Distribution Network Electricity Market Clearing: Parallelized PMP Algorithms with Minimal Coordination

Elli Ntakou and Michael Caramanis†

***Abstract*— The socially optimal power market clearing problem with diverse, complex-utility-structure participants at the distribution level, poses computational challenges that are exacerbated by the associated non-convex load flow constraints. We investigate variations and extensions of Proximal Message Passing algorithms proposed in the literature and implemented on similar, though simpler, social welfare function instantiations. Numerical results demonstrate that (i) in comparison to solving a computationally demanding, yet exact, centralized market clearing problem, significant computational improvements are possible with the proposed PMP algorithm extensions, (ii) comparison to the benchmark results obtained by solving the centralized formulation reveals that excellent accuracy is attainable by the PMP algorithms and (iii) in comparison to the PMP algorithms existing in the literature, the proposed extension reduces significantly the communication requirements for sub-problem coordination and convergence verification and enables faster converging asynchronous sub-problem iterations.**

I. Introduction

Increased computation and communication capabilities in today's cyber-enabled smart power grid, and the promise of distributed generation, demand response and other distributed resources such as EVs that will be able to mitigate the cost of renewable generation volatility, have peaked interest in distribution network power markets [1,8].

In [1], we formulated and solved numerically the full AC optimal power flow (OPF) problem for a moderate size distribution network with complex participants. More specifically, we used a radial distribution network consisting of two feeders: a primary voltage industrial feeder with 47 busses published in [5] and a secondary voltage commercial/residential feeder with 206 busses that we created by expanding the industrial feeder to include primary to secondary voltage transformers and secondary voltage lines serving commercial and residential loads. Both feeders have distributed PV generation with power electronics rectifiers that can be put to dual use providing reactive power compensation and voltage control as needed. They also feature capacitors dedicated to reactive power compensation, and conventional loads with fixed load factors. Finally, the cost of loss of transformer life from overloading, the opportunity cost of providing reactive power and higher voltage at the substation are also modeled.

†Boston University Division of Systems Engineering, entakou@bu.edu, mcaraman@bu.edu. Research reported here was supported by NSF Grant 1038230

The non-convexities inherent in the AC load flow equations are major contributors to the computational complexity of the AC OPF problem that is essentially equivalent to the socially optimal market clearing problem [3,4,8]. The computational burden of the numerical solutions reported in [1] is further exacerbated by the detailed modeling of the diverse and complex cost and capability structures of the distribution market participants, such as encountered in loads that involve power electronics (EVs, PV, variable speed drive HVAC) and have storage like degrees of freedom. At a real size urban scale application, where thousands of active distributed market participants will be involved, centrally solved market clearing problems become intractable.

In this paper, we investigate an alternate, decentralized yet iterative problem formulation based on the Proximal Message Passing (PMP) algorithm by Boyd et al, [2] which is in principle scalable under parallel computation implementations. We borrow most of our notation from [2]. Distributed optimization methods applied to power systems have focused either on dual decomposition to take advantage of separability, or on augmenting the Lagrangian in order to take advantage of milder convergence conditions. The Proximal Message Passing formulation uses the Alternating Method of Multipliers [11] which in essence combines decomposability and convergence robustness. In the PMP algorithms, market participants (loads, generators, etc) as well as distribution network components (lines, transformers, etc), denoted collectively as devices, solve in parallel device-specific optimization sub-problems reflecting their own cost and capability structure as well as imbalances with neighboring devices. Devices exchange messages with their neighbors at each iteration. The iterative PMP problem is completely decentralized allowing solution of all sub-problems in parallel. If the iterations are implemented in a synchronous fashion, a global scale coordination is needed in the form of a common clock controlling the iteration count and determining global convergence. Under such global coordination and the condition that individual sub-problems are convex, the algorithm is guaranteed to converge [2,6].

In this paper we consider modifications and extensions to both the market complexity that is modeled and the PMP algorithm specifics used in [2]. While [2] is mostly a proof of concept that PMP can be applied to the solution of the AC load flow based power market clearing problem, we employ, as noted, a detailed and more realistic diverse market participant model. For example, we add variables (like reactive power profiles) and relax simplifying assumptions such as constant voltage and line flow approximations. Our market clearing PMP formulation is able to capture the exact same solution obtained by our centralized non-convex AC OPF based benchmark in [1].

We note related work in [7] which applies distributed algorithms to a market model that is closest to ours. Contrary to our model, [7] decomposes the centralized problem to bus-specific subproblems, which are fewer than the device specific subproblems in the PMP algorithm we use. In doing so, [7] does not model continuously schedulable real and reactive power providing resources. To do so, more optimization problems would have to be solved externally at the level of the bus aggregator. In addition, [7] does not model voltage decisions at the substation bus, and its convexification requirements impose voltage magnitude constraints only from below. It also restricts power flow reversal.

This paper's contribution is the proposal and performance evaluation (w.r.t. computational as well as communication requirements) of Proximal Message Passing (PMP) algorithm extensions and variations relative to published distributed algorithms, most notably, among others, in [2] and [7]. The PMP extensions proposed here include:

- The introduction of net specific penalty factors that drive discrepancies in real and reactive power, voltage and voltage phase angles to zero at each net. Three alternative penalty adaptation methods are used: (i) centrally with a single penalty used across the whole network , (ii) in a distributed fashion for each net, and (iii) finally in a distributed fashion for each net and each of the key quantities, namely real power, reactive power, voltage level and phase angle.
- Using a convergence criterion based on (i) either a network-wide, centrally evaluated tolerance or (ii) a location specific tolerance, evaluated in a distributed manner.
- Although not implemented here, the above enables asynchronous parallel processing.

