found for any $\boldsymbol{\xi}_t$.

Back to finding the optimal primal and dual slow-timescale variables, note that the dual problem associated with (17) is

$$\boldsymbol{\nu}^* := \arg\max_{\boldsymbol{\nu} \geq \mathbf{0}} \; D_a(\boldsymbol{\nu}; \mathbf{z}). \tag{20}$$

Duality theory asserts that (20) is a convex problem. Moreover, assuming a strictly feasible point exists for (15), strong duality implies that $G_a(\mathbf{z}) = D_a(\boldsymbol{\nu}^*, \mathbf{z})$. Due to the latter, the original problem in (14) can be transformed to:

$$\min_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z}) + G_a(\mathbf{z}) = \min_{\mathbf{z} \in \mathcal{Z}} \{f(\mathbf{z}) + \max_{\boldsymbol{\nu} \geq \mathbf{0}} D_a(\boldsymbol{\nu}; \mathbf{z})\} \tag{21a}$$

$$= \min_{\mathbf{z} \in \mathcal{Z}} \max_{\boldsymbol{\nu} \geq \mathbf{0}} \tilde{f}_a(\boldsymbol{\nu}, \mathbf{z}) \tag{21b}$$

where the auxiliary function $\tilde{f}_a$ is defined as:

$$\tilde{f}_a(\boldsymbol{\nu}, \mathbf{z}) := f(\mathbf{z}) + D_a(\boldsymbol{\nu}; \mathbf{z}). \tag{22}$$

Being a dual function, $D_a(\boldsymbol{\nu}; \mathbf{z})$ is a concave function of $\boldsymbol{\nu}$. At the same time, $D_a(\boldsymbol{\nu}; \mathbf{z})$ is a perturbation function with respect to $\mathbf{z}$; and hence, it is a convex function of $\mathbf{z}$ [39]. Recall that $f(\mathbf{z})$ is a convex function of $\mathbf{z}$ too. Therefore, the auxiliary function $\tilde{f}_a(\boldsymbol{\nu}, \mathbf{z})$ is convex in $\mathbf{z}$ and concave in $\boldsymbol{\nu}$. Because of the randomness of $\{\boldsymbol{\xi}_t\}$, function $D_a(\boldsymbol{\nu}; \mathbf{z})$ in (17) is stochastic. Consequently, problem (21b) is a stochastic convex-concave saddle point problem [39], [30].

Algorithm 1 Average Dispatch Algorithm (ADA)

1: Initialize $(\mathbf{z}^0, \boldsymbol{\nu}^0)$.
2: **repeat** for $k = 0, 1, \ldots$
3: Draw sample $\boldsymbol{\xi}_k$.
4: Find $(\mathbf{y}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k), \boldsymbol{\lambda}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))$ by solving (18).
5: Update $(\mathbf{z}^{k+1}, \boldsymbol{\nu}^{k+1})$ via (25).
6: Compute sliding averages $(\tilde{\mathbf{z}}^k, \tilde{\boldsymbol{\nu}}^k)$ through (26).
7: **until** convergence of $(\tilde{\mathbf{z}}^k, \tilde{\boldsymbol{\nu}}^k)$.
8: **Output** $\mathbf{z}^* = \tilde{\mathbf{z}}^k$ and $\boldsymbol{\nu}^* = \tilde{\boldsymbol{\nu}}^k$.

To solve (21b), we rely on the stochastic saddle-point approximation method of [30]. The method involves the subgradient of $\tilde{f}_a$ with respect to $\mathbf{z}$, and its supergradient with respect to $\boldsymbol{\nu}$. Upon viewing $D_a(\boldsymbol{\nu}, \mathbf{z})$ in (17) as a perturbation function of $\mathbf{z}$, the subgradient of $\tilde{f}_a$ with respect to $\mathbf{z}$ is [39]

$$\partial_{\mathbf{z}} \tilde{f}_a = \partial_{\mathbf{z}} f(\mathbf{z}) + \mathbf{K}^\top \mathbb{E}_t[\boldsymbol{\lambda}_t^*(\boldsymbol{\nu}, \mathbf{z})]. \tag{23}$$

By definition of the dual function, the supergradient of $\tilde{f}_a$ with respect to $\boldsymbol{\nu}$ is

$$\partial_{\boldsymbol{\nu}} \tilde{f}_a = \mathbb{E}_t[\mathbf{h}(\mathbf{y}_t^*(\boldsymbol{\nu}, \mathbf{z}))]. \tag{24}$$

The stochastic saddle point approximation method of [30] involves primal-dual subgradient iterates with the expectations in (23)–(24) being replaced by their instantaneous estimates based on a single realization $\boldsymbol{\xi}_k$. Precisely, the method involves the iterates over k:

$$\boldsymbol{\nu}^{k+1} := [\boldsymbol{\nu}^k + \operatorname{dg}(\boldsymbol{\mu}_k)\mathbf{h}(\mathbf{y}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))]_+ \tag{25a}$$

$$\mathbf{z}^{k+1} := [\mathbf{z}^k - \operatorname{dg}(\boldsymbol{\epsilon}_k)(\partial_z f(\mathbf{z}^k) + \mathbf{K}^\top \boldsymbol{\lambda}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))]_{\mathcal{Z}} \tag{25b}$$

