Stochastic Day Ahead Load Scheduling for Aggregated Distributed Energy Resources

Bertrand Travacca[1], Sangjae Bae[1], Jiacheng Wu[2], Scott Moura[1]

***Abstract*— This paper presents an optimal Day-Ahead Electricity Market (DAM) bidding strategy for an aggregator leveraging a pool of residential prosumers: residential customers with local photovoltaic (PV) production and plug-in electric vehicle (PEV) charging flexibility. The aggregator's point-of-view differs from the social planner angle that is taken in the majority of the existing literature, mainly the aggregator is considered to be a private entity (e.g. an electricity retailer). We propose a novel approach to tackling this optimization problem, by including risk management in the objective function and chance constraints on the aggregated PEV mobility constraints. In a first step, we model local system constraints and define a stochastic optimization scheme that exploits the problem structure to distribute the objective among prosumers via dual-splitting. Dual splitting is achieved with two consensus variables: a shadow price for energy and for PEV charging. In a second step, we propose a projected gradient ascent algorithm to solve the dual problem and we prove its corresponding rate of convergence (upper-bound). Finally, we implement a case study, with a model of 100 prosumers, to illustrate the convergence rate of our algorithm. We show that we reach an acceptable level of precision with less than 50 iterations.**

I. INTRODUCTION

This paper derives and analyzes distributed optimization algorithms for scheduling an aggregate of distributed energy resources (DERs), with specific considerations for uncertainties in electricity prices and DER availability.

A. Context and Motivation

In current electricity systems, residential electricity demand is typically inelastic and, most of the time, end-users are subject to flat rates. That said, the recent massive integration of renewable electricity supply has lead to increasing power rampings in the net electricity demand. In many electricity markets (e.g. Germany, Italy, California) peak demand occurs after sunset, when solar power is no longer available, this phenomenon is commonly referred to as the 'duck curve' [1]. Peak power demand creates a need for more flexible power supply [2] which could be leveraged on the residential demand-side as often suggested in the literature [3]–[5].

In this paper, residential electricity end-users with controllable plug-in electric vehicle (PEVs) chargers and photovoltaic (PV) systems are referred to as prosumers. An aggregator is a company pooling prosumers to bring them to the Day-Ahead energy Market for electricity (DAM). The DAM takes place one day before the operating day, and consists in various market entities bidding prices and quantities for each hour of the next day. The objective of the aggregator is to maximize the total electricity profits it delivers in the DAM. It is important to highlight the fact that a single prosumer cannot participate in the DAM because it does not fulfill the power threshold requirements of the market regulator. It is true that Time-of-Use tariffs (TOU) and Real Time Prices (RTP) aim to bring the market to the prosumer level. Nevertheless, there might arguably be at this level a low acceptance of RTP [6]. Moreover TOU rates are not able to capture entirely the state of the electricity system and can even create higher peak loads as [7] suggests.

Two main challenges for managing a large population of flexible resources are (i) uncertainty and (ii) computational scalability. (i) First, we must ensure consumer comfort (e.g. mobility) in the face of uncertain electricity consumption and limit financial risk in the face of uncertain DAM prices. (ii) Second, we require computationally scalable scheduling algorithms that guarantee delivered power from the aggregator to the power system operator.

B. Literature Review

A growing body of literature addresses optimal charging of PEV populations and residential demand response. This research can be classified as having centralized or distributed protocols. Centralized algorithms [8]–[10], use a central infrastructure to communicate with each agent, gather their information, and compute the optimal aggregated load profile. The challenges for centralized methods are scalability with respect to communication and computation, as well as privacy issues. In distributed or decentralized optimization algorithms [11], [12], each local agent solves its own problem and communicates information to its neighbors or the aggregator.

Market bidding strategies and market uncertainty for aggregated PEVs have been studied in [13]–[15]. The aforementioned methods could successfully address uncertainty in aggregated load scheduling, but do not provide a rigorous convergence analysis (except [11]). In particular, finding the necessary number of iterations to reach a specific precision is crucial if we seek to assess implementation burdens for the aggregator.

In this paper, we construct a tailored optimization method for scheduling uncertain electricity resources in the uncertain DAM using our aggregated resources. Leveraging the particular structure of the problem, we derive a distributed

*This work was was supported by TOTAL SA. and Lawrence Berkeley National Laboratory

[1] Energy, Controls, and Applications Lab (eCAL), Civil & Environmental Engineering, University of California, Berkeley, USA. Corresponding author: bertrand.travacca@berkeley.edu

[2] Mechanical Engineering, University of California, Berkeley, USA.

dual-optimization scheme. We then perform a convergence analysis to yield an explicit upper-bound on the rate of convergence, for the projected gradient ascent algorithm. Finally, a case study is implemented to illustrate the performance of our algorithm. Ultimately, the contributions of this paper to ensemble DER control are twofold:

1) Formulation of a convex optimization scheduling problem and distributed algorithm that accounts for uncertain DAM prices and PEV availability.
2) Proof of an upper-bound for convergence (without strong convexity properties).

C. Paper Structure

The report is structured as follows:

- In Section II, we formulate the DER model with local constraints on power, energy, and availability. Next, we define an optimization model to address risk management for DAM bidding, in the face of uncertain DAM prices and uncertain PEV mobility aggregation.
- In Section III, we show how to exploit the problem structure to enable distributed computations.
- In Section IV, we propose a distributed gradient ascent method to solve the problem and derive an explicit bound for convergence.
- In Section V, we illustrate our model for price prediction and corresponding risk (covariance matrix). We then show how our algorithm performs with a case study of 100 prosumers.

