A Privacy-preserving Decentralized Algorithm for Distribution Locational Marginal Prices

Olivier Bilenne\textsuperscript{1}, Barbara Franci\textsuperscript{1}, Paulin Jacquot\textsuperscript{2}, Nadia Oudjane\textsuperscript{2}, Mathias Staudigl\textsuperscript{1}, and Cheng Wan\textsuperscript{2}

Abstract—A major challenge in today’s electricity system is the management of flexibilities offered by new usages, such as smart home appliances or electric vehicles. By incentivizing energy consumption profiles of individuals, demand response seeks to adjust the power demand to the supply, for increased grid stability and better integration of renewable energies. This optimization of flexibility is typically managed by Load Aggregators, independent entities which aggregate and optimize numerous flexibility providers. The consideration of the underlying distribution network constraints, which couple the different actors, leads to a complex multi-agent problem. To address it, we propose a new decentralized algorithm that solves a convex relaxation of the classical Alternative Current Optimal Power Flow (ACOPF) problem, and which relies on local information only. Each computational step is performed in a privacy-preserving manner, and system-wide coordination is achieved via node-specific distribution locational marginal prices (DLMPs). We demonstrate the efficiency of our approach on a 15-bus radial distribution network.

I. INTRODUCTION

The modern distribution network is undergoing an unprecedented reformation, thanks to the increased deployment of Distributed Energy Resources (DERs) in the form of distributed generators, distributed storage, microgrids, aggregators managing fleets of electric vehicles or groups of prosumers [1]. While the potential benefits of DERs are globally accepted, reaching those benefits requires smart management methods. Specifically, wrong control strategies could lead to drastic voltage fluctuations and supply-demand imbalances. With this in mind, a replication of the transmission-level Locational Marginal Price (LMP) (defined as the marginal cost induced by an additional unit of demand at a particular bus) is much desired. The price signals differ spatially and temporally, and are used to incentivize DERs to balance supply-demand, support voltage, and minimize system losses. The necessary extension to Distribution LMPs—abbreviated as DLMPs—has been developed in [2], [3]. A key question in the Distribution Locational Marginal Price (DLMP) approach is their effective computation. According to [4], [5], the procedure of using DLMPs is as follows: the Distribution System Operator (DSO) obtains the flexible demand and supply data, such as active and reactive power generation at the buses, from the Load Aggregators (LAs). Having complete information about the distribution network and the predicted spot prices at the relevant distribution buses, the DLMPs are calculated by solving a network optimization problem. Specifically, DLMPs are obtained as dual variables, measuring the sensitivity of the network flow constraints describing the physics of the problem, and are then announced to aggregators. Considering the received DLMPs and predicted spot prices, each aggregator makes its optimal plans and submits its energy schedules to the spot market. Our distributed coordination mechanism follows this approach line-by-line. We develop a new distributed block-coordinate descent algorithm, designed to effectively compute DLMPs in a decentralized and privacy-preserving way. Our computational architecture involves direct communication between the LAs and the feeding bus, who acts as a central computational unit which updates the DLMPs. However, the data communicated to the center are not containing any information on the local cost function or power profiles managed by the LA. In that sense, our notion of privacy should not be confused with the influential concept of differential privacy in computer science. Our main theoretical result (Theorem 1) states the convergence of a general primal-dual block-coordinate descent algorithm, that extends recent block-coordinate primal-dual splitting methods for solving linearly constrained composite convex optimization as presented in [6], [7], [8].

A. Related Literature

The increased importance of effective management of DERs as been supported by an active literature on decentralized control strategies. [9] introduced DLMPs for the distributed management of fleets of electric vehicles. [10] proposed a quadratic programming approach to solve DLMPs for decentralized congestion management. The Optimal Power Flow model (OPF) is the standard approach for power flow analysis and optimization of power systems. Using the convex relaxation derived in [11] (see also [12]), Papavasiliou [13] derived DLMPs based on the KKT conditions, yet he does not provide any algorithm to effectively compute DLMPs. [14] proposes to compute DLMPs via semi-definite programming but their algorithm is not distributed. [15] proposes a practical distributed algorithm for optimizing DERs, but their approach differs from us in that it relies on a linearization of Optimal Power Flow (OPF), and uses dual decomposition and gradient descent, which requires the exchange of local primal variables.
II. DISTRIBUTED OPTIMAL POWER FLOW

Consider a power system with $N$ buses $\mathcal{N} = \{1, \ldots, N\}$ on a radial distribution network, modeled as a tree graph $\mathcal{G} = (\mathcal{N}_{+}, \mathcal{E})$, where $\mathcal{N}_{+} = \mathcal{N} \cup \{0\}$. The root node 0 is selected as the reference bus. The network is optimized over a time window $\mathcal{T} = \{1, \ldots, T\}$.