Numerical results demonstrate that (i) in comparison to solving a computationally demanding, yet exact, centralized market clearing problem, significant computational improvements are possible with the proposed PMP algorithm extensions, (ii) comparison to the centralized solution considered as a benchmark shows that the PMP algorithm converges to an arbitrarily close result, and (iii) relative to PMP algorithms reported in the literature, the PMP extension proposed in this work has significantly reduced global coordination and communication requirements.

The rest of this paper is organized as follows: Section II formulates the market clearing problem, Section III presents the proposed algorithm. Section IV reports on computational results and the demonstrated properties, and Section V concludes.

II. Problem Formulation

We solve the hour-ahead distribution market clearing problem which schedules real and reactive nodal power injections that minimize:

(*i*) the cost of real power procured at the substation, plus

(*ii*) the cost of substation voltage increase that may be required for compliant voltage magnitudes throughout the network, minus

(*iii*) real power consumer utility, plus

(*iv*) substation generator opportunity cost associated with providing the reactive power reaching the substation, plus

(*v*) the cost of transformer loss of life, plus

(*vi*) distributed generation costs,

subject to:

(*i*) AC load flow relationships,

(*ii*) real and reactive power injections by loads and generators,

(*iii*) power conditioning assets accompanying loads such as asynchronous motor HVAC systems, PV installations, elevator banks, and lastly

(*iv*) voltage magnitude constraints.

This centralized market clearing problem was solved for the 253 bus distribution network in [1] and is used here as a benchmark. As mentioned already, the centralized solution of [1] is not scalable to real distribution systems. The remainder of this section presents the detailed formulation of several parallelizable PMP algorithm that are scalable.

A. Parallelizable Network Representation

First, we need to redefine the network elements such that their properties can be easily used in the concept of proximal message passing algorithms.

We follow [2], in decomposing the distribution network and market participants to devices *a*, terminals *t* and nets *n*. Each device and each net has a set of terminals associated with it. Each terminal is associated with exactly one device and exactly one net. Devices model market participants (such as generators, loads, power electronics, etc.) and network components, (such as lines, transformers, etc). Nets are lossless energy carriers, the equivalent of busses. At each net, real and reactive power balance, phase consistency and voltage consistency are required. We visualize message passing, or communication of information across the network as taking place over a bipartite graph, where, the devices and the nets are the two classes of vertices while the terminals are the edges connecting them. This enables the distributed nature of the PMP algorithm where devices solve their own sub-problems, communicate their tentative results to their neighbors by passing them to nets that they are associated with, and iterate until imbalances and inconsistencies at the nets are eliminated. Notation is summarized next.

B. Notation Conventions

d_i, g_i, f_i, e_i: Subscripts denoting respectively a specific distributed load, distributed generation, shunt capacitor or distributed power electronics.

c_{g_i}, u_{d_i}: Marginal Cost and marginal Utility, respectively, associated with generation type/load type g_i, d_i .

φ_{d_i}: fixed current/voltage phase shift introduced by load d_i

c_∞^V: cost of substation voltage rise

π_∞: Substation Locational Marginal Price. We consider a decoupling of the transmission system from distribution feeders that are connected to a transmission bus with a known Locational Marginal Price (LMP). As such, we treat the LMP at the substation bus as an exogenously specified quantity.

$C_\infty, C_{g_i}, C_{f_i}, C_{e_i}$: Capacity of reactive power compensating generator at the substation bus, distributed generator g_i , capacitor f_i , and power electronics e_i .

k: iteration count

t, a, n: terminal, device and net subscript

T: total number of terminals in the network

$|n|, |a|$: number of terminals associated with net n or device a

t_n, t_a : set of terminals t associated with net n or device a

R_{t_1,t_2}, X_{t_1,t_2} : per unit series resistance, reactance of line t_1, t_2

$I_{t_1,t_2}, \ell_{t_1,t_2}$: line t_1, t_2 current and current magnitude squared.

$c^{tr}_{t_1,t_2}$: Cost of one hour of transformer t_1, t_2 economic life

$S^N_{t_1,t_2}$: Apparent flow rating of transformer t_1, t_2

$k_{1,t_1,t_2}, k_{2,t_1,t_2}$: Transformer t_1, t_2 hottest spot temperature coefficients.

ρ^k : Penalty factor for the network at iteration k

ρ^k_n : Penalty factor specific to net n at iteration k

$\rho^k_{n,P}, \rho^k_{n,Q}, \rho^k_{n,V}$: Penalty factor specific to net n and to real power, reactive power or voltage respectively at iteration k

p_t : real power injected at terminal t. $p_t > 0$ refers to real power consumption, while $p_t < 0$ refers to real power generation.

p_{e_i} : real power generated by distributed PV generator e_i (i.e. $p_{e_i} < 0$). We consider this to be exogenously specified, i.e. p_{e_i} is a known constant.

p_{d_i} : fixed real power demand of constant load d_i (i.e. $p_{d_i} > 0$)

$\mathbf{p}_a$: $|a| \times 1$ vector with elements $p_t, \forall t \in t_a$

$\mathbf{p}_n$: $|n| \times 1$ vector with elements $p_t, \forall t \in t_n$

$\hat{p}_t = \frac{1}{|n|} \sum_{t' \in t_n} p_{t'}, t \in t_n$: real power imbalance at t