D. Notation and Nomenclature

This paper uses the following notation: $(x, y) \in \mathbb{R}^d$, $x^T y$ refers to the euclidean scalar product of x and y, $\|\cdot\|_2^2$ refers to the corresponding euclidean norm. $x \leq y$ refers to element-wise inequality and $A \succeq B$ for $A, B \in \mathbb{R}^{d\times d}$ symmetric matrices means that $A - B$ is positive semi-definite. If A is symmetric positive definite matrix, then $M = A^{1/2}$ denotes the unique matrix square root of A (i.e. $M^2 = A$). For $A \in \mathbb{R}^{q\times d}$, $\|A\|_2$ refers to the operator norm defined by: $\|A\|_2 = \max_{x\in\mathbb{R}^d \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$.

The following notation is specific to this paper. Uppercase letters refer to variables with units of power (kW) while lower case letters refer to variables with units of energy (kWh). Symbol x^t refers to the value taken by variable x at time t. In the absence of the exponent we will consider the variable x as a vector $\in \mathbb{R}^{24}$. Symbol x_i refers to a local prosumer variable $i \in \{1, ...N\}$ and $x_\Sigma = \sum_{i=1}^N x_i$ to the sum of these local variables. Symbol $\hat{X}$ refers to the average or estimate of a random or unknown variable X. Finally, $\overline{x}$ (respectively $\underline{x}$) refers to an upper (lower) bound of the variable x.

TABLE I
NOMENCLATURE

N	Number of prosumers
Δt	Time-step: 1 hour
L_i	Residential Load of prosumer i (not including PEV charge)
EV_i	Charging rate of PEV i
ev_i	State of Energy of PEV i
S_i	Solar PV production of prosumer i
G_i	Power imported from the grid for prosumer i

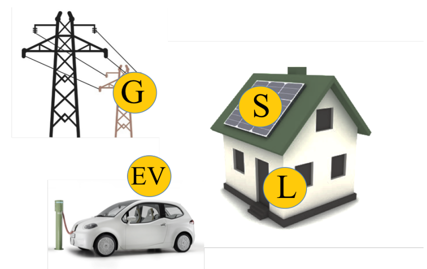

Fig. 1. Local system representation

II. Local System Model and Problem Formulation

A. Local System Model

1) Local Power Balance: The power balance (1) establishes the link between the control variables G_i and EV_i.

$$L_i + EV_i \leq S_i + G_i \tag{1}$$

Note that the power balance has been relaxed from an equality constraint to an inequality constraint representing the fact that PV production can be curtailed.

2) Local Grid Constraints: At any given node in the distribution network, there is a limit on power import or export (2). Typically, for residential customers, $\overline{G}_i \simeq 8\,kW$.

$$\underline{G}_i \leq G_i \leq \overline{G}_i \tag{2}$$

3) Local PEV Constraints: Equation (3) governs the PEV battery state of energy (SOE) dynamics. The subtle difference with the model used in [11] is the fact that the PEVs are not in a closed system. When the i-th PEV leaves its house, it is no longer in the aggregator's perimeter. For instance, a PEV can leave with $SOC = 0.5$ and come back fully charged (because it has charged elsewhere). Therefore, EV_i represents the on-site charge only, while $EV_{D,i}$ is an algebraic value that represents the observed and uncontrolled charge or discharge of the PEV while off-site.

$$ev_i^t = ev_i^{t-1} + EV_i^t \Delta t - EV_{D,i}^t \Delta t \tag{3}$$

Considering a SOE constraint for each hour, we rewrite (3) as an on-site cumulative energy consumption constraint, which is expressed as:

$$\underline{ev}_i \leq A \cdot EV_i \leq \overline{ev}_i \tag{4}$$

With A the discrete integration matrix:

$$A = \Delta t \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}$$

Finally, the PEVs charging power constraint is given by:

$$\underline{EV}_i \leq EV_i \leq \overline{EV}_i \tag{5}$$

Note that when the PEV is un-plugged at a given time t, then $\underline{EV}_i^t = \overline{EV}_i^t = 0$ is required.
For the purpose of conciseness, the convex feasibility set generated by the local constraints (1), (2), (4) and (5) is hereby referred to as $local_i$.

B. Optimization problem formulation

1) Aggregator's objective: As described in Section V-A, the DAM price p is considered as a random variable $\in \mathbb{R}^{24}$ with multivariate Gaussian distribution:

$$p \sim \mathcal{N}(\hat{p}, \Sigma_p), \text{ with } \Sigma_p \succ 0 \tag{6}$$

The total power bought or sold by the aggregator in the DAM is the vector $G_\Sigma = \sum_{i=1}^N G_i$. The choice of the quantity G_Σ^t for $t \in \{1, ...24\}$ can be considered as a portfolio problem where the assets are the DERs, the returns are the DAM prices and the budget constraint is the flexibility constraint described in the previous section. The choice of the portfolio G_Σ involves a trade-off between the expected DAM cost $\hat{p}^T G_\Sigma$ and its variance $G_\Sigma^T \Sigma_p G_\Sigma$.

We define the optimization objective for the aggregator to be given by (7). With α representing the risk aversion propensity of the aggregator. This objective corresponds to a scalar bi-criterion objective. This produces Pareto optimal portfolios except for the two limiting cases: $\alpha \to 0$ or $\alpha \to \infty$ [16].

$$\hat{p}^T G_\Sigma + \alpha G_\Sigma^T \Sigma_p G_\Sigma + \frac{\delta}{2}\left[\sum_{i=1}^N \|EV_i\|_2^2 + \sum_{i=1}^N \|G_i\|_2^2\right] \tag{7}$$

Remark: *If one wishes to consider a risk-seeking aggregator (i.e.* $\alpha < 0$*), then the objective would be concave and therefore the problem would be non convex. This changes the problem structure dramatically, and is not considered here.*

Tikhonov (a.k.a. L_2 or "ridge") regularization terms are also included in the objective. The term $\frac{\delta}{2}\sum_{i=1}^N \|EV_i\|_2^2$ penalizes PEV battery degradation as described in [11], [17]. Nevertheless, battery degradation can also depend on other factors than charging power magnitude. The term $\frac{\delta}{2}\sum_{i=1}^N \|G_i\|_2^2$ penalizes the magnitude of the G_i's. This term can be interpreted as a cost linked to local stability of the distribution grid. The parameter δ is tuned by the practitioner.