A. Branch flow equations

We use $p_n = (p_{n,1}, \ldots, p_{n,T})$ and $q_n = (q_{n,1}, \ldots, q_{n,T})$ to denote active and reactive power consumption at bus $n$ at each time point $t \in \mathcal{T}$. Thus, $p_{n,t} < 0$ means that there is production of energy at bus $n$ at time $t$. At $n = 0$, we assume that power will only be generated and there is no consumption, i.e., $p_{0,t} \leq 0$.

In deriving the power flow equations, we follow [16]. Specifically, after elimination of phase angles and convex relaxation, the AC branch flow equations for a node $n \in \mathcal{N}$ and its (unique) ancestor on the graph, denoted by $n_-$, are:

$$f_n - \sum_{m:n,m=n} (f_m - R_m l_m) + p_n + G_n v_n = 0 \quad [g^p_n] \quad (1a)$$

$$g_n - \sum_{m:n,m=n} (g_m - R_m l_m) + q_n - B_n v_n = 0 \quad [g^q_n] \quad (1b)$$

$$v_n - 2(R_n f_n + X_n g_n) + (R_n^2 + X_n^2) l_n = v_{n_-} \quad (1c)$$

$$f_{n,t}^2 + g_{n,t}^2 \leq v_{n,t} l_{n,t} \quad \forall t \in \mathcal{T} \quad (1d)$$

$$f_{n,t}^2 + g_{n,t}^2 \leq S_n \quad \forall t \in \mathcal{T} \quad (1e)$$

$$(f_{n,t} - R_n l_{n,t})^2 + (g_{n,t} - X_n l_{n,t})^2 \leq S_n^2 \quad \forall t \in \mathcal{T} \quad (1f)$$

$$\bar{V}_n \leq v_{n,t} \leq \overline{V}_n, \quad \forall t \in \mathcal{T} \quad (1g)$$

where

- $v_n = (v_{n,1}, \ldots, v_{n,T})$ and $v_{n_-}$ are the squared voltage magnitudes at buses $n$ and $n_-$.
- $l_n$ is the squared current magnitude on branch $(n, n_-)$.
- $f_n$ and $g_n$ are the active and the reactive parts of the power flow over line $(n, n_-)$.
- $R_n$ and $X_n$ are the resistance and the reactance of branch $(n, n_-)$.
- $G_n$ and $B_n$ are the line conductance and susceptance at $n$.

Equation (1a) and (1b) are an expression of Ohm’s law for the branch $(n, n_-)$, and (1d) is a SOCP relaxation of the the power flow equation [17]. Equations (1c) and (1f) are limitations on the squared power flow magnitude on $(n, n_-)$, and (1g) gives lower and upper bounds on the voltage at $n$.

For the coupling flow conservation laws, dual variables are attached, which are the DLMPs corresponding to active and reactive power. There exist theoretical sufficient conditions under which the relaxation (1) is exact [11], [17].

For later reference, we point out that the network flow constraints (1a)-(1b) can be compactly summarized as

$$A_0 x_0 + \sum_{a} A_a x_a = b$$

for suitably defined matrices $A_0$, $A_a$ and vector $b$.

B. Load aggregators

The set of buses $\mathcal{N}$ is partitioned into a collection $(\mathcal{N}_a)_{a \in \mathcal{A}}$ of subsets, such that each node subset $\mathcal{N}_a$ is managed by a Load Aggregator $a \in \mathcal{A}$. Each LA controls the flexible net power consumption ($p_{n,t}$) and generation at each node $n \in \mathcal{N}_a$, given at time $t$ by

$$p_{n,t} = p_{n,t}^c - p_{n,t}^p, \quad q_{n,t} = q_{n,t}^c - q_{n,t}^p, \quad (2a)$$

for all $n \in \mathcal{N}$ and $t \in \mathcal{T}$. $p_{n,t}^c \geq 0$ is the consumption part and $p_{n,t}^p \geq 0$ is the production part of the power profile. Power consumption and production at the nodes are made flexible by the presence of deferrable loads (electric vehicles, water heaters) and Distributed Energy Resources (DERs). The consumption at each node $n \in \mathcal{N}$ must satisfy a global energy demand $E_n$ over the full time window, 

$$\sum_{t \in \mathcal{T}} p_{n,t}^c \geq E_n, \quad \forall n \in \mathcal{N} \quad (2b)$$

Consumption and production are also constrained by power bounds and active to reactive power ratio:

$$\frac{p_{n,t}}{q_{n,t}} \leq P_{n,t} \leq \overline{P}_{n,t}, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (2c)$$

$$q_{n,t} = \tau c q_{n,t}^c, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (2d)$$

$$0 \leq p_{n,t}^p \leq \overline{P}_{n,t}, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (2e)$$

$$p_{n,t}^p \leq q_{n,t}^p \leq \overline{P}_{n,t} p_{n,t}^p, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}. \quad (2f)$$

Constraints (2a)-(2f) define the feasible set $\mathcal{X}_a$ of LA decisions, containing vectors $x_a = (p_n, q_n)_{n \in \mathcal{N}_a}$. Both, consumption and production, must be scheduled by the LA, taking into account the current spot market prices, and other specific local factors characterizing the private objectives of the LA. Formally, there is a convex cost function $\phi_a(x_a)$ which the LA would like to unilaterally minimize, subject to private feasibility $x_a \in \mathcal{X}_a$.