where the operator $[\cdot]_{\mathcal{Z}}$ projects its argument onto $\mathcal{Z}$; and vectors $\boldsymbol{\mu}_k = \boldsymbol{\mu}_0/\sqrt{k}$ and $\boldsymbol{\epsilon}_k = \boldsymbol{\epsilon}_0/\sqrt{k}$ collect respectively the primal and dual step sizes for positive $\boldsymbol{\mu}_0$ and $\boldsymbol{\epsilon}_0$. At every iteration k, the method draws a realization $\boldsymbol{\xi}_k$ and solves (18) for the tuple $(\boldsymbol{\xi}_k, \boldsymbol{\nu}^k, \mathbf{z}^k)$ to acquire $(\mathbf{y}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k), \boldsymbol{\lambda}_k^*(\boldsymbol{\nu}^k, \mathbf{z}^k))$ and perform the primal-dual updates in (25). The method finally outputs the sliding averages of the updates as:

$$\tilde{\mathbf{z}}^k := \big(\textstyle\sum_{i=\lceil k/2 \rceil}^k \mathbf{z}^i/\sqrt{i}\big)/\big(\sum_{i=\lceil k/2 \rceil}^k 1/\sqrt{i}\big) \tag{26a}$$

$$\tilde{\boldsymbol{\nu}}^k := \big(\textstyle\sum_{i=\lceil k/2 \rceil}^k \boldsymbol{\nu}^i/\sqrt{i}\big)/\big(\sum_{i=\lceil k/2 \rceil}^k 1/\sqrt{i}\big). \tag{26b}$$

The proposed scheme converges to the value $\tilde{f}_a(\boldsymbol{\nu}^*, \mathbf{z}^*)$ obtained at a saddle point $(\boldsymbol{\nu}^*, \mathbf{z}^*)$ asymptotically in the number of iterations k [30, Sec. 3.1].

Upon convergence of the iterates in (26), the slow-timescale variables $\mathbf{z}^*$ have been derived together with the optimal Lagrange multiplier $\boldsymbol{\nu}^*$ related to constraint (15c). The grid operator can implement $\mathbf{z}^*$, and the fast-timescale decisions $\mathbf{y}_t^*$ for a realization $\boldsymbol{\xi}_t$ can be found by solving (19). The average dispatch algorithm (ADA) is summarized as Alg. 1.

V. Probabilistic Dispatch Algorithm

The probabilistic version of problem (14) is considered next. Here, the ergodic constraint (14e) reads $h(\mathbf{y}_t) = \mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} - \alpha$. Despite the non-convexity of the probabilistic constraint, (15) can still be solved optimally. However, optimality for (14) cannot be guaranteed. A heuristic solution is detailed next by adapting the solution of Sec. IV.

To that end, dual decomposition is used here as well. If ν is the scalar Lagrange multiplier associated with constraint (14e), the partial Lagrangian function for (15) is now $L_p(\{\mathbf{y}_t\},\nu) := \mathbb{E}_t\left[g_t(\mathbf{y}_t) + \nu(\mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} - \alpha)\right]$. The corresponding dual function, fast-timescale problem, and dual problem are defined analogously to (17), (18), and (20). The indicator function renders $L_p(\{\mathbf{y}_t\},\nu)$ non-convex. Surprisingly enough though, under the practical assumption that $\{\boldsymbol{\xi}_t\}$ follows a continuous pdf, problem (15) enjoys zero duality gap; see [31, Th. 1].

The additional challenge here is the non-convexity of the Lagrangian minimization:

$$\mathbf{y}_t^*(\nu,\mathbf{z}) \in \arg\min_{\mathbf{y}_t\in\mathcal{Y}_t} \; g_t(\mathbf{y}_t) + \nu\mathbb{1}\{\mathbf{v}_t \notin \mathcal{V}_A\} \quad \text{s.to:} \;\; \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t. \tag{27}$$

Because the indicator function takes only the values $\{0,1\}$ however, the solution to (27) can be found by solving a pair of slightly different convex problems. The first problem is

$$\mathbf{y}_{t,A}^*(\mathbf{z}) \in \arg\min_{\mathbf{y}_t\in\mathcal{Y}_t} \; g_t(\mathbf{y}_t) \tag{28a}$$

$$\text{s.to:} \;\; \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t \tag{28b}$$

$$\mathbf{v}_t \in \mathcal{V}_A \tag{28c}$$

whereas the second problem ignores constraint $\mathbf{v}_t \in \mathcal{V}_A$ as

$$\mathbf{y}_{t,B}^*(\mathbf{z}) \in \arg\min_{\mathbf{y}_t\in\mathcal{Y}_t} \; g_t(\mathbf{y}_t) \tag{29a}$$

$$\text{s.to:} \;\; \mathbf{K}\mathbf{z} + \mathbf{B}\mathbf{y}_t = \mathbf{H}\boldsymbol{\xi}_t. \tag{29b}$$

From the point of view of (27), if the voltages in $\mathbf{y}_{t,B}^*(\mathbf{z})$ do not belong to $\mathcal{V}_A$, the solution to the second problem will incur an additional cost quantified by ν. Observe that neither problem (28) nor (29) depend on ν, while their complexity is similar to the one problem (18). Suppose that (28) and (29) have been solved and let $\boldsymbol{\lambda}_{t,A}^*(\mathbf{z})$ and $\boldsymbol{\lambda}_{t,B}^*(\mathbf{z})$ denote the optimal multipliers associated with (28b) and (29b), respectively. Then, problem (27) can be neatly tackled by identifying two cases:

(c1) If $g_t(\mathbf{y}_{t,A}^*(\mathbf{z})) > g_t(\mathbf{y}_{t,B}^*(\mathbf{z})) + \nu$, then $\mathbf{y}_{t,B}^*(\mathbf{z})$ is a minimizer of (27) as well and voltages are allowed to lie outside $\mathcal{V}_A$. In this case, set $\mathbf{y}_t^*(\nu,\mathbf{z}) := \mathbf{y}_{t,B}^*(\mathbf{z})$ and $\boldsymbol{\lambda}_t^*(\nu,\mathbf{z}) := \boldsymbol{\lambda}_{t,B}^*(\mathbf{z})$. This case includes instances where problem (28) is infeasible for which $g_t(\mathbf{y}_{t,A}^*(\mathbf{z})) = \infty$.