2) Aggregated PEV mobility uncertainty: It is extremely difficult to perfectly predict the total hourly day-ahead mobility energy demand profile. Modeling PEV mobility uncertainty at the local level is not tractable, as the individual PEV use is highly unpredictable and may require access to private information. This is not the case at the aggregated level, where behaviors are 'smoothed' via the law of large numbers. Therefore, PEV mobility uncertainty is modeled at the aggregate level via chance constraints – explained next.

We hypothesize that the aggregated energy and power bounds $\underline{ev}_\Sigma$, $\overline{ev}_\Sigma$, $\underline{EV}_\Sigma$, $\overline{EV}_\Sigma$ follow multivariate normal distributions with diagonal covariance matrices. We require these bounds to be respected for each hour with a probability of at least η. Let us denote $\Phi(\cdot)$ the cumulative distribution function for a zero-mean, unit variance Gaussian random variable. Consider the first lower bound. We can write:

$$\underline{ev}_\Sigma^t \sim \hat{\underline{ev}}_\Sigma^t + \sigma_{\underline{ev}}^t \mathcal{N}(0, 1) \tag{8}$$

where $\hat{\underline{ev}}_\Sigma^t$ and $\sigma_{\underline{ev},t}$ are the mean and standard deviation of $\underline{ev}_\Sigma^t$, respectively. Instead of requiring this lower bound to be satisfied for all realizations of $\hat{\underline{ev}}_\Sigma^t$, we require the lower bound is satisfied with probability η. In our particular case:

$$\mathbb{P}(\underline{ev}_\Sigma^t \leq (A \cdot EV_\Sigma)^t) \geq \eta \equiv (A \cdot EV_\Sigma)^t \geq \hat{\underline{ev}}_\Sigma^t + \sigma_{\underline{ev}}^t \Phi^{-1}(\eta) \tag{9}$$

The same procedure is applied to $\overline{ev}_\Sigma$, $\underline{EV}_\Sigma$, $\overline{EV}_\Sigma$. The interpretation of chance constraints, like (9), is that they provide tighter deterministic bounds on mobility constraints that are more robust to error on mobility prediction. The right hand-side of the second inequality in (9) are denoted as $\tilde{\underline{ev}}_\Sigma$, $\tilde{\overline{ev}}_\Sigma$,$\tilde{\underline{EV}}_\Sigma$,$\tilde{\overline{EV}}_\Sigma$. In the following, for legibility, the symbols $\hat{}$ and $\tilde{}$ are dropped for all variables.

3) Stochastic Optimization Problem Formulation: We are now positioned to formulate the *stochastic* optimization problem:

$$\min_{EV_i, G_i, EV_\Sigma, G_\Sigma} p^T G_\Sigma + \alpha G_\Sigma^T \Sigma_p G_\Sigma + ... \tag{10}$$

$$\frac{\delta}{2}\left[\sum_{i=1}^N \|EV_i\|_2^2 + \sum_{i=1}^N \|G_i\|_2^2\right]$$

subject to: Local constraints : $\forall i \quad local_i$

Aggregated variable constraints :

$$EV_\Sigma = \sum_{i=1}^N EV_i, \quad G_\Sigma = \sum_{i=1}^N G_i$$

Aggregated EV chance constraints :

$$\underline{ev}_\Sigma \leq A \cdot EV_\Sigma \leq \overline{ev}_\Sigma$$

$$\underline{EV}_\Sigma \leq EV_\Sigma \leq \overline{EV}_\Sigma \tag{11}$$

Let us denote c^* the optimal cost for (11). The optimization problem (11) couples the individual EV_i, G_i variables in the objective via the variance term $G_\Sigma^T \Sigma_p G_\Sigma$. Additionally, coupling terms appear with the aggregated PEV mobility constraints. Without these coupling terms, the optimization problem is sum-separable. That said, the problem contains independent local constraints $local_i$, and this structure can be exploited to distribute the objective – discussed next.

III. Distributed Optimization Scheme

Dual splitting is used to exploit the inherent problem structure described previously. This enables us to achieve the following key result, which enables scalability:

Theorem 3.1: Solving (11) is equivalent to solving:

$$\max_{\mu\in\mathbb{R}^{96},\ \mu\geq 0,\ \nu\in\mathbb{R}^{24}} \Bigg[-\frac{1}{4\alpha}\nu^T\Sigma_p^{-1}\nu + c^T\mu + \ldots$$
$$\sum_{i=1}^{N} \min_{EV_i,G_i,\ \textbf{s.t.}\ local_i} \Big[G_i^T(p-\nu) + EV_i^T B\mu + \ldots \tag{12}$$
$$\frac{\delta}{2}(\|EV_i\|_2^2 + \|G_i\|_2^2)\Big]\Bigg]$$

With $B \in \mathbb{R}^{24\times 96}$, and $c \in \mathbb{R}^{96}$ given (cf. Proof).