C. The distribution system operator

In order to guarantee stability of the distribution network, the DSO takes the individual aggregators’ decisions into account and adjusts the power flows so that the flow conservation constraints (1a)-(1b), together with the SOCP constraints (1c)-(1g), are satisfied. Let $x_0 = (p_0, q_0, f, g, v, l)$ denote the vector of the variables controlled by the DSO, and define the DSO’s feasible set $\mathcal{X}_0 = \{x_0 | (1a)-(1g) \text{ hold for } n \in \mathcal{N}\}$. Then, the set of DSO decision variables inducing a physically meaningful network flow for a given tuple of LA decisions $x_A$ is described as

$$\mathcal{F}(x_A) = \{x_0 \in \mathcal{X}_0 | (1a)-(1b) \text{ hold for } x_A\}.$$

Denoting the DSO cost function $\phi_0(x_0)$, we arrive at the DSO’s decision problem

$$\Psi(x_A) = \min \{\phi_0(x_0) | x_0 \in \mathcal{F}(x_A)\}, \quad (3)$$
Algorithm 1: privacy-preserving DLMP solver (PPDLMP)

Parameters: \( p = |A|, \sigma > 0, T_0, T_A \)

Initialization at each LA \( a \in A \):

\[
x_0^a \in X_A
\]

send bid \( u_a = A_0x_0^a - b_a \) to the DSO

Initialization at the DSO:

\[
v^0 = \sigma \sum_{a \in A} u_a, \quad y^0 = v^0 + \sigma (A_0x_0^a - b_0)
\]

Output: \( x^k, s^k = \frac{1}{k} \sum_{i=1}^k x^i \)

for \( k = 0, 1, \ldots \) do

at the DSO do

1. \( x_{0}^{k+1} = \arg \min_{x_0 \in X_0} \{ (\nabla \phi(x_0^k) + A_0^T y^k, \tilde{x}_0) + \frac{1}{2} \| \tilde{x}_0 - x_0^k \|_2^2 \} \)

2. \( x_{a}^{k+1} = \arg \min_{x_a \in A_a} \{ (\nabla \phi_a(x_a^k) + A_a^T y_a^k, \tilde{x}_a) + \frac{1}{2} \| \tilde{x}_a - x_a^k \|_2^2 \} \)

3. \( w^k = A_a(x_a^{k+1} - x_a^k) \)

at each other LA \( a' \neq a \) do

4. \( x_{a'}^{k+1} = x_{a'}^k \)

at the DSO do

5. receive bid \( w^k \) from LA \( a \)

6. \( y^{k+1} = y^k + \sigma (A_0(2x_0^{k+1} - x_0^k) - b_0) + r^k + \sigma (p+1)w^k \)

7. \( y^{k+1} = y^k + \sigma w^k \)

III. PRIMAL-DUAL BLOCK COORDINATE DESCENT

We are facing a multi-agent optimization problem, in which LAs and a single DSO aim for solving the AC-OPF problem by unilaterally solving their individual cost minimization problem. All these decision problems are coupled thanks to the network flow constraints (1a)-(1b). Algorithm 1 proposes a privacy-preserving DLMP solver (PPDLMP), in which the DSO influences the decentralized decisions of the LAs by sending out information about prevailing DLMPs, and iteratively updates DLMPs based on the power profiles in the local markets. PPDLMP asks the DSO to adjust DLMPs based on the prevailing plans reported by the LAs. Once the price update is completed, a single LA is appointed at random to adapt the power profile within the subnetwork this LA manages. The local update of the LA results in bid vector \( w^k \), which will be fed into the DSO final computational step to perform dispatch. Hence, PPDLMP is based on block-coordinate primal updates, involving pairs of the type \((x_0, x_a)\) picked randomly with probability \(1/|A|\) for every \( a \in A \). The optimization problem solved by this algorithm can be described compactly in terms of the convex master problem

\[
\begin{align*}
\text{minimize}_{x \in \mathbb{R}^m} & \quad \phi(x) + r(x) = \Phi(x) \\
\text{subject to} & \quad x \in \arg \min_{x \in \mathbb{R}^m} h(x).
\end{align*}
\] (4)