(c2) If $g_t(\mathbf{y}_{t,A}^*(\mathbf{z})) \leq g_t(\mathbf{y}_{t,B}^*(\mathbf{z})) + \nu$, then $\mathbf{y}_{t,A}^*(\mathbf{z})$ minimizes (27) too and voltages lie within $\mathcal{V}_A$. In this case, set $\mathbf{y}_t^*(\nu,\mathbf{z}) := \mathbf{y}_{t,A}^*(\mathbf{z})$ and $\boldsymbol{\lambda}_t^*(\nu,\mathbf{z}) := \boldsymbol{\lambda}_{t,A}^*(\mathbf{z})$.

Case (c2) covers also instances where $\mathbf{v}_{t,B}^*(\mathbf{z})$ happens to lie in $\mathcal{V}_A$. In these particular instances, $\mathbf{y}_{t,B}^*(\mathbf{z})$ serves as a minimizer of (28) too. Then, it follows that $g_t(\mathbf{y}_{t,A}^*(\mathbf{z})) = g_t(\mathbf{y}_{t,B}^*(\mathbf{z})) \leq g_t(\mathbf{y}_{t,B}^*(\mathbf{z})) + \nu$ for $\nu \geq 0$. This implies that one can solve (29) first and, if $\mathbf{v}_{t,B}^*(\mathbf{z}) \in \mathcal{V}_A$, there is no need to solve problem (28).

Algorithm 2 Probabilistic Dispatch Algorithm (PDA)

1: Initialize $(\mathbf{z}^0,\nu^0)$.
2: **repeat** for $k = 0, 1, \ldots$
3: Draw sample $\boldsymbol{\xi}_k$.
4: Find $(\mathbf{y}_{k,B}^*(\nu^k,\mathbf{z}^k), \boldsymbol{\lambda}_{k,B}^*(\nu^k,\mathbf{z}^k))$ by solving (29).
5: Set $\mathbf{y}_t^*(\nu,\mathbf{z}) := \mathbf{y}_{t,B}^*(\mathbf{z})$ and $\boldsymbol{\lambda}_t^*(\nu,\mathbf{z}) := \boldsymbol{\lambda}_{t,B}^*(\mathbf{z})$.
6: **if** $\mathbf{v}_{k,B}^*(\mathbf{z}) \notin \mathcal{V}_A$, **then** find $\mathbf{y}_{k,A}^*(\nu^k,\mathbf{z}^k)$ and $\boldsymbol{\lambda}_{k,A}^*(\nu^k,\mathbf{z}^k)$ by solving (28).
7: **if** $g_t(\mathbf{y}_{t,A}^*(\mathbf{z})) \leq g_t(\mathbf{y}_{t,B}^*(\mathbf{z})) + \nu$, **then** set $\mathbf{y}_t^*(\nu,\mathbf{z}) := \mathbf{y}_{t,A}^*(\mathbf{z})$ and $\boldsymbol{\lambda}_t^*(\nu,\mathbf{z}) := \boldsymbol{\lambda}_{t,A}^*(\mathbf{z})$.
8: **end if**
9: **end if**
10: Update $(\mathbf{z}^{k+1},\nu^{k+1})$ via (30).
11: Compute sliding averages $(\tilde{\mathbf{z}}^k,\tilde{\nu}^k)$ through (26).
12: **until** convergence of $(\tilde{\mathbf{z}}^k,\tilde{\nu}^k)$.
13: **Output** $\mathbf{z}^* = \tilde{\mathbf{z}}^k$ and $\nu^* = \tilde{\nu}^k$.

To find the optimal slow-timescale variables under the probabilistic dispatch, the stochastic primal-dual iterations of Sec. IV are adapted here as

$$\nu^{k+1} := [\nu^k + \mu_k(\mathbb{1}\{\mathbf{v}_k^*(\nu^k,\mathbf{z}^k) \notin \mathcal{V}_A\} - \alpha)]_+ \tag{30a}$$

$$\mathbf{z}^{k+1} := [\mathbf{z}^k - \mathrm{dg}(\boldsymbol{\epsilon}_k)(\partial_z f(\mathbf{z}^k) + \mathbf{K}^\top\boldsymbol{\lambda}_k^*(\nu^k,\mathbf{z}^k))]_{\mathcal{Z}}. \tag{30b}$$

The probabilistic dispatch algorithm (PDA) is tabulated as Alg. 2. At every fast-timescale iteration, PDA solves (29) and possibly (28). Since the optimizations tasks (28)–(29) are structurally similar to (18), PDA has at most twice the per-iteration complexity of ADA. Because function $G_p(\mathbf{z})$ is not necessarily convex, the iterates in (30) are not guaranteed to converge to a minimizer of (14). The practical performance of PDA in finding $\mathbf{z}^*$ is numerically validated in Sec. VI.