$\hat{\mathbf{p}}_a$: $|a| \times 1$ vector with elements $\hat{p}_t, \forall t \in t_a$

$\hat{\mathbf{p}}_n$: $|n| \times 1$ vector with elements $\hat{p}_t, \forall t \in t_n$

θ_t : phase of terminal t

$\boldsymbol{\theta}_a$: $|a| \times 1$ vector with elements $\theta_t, \forall t \in t_a$

$\boldsymbol{\theta}_n$: $|n| \times 1$ vector with elements $\theta_t, \forall t \in t_n$

$\widehat{\theta}_t = \theta_t - \frac{1}{|n|} \sum_{t' \in t_n} \theta_{t'}, t \in t_n$: phase residual at terminal t

$\widehat{\boldsymbol{\theta}}_a$: $|a| \times 1$ vector with elements $\widehat{\theta}_t, t \in t_a$

$\widehat{\boldsymbol{\theta}}_n$: $|n| \times 1$ vector with elements $\widehat{\theta}_t, t \in t_n$

q_t : reactive power schedule of terminal t. $q_t > 0$ refers to reactive power consumption, while $q_t < 0$ refers to reactive power generation.

q_{d_i} : fixed reactive power demand of constant load d_i (i.e. $q_{d_i} > 0$)

$\mathbf{q}_a$: $|a| \times 1$ vector with elements $q_t, \forall t \in t_a$

$\mathbf{q}_n$: $|n| \times 1$ vector with elements $q_t, \forall t \in t_n$

$\hat{q}_t = \frac{1}{|n|} \sum_{t' \in t_n} q_{t'}, t \in t_n$: reactive power imbalance at

$\hat{\mathbf{q}}_a$: $|a| \times 1$ vector with elements $\hat{q}_t, \forall t \in t_a$

$\hat{\mathbf{q}}_n$: $|n| \times 1$ vector with elements $\hat{q}_t, \forall t \in t_n$

V_t : voltage at terminal t

v_t : voltage magnitude squared of terminal t

$\mathbf{v}_a$: $|a| \times 1$ vector with elements $v_t, \forall t \in t_a$

$\mathbf{v}_n$: $|n| \times 1$ vector with elements $v_t, \forall t \in t_n$

$\widehat{v}_t = v_t - \frac{1}{|n|} \sum_{t' \in t_n} v_{t'}, t \in t_n$: voltage residual at terminal t

$\widehat{\mathbf{v}}_a$: $|a| \times 1$ vector with elements $\widehat{v}_t, \forall t \in t_a$

$\widehat{\mathbf{v}}_n$: $|n| \times 1$ vector with elements $\widehat{v}_t, \forall t \in t_n$

$\mathbf{r}^k = (\hat{\mathbf{p}}^k, \hat{\mathbf{q}}^k, \widehat{\mathbf{V}}^k, \widehat{\boldsymbol{\theta}}^k)$: $4|T| \times 1$ vector of imbalances of real power and reactive power, and voltage magnitude and phase angle residuals across all terminals at iteration k

$\mathbf{r}^k_n = (\hat{\mathbf{p}}^k_n, \hat{\mathbf{q}}^k_n, \widehat{\mathbf{V}}^k_n, \widehat{\boldsymbol{\theta}}^k_n)$: $4|n| \times 1$ vector of imbalances of real power and reactive power, and voltage magnitude and angle residuals at net n and iteration k

$\mathbf{s}^k = \rho[(\mathbf{p}^k - \hat{\mathbf{p}}^k) - (\mathbf{p}^{k-1} - \hat{\mathbf{p}}^{k-1}), (\mathbf{q}^k - \hat{\mathbf{q}}^k) - (\mathbf{q}^{k-1} - \hat{\mathbf{q}}^{k-1}), \hat{\mathbf{V}}^k - \hat{\mathbf{V}}^{k-1}, \hat{\boldsymbol{\theta}}^k - \hat{\boldsymbol{\theta}}^{k-1}]$: $4|T| \times 1$ vector of the change of imbalance across all terminals at iteration k

$\mathbf{s}^k_n = \rho[(\mathbf{p}^k_n - \hat{\mathbf{p}}^k_n) - (\mathbf{p}^{k-1}_n - \hat{\mathbf{p}}^{k-1}_n), (\mathbf{q}^k_n - \hat{\mathbf{q}}^k_n) - (\mathbf{q}^{k-1}_n - \hat{\mathbf{q}}^{k-1}_n), \hat{\mathbf{V}}^k_n - \hat{\mathbf{V}}^{k-1}_n, \hat{\boldsymbol{\theta}}^k_n - \hat{\boldsymbol{\theta}}^{k-1}_n]$: $4|n| \times 1$ vector of the change of imbalance at net n, iteration k

σ_a : device a objective function with all device specific constraints appended to it

$\upsilon, \lambda, \mu, \varsigma$: scaled shadow prices of real and reactive energy balance, and voltage and phase consistency constraints.