where \( h(x) = \frac{1}{2} \|Ax - b\|^2 \), in which \( A \in \mathbb{R}^{q \times m} \) and \( b \in \mathbb{R}^q \). We assume that the decision variable is partitioned into \( d \) blocks \( x = (x_1, \ldots, x_d) \) with \( x_i \in \mathbb{R}^{m_i} \) and \( \sum_{i=1}^d m_i = m \). The separable cost function \( \phi(x) = \sum_i \phi_i(x_i) \) is convex and smooth in each block. The non-smooth component \( r(x) = \sum_i r_i(x_i) \) is additively separable with respect to the \( d \) block-coordinate directions, and write \( A = (A_1 \ldots A_d) \) with \( A_i \in \mathbb{R}^{q \times m_i} \) for \( i = 1, \ldots, d \). We assume that \( r : \mathbb{R}^m \to (-\infty, +\infty) \) is a proper closed lower semi-continuous and prox-friendly function. In order to recover the OPF problem, we identify each function \( \phi_i \) with a cost function of the DSO or LA, and \( r_i \) is an indicator function of the feasible set \( X_a \) and \( X_0 \), respectively. We also assume that there exists a positive semidefinite matrix \( \Lambda \in \mathbb{R}^{m \times m} \) such that, for every \( x, \tilde{x} \in \text{dom}(r) \), it holds that

\[
\phi(\tilde{x}) \leq \phi(x) + (\nabla \phi(x), \tilde{x} - x) + \frac{1}{2} \| \tilde{x} - x \|_2^2,
\] (5)

where \( \| \cdot \|_{\Lambda} \) is the induced \( \| \cdot \|_\Lambda \) norm.

Define the \( m \times m \) weighting matrices \( P \) and \( \Sigma \) by

\[
P_{i} \text{ def } \text{diag}(1/\pi_1, \ldots, 1/\pi_d) \quad \text{and} \quad P_{i} \text{ def } \text{diag}(1/\pi_i, \ldots, 1/\pi_d)
\]

for all \( i \in I \). Similarly, let \( T \text{ def } \text{diag}(T_1, \ldots, T_d) \geq 0 \) be a block diagonal matrix and, for each \( i \in I \), define \( T_{i} \text{ def } \text{diag}(T_{i1}, \ldots, T_{id}) \). If coordinate \( i \) is selected for updating, a proximal-based update step, based on the linearization

\[
\xi_i^k(\tilde{x}_i) \text{ def } \langle \nabla \phi_i(x_i^k) + A_i^T y_i^k, \tilde{x}_i \rangle.
\]

This delivers the next iterate

\[
x_i^{k+1} = \arg \min_{u_i} \{ \xi_i^k(u_i) + r_i(u_i) + \frac{1}{2} \| u_i - x_i^k \|_{T_i}^2 \}.
\]

In Appendix I we show that PPDLMP (Algorithm 1) is a special case of the more general primal-dual Algorithm 2.

In Algorithm 2, a sensible choice for \( T \) is to set

\[
T_i = I_{m_i} / \tau_i + \pi_i A_i + \sigma A_i^T A_i \quad (i = 1, \ldots, d)
\] (6)

with the constraint \( T \succ \sigma \Sigma_0 \), where \( \Sigma = \Sigma_{ij} \) and \( \tau_i = \text{Prob}(i, j) \). A detailed analysis of the sequence generated by Algorithm 2 yields our main result.

Theorem 1: Let \( x^* \) denote the solution set of Problem (4), and let \( (x^k)_k \) and \( (s^k)_k \) be issued by Algorithm 2 with \( s^k = \frac{1}{k} \sum_{l=1}^k x^l \) and with \( T, \sigma \) satisfying (6). Then,

(i) If there exists a Lagrange multiplier for Problem (4), then \( (x^k)_k \) and \( (s^k)_k \) converge a.s. to a solution of (4) and

\[
\lim_{k \to \infty} (h(x^k) - h^*) = o(1/k), \quad h(s^k) - h^* = O(1/k^2)
\] a.s.
Algorithm 2: Primal-dual Block Coordinate Descent

Parameters: $P, \sigma > 0, T, (\theta_k)_{k \geq 0}$

Initialization: $x^0 \in \mathbb{R}^m, u^0 = \sigma(Ax^0 - b)$

Output: $x^k, s^k = \frac{1}{k} \sum_{i=1}^k x^i$

for $k = 0, 1, 2, \ldots$ do
1. draw block $i \in \mathcal{I}$ at random according to $\Pi$
2. $x^k_{i+1} = \arg \min_{x_{i+1}} \{ \nabla \phi_i(x_{i+1}) + A_i y_{i+1} + r_i(x_{i+1}) + \frac{1}{2} \|x_{i+1} - x_i\|_2^2 \}$
3. $x^{k+1} = x^{k+1}_{i+1}$
4. $u^{k+1} = u^k + \sigma A(x^{k+1} - x^k) + u^{k+1}$
5. $y^{k+1} = y^k + \sigma AP(x^{k+1} - x^k) + u^{k+1}$

(ii) If $\mathcal{X}^*$ is a bounded set and $\phi + r$ is bounded from below, then a.s. all limit points of $(s^k)_k$ belong to $\mathcal{X}^*$ and $h(s^k) - h^* = o(1/k)$.