Proof: This proof utilizes duality theory. Namely, the aggregated decision variables and EV chance constraints from (11) are used to form the dual problem and the corresponding Lagrangian $\mathcal{L}$ given by (13). Slater condition holds, henceforth min and max symbols can be interchanged without generating a dual gap.

$$\mathcal{L} = \sum_i p^T G_i + \alpha G_\Sigma^T \Sigma_p G_\Sigma + \frac{\delta}{2}(\sum_i \|EV_i\|_2^2 + \sum_i \|G_i\|_2^2)$$
$$+ \mu_{\underline{ev}}^T(-A\cdot EV_\Sigma + \underline{ev}_\Sigma) + \mu_{\overline{ev}}^T(A\cdot EV_\Sigma - \overline{ev}_\Sigma) +$$
$$\mu_{\underline{EV}}^T(-EV_\Sigma + \underline{EV}_\Sigma) + \mu_{\overline{EV}}^T(EV_\Sigma - \overline{EV}_\Sigma) +$$
$$\nu_{EV}^T(EV_\Sigma - \sum_i EV_i) + \nu_G^T(G_\Sigma - \sum_i G_i) \tag{13}$$

where the μ's≥ 0 and ν's are dual variables associated with corresponding inequality and equality constraints, respectively. Using the facts that:

$$\min_{G_\Sigma}\left[\alpha G_\Sigma^T\Sigma_p G_\Sigma + \nu_G^T G_\Sigma\right] = -\frac{1}{4\alpha}\nu_G^T\Sigma_p^{-1}\nu_G$$
$$\min_{EV_\Sigma} EV_\Sigma^T\left(A^T(\mu_{\overline{ev}} - \mu_{\underline{ev}}) + \mu_{\overline{EV}} - \mu_{\underline{EV}} + \nu_{EV}\right) = 0$$

The following equality constraint on the dual variables ensures the problem is bounded:

$$A^T(\mu_{\overline{ev}} - \mu_{\underline{ev}}) + \mu_{\overline{EV}} - \mu_{\underline{EV}} + \nu_{EV} = 0$$

This leads to the dual function:

$$\mathcal{L}^*(\mu,\nu) = -\frac{1}{4\alpha}\nu^T\Sigma_p^{-1}\nu + c^T\mu + \ldots$$
$$\sum_{i=1}^{N} \min_{EV_i,G_i,\ \textbf{s.t.}\ local_i} G_i^T(p-\nu) + EV_i^T B\mu + \ldots \tag{14}$$
$$\frac{\delta}{2}(\|EV_i\|_2^2 + \|G_i\|_2^2)$$

with,

$$\mu = [\mu_{\underline{ev}};\ \mu_{\overline{ev}};\ \mu_{\underline{EV}};\ \mu_{\overline{EV}}] \in \mathbb{R}^{96} \geq 0,$$
$$\nu = \nu_G,$$
$$c = [\underline{ev}_\Sigma;\ -\overline{ev}_\Sigma;\ \underline{EV}_\Sigma;\ -\overline{EV}_\Sigma] \in \mathbb{R}^{96},$$
$$B = \begin{bmatrix} -A^T & A^T & -I & I \end{bmatrix} \in \mathbb{R}^{24\times 96}$$

Since all these steps are necessary and sufficient, we conclude (11) $\iff$ $\max_{\mu\geq 0,\nu}\mathcal{L}^*(\mu,\nu)$. ■

Theorem 3.1 shows that only the dual variables (μ^*,ν^*) need to be sent to the prosumers in order to reach the primal optimal objective c^* described by (11). With this distributed scheme the aggregator partakes the primal objective and the computational burden between the prosumers (each prosumer solves its own linear constrained quadratic program) while enforcing the chance constraint on aggregated PEV mobility. Figure 2 depicts this distributed scheme. Note that, how to find the optimal dual variables (μ^*,ν^*) has not been tackled yet– discussed in the next section.

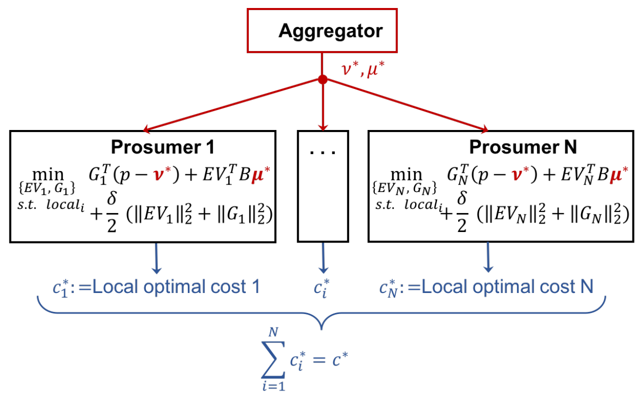

Fig. 2. Distributed scheme visualization

Remark *In the local objective the term* $-\nu^*$ *can be interpreted as a shadow price for power import for the grid and* $B\mu^*$ *as a shadow price for PEV charging.*

IV. Distributed Ascent Method

A. Algorithm Objective

In this section, a projected gradient ascent algorithm with constant step-size is outlined to solve the Lagrange dual problem in Theorem 3.1. The algorithm's rate of convergence and its so-called "oracle complexity" $K(\epsilon)$ are proven. The oracle complexity guarantees that for a number of iterations $k > K(\epsilon)$:

$$|\mathcal{L}^*(\mu^*,\nu^*) - \mathcal{L}^*(\mu^k,\nu^k)| < \epsilon \tag{15}$$

with $\mathcal{L}^*$ the optimal dual function defined in the proof of Thm 3.1, $\epsilon > 0$ is the desired algorithm solution precision, and μ^* and ν^* denote the optimal dual variables. Note that $\mathcal{L}^*(\mu^*,\nu^*) = c^*$. A unique contribution of this result is quantifying the rate of convergence without strong convexity. In contrast, previous work [11] required strong convexity to bound convergence.