C. Proximal Message Passing

Given the definitions above, and the classification of network components and market participants to nets, devices and terminals, we next describe device instances or types:

Substation Generator: Single terminal device. Assuming that this generator's variable costs are negligible relative to the LMP, π_∞, the cost of the substation generator is related to its revenues from whole sale market sales at the prevailing LMP plus the opportunity cost of the requisite reactive power compensation, plus voltage rise costs, $\pi_\infty(-p_t)+\pi_\infty(C_\infty-\sqrt{C_\infty^2-q_t^2})+c_\infty^V(v_t-1)^2$ while constraints are the generator's capacity limits, $p_t^2+q_t^2\le C_\infty^2$ and the voltage magnitude limits, $\underline{V}\le V_t\le\overline{V}$.

Fixed load: Single terminal device. No cost for this device. The constraints are the exogenously specified real and reactive power injected at its terminal, $p_t=p_{d_i}, q_t=q_{d_i}$, and the voltage magnitude limits $\underline{V}\le V_t\le\overline{V}$.

Curtailable load: Single terminal device. The cost of this device is the negative utility derived when it is scheduled to consume at p_t, namely, $-u_{d_i}p_t$. Voltage magnitude constraints, $\underline{V}\le V_t\le\overline{V}$, fixed power factor reactive power consumption, $q_t=p_t\tan(\varphi_{d_i})$ and capacity limits $\underline{p}_{d_i}\le p_t\le\overline{p}_{d_i}$ constitute the associated constraints.

AC lines: Two terminal devices. No cost for this device. The constraints are voltage magnitude constraints $\underline{V}\le V_t\le\overline{V}$ and typical load flow constraints:

$$P_{t_1,t_2}=V_{t_1}^2G_{t_1,t_2}-V_{t_1}V_{t_2}G_{t_1,t_2}\cos(\theta_{t_1}-\theta_{t_2})-V_{t_1}V_{t_2}B_{t_1,t_2}\sin(\theta_{t_1}-\theta_{t_2})$$

$$Q_{t_1,t_2}=-V_{t_1}^2B_{t_1,t_2}+V_{t_1}V_{t_2}B_{t_1,t_2}\cos(\theta_{t_1}-\theta_{t_2})-V_{t_1}V_{t_2}G_{t_1,t_2}\sin(\theta_{t_1}-\theta_{t_2}).$$

Transformers: Two terminal devices. The cost of this device is the cost of transformer loss of life, $c_{t_1,t_2}^{tr}\exp\left(\frac{15000}{383}-\frac{15000}{273+k_{1,t_1,t_2}+k_{2,t_1,t_2}\frac{p_{t_1}^2+q_{t_1}^2}{(S_{t_1,t_2}^N)^2}}\right)$. The constraints are the same as in AC lines, i.e. load flow and voltage magnitude constraints.

Photovoltaics: Single terminal devices. No cost for this device. Constraints involve capacity limits given the exogenously specified real power output, p_{e_i}, yielding $p_{e_i}^2+q_t^2\le C_{e_i}^2$, and the voltage magnitude constraints $\underline{V}\le V_t\le\overline{V}$.

Capacitors: Single terminal devices. No cost for this device. Constraints involve its capacity $p_t=0, -C_{f_i}\le q_t\le 0$ and the voltage magnitude limits $\underline{V}\le V_t\le\overline{V}$.

We proceed by using a generalized notion of the device objective function by appending the individual device constraints to its costs. For example, the generalized substation generator cost function is:

$$\sigma_a(\mathbf{p}_a,\mathbf{q}_a,\mathbf{V}_a,\boldsymbol{\theta}_a)=\pi_\infty(-p_t)+\pi_\infty(C_\infty-\sqrt{C_\infty^2-q_t^2})$$
$$+c_\infty^V(V_t-1)^2+1_{p_t^2+q_t^2\ge C_\infty^2}M+1_{\underline{V}\ge V_t}M+1_{V_t\ge\overline{V}}M$$

Where M is a large number increasing the cost to infinity when a constraint is violated.

The generalized cost function and the classification of distribution market participants and network components to nets, devices and terminals, gives the following equivalent formulation of the AC Optimal Power Flow problem:

$$\left.\begin{array}{l}\min\sigma(\mathbf{p},\mathbf{q},\mathbf{V},\boldsymbol{\theta})=\sum_a\sigma_a(\mathbf{p}_a,\mathbf{q}_a,\mathbf{V}_a,\boldsymbol{\theta}_a)\\ \textit{subject to}\ \hat{\mathbf{p}}_\mathbf{n}=\mathbf{0},\forall n;\qquad \hat{\mathbf{q}}_\mathbf{n}=\mathbf{0},\forall n\\ \qquad\qquad \widehat{\mathbf{V}}_\mathbf{n}=\mathbf{0},\forall n;\qquad \widehat{\boldsymbol{\theta}}_\mathbf{n}=\mathbf{0},\forall n\end{array}\right\}\Rightarrow$$

$$\min\sigma(\mathbf{p},\mathbf{q},\mathbf{V},\boldsymbol{\theta})+\sum_n\{g_n(z_n)+f_n(w_n)+k_n(\zeta_n)+h_n(\xi_n)\}$$

$$\textit{subject to}\ \mathbf{p}=\mathbf{z}\rightarrow\mathbf{y}^\mathbf{P};\qquad \mathbf{q}=\mathbf{w}\rightarrow\mathbf{y}^\mathbf{Q}$$
$$\mathbf{V}=\boldsymbol{\zeta}\rightarrow\mathbf{y}^\mathbf{V};\qquad \boldsymbol{\theta}=\boldsymbol{\xi}\rightarrow\mathbf{y}^{\boldsymbol{\theta}}$$