VI. Numerical Tests

The proposed grid dispatches were tested on a 56-bus Southern California Edison (SCE) distribution feeder [11]. 5-MW PVs were added on buses 44 and 50; both with 6-MVA inverters enabling power factors as low as 0.83 (leading or lagging) at full solar generation. The prices for the energy exchange with the main grid were $\beta = 37$ \$/MWh; $\gamma_b = 45$ \$/MWh, and $\gamma_s = 19$ \$/MWh. Diesel generators with capacity $\overline{p}_n^d = 0.5$ MW were sited on buses 10, 18, 21, 30, 36, 43, 51, and 55. The cost of diesel generation was $C_{\mathrm{D}}(\mathbf{p}^d) = \sum_{n=1}^{N}(30p_n^d + 15(p_n^d)^2)$ \$/h with $\mathbf{p}^d$ expressed in MW. Apparent power flows were limited to 7 MVA. The voltage operation limits were set to $\underline{v}_A = 0.98^2$, $\overline{v}_A = 1.02^2$, $\underline{v}_B = 0.97^2$, and $\overline{v}_B = 1.03^2$, expressed in pu with respect to a voltage base of 12 kV. (Re)active nodal loads were Gaussian distributed with the nominal load of the SCE benchmark as mean value, and standard deviation of 0.2 times the nominal load. The solar energy generated at each PV was drawn uniformly between 0.5 and 1 times the actual power PV rating.

The simulations presented next have been run using the LDF model. The LDF model is computationally less complex than the relaxed AC grid model of Section II-D, which is advantageous when many instances of the fast-variation scale have to be solved. Our tests show the LDF model is 33%

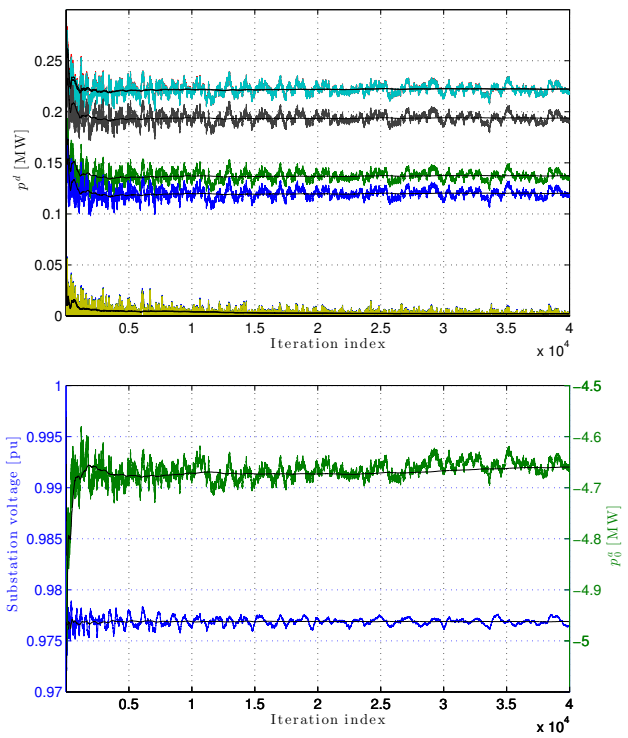


Fig. 1. Convergence of primal variables for ADA: (top) diesel generation; (bottom) substation voltage v_0 (left y-axis) and energy exchange p_0^a (right y-axis). Sliding averages of optimization variables are depicted too.

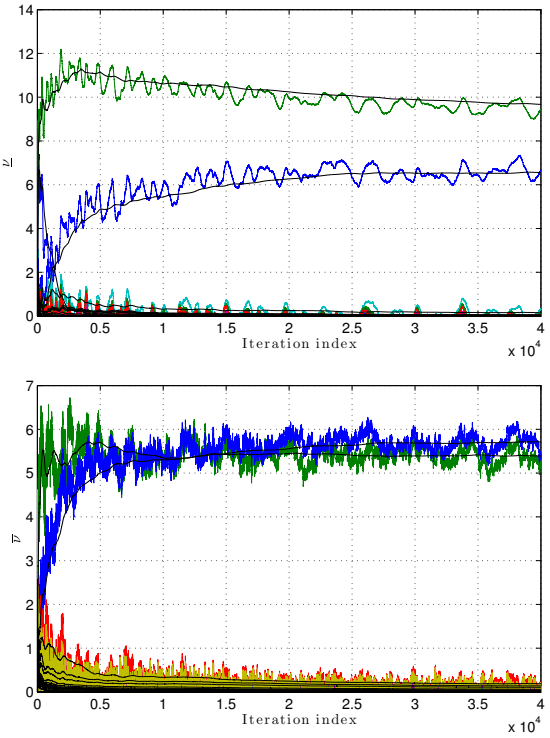


Fig. 2. Convergence of dual variables for ADA: (top) dual variables associated with average lower voltage limits for all buses; and (bottom) dual variables associated with average upper voltage limits for all buses. Sliding averages of optimization variables are depicted too.

faster, while it incurs 10% higher cost compared to the SOCP relaxation. Such numbers are consistent with those observed for other problems [6], [28]. In any case, the findings presented next are valid for both models and also for cases where the LDF model is adopted only for finding the slow-timescale variables, while the exact/relaxed AC grid model is employed during the actual fast-timescale dispatch.

ADA was run with step sizes proportional to $1/\sqrt{k}$ with initial values $\epsilon_0^{v_0} = 4 \cdot 10^{-5}$, $\epsilon_0^{p_0} = 4 \cdot 10^{-1}$, $\epsilon_0^{p_d} = 6 \cdot 10^{-3}$, and $\mu_0 = 225$, to account for different dynamic ranges. The iterates for primal and dual variables as well as their corresponding sliding averages are depicted in Figs. 1 and 2. Primal and dual slow-timescale variables hover in a small range whose width diminishes with time. Their sliding averages converge asymptotically. The algorithm reaches a practically meaningful solution within 5,000 iterations. Buses 44 and 50 are prone to overvoltages since they host PVs, and buses 2 and 15 are prone to under-voltages; thus yielding non-zero dual variables for the average upper and lower voltage constraints, respectively.