The regularization parameter δ in (7) can be tuned by the practitioner. These local regularization terms are central to prove response continuity, differentiability and smoothness of the local optimizer with respect to the dual variables. Without these regularization terms, the local optimizations are linear programs, where the local optimal solution EV_i and G_i lie on the polyhedron edge defined by the linear constraints. As highlighted in [16], a small change in the dual variable can lead to a 'jump' from one edge to another. The regularization terms serve to smooth this behavior.

B. Projected Gradient Ascent Method

Denote $\mathcal{P}_+$ the projection map onto the positive orthant of $\mathbb{R}^{24}$. Parameter γ denotes the step-size of the projected gradient ascent. Our proposed projected gradient ascent is given by Algorithm 1 and consists in updating the dual variables at iteration k with a linear feedback of the aggregated optimal local response $\sum_{i=1}^{N} EV_i^k$ and $\sum_{i=1}^{N} G_i^k$ (here the exponent k does not refer to time but the number of iterations, cf. algorithm). The structure of this linear feedback to the aggregator shows that private consumption profiles do not have to be sent: the aggregator only has to have access to the sum of these profiles, therefore privacy is respected with this scheme. Figure 3 depicts this iterative dynamic that allows to converge to the optimal dual variables.

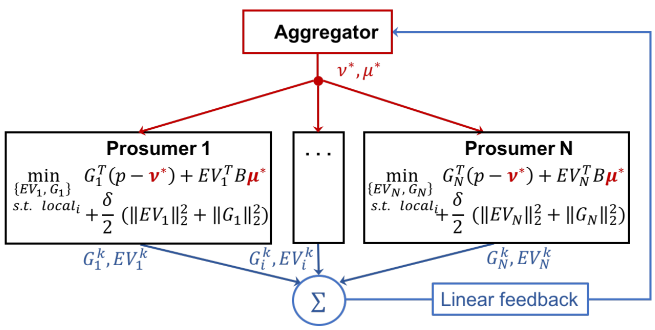

Fig. 3. Project gradient ascent method: a linear 'time' variant dynamic system

Algorithm 1 Projected Gradient Ascent Method

Initialization: $k = 0$, $\mu := \mu_0 \geq 0$ and $\nu := \nu_0$
while $\mathcal{L}^*(\nu^*, \mu^*) - \mathcal{L}^*(\nu^k, \mu^k) \geq \epsilon$ **do**
    $k + 1 \leftarrow k$
    (1) Find the optimal local solutions EV_i and G_i
    for $i = 1$ to N **do**
        $EV_i^k, G_i^k = \text{argmin}\Big[G_i^T(p - \nu^k) + EV_i^T B\mu^k + \frac{\delta}{2}(\|EV_i\|_2^2 + \|G_i\|_2^2)\Big]$ s. to: $local_i$
    end for
    (2) Update dual variables μ and ν
    $\mu^{k+1} := \mathcal{P}_+(\mu^k + \gamma c + \gamma \sum_i B^T EV_i^k)$
    $\nu^{k+1} := \nu^k - \gamma \frac{1}{2\alpha}\Sigma_p^{-1}\nu^k - \gamma \sum_i G_i^k$
end while

Next we quantify the convergence rate of the projected gradient ascent algorithm. Let us re-write the Lagrangian:

$$\mathcal{L}^*(\nu, \mu) = \Phi_0(\nu, \mu) + \sum_{i=1}^{N} \Phi_i(\nu, \mu)$$

with:

$$\Phi_0 = -\frac{1}{4\alpha}\nu^T \Sigma_p^{-1} \nu + c^T \mu \tag{16}$$

$$\Phi_i = \min_{EV_i, G_i, \textbf{ s.t. } local_i} \Big[G_i^T(p - \nu) + EV_i^T B\mu + ... \frac{\delta}{2}(\|EV_i\|_2^2 + \|G_i\|_2^2)\Big] \tag{17}$$

The gradient of Φ_0 w.r.t. (ν, μ) is given by:

$$\nabla \Phi_0 = \Big[-\frac{1}{2\alpha}\Sigma_p^{-1}\nu \; ; \quad c\Big] \tag{18}$$

Therefore, the Hessian of Φ_0 respects the following bounds that characterize smoothness condition:

$$M_0 I \succeq -\nabla^2 \Phi_0 \tag{19}$$

with $M_0 = \frac{1}{2\alpha\lambda_{min}}$, λ_{min} being the smallest eigenvalue of Σ_p. Using Danskin's Theorem [16] to derive the gradient of Φ_i w.r.t. (ν, μ) yields:

$$\nabla \Phi_i = \Big[-G_i^k \; ; \quad +B^T EV_i^k\Big] \tag{20}$$

Where:

$$EV_i^k, G_i^k = \underset{s.t.\ local_i}{\text{argmin}} \Big[G_i^T(p - \nu^k) + EV_i^T B\mu^k ... + \frac{\delta}{2}(\|EV_i\|_2^2 + \|G_i\|_2^2)\Big] \tag{21}$$

Note that the local optimization problems are strictly convex and therefore the local solutions are unique for a given $\{\nu, \mu\}$. Using the first order optimality condition for convex constrained problems on the local optimization problems and the above gradient formulae allows us to prove that Φ_i is $\frac{1}{\delta}\|\Theta\|_2^2$-smooth, where:

$$\Theta = \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix} \in \mathbb{R}^{48 \times 120}$$

Since Φ_i is the point-wise minimum of a jointly concave function of (ν, μ), we conclude Φ_i is concave w.r.t. (ν, μ). Using the concavity of Φ_i, (19) and (20), we conclude that $\mathcal{L}^*$ is a M-smooth differentiable concave function with:

$$M = M_0 + \frac{N}{\delta}\|\Theta\|_2^2 \tag{22}$$

This M-smoothness of the Lagrangian $\mathcal{L}^*$ plays an important role in the following convergence result.