Because of space limitations, we are not able to give the proof of Theorem 1 in full detail here, but we present a sketch of it in Appendix II. The interested reader can find a detailed proof in the arXiv Version of this paper [21].

IV. NUMERICAL RESULTS

We apply Algorithm 1 to a realistic 15-bus network example based on the instance proposed in [13], over a time horizon $T = \{0, 1\}$. The network parameters are specified in Table I. Lines physical parameters $(R_n, X_n, S_n, B_n, V_n)$ are those used in [13]. While [13] considers fixed loads, here we consider variable, flexible active and reactive loads as specified in (2) and with parameters $(P_n^0, P_n, E_n, \tau_n^c)_n$ generated based on the values of [13]; see also [22].

Bus 11 is the only bus to offer renewable production, with $P_{11}^p \overset{\text{def}}{=} [0.438, 0.201]$ and $p^p = \bar{p}^p = 0$ (the renewable production is purely active power). The bounds $(V_n, \bar{V}_n)$ are set to 0.81 and 1.21 for each $n \in \mathcal{N}$, while $V_0 = 1.0$.

We consider a zero cost function for each LA ($\phi_a = 0$ for each $a \in \mathcal{A}$), meaning that LAs are indifferent to consumption profiles for as long as their feasibility constraints are satisfied. This is a reasonable assumption in practice for certain types of consumption flexibilities (electric vehicles, batteries). We consider the DSO objective

$$\phi(x) = \phi_0(x_0) = \sum_{t \in T} c_t(p_{0t}) + k^\text{loss} \sum_{n, t} R_n \epsilon_{nt},$$

with loss penalization factor $k^\text{loss} = 0.001$ and with:

$$c_0 : p \mapsto 2p + p^2, \quad c_1 : p \mapsto p,$$

giving an expensive time period and a cheap one, which can be interpreted as peak and offpeak pricing periods.

The solution obtained by Algorithm 1 after 2000 iterations is illustrated in Figure 1, which displays the active flow directions as well as the DLMP values. The solutions show that the active (and reactive) DLMPs obtained for each time period are close to the DLMPs at the root node $(y^0_0, y^0_0)$, with the following exceptions:

- For the branch composed of nodes 8, 7, 9, 10, 11, active DLMPs are close to 0.0 due to the presence of renewable production (at null cost) at node 11, and of negative load at node 7, which together fully compensate for the demand on this branch. Since Line (3, 8) is saturated, no energy can be exported further.
- Active DLMPs on the branch composed of nodes (12, 13, 14) at $t = 1$ are much larger than on other nodes; this is explained by the congestion of line (0, 12).
- The DLMP for node 7 and $t = 0$ is strictly negative: the (negative) consumption for this node is at its upper bound $p_{7,0} = P_{7,0} = -0.173$. The negative DLMP suggests that the system will be better off if less power is injected by node 7.

Convergence of Algorithm 1 for the 15-bus network is shown in Figures 2 and 3. Figure 2 displays the convergence of the last iterate with respect to various criteria: convergence of $\phi(x^k)$ to the optimal cost $\phi(x^*)$, convergence to zero of the primal residuals $h(x^k)$ and convergence to zero of the

---

TABLE I: Parameters for the 15 buses network based on [13]

| $n$ | $S_n$ | $R_n \cdot 10^4$ | $X_n \cdot 10^3$ | $B_n \cdot 10^6$ | $P_n$ | $P_n$ | $E_n$ | $\tau_n^c$ |
|-----|-------|-----------------|-----------------|-----------------|------|------|------|-------|
| 1   | 2.000 | 1.0            | 120.0           | 1.1             | 0.593| 0.256| 1.566| 1.539 |
| 2   | 0.256 | 88.3           | 126.2           | 2.8             | 0.000| 0.000| 0.000| 0.000 |
| 3   | 0.256 | 138.4          | 197.8           | 2.4             | 0.003| 0.011| 0.020| 0.035 |
| 4   | 0.256 | 19.1           | 27.3            | 0.4             | 0.015| 0.013| 0.027| 0.019 |
| 5   | 0.256 | 17.3           | 25.1            | 0.8             | 0.021| 0.024| 0.043| 0.053 |
| 6   | 0.256 | 48.2           | 68.9            | 0.6             | 0.017| 0.001| 0.032| 0.037 |
| 7   | 0.256 | 40.7           | 55.2            | 1.2             | 0.021| 0.009| 0.040| 0.039 |
| 8   | 0.256 | 52.3           | 74.7            | 0.6             | -0.233| -0.210| -0.173| -0.115 |
| 9   | 0.256 | 10.0           | 14.3            | 0.4             | 0.008| 0.002| 0.032| 0.028 |
| 10  | 0.256 | 24.1           | 34.5            | 0.4             | 0.004| 0.001| 0.024| 0.040 |
| 11  | 0.256 | 10.3           | 14.8            | 0.1             | 0.010| 0.010| 0.015| 0.024 |
| 12  | 0.600 | 1.0            | 120.0           | 0.1             | 0.243| 0.057| 0.642| 0.625 |
| 13  | 0.204 | 155.9          | 111.9           | 0.2             | 0.001| 0.000| 0.003| 0.003 |
| 14  | 0.204 | 95.3           | 68.4            | 0.1             | 0.015| 0.012| 0.032| 0.042 |

Fig. 1: Directions of active flows $f$ and DLMPs $(y^p, y^q)$ at the solution given by Algorithm 1. Saturated lines are dashed. 