PDA was tested using the same simulation setup for $\alpha = 0.05$ and $\mu_0 = 1$. Figure 3 shows the convergence of primal and dual variables, and the probability of voltages deviating from $\mathcal{V}_A$. Since we know that the per-iteration computation of PDA is at most twice that of ADA and the simulations show that the number of iterations required for PDA and ADA is similar, it then follows that the total computation time for PDA is at most twice that for ADA. Granted that the probabilistic constraint in (10) applies collectively to all buses, the under-/over-voltage probabilities on a per-bus basis is depicted in Fig. 4. The occurrences of overvoltage seem to be shared primarily among buses 40–56 which are neighboring to the PV buses 40 and 55. On the contrary, buses 10–16 being electrically far from both the substation and PVs, experience under-voltage with a small probability.

The effect of the average versus the probabilistic constraint on voltage magnitudes was evaluated next. After slow-timescale variables $\mathbf{z}$ had converged, fast-timescale variables $\mathbf{y}_t$ were calculated for 6,000 instances of $\boldsymbol{\xi}_t$ using both ADA and PDA. The histograms of the voltage magnitudes on two representative buses are presented in Fig. 5. Under PDA, the average voltage on bus 15 is slightly higher than the average voltage obtained by ADA. In exchange, the instantaneous value of the voltage on bus 15 stays within $\mathcal{V}_A$ with higher probability. A similar behavior is observed for the overvoltage instances on PV bus 40.

ADA and PDA were finally compared to three alternative schemes. The first two, henceforth called *approximate average* and *approximate probabilistic* dispatches, obtained $\mathbf{z}$ by setting loads and solar generation to their expected values, while variables $\boldsymbol{\nu}$ were calculated via dual stochastic subgradient, and $\{\mathbf{y}_t\}_{t=1}^T$ were found by solving either (18) or (27), depending on whether the setting is average or probabilistic. The third *deterministic* dispatch found $\mathbf{z}$ as the approximate schemes do, and $\{\mathbf{y}_t\}_{t=1}^T$ by enforcing $\mathbf{v}_t \in \mathcal{V}_A$ at all times. Note that the three proposed alternatives provide *feasible* solutions satisfying voltage regulation constraints. The five dispatches were tested under five scenarios: Scenario 1 is the setup described earlier. Scenario 2 involved the tighter

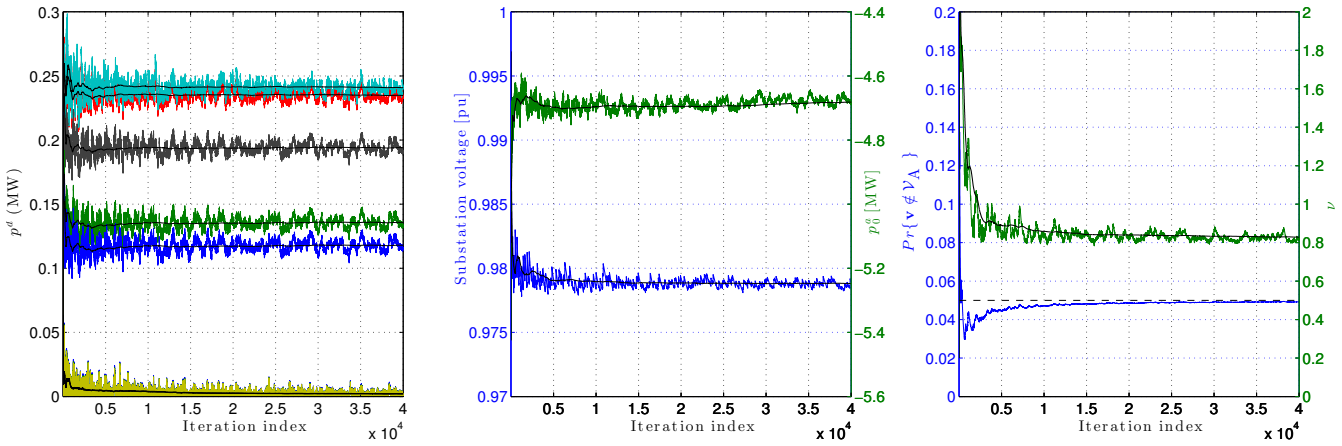


Fig. 3. Convergence for PDA: (left) diesel generation; (middle) substation voltage (left y-axis) and energy exchange p_0^a (right y-axis); and (right) dual variable related to probabilistic constraint (left y-axis) and under-/over-voltage probability (right y-axis). Sliding averages of optimization variables are shown too.

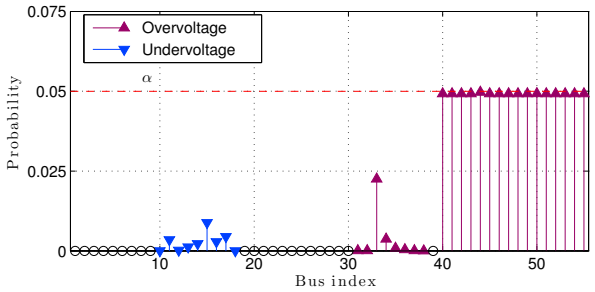


Fig. 4. Per-bus probability of under-/over-voltages.