Theorem 4.1: Consider Algorithm 1. A step-size of $\gamma = \frac{1}{M}$ leads to:

$$\mathcal{L}^*(\mu^*, \nu^*) - \mathcal{L}^*(\mu^k, \nu^k) < \frac{M}{2k}(\|\nu_0 - \nu^*\|_2^2 + \|\mu_0 - \mu^*\|_2^2) \tag{23}$$

Additionally, the oracle complexity of Algorithm 1 is:

$$K(\epsilon) = \frac{M}{2\epsilon}(\|\nu_0 - \nu^*\|_2^2 + \|\mu_0 - \mu^*\|_2^2) \tag{24}$$

Remark: *Equation (23) says the convergence precision is upper-bounded by a value proportional to M/k. Previous* *work [11] derived linear convergence rate, but leveraged strong convexity. In this case, we have $1/k$ convergence rates since the objective is not strongly convex.*

Proof: Let $k \in \mathbb{N}^*$ and denote $\omega^k := [\nu^k, \mu^k]$, then by M-smoothness of $\mathcal{L}^*$:

$$\mathcal{L}^*(\omega^{k+1}) - \mathcal{L}^*(\omega^k) \leq \nabla \mathcal{L}^*(\omega^k)^T \left(\omega^{k+1} - \omega^k\right) + \frac{M}{2}\|\omega^{k+1} - \omega^k\|_2^2 \quad (25)$$

Noting that $\omega^{k+1} = \mathcal{P}_+(\omega^k + \gamma \nabla \mathcal{L}^*(\omega^k))$, with $\mathcal{P}_+$ denoting the projection on the convex set $\{(\nu, \mu) \; |\mu \geq 0\}$, using the non-expansiveness property of $\mathcal{P}_+$ and Cauchy-Schwartz inequality yields:

$$\mathcal{L}^*(\omega^{k+1}) \geq \mathcal{L}^*(\omega^k) + (\gamma - \frac{M\gamma^2}{2})\|\nabla \mathcal{L}^*(\omega_k)\|_2^2 \quad (26)$$

The maximum of the quadratic function $t \mapsto t - \frac{Mt^2}{2}$ is achieved at $t = \frac{1}{M}$. Substituting this step-size choice, taking the opposite of the inequality and adding $\mathcal{L}^*(\omega^*)$, with $\omega^* \in \operatorname{argmax} \mathcal{L}^*(\omega)$ from both sides yields:

$$\mathcal{L}^*(\omega^*) - \mathcal{L}^*(\omega^{k+1}) \leq \mathcal{L}^*(\omega^*) - \mathcal{L}^*(\omega^k) - \frac{1}{2M}\|\nabla \mathcal{L}^*(\omega_k)\|_2^2 \quad (27)$$

Consequently, the error sequence $\{\mathcal{L}^*(\omega^*) - \mathcal{L}^*(\omega^{k+1})\}_{k=0}^{+\infty}$ is decreasing in k. By concavity of $\mathcal{L}^*$:

$$\mathcal{L}^*(\omega^*) \leq \mathcal{L}^*(\omega^k) + \nabla \mathcal{L}^*(\omega^k)^T \left(\omega^* - \omega^k\right) \quad (28)$$

Combining (27) and (28) leads to:

$$\begin{aligned}\mathcal{L}^*(\omega^*) - \mathcal{L}^*(\omega^{k+1}) \leq & \nabla \mathcal{L}^*(\omega^k)^T \left(\omega^* - \omega^k\right) ... \\ & - \frac{1}{2M}\|\nabla \mathcal{L}^*(\omega_k)\|_2^2 \\ & = \frac{M}{2}(\|\omega^k - \omega^*\|_2^2 - \|\omega^{k+1} - \omega^*\|_2^2)\end{aligned}$$

Using the fact that the error is decreasing, it is possible to write this bound as a telescopic sum in the following way:

$$\begin{aligned}\mathcal{L}^*(\omega^*) - \mathcal{L}^*(w^k) &\leq \frac{1}{k}\sum_{j=0}^{k-1}(\mathcal{L}^*(\omega^*) - \mathcal{L}^*(w^j)) \\ &\leq \frac{M}{2k}\sum_{j=0}^{k}(\|\omega^j - \omega^*\|_2^2 - \|\omega^{j+1} - \omega^*\|_2^2) \\ &\leq \frac{M}{2k}\|\omega^0 - \omega^*\|_2^2\end{aligned}$$

Which establishes the claim. ■

V. Study Case for Convergence Analysis

Next we illustrate the proposed algorithm and convergence analysis via a case study on 100 prosumers. To begin, we describe our price forecast model.

A. Day-Ahead Energy Market price Model

Three years of data (Jan 2013- Dec 2015) have been collected from the California Independent System Operator (CAISO) across the Pacific Gas and Electricity (PG&E) service territory. These data are publicly available [21]. An online prediction model for DAM prices is built using random forest regression with 10 trees [23]. The features for prediction are:

- hourly demand forecast (provided by CAISO)
- year, month, day, and hour

The model is initially trained with data from 2013, and then recursively updated as follow. Let us denote $\overline{p_d} = \{p_1, ...p_d\}$ the available price history. Let us denote $\mathcal{M}_d$ the random forest regression model corresponding to this historical data. In order to predict the next day price $\hat{p}_{d+1} \in \mathbb{R}^{24}$, the model $\mathcal{M}_d$ is used. For the next day, an updated model $\mathcal{M}_{d+1}$ using historical data $\overline{p_{d+1}}$ is used to predict the price $\hat{p}_{d+2}$, etc. In other words, it is an online model with a retrospective rolling horizon of d days. The obtained DAM price prediction model has a root mean square error (RMSE) of 3.5 USD and a mean absolute percentage error (MAPE) of 8.4 %.