Where $g_n(z_n)$, $f_n(w_n)$, $k_n(\zeta_n)$, and $h_n(\xi_n)$ are the indicator functions respectively on the sets $\{z_n\,|\,\hat{z}_n=0\}$, $\{w_n\,|\,\hat{w}_n=0\}$, $\{\zeta_n\,|\,\widehat{\zeta}_n=0\}$ and $\{\xi_n\,|\,\widehat{\xi}_n=0\}$. This allows us to write the Lagrangian as:

$$L(\mathbf{p},\mathbf{q},\mathbf{V},\boldsymbol{\theta},\mathbf{z},\mathbf{w},\boldsymbol{\zeta},\boldsymbol{\xi},\mathbf{y}^\mathbf{P},\mathbf{y}^\mathbf{Q},\mathbf{y}^\mathbf{V},\mathbf{y}^{\boldsymbol{\theta}})=\sum_a\sigma_a(\mathbf{p}_a,\mathbf{q}_a,\mathbf{V}_a,\boldsymbol{\theta}_a)$$
$$+\sum_n\{g_n(z_n)+f_n(w_n)+k_n(\zeta_n)+h_n(\xi_n)\}$$
$$+\mathbf{y}^\mathbf{P}\times(\mathbf{p}-\mathbf{z})+\mathbf{y}^\mathbf{Q}\times(\mathbf{q}-\mathbf{w})+\mathbf{y}^\mathbf{V}\times(\mathbf{V}-\boldsymbol{\zeta})+\mathbf{y}^{\boldsymbol{\theta}}\times(\boldsymbol{\theta}-\boldsymbol{\xi})$$
$$+\frac{\rho}{2}\|\mathbf{p}-\mathbf{z}\|_2^2+\frac{\rho}{2}\|\mathbf{q}-\mathbf{w}\|_2^2+\frac{\rho}{2}\|\mathbf{V}-\boldsymbol{\zeta}\|_2^2+\frac{\rho}{2}\|\boldsymbol{\theta}-\boldsymbol{\xi}\|_2^2$$

Where the last line includes additional penalty terms used for smoothness. These terms can be interpreted as the cost of the market not clearing.

With standard algebraic manipulation the augmented Lagrangian can be rewritten more concisely as:

$$L(\mathbf{p},\mathbf{q},\mathbf{V},\boldsymbol{\theta},\mathbf{z},\mathbf{w},\boldsymbol{\zeta},\boldsymbol{\xi},\mathbf{y}^\mathbf{P},\mathbf{y}^\mathbf{Q},\mathbf{y}^\mathbf{V},\mathbf{y}^{\boldsymbol{\theta}})=\sum_a\sigma_a(\mathbf{p}_a,\mathbf{q}_a,\mathbf{V}_a,\boldsymbol{\theta}_a)$$
$$+\sum_n\{g_n(z_n)+f_n(w_n)+k_n(\zeta_n)+h_n(\xi_n)\},$$
$$+\frac{\rho}{2}\|\mathbf{p}-\mathbf{z}+\boldsymbol{\upsilon}\|_2^2+\frac{\rho}{2}\|\mathbf{q}-\mathbf{w}+\boldsymbol{\lambda}\|_2^2+\frac{\rho}{2}\|\mathbf{V}-\boldsymbol{\zeta}+\boldsymbol{\mu}\|_2^2+\frac{\rho}{2}\|\boldsymbol{\theta}-\boldsymbol{\xi}+\boldsymbol{\varsigma}\|_2^2$$

where $\boldsymbol{\upsilon}=\frac{\mathbf{y}^\mathbf{P}}{\rho}$, $\boldsymbol{\lambda}=\frac{\mathbf{y}^\mathbf{Q}}{\rho}$, $\boldsymbol{\mu}=\frac{\mathbf{y}^\mathbf{V}}{\rho}$ and $\boldsymbol{\varsigma}=\frac{\mathbf{y}^{\boldsymbol{\theta}}}{\rho}$. Noting the association with the constraint shadow prices, we call $\upsilon, \lambda, \mu, \varsigma$ "scaled prices".

Using devices and nets to group terminals and using the properties of indicator functions, the individual device problems may be written as [2]:

$$(\mathbf{p}_a^{k+1}, \mathbf{q}_a^{k+1}, \mathbf{V}_a^{k+1}, \boldsymbol{\theta}_a^{k+1}) = \arg\min_{\mathbf{p}_a, \mathbf{q}_a, \mathbf{V}_a, \boldsymbol{\theta}_a} \{\sigma_a(\mathbf{p}_a, \mathbf{q}_a, \mathbf{V}_a, \boldsymbol{\theta}_a)$$

$$+\frac{\rho}{2}[\|\mathbf{p}_a - \mathbf{p}_a^k + \hat{\mathbf{p}}_a^k + \boldsymbol{\upsilon}_a^k\|_2^2 + \|\mathbf{q}_a - \mathbf{q}_a^k + \hat{\mathbf{q}}_a^k + \boldsymbol{\lambda}_a^k\|_2^2,$$

$$+\|\mathbf{V}_a - \hat{\mathbf{V}}_a^k - \boldsymbol{\mu}_a^k\|_2^2 + \|\boldsymbol{\theta}_a - \hat{\boldsymbol{\theta}}_a^k - \boldsymbol{\varsigma}_a^k\|_2^2]\}$$

where $\upsilon_n^{k+1} = \upsilon_n^k + \hat{p}_n^{k+1}$, $\lambda_n^{k+1} = \lambda_n^k + \hat{q}_n^{k+1}$, $\mu_n^{k+1} = \mu_n^k + \hat{V}_n^{k+1}$ and $\varsigma_n^{k+1} = \varsigma_n^k + \hat{\theta}_n^{k+1}$.