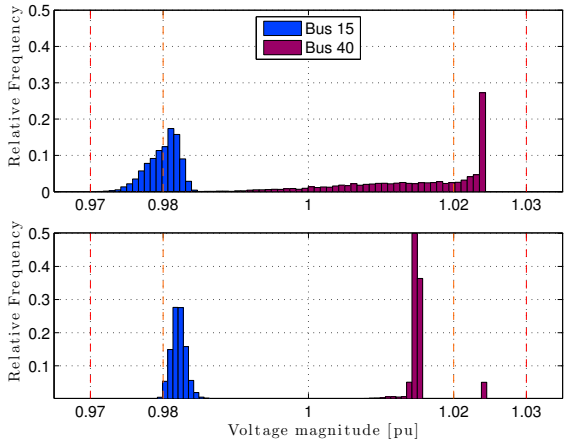


Fig. 5. Histograms of voltage magnitudes on buses 15 and 40 under ADA (top) and PDA (bottom). Dashed lines show regulation limits $\mathcal{V}_A$ and $\mathcal{V}_B$.

voltage limits $\underline{v}_A = 0.99^2$ and $\overline{v}_A = 1.01^2$. Scenarios 3, 4, and 5 were generated by scaling the mean value and the standard deviation for loads of scenario 1 by 0.5, 1.5, and 2, respectively. Figure 6 shows the expected operation costs for all five scenarios. ADA (PDA) yielded the lowest cost under all scenarios in the average (probabilistic) setting as expected. In all test cases, ADA yielded a slightly lower objective than PDA for $\alpha = 0.05$. The loss of optimality entailed by the approximate average and probabilistic schemes is due to the suboptimal choice of $\mathbf{z}$. The deterministic scheme entailed

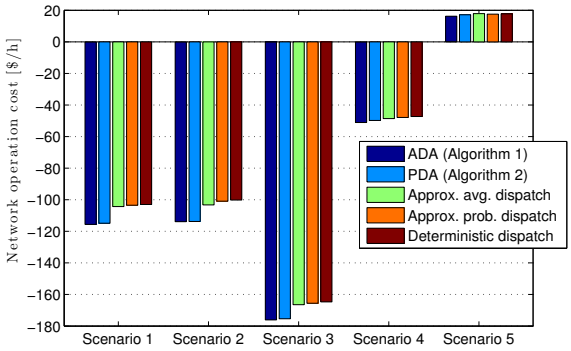


Fig. 6. Performance for ADA, PDA, approximate average, approximate probabilistic, and deterministic scheme.

an additional loss of optimality by preventing the occasional violation of $\mathcal{V}_A$.

To gain insights on the algorithm scalability, numerical tests were also performed using the IEEE 123-bus feeder [40]. PV systems were added at buses 92, 103, 119 and 122; and diesel generators at buses 3, 7, 32, 37, 39, 44, 51, 54, 56, 70, 74, 85, 92, 103, 119, and 122. Diesel generation costs and limits, and PV generation pdfs remained similar to the previous test. The nominal (re)active loads were perturbed by zero-mean Gaussian random variables having a standard deviation of 0.2 times the nominal value. The voltage operation limits were set to $\underline{v}_A = 0.99^2$, $\overline{v}_A = 1.01^2$, $\underline{v}_B = 0.98^2$, and $\overline{v}_B = 1.02^2$ (pu). ADA was run with step sizes proportional to $1/\sqrt{k}$, $\epsilon_0^{v_0} = 10^{-4}$, $\epsilon_0^{p_0} = 2 \cdot 10^{-2}$, $\epsilon_0^{p_d} = 10^{-3}$, and $\mu_0 = 400$. Figs. 7 and 8 show the convergence of the primal and dual variables. For this larger feeder, the algorithm reaches a practically meaningful solution after around 10,000 iterations and the average per-iteration computation increases by 90%.

VII. Conclusions

By nature of renewable generation, electromechanical component limits, and the manner markets operate, energy management of smart distribution grids involves decisions at

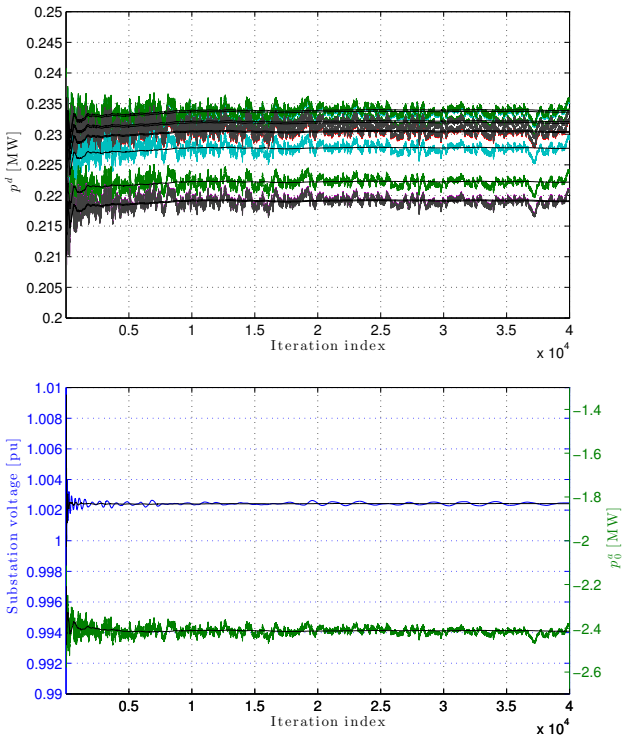


Fig. 7. Convergence of primal variables for ADA on the IEEE 123-bus feeder: (top) diesel generation; (bottom) voltage v_0 (left y-axis) and energy exchange p_0^a (right y-axis). Sliding averages of optimization variables are depicted too.