A Gaussian mixture model (GMM) is then implemented on the ex-post prediction error. It shows that the lowest Bayesian information criterion score is attained for a single GMM component. This justifies the use of a simple multi-variate Gaussian distribution to model price uncertainty as it is stated in II-B.1. A sparse covariance matrix estimator [23] is then fitted to the ex-post DAM price forecasting error with a L_1 regularization parameter equal to 5. This estimator allows to delete covariance between hours that are less significant while empirically increasing the covariance matrix condition number (which has a favorable impact on the projected gradient ascent convergence).

B. Prosumer Modeling and Parameters

A total of 100 prosumers are modeled. Each prosumer is considered to have a PEV with identical battery size of 24 kWh and itineraries based on National Household Travel Survey [20]. We assume PEVs cannot provide power to the grid: $\underline{EV_i} = 0$. As the mobility data [20] does not allow to fit a stochastic model as described in (8), and because we showed that the chance constraint on the aggregate PEVs can be expressed using (9), we make the assumption that $(A.EV_\Sigma)^t \geq (1+0.05)\underline{\hat{ev}}_\Sigma^t$ and $(A.EV_\Sigma)^t \leq (1-0.05)\overline{\hat{ev}}_\Sigma^t$ (similar assumptions for EV_Σ^t). Table II details parameter values used for the simulation.

C. Simulation Results

The projected gradient ascent algorithm described in Section IV-B is implemented in Matlab with the parameters displayed in Table II. In Figure 4, $\mathcal{L}^*(\omega_k)$ is plotted with respect to k (the number of iterations). We compare it to the optimal primal cost $\mathcal{L}^*(\omega^*) = c^*$ depicted by the solid red line. The primal optimal cost was obtained by solving the primal problem with CVX [24]. In the proof of Theorem 4.1 we show that the choice of $\gamma = 1/M$ with M given by

TABLE II
PARAMETER VALUES FOR THE MODEL

N	100
Δt	1 hour
α	1
δ	10^{-2}
L_i	Heterogeneous, same load data scaled from [21] with uniform independent noise added for each prosumer
S_i	Heterogeneous, same production data generated via [19] with uniform independent noise added for each prosumer
η	90 %
$\overline{EV}_i$	1.4 kW $\forall i$
$\underline{ev_i}, \overline{ev_i}$	Heterogeneous, generated from [20]
$\overline{G_i}, \underline{G_i}$	$+/-10$ kW $\forall i$
p, Σ_p	cf. Section V-A

(22) guarantees convergence to c^*. From theory we have that $\gamma = 2.10^{-7}$, but in practice we find that $\gamma = 10^{-5}$ displays better convergence behavior because the derived smoothness value M is conservative. From Theorem 4.1 we have that:

$$\mathcal{L}^*(\omega^*) - \frac{M}{2k}\|\omega_0 - \omega^*\|_2^2 < \mathcal{L}^*(\omega^k) \tag{29}$$

The left hand side of the inequality is called an upper-bound for the projected gradient ascent algorithm convergence to the optimum. This upper-bound is represented by the solid blue curve in Figure 4. For this simulation we take $\frac{1}{M} = \gamma = 10^{-5}$ to illustrate the fact that the derived M-smoothness bound is indeed conservative. We find that our simulation is consistent with theorem 4.1:

- convergence is sub-linear because the problem is not strictly convex
- derived upper-bound is verified in practice

From Fig. 4, it can be seen that the scheme converges to a rough consensus in less than 50 iterations. It practice, this means that the aggregator needs to communicate less than 50 times with all the prosumers. Nevertheless, as the convergence is not linear, extra precision requires more iterations. However, since our main objective is to schedules the aggregated load G_Σ in the DAM, this algorithm offers an appropriate convergence rate that is robust and theoretically guaranteed.

The algorithm is scalable with N in terms of computational burden for the aggregator: the aggregator only has to update the dual variable (and this update complexity does not depend on the number of prosumers). Nevertheless for a given precision ϵ, the number of iterations required increase linearly with the number of prosumers. We argue that in practice, the required precision should be relative to the total cost c^* which also scales (more than) linearly with N. It is important to remember that the dual variables obtained from this algorithm allow to minimize the risk taken on the DAM while guaranteeing chance constraints/robust bounds on EV_Σ and AEV_Σ.

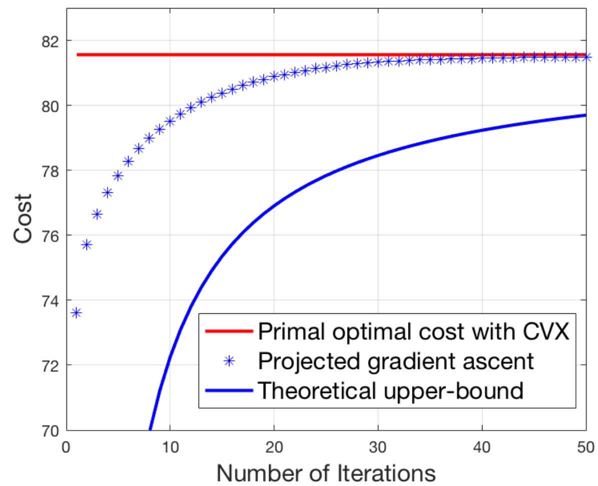

Fig. 4. Simulated convergence of $\mathcal{L}^k(\omega^k)$ to the optimal primal cost $c^*(\omega^*)$, using the Projected Gradient Ascent Algorithm.