D. Individual Device Problems

As stated above, a sufficient condition for the algorithm to converge, is that the individual device problems are convex. This raises an issue for line devices where the non-convex AC flow equations must be incorporated in the generalized Lagrangian. To address this issue, we use the relaxed branch flow model introduced in [9] and extended in [5], where (i) the voltage drop and complex power flow relations are squared eliminating angles and resulting in voltage magnitude appearing only in its square form [9], and (ii) the only remaining non-convex line current equality constraint, $\ell_{t_1,t_2} = \frac{P_{t_1}^2 + Q_{t_1}^2}{v_{t_1}}$, is relaxed to an inequality constraint [5]. Since lines are modeled as individual devices, the constraints on the injections of the sending or the receiving ends of the line are constraints of the individual distributed participant subproblem and are not included in the line subproblem. As such the above relaxation is shown in [6] to be exact.

We explicitly show below the individual device problems:

Generators:

$$\min_{p_t, q_t, v_t} \left\{ \begin{array}{l} \pi_\infty(-p_t) + \pi_\infty(C_\infty - \sqrt{C_\infty^2 - q_t^2}) + c_\infty^V(v_t - 1)^2 + \\ \frac{\rho}{2}[\|p_t - p_t^k + \hat{p}_t^k + \upsilon_t^k\|_2^2 + \|q_t - q_t^k + \hat{q}_t^k + \lambda_t^k\|_2^2 \\ +\|v_t - \hat{v}_t^k - \mu_t^k\|_2^2] \end{array} \right\}$$

$$\textit{subject to } p_t^2 + q_t^2 \le C_\infty^2, p_t \le 0, \underline{v} \le v_t \le \overline{v}$$

Fixed load:

$$\min_{p_t, q_t, v_t} \left\{ \begin{array}{l} \frac{\rho}{2}[\|p_t - p_t^k + \hat{p}_t^k + \upsilon_t^k\|_2^2 + \|q_t - q_t^k + \hat{q}_t^k + \lambda_t^k\|_2^2 \\ +\|v_t - \hat{v}_t^k - \mu_t^k\|_2^2] \end{array} \right\}$$

$$\textit{subject to } p_t = p_{d_i}, q_t = q_{d_i}, \underline{v} \le v_t \le \overline{v}$$

Curtailable load:

$$\min_{p_t, q_t, v_t} \left\{ \begin{array}{l} -u_t p_t + \frac{\rho}{2}[\|p_t - p_t^k + \hat{p}_t^k + \upsilon_t^k\|_2^2 + \|q_t - q_t^k + \hat{q}_t^k + \lambda_t^k\|_2^2 \\ +\|v_t - \hat{v}_t^k - \mu_t^k\|_2^2] \end{array} \right\}$$

$$\textit{subject to } \underline{p}_{d_i} \le p_t \le \overline{p}_{d_i}, q_t = p_t \tan(\varphi_{d_i}), \underline{v} \le v_t \le \overline{v}$$

Capacitors:

$$\min_{p_t, q_t, v_t} \left\{ \begin{array}{l} \frac{\rho}{2}[\|p_t - p_t^k + \hat{p}_t^k + \upsilon_t^k\|_2^2 + \|q_t - q_t^k + \hat{q}_t^k + \lambda_t^k\|_2^2 \\ +\|v_t - \hat{v}_t^k - \mu_t^k\|_2^2] \end{array} \right\}$$

$$\textit{subject to } p_t = 0, -C_{f_i} \le q_t \le 0, \underline{v} \le v_t \le \overline{v}$$

Photovoltaics:

$$\min_{p_t, q_t, v_t} \left\{ \begin{array}{l} \frac{\rho}{2}[\|p_t - p_t^k + \hat{p}_t^k + \upsilon_t^k\|_2^2 + \|q_t - q_t^k + \hat{q}_t^k + \lambda_t^k\|_2^2 \\ +\|v_t - \hat{v}_t^k - \mu_t^k\|_2^2] \end{array} \right\}$$

$$\textit{subject to } p_t = p_{e_i}; p_t^2 + q_t^2 \le C_e^2; \underline{v} \le v_t \le \overline{v}$$

AC lines:

$$\min_{\mathbf{p}_a, \mathbf{q}_a, \mathbf{v}_a} \left\{ \begin{array}{l} \frac{\rho}{2}[\|\mathbf{p}_a - \mathbf{p}_a^k + \hat{\mathbf{p}}_a^k + \boldsymbol{\upsilon}_a^k\|_2^2 + \|\mathbf{q}_a - \mathbf{q}_a^k + \hat{\mathbf{q}}_a^k + \boldsymbol{\lambda}_a^k\|_2^2 \\ +\|\mathbf{v}_a - \hat{\mathbf{v}}_a^k - \boldsymbol{\mu}_a^k\|_2^2] \end{array} \right\}$$

$$\begin{aligned} \textit{subject to } & \ell_{t_1,t_2} \ge \frac{P_{t_1}^2 + Q_{t_1}^2}{v_{t_1}} \\ & p_{t_1} + p_{t_2} = R_{t_1,t_2}\ell_{t_1,t_2} \\ & q_{t_1} + q_{t_2} = X_{t_1,t_2}\ell_{t_1,t_2} \\ & v_{t_2} = v_{t_1} - 2(P_{t_1}R_{t_1,t_2} + Q_{t_1}X_{t_1,t_2}) + (R_{t_1,t_2}^2 + X_{t_1,t_2}^2)l_{t_1,t_2} \\ & \underline{\mathbf{v}} \le \mathbf{v}_a \le \overline{\mathbf{v}} \end{aligned}$$