---
Fig. 2: Convergence of last iterate \( x^k \)

![Image](image_url)

Fig. 3: Convergence of ergodic average \( s^k \)

KKT residual [7]:

\[
R^{\text{KKT}}(x^k, y^k) = \text{dist}_\infty((\partial_x, -\partial_y)L(x^k, y^k), 0),
\]

where \( L(x, y) = \Phi(x) + (y, Ax - b) \) denotes the Lagrangian of (4). Figure 3 shows the convergence to zero of the primal infeasibility in the ergodic average \( s^k \), as predicted by Theorem 1.

V. CONCLUSION

We developed a novel distributed and privacy-preserving algorithm for the computation of distribution locational marginal prices. Our computational strategy builds on extends state-of-the-art block coordinate descent algorithms for convex optimization problems with affine coupling constraints. Non-convex versions of PPDLMP will be investigated in the future. We also plan to conduct extensions of this work where the electric network is exposed to stochastic uncertainty.

APPENDIX I

THE RELATION OF ALGORITHM 2 TO ALGORITHM 1

In this section we show that Algorithm 2 contains Algorithm 1 as a special case. In the latter algorithm, the sampling takes values from subsets of the form \( i = \{0, a\} \), where \( a \in \{1, \ldots, |A|\} = \{1, \ldots, p\} \), with probability \( \Pi_i = 1/p \) for each \( i \). Thus, \( d = p + 1, \pi_0 = 1, \) and \( \pi_i = 1/p \) if \( i \in \{1, \ldots, p\} \). The weighting matrix \( P \) is given by \( P = \text{diag}(I_{m_0}, pI_{m_1}, \ldots, pI_{m_p}) \), where \( m_0 \) is the dimension of the feasible set of the DSO, and \( m_a \) is the dimension of the feasible set of aggregator \( a \in A \). Now, define the function \( r \) in (4) as \( r(x) = r_0(x_0) + \sum_{a \in A} r_a(x_a) \), where \( r_0 = 1_{x_0 \in X_0} \) and \( r_a = 1_{x_a \in X_a} \), in which \( 1 \) denotes the indicator function. If the load aggregator \( a \) is chosen at step \( k \), Line 5 in Algorithm 2 becomes

\[
y^{k+1} = y^k + \sigma A_P(x^{k+1} - x^k) + v^{k+1} = y^k + \sigma A_0(x^{k+1}_a - x^k_a) + \sigma p_A(x^{k+1}_a - x^k_a) + v^{k+1} = y^k + \sigma |A_0(2x^{k+1}_a - x^k_a) - b_0| + \sigma p_A(x^{k+1}_a - x^k_a) + v^{k+1}
\]

where we define \( v^k = u^k - [A_0 x^k_0 - b_0] \). Exploiting Line 4 in Algorithm 2, we find that \( v^k \) can be computed locally and inductively by choosing the initial condition \( v^0 = \sigma \sum_{a \in A} (A_a x^0_a - b_a) \) initially, then by letting

\[
v^{k+1} = v^k + \sigma A_0(x^{k+1}_a - x^k_a) \equiv v^k + \sigma u^k
\]

where \( u^k = A_0(x^{k+1}_a - x^k_a) \), and Line 5 rewrites as

\[
y^{k+1} = y^k + \sigma |A_0(2x^{k+1}_a - x^k_a) - b_0| + \sigma (p + 1) u^k + v^k.
\]

APPENDIX II

CONVERGENCE ANALYSIS

The proof for Theorem 1 uses the reduction of Algorithm 2 to the simpler Algorithm 3. It is straightforward to show the equivalence between these two schemes if we set \( \theta_k = 1/(k + 1) \) and \( y^k = (\sigma/\theta_k)(Az^k - b) \) [6].

For analysis purposes we introduce the auxiliary sequence

\[
\hat{x}^{k+1} = \arg\min_{\hat{x}} \left\{ \langle \nabla \phi(x^k) + \frac{\sigma}{\theta_k} \nabla h(z^k), \hat{x} \rangle + r(\hat{x}) + \frac{1}{2} \| \hat{x} - x^k \|_{L^2}^2 \right\}.
\]

The iterate \( \hat{x}^{k+1} \) corresponds to the next fictitious state if all coordinates were to perform an update in parallel. We now illustrate the main steps involved in proving convergence of the iterates produced by running Algorithm 3.