VI. CONCLUSION

This article studies a scalable scheduling optimization algorithm for an aggregation of residential energy prosumers in the DAM, taking into account uncertainty in market prices and PEV flexibility. More precisely a model for prosumers with PEVs and PV panels is considered. We consider stochastic constraints on PEV mobility. The presented methodology can be extended to various sources of electricity production, flexibility and uncertainties.

The structure of the primal optimization problem is exploited to distribute the objective among the prosumers. In the distributed scheme, the aggregator broadcasts price signals (i.e. dual variables) to the prosumers, and the prosumers send back their corresponding energy consumption profile. These price signals are then updated by the aggregator until a consensus is reached.

It is shown that the price signal updates can be generated by the aggregator using a projected gradient ascent method. An upper-bound on the number of iterations needed to reach a given precision is then derived. The convergence rate is slower than linear, but well-suited for an aggregator bidding into the DAM. Finally, the theory is illustrated in a case study with real-world price and PEV mobility data. We show that an acceptable level of precision is reached for less than 50 iterations.

REFERENCES

[1] P. Denholm, M. O'Connell, G. Brinkman and J. Jorgenson, '"Overgeneration from solar energy in california: A field guide to the duck chart," National Renewable Energy Laboratory, Tech.Rep.NREL/TP-6A20-65023.

[2] M. Huber, D. Dimkova and T. Hamacher, '"Integration of wind and solar power in Europe: Assessment of flexibility requirements," Energy, vol. 69, pp. 236-246.

[3] Z. Chen, L. Wu and Y. Fu, '"Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization," IEEE Transactions on Smart Grid, vol. 3, no. 4, pp. 1822-1831.
[4] M.H. Albadi and E. El-Saadany, '"A summary of demand response in electricity markets," Electr.Power Syst.Res., vol. 78, no. 11, pp. 1989-1996.
[5] J.H. Yoon, R. Baldick and A. Novoselac, '"Dynamic demand response controller based on real-time retail price for residential buildings," IEEE Transactions on Smart Grid, vol. 5, no. 1, pp. 121-129.
[6] S. Annala, S. Viljainen, J. Tuunanen and S. Honkapuro, '"Does knowledge contribute to the acceptance of demand response?" Journal of Sustainable Development of Energy, Water and Environment Systems, vol. 2, no. 1, pp. 51-60.
[7] M. Ruth, A. Pratt, M. Lunacek, S. Mittal, H. Wu and W. Jones, '"Effects of home energy management systems on distribution utilities and feeders under various market structures,".
[8] K. Clement, E. Haesen and J. Driesen, '"Coordinated charging of multiple plug-in hybrid electric vehicles in residential distribution grids,", pp. 1-7.
[9] E. Sortomme, M.M. Hindi, S.J. MacPherson and S. Venkata, '"Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses," IEEE transactions on smart grid, vol. 2, no. 1, pp. 198-205.
[10] C. Le Floch, F. Di Meglio and S. Moura, '"Optimal charging of vehicle-to-grid fleets via pde aggregation techniques,", pp. 3285-3291.
[11] C. Le Floch, F. Belletti and S. Moura, '"Optimal Charging of Electric Vehicles for Load Shaping: A Dual-Splitting Framework With Explicit Convergence Bounds," IEEE Transactions on Transportation Electrification, vol. 2, no. 2, pp. 190-199.
[12] P. Richardson, D. Flynn and A. Keane, '"Local versus centralized charging strategies for electric vehicles in low voltage distribution systems," IEEE Transactions on Smart Grid, vol. 3, no. 2, pp. 1020-1028.
[13] S.I. Vagropoulos and A.G. Bakirtzis, '"Optimal bidding strategy for electric vehicle aggregators in electricity markets," IEEE Trans.Power Syst., vol. 28, no. 4, pp. 4031-4041.
[14] M.G. Vay and G. Andersson, '"Optimal bidding strategy of a plug-in electric vehicle aggregator in day-ahead electricity markets under uncertainty," IEEE Trans.Power Syst., vol. 30, no. 5, pp. 2375-2385.
[15] R.J. Bessa, M.A. Matos, F.J. Soares and J.A.P. Lopes, '"Optimized bidding of a EV aggregation agent in the electricity market," IEEE Transactions on Smart Grid, vol. 3, no. 1, pp. 443-452.
[16] S. Boyd and L. Vandenberghe, '"Convex optimization,", 2004.
[17] S.J. Moura, J.L. Stein and H.K. Fathy, '"Battery-health conscious power management in plug-in hybrid electric vehicles via electrochemical modeling and stochastic control," IEEE Trans.Control Syst.Technol., vol. 21, no. 3, pp. 679-694.
[18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg, '"Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, no. Oct, pp. 2825-2830.
[19] D.L. King, J.K. Dudley and W.E. Boyson, '"PVSIM/sub C: a simulation program for photovoltaic cells, modules, and arrays,", pp. 1295-1297.
[20] A. Santos, N. McGuckin, H.Y. Nakamoto, D. Gray and S. Liss, Summary of travel trends: 2009 national household travel survey.
[21] LCG Consulting :: EnergyOnline, CAISO: Average Price and Demand, accessed March 2017.
[22] H. Zhang, Z. Hu, Z. Xu and Y. Song, '"Evaluation of Achievable Vehicle-to-Grid Capacity Using Aggregate PEV Model," IEEE Trans.Power Syst., vol. 32, no. 1, pp. 784-794.
[23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg, '"Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, no. Oct, pp. 2825-2830.
[24] M. Grant, S. Boyd and Y. Ye, CVX: Matlab software for disciplined convex programming.