Transformer Lines:

$$\min_{\mathbf{p}_a, \mathbf{q}_a, \mathbf{v}_a} \left\{ \begin{array}{l} c_{t_1,t_2}^{tr} \exp\left(\frac{15000}{383} - \frac{15000}{273 + k_{1,t_1,t_2} + k_{1,t_1,t_2}\frac{p_{t_1}^2 + q_{t_1}^2}{(S_{t_1,t_2}^N)^2}}\right) \\ +\frac{\rho}{2}[\|\mathbf{p}_a - \mathbf{p}_a^k + \hat{\mathbf{p}}_a^k + \boldsymbol{\upsilon}_a^k\|_2^2 + \|\mathbf{q}_a - \mathbf{q}_a^k + \hat{\mathbf{q}}_a^k + \boldsymbol{\lambda}_a^k\|_2^2 \\ +\|\mathbf{v}_a - \hat{\mathbf{v}}_a^k - \boldsymbol{\mu}_a^k\|_2^2] \end{array} \right\}$$

$$\begin{aligned} \textit{subject to } & \ell_{t_1,t_2} \ge \frac{p_{t_1}^2 + q_{t_1}^2}{v_{t_1}} \\ & p_{t_1} + p_{t_2} = R_{t_1,t_2}\ell_{t_1,t_2} \\ & q_{t_1} + q_{t_2} = X_{t_1,t_2}\ell_{t_1,t_2} \\ & v_{t_2} = v_{t_1} - 2(p_{t_1}R_{t_1,t_2} + q_{t_1}X_{t_1,t_2}) + (R_{t_1,t_2}^2 + X_{t_1,t_2}^2)l_{t_1,t_2} \\ & \underline{\mathbf{v}} \le \mathbf{v}_a \le \overline{\mathbf{v}} \end{aligned}$$

III. Algorithm

A. Iterative Process

During each iteration cycle k, every device minimizes its own generalized objective function augmented by an additional term whose value depends on the messages the device receives[1] from its neighboring (or proximal) nets through its terminals , $\hat{\mathbf{p}}_a^k, \hat{\mathbf{q}}_a^k, \hat{\mathbf{v}}_a^k, \boldsymbol{\upsilon}_a^k, \boldsymbol{\lambda}_a^k, \boldsymbol{\mu}_a^k$. The new power and voltage schedules $\mathbf{p}_a^{k+1}, \mathbf{q}_a^{k+1}, \mathbf{v}_a^{k+1}$ that the device calculates are passed on to the nets through the device terminals. At the end of each iteration cycle, all nets calculate the new real and reactive power imbalance $\hat{\mathbf{p}}_n^{k+1}, \hat{\mathbf{q}}_n^{k+1}$ and voltage residuals $\hat{\mathbf{v}}_n^{k+1}$, update the scaled price variables $\boldsymbol{\upsilon}_n^{k+1}, \boldsymbol{\lambda}_n^{k+1}, \boldsymbol{\mu}_n^{k+1}$ and pass them on to each terminal associated with the net. Since each terminal is associated with a device, the corresponding devices receive this new information. Depending on the version of the algorithm, the nets may be also responsible for the update of net specific penalties (see next section). Following the data update, each device reoptimizes its sub-problem. The cycle repeats until the local or global stopping criterion is satisfied.

Proximal message passing (PMP) algorithms guarantee [11]:

- As $k \to \infty$, $\hat{\mathbf{p}}^k \to \mathbf{0}$, $\hat{\mathbf{q}}^k \to \mathbf{0}$, $\hat{\mathbf{V}}^k \to \mathbf{0}$, i.e. power balance and voltage consistency are achieved.
- As $k \to \infty, \sum_a \sigma_a(\mathbf{p}_a^k, \mathbf{q}_a^k, \mathbf{V}_a^k) \to \sigma^*$, i.e. operation is optimal.
- As $k \to \infty$, $\boldsymbol{\rho\upsilon}^k \to \boldsymbol{\pi}^{\mathbf{P}}, \boldsymbol{\rho\lambda}^k \to \boldsymbol{\pi}^{\mathbf{Q}}$, i.e. optimal Location Marginal Prices are found.