For \( i \in \{1, \ldots, d\} \) let \( U_i \) be the \( m \times m \) block unitary matrix of the form \( U_i = \text{diag}(0, \ldots, I_{m_i}, 0, \ldots, 0) \). Clearly \( \sum_{i=1}^d U_i = I_m \), and applying the matrix \( U_i \) to the left of a vector \( t = (t_1, \ldots, t_d)^T \) gives \( U_i t = (0, \ldots, t_i, \ldots, 0)^T \in \mathbb{R}^m \). For \( i \in I \), define the \( m \times m \) matrix \( U_i \equiv \sum_{i \in I} U_i \). We have \( \mathbb{E}[U_i P] = I_m \), with \( i \sim \mathcal{U}(I) \), and we define \( \Sigma = \mathbb{E}[U_i PA^TAPU_i] \). It follows from the quadratic form of \( h \) that

\[
\mathbb{E}[h(x + U_i Pt)] = h(x) + \langle \nabla h(x), t \rangle + \frac{1}{2} \| t \|_{L^2}^2.
\]

4147
Let $F_k := \sigma(x_0, s_0, z_0, \ldots, x_k, s_k, z_k)$ denote the history of the process up to step $k$. We infer the following result for Algorithm 3, which corresponds to an Expected Separable Overapproximation (ESO), as introduced in [23], [24], [25].

**Lemma 1:** In Algorithm 3,
\[
E[h(s^{k+1})|F_k] = h(x_k) + \theta_k (\nabla h(x_k), \hat{x}^{k+1} - x_k) + \frac{\theta_k^2}{2} \|\hat{x}^{k+1} - x_k\|^2_\Sigma
\]

where $\hat{x}_k := \hat{\phi}(\hat{e}_k, \hat{u}_k, \hat{p}_k)$ denotes the next iterate. This result characterizes the sequence $(s^k)$ as a linear combination of the past primal iterates. This characterization is a generalization of [24, Lemma 2], and its proof is similar to that work.

**Lemma 2 (Proximal step):** In Algorithm 3, $\forall x \in \mathbb{R}^m$,
\[
r(\hat{x}^{k+1}) + \zeta(\hat{x}^{k+1}) \leq r(x) + \zeta(x) - \frac{1}{2} \|x - \hat{x}^{k+1}\|^2_{P_T},
\]
where
\[
\zeta(x) = \phi(x_k) + \langle \nabla \phi(x_k), x - x_k \rangle + \frac{1}{2} \|x - x_k\|^2_{P_T}.
\]

**Lemma 3:** In Algorithm 3, for any $x^* \in X^*$,
\[
E[h(s^{k+1})] = \theta_k (\nabla h(x_k), x_k^* - x_k) + \frac{1}{2} \|x_k^* - x_k\|^2_{P_T}.
\]

**Lemma 4:** In Algorithm 3, we have
\[
s^k = \sum_{l=0}^k \Gamma_k x^l, \quad k \geq 1,
\]
where $(\Gamma_k)$ is a collection of diagonal matrices defined by
\[
\Gamma_0^l = I - \theta_0 P, \quad \Gamma_1^l = \theta_0 P, \quad \text{and}, \quad \Gamma_k^l = \left\{ \begin{array}{ll}
(1 - \theta_k) \Gamma_k^l & \text{for } l = 0, \ldots, k-1, \\
(1 - \theta_k) \theta_{k-1} P - \theta_k (P - I) & \text{if } l = k, \\
\theta_k P & \text{if } l = k+1.
\end{array} \right.
\]

Besides, $\Gamma_k^l + \Gamma_k^{l+1} = (1 - \theta_k) \Gamma_k^l - \theta_k (P - I)$.

Now, define $\Phi = (\Phi_1, \ldots, \Phi_d)$ and $\hat{\Phi}_k = 1 \hat{\Phi}_k$, where $\hat{\Phi}_k = \sum_{l=0}^k \Gamma_k \Phi(x^l)$. By convexity, we can show that $\hat{\Phi}_k \geq \Phi(s_k)$ and $\hat{\Phi}_k \geq \Phi(s^k)$. Choosing $\theta_k = 1/(k+1)$, we obtain a Lyapunov function
\[
L_k = k[\hat{\Phi}_k - \Phi(x^*)] + \sigma k^2 (h(x^*) - h(x^*)) + \frac{1}{2} \|x^* - x^*\|^2_{P_T},
\]
satisfying $E[L_k+1|F_k] \leq L_k - \sigma h(x^*) - h(x^*) - \frac{1}{2} \|x^{k+1} - x_k\|^2_{\Sigma}$. The rest of the proof relies on arguments due to [26], where it follows the exact steps [7, pp. 13-15]. We omit the details here.

**REFERENCES**

[1] National Academy of Sciences, Engineering, and Medicine, Analytic Research Foundations for the Next-Generation Electric Grid. Washington, DC: The National Academies Press, 2016.