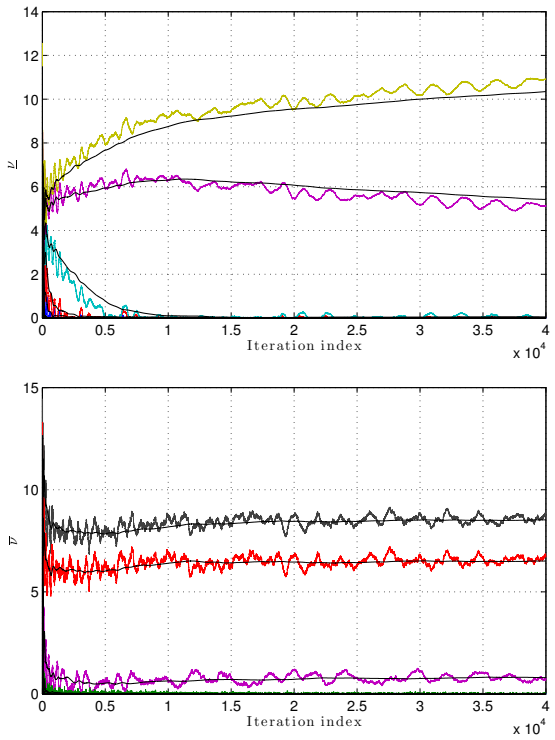


Fig. 8. Convergence of dual variables for ADA on the IEEE 123-bus benchmark: (top) dual variables related to average lower voltage limits for all buses; and (bottom) dual variables related to average upper voltage limits for all buses. Sliding averages of optimization variables are depicted too.

slower and faster timescales. Since slow-timescale controls remain fixed over multiple PV operation slots, decisions are coupled across time in a stochastic manner. To accommodate solar energy fluctuations, voltages have been allowed to be sporadically overloaded; hence introducing coupling of fast-timescale variables on the average or in probability. Average voltage constraints have resulted in a stochastic convex-concave problem, whereas non-convex probabilistic constraints were tackled using dual decomposition and convex optimization. Efficient algorithms for finding both slow and fast controls using only random samples have been put forth. Our two novel solvers converge in terms of the primal and dual variables, and have attained lower operational costs compared to deterministic alternatives. Although probabilistic constraints have been applied grid-wise, voltages on individual buses remained within limits. Enforcing probabilistic constraints on a per-bus basis, developing decentralized implementations, and including voltage regulators are interesting research directions.

REFERENCES

[1] E. Liu and J. Bebic, "Distribution system voltage performance analysis for high-penetration photovoltaics," National Renewable Energy Laboratory, Tech. Rep., Feb. 2008. [Online]. Available: https://www1.eere.energy.gov/solar/pdfs/42298.pdf

[2] P. M. S. Carvalho, P. F. Correia, and L. A. Ferreira, "Distributed reactive power generation control for voltage rise mitigation in distribution networks," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 766–772, May 2008.

[3] P. P. Varaiya, F. F. Wu, and J. W. Bialek, "Smart operation of smart grid: Risk-limiting dispatch," *Proc. IEEE*, vol. 99, no. 1, pp. 40–57, Jan. 2011.

[4] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proc. IEEE*, vol. 99, no. 6, pp. 1063–1073, Jun. 2011.

[5] B. Zhang, A. Dominguez-Garcia, and D. Tse, "A local control approach to voltage regulation in distribution networks," in *Proc. North American Power Symposium*, Manhattan, KS, Sep. 2013.

[6] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Voltage regulation algorithms for multiphase power distribution grids," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3913–3923, Sep. 2016.

[7] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Trans. Automat. Contr.*, vol. 60, no. 4, pp. 966–981, Apr. 2015.

[8] H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sep. 2016.

[9] S. Paudyal, C. A. Canizares, and K. Bhattacharya, "Optimal operation of distribution feeders in smart grids," *IEEE Trans. Ind. Applicat.*, vol. 10, no. 58, pp. 4495—4503, Oct. 2011.

[10] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *Proc. IEEE Power & Energy Society General Meeting*, San Diego, CA, Jul. 2012.

[11] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Trans. Automat. Contr.*, vol. 60, no. 1, pp. 72–87, Jan. 2015.

[12] E. Dall'Anese, S. V. Dhople, and G. B. Giannakis, "Optimal dispatch of photovoltaic inverters in residential distribution systems," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 487–497, Dec. 2014.

[13] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

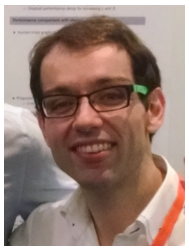
[14] B. Zhang, A. Y. S. Lam, A. D. Domínguez-García, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1714–1726, Jul. 2015.