[2] P. M. Sokoliewicz and J. M. Vignolo, “Nodal pricing for distribution networks: efficient pricing for efficiency enhancing dg.” IEEE Transactions on Power Systems, vol. 21, no. 2, pp. 1013–1014, 2006.

[3] G. T. Heydt, B. H. Chowdhury, M. L. Crow, D. Haughton, B. D. Kiefer, F. Meng, and B. R. Sathyaranayana, “Pricing and control in the next generation power distribution system,” IEEE Transactions on Smart Grid, vol. 3, no. 2, pp. 907–914, 2012.

[4] N. O’Connell, Q. Wu, J. Østergaard, A. H. Nielsen, S. T. Cha, and Y. Ding, “Day-ahead tariffs for the alleviation of distribution grid congestion from electric vehicles,” Electric Power Systems Research, vol. 92, pp. 106–114, 2012.

[5] R. Singh and S. Gowsain, “Optimum allocation of distributed generations based on nodal pricing for loss reduction, and voltage improvement including voltage rise issue,” International Journal of Electrical Power & Energy Systems, vol. 32, no. 6, pp. 637–644, 2010.

[6] Y. Malitsky, “Chambolle-pock and tseng’s methods: relationship and extension to the bilevel optimization,” 06 2017.

[7] D. R. Luke and Y. Malitsky, Block-Coordinate Primal-Dual Method for Nonsmooth Minimization over Linear Constraints. Cham: Springer International Publishing, 2018, pp. 121–147.

[8] Y. Malitsky, “The primal-dual hybrid gradient method reduces to a primal method for linearly constrained optimization problems,” 2019.

[9] R. Li, Q. Wu, and S. S. Oren, “Distribution locational marginal pricing for optimal electric vehicle charging management,” IEEE Transactions on Power Systems, vol. 29, no. 1, pp. 203–211, 2014.

[10] S. Huang, Q. Wu, S. S. Oren, R. Li, and Z. Liu, “Distribution locational marginal pricing through quadratic programming for congestion management in distribution networks,” IEEE Transactions on Power Systems, vol. 30, no. 4, pp. 2170–2178, 2014.

[11] M. Farivar and S. H. Low, “Branch flow model: Relaxations and convexification—part i: ieee Transactions on Power Systems, vol. 28, no. 3, pp. 2554–2564, 2013.

[12] B. Kocuk, S. S. Dey, and X. A. Sun, “Strong socp relaxations for the optimal power flow problem,” Operations Research, vol. 64, no. 6, pp. 1177–1196, 2020/02/20 2016.

[13] A. Papavasiliou, “Analysis of distribution locational marginal prices,” IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 4872–4882, 2018.

[14] I. Alsaheh and L. Fan, “Distribution locational marginal pricing (dlmp) for multiphase systems,” in 2018 North American Power Symposium (NAPS), 2018, pp. 1–6.

[15] X. Zhou, E. Dall’Anese, L. Chen, and A. Simonetto, “An incentive-based online optimization framework for distribution grids,” IEEE Transactions on Automatic Control, vol. 63, no. 7, pp. 2019–2031, 2017.

[16] P. Peng and S. H. Low, “Distributed optimal power flow algorithm for radial networks, i: Balanced single phase case,” IEEE Transactions on Smart Grid, vol. 9, no. 1, pp. 111–121, 2018.

[17] D. K. Molzahn, I. A. Hiskens et al., “A survey of relaxations and approximations of the power flow equations,” Foundations and Trends® in Electric Energy Systems, vol. 4, no. 1-2, pp. 1–221, 2019.

[18] Y. Nesterov, Lectures on Convex Optimization, ser. Springer Optimization and Its Applications. Springer International Publishing, 2018, vol. 137.

[19] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” Journal of Mathematical Imaging and Vision, vol. 40, no. 1, pp. 120–145, May 2011.

[20] P. Tseng, “On accelerated proximal gradient methods for convex-concave optimization,” 2008.

[21] O. Bilenne, P. Jacquot, and C. Wan, “A privacy-preserving distributed computational approach for distributed locational marginal pricing,” arXiv preprint arXiv:2103.14094, 2021.

[22] P. Jacquot, “DLMP-based Coordination Procedure for Decentralized Demand Response under Distribution Network Constraints,” Aug. 2020, working paper or preprint.

[23] P. Richtárik and M. Takáč, “Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function,” Mathematical Programming, vol. 144, no. 1, 2014.

[24] G. Fercoq and P. Richtárik, “Accelerated, parallel, and proximal coordinate descent.” SIAM Journal on Optimization, vol. 25, no. 4, pp. 1997–2023, 2015.

[25] P. Richtárik and M. Takáč, “Parallel coordinate descent methods for big data optimization,” Mathematical Programming, vol. 156, no. 1-2, pp. 433–484, 2016.

[26] M. Solodov, “An explicit descent method for bilevel convex optimization,” Journal of Convex Analysis, vol. 14, pp. 227–238, 04 2007.