[15] M. Bazrafshan and N. Gatsis, "Decentralized stochastic optimal power

flow in radial networks with distributed generation," *IEEE Trans. Smart Grid*, 2016, to appear.
[16] F. Dorfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Trans. Control of Network Systems*, vol. 3, no. 3, pp. 241–253, Sep. 2016.
[17] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *IEEE Trans. Smart Grid*, 2016 (to appear).
[18] A. J. Conejo, M. Carrión, and J. M. Morales, *Decision making under uncertainty in electricity markets*. Springer, 2010.
[19] Y. Zhang, N. Gatsis, and G. B. Giannakis, "Robust energy management for microgrids with high-penetration renewables," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 944–953, Oct. 2013.
[20] F. Bouffard, F. D. Galiana, and A. J. Conejo, "Market-clearing with stochastic security–Part I: Formulation," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1818–1826, Nov. 2005.
[21] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Rev.*, vol. 56, no. 3, pp. 461–495, Aug. 2014.
[22] B. Zhang, R. Rajagopal, and D. Tse, "Network risk limiting dispatch: Optimal control and price of uncertainty," *IEEE Trans. Automat. Contr.*, vol. 59, no. 9, pp. 2442–2456, Sep. 2014.
[23] F. Bouffard and F. D. Galiana, "Stochastic security for operations planning with significant wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 306–316, May 2008.
[24] T. Summers, J. Warrington, M. Morari, and J. Lygeros, "Stochastic optimal power flow based on convex approximations of chance constraints," in *Power Systems Computation Conf.*, Wroclaw, Poland, May 2014.
[25] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Trans. Automat. Contr.*, vol. 51, no. 5, pp. 742–753, May 2006.
[26] S. Bolognani and F. Dorfler, "Fast scenario-based decision making in unbalanced distribution networks," in *Power Systems Computation Conf.*, Genoa, Italy, Jun. 2016.
[27] M. S. El-Moursi, W. Xiao, and J. L. Kirtley, "Fault ride through capability for grid interfacing large scale PV power plants," *IET Gener. Transm. Dis.*, vol. 7, no. 9, pp. 1027–1036, Sep. 2013.
[28] G. Wang, V. Kekatos, A.-J. Conejo, and G. B. Giannakis, "Ergodic energy management leveraging resource variability in distribution grids," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4765–4775, Nov. 2016.
[29] L. M. Lopez-Ramos, V. Kekatos, A. G. Marques, and G. B. Giannakis, "Microgrid dispatch and price of reliability using stochastic approximation," in *Proc. IEEE Global Conf. on Signal and Inform. Process.*, Orlando, FL, Dec. 2015.
[30] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, "Robust stochastic approximation approach to stochastic programming," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1574–1609, 2009.
[31] A. Ribeiro and G. B. Giannakis, "Separation principles in wireless networking," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4488–4504, Sep. 2010.
[32] *C84.1-1995 Electric Power Systems and Equipment Voltage Ratings (60 Herz)*, ANSI Std., 2011.
[33] *EN 50160: Voltage Characteristics of Public Distribution Systems*, DIN Std., Jul. 2004.
[34] M. Baran and F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.
[35] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning and statistical estimation in distribution networks — Part I," 2015, (submitted). [Online]. Available: http://arxiv.org/abs/1501.04131
[36] S. Bolognani and F. Dorfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *Proc. Allerton Conf. on Comm., Control, and Computing*, Allerton, IL, Sep. 2015, pp. 402–409.
[37] P. Sulc, S. Backhaus, and M. Chertkov, "Optimal distributed control of reactive power via the alternating direction method of multipliers," *IEEE Trans. Energy Conversion*, vol. 29, no. 4, pp. 968–977, Dec. 2014.
[38] S. Low, "Convex relaxation of optimal power flow — Part II: exactness," *IEEE Trans. Control of Network Systems*, no. 99, pp. 1–1, 2014.
[39] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge University Press, 2004.
[40] W. H. Kersting, "Radial distribution test feeders," in *Proc. Power Engineering Society Winter Meeting*, vol. 2, 2001, pp. 908–912.

Luis M. Lopez-Ramos (M'16) received the B.Sc. degree (with highest honors) in Telecommunications Engineering in 2010 from King Juan Carlos University (URJC), Madrid, Spain; and the M.Sc. and Ph.D. degrees in multimedia and communications in 2012 and 2016, respectively, from URJC and Carlos III University of Madrid, Spain. During the fall of 2013 and 2014, he was a visiting scholar at the ECE Dept. of the Univ. of Minnesota. In February 2017, he joined the Univ. of Agder, Grimstad, Norway as a post-doctoral research fellow. Dr. Lopez-Ramos' research currently focuses on stochastic nonlinear programming, stochastic approximation and signal processing techniques, and their applications in wireless networks and power grids.

Vassilis Kekatos (SM'16) obtained his Diploma, M.Sc., and Ph.D. in Computer Science and Engr. from the Univ. of Patras, Greece, in 2001, 2003, and 2007, respectively. He was a recipient of a Marie Curie Fellowship during 2009-2012. During the summer of 2012, he worked for Windlogics Inc. After that, he was a research associate with the Dept. of Electrical and Computer Engr. of the Univ. of Minnesota. During 2014, he stayed with the Univ. of Texas at Austin and the Ohio State Univ. as a visiting researcher, and he received the postdoctoral career development award (honorable mention) by the Univ. of Minnesota. In August 2015, he joined the Dept. of Electrical and Computer Engr. of Virginia Tech as an Assistant Professor. His research focus is on optimization, learning, and management of future energy systems. He is currently serving as an Associate Editor of the IEEE Trans. on Smart Grid.



Antonio G. Marques (SM'13) received the Telecommunications Engineering degree and the Doctorate degree, both with highest honors, from the Carlos III Univ. of Madrid, Spain, in 2002 and 2007, respectively. In 2007, he became a faculty of the Dept. of Signal Theory and Communications, King Juan Carlos Univ., Madrid, Spain, where he currently develops his research and teaching activities as an Associate Professor. From 2005 to 2015, he held different visiting positions at the Univ. of Minnesota, Minneapolis. In 2015 and 2016 he was a Visiting Scholar at the Univ. of Pennsylvania.

His research interests lie in the areas of signal processing, communication theory, and networking. His current research focuses on stochastic resource allocation for wireless networks and smart grids, nonlinear network optimization, and signal processing for graphs. Dr. Marques has served the IEEE in a number of posts (currently, he is an Associate Editor of the IEEE Signal Process. Letters and a member of the Signal Process. Theory and Methods Technical Committee), and his work has been awarded in several conferences.

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. He was with the Univ. of Virginia from 1987 to 1998, and since 1999 he has been a professor with the Univ. of Minnesota, where he holds an Endowed Chair in Wireless Telecommunications, a Univ. of Minnesota McKnight Presidential Chair in ECE, 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 400 journal papers, 680 conference papers, 25 book chapters, two edited books and two research monographs (h-index 119). 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 28 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.