B. Penalty Updates

i. The same penalty is used at all nets and for all imbalances. It is updated after each iteration

In [2], the use of an iteratively updated penalty is implemented to speed up the algorithm. We define the vector of imbalances $\mathbf{r}^k = (\hat{\mathbf{p}}^k, \hat{\mathbf{q}}^k, \hat{\mathbf{v}}^k)$ and the change of the imbalances

$\mathbf{s}^k = \rho[(\mathbf{p}^k - \hat{\mathbf{p}}^k) - (\mathbf{p}^{k-1} - \hat{\mathbf{p}}^{k-1}), (\mathbf{q}^k - \hat{\mathbf{q}}^k) - (\mathbf{q}^{k-1} - \hat{\mathbf{q}}^{k-1}), \hat{\mathbf{v}}^k - \hat{\mathbf{v}}^{k-1}]$. We use the following penalty update rule:

$$\rho^{k+1} = \begin{cases} \rho^k(1 + \|\mathbf{r}^k\| + \|\mathbf{s}^k\|), \text{if } \|\mathbf{r}^k\| > 5\|\mathbf{s}^k\| \,\&\, \|\mathbf{r}^k\| + \|\mathbf{s}^k\| < 0.3 \\ 1.3\rho^k, \text{if } \|\mathbf{r}^k\| > 5\|\mathbf{s}^k\| \,\&\, \|\mathbf{r}^k\| + \|\mathbf{s}^k\| \geq 0.3 \\ \rho^k[1 - (\|\mathbf{r}^k\| + \|\mathbf{s}^k\|)], \text{if } \|\mathbf{s}^k\| > 5\|\mathbf{r}^k\| \,\&\, \|\mathbf{r}^k\| + \|\mathbf{s}^k\| < 0.3 \\ 0.7\rho^k, \text{if } \|\mathbf{s}^k\| > 5\|\mathbf{r}^k\| \,\&\, \|\mathbf{r}^k\| + \|\mathbf{s}^k\| \geq 0.3 \\ \rho^k, \text{else} \end{cases}$$

Adapting the penalty means that the scaled prices also have to be adapted accordingly, therefore we have:

$$\boldsymbol{\upsilon}^{k+1} = \frac{\rho^k}{\rho^{k+1}}\boldsymbol{\upsilon}^{k+1}, \boldsymbol{\lambda}^{k+1} = \frac{\rho^k}{\rho^{k+1}}\boldsymbol{\lambda}^{k+1} \text{ and } \boldsymbol{\mu}^{k+1} = \frac{\rho^k}{\rho^{k+1}}\boldsymbol{\mu}^{k+1}.$$

ii. Net Specific Penalty Updates

The update of a common, network-wide penalty, requires central coordination and the requisite information communication. To decrease the communication requirement and to possibly increase the convergence speed, we consider using a net specific penalty and update it locally, i.e. within each net. We call this the net-specific penalty PMP algorithm. We define the vector of imbalances and the change of the imbalances per net as $\mathbf{r}_n^k = (\hat{\mathbf{p}}_n^k, \hat{\mathbf{q}}_n^k, \hat{\mathbf{v}}_n^k)$ and

$\mathbf{s}_n^k = \rho[(\mathbf{p}_n^k - \hat{\mathbf{p}}_n^k) - (\mathbf{p}_n^{k-1} - \hat{\mathbf{p}}_n^{k-1}), (\mathbf{q}_n^k - \hat{\mathbf{q}}_n^k) - (\mathbf{q}_n^{k-1} - \hat{\mathbf{q}}_n^{k-1}), \hat{\mathbf{v}}_n^k - \hat{\mathbf{v}}_n^{k-1}]$,

and implement a net-specific penalty update similar to the network-wide rule described above.

iii. Net-and-quantity Specific Penalty Updates

We further define a net-and-quantity specific PMP algorithm where a different net specific penalty is used to drive the imbalance of real power, reactive power, and voltage imbalances to zero. The resulting PMP algorithm with the use of net-and-quantity specific penalties is faster and more accurate since the primal quantities that the algorithm is solving for are not necessarily of the same order of magnitude. We hence benefit from quantity specific penalties.Net-and-quantity specific update rules follow naturally from the above penalty update definitions.

C. Stopping Criterion

i. Central Stopping Criterion

The stopping criterion used in [2] for a single hour problem is $\|\mathbf{r}^k\| \leq \varepsilon^{abs}\sqrt{T}, \|\mathbf{s}^k\| \leq \varepsilon^{abs}\sqrt{T}$. However, its implementation requires a central coordinator and the accompanying information communication. To lessen the PMP algorithm's requisite communication burden, we propose a net specific stopping criterion.

ii. Net Specific Stopping Criterion

Since the imbalance and the change in the imbalance at a given net is not known at other nets, an individual net determines the following binary flag after each iteration.

$$flag(n,k) = \begin{cases} 1, \text{if } \|\mathbf{r}_n^k\| \leq 5 \cdot 10^{-3} \text{ and } \|\mathbf{s}_n^k\| \leq 5 \cdot 10^{-3} \\ 0, \text{else} \end{cases}.$$

Moreover, at the beginning of iteration cycle $k+1$ each net communicates the value its flag had at the end of iteration cycle k to its direct upstream nets. The upstream net receives the message and at the end of iteration $k+1$ adds to it the value of its own flag and then communicates the sum to its direct upstream net. If the value of the sum at the substation or root node, equals the number of nets in the network, and if that value persists for as many sequential iteration cycles as the number of nets in the longest line of our tree network, then it follows logically that all the flags were 1 at all nets at the same time [10] and that the algorithm has converged.

D. Decreased Reliance on Global Coordination

The use of the local penalty updates III.B.ii and III.B.iii together with the local stopping criterion, enable the algorithm to work essentially without global coordination, a fact that allows asynchronous sub-problem solution that do not conform to the beat of a common clock. This makes more sub-problem solutions possible and does not limit the convergence time to the bottleneck of the slowest device sub-problem solution. Moreover, the information communication burden is significantly lower.

We have noticed that when local penalties are used, several numerical issues appear with the use of the exponential penalty update rule used in [2]. These appear when the ratio of the norms is either too small or too large. This is also affected by the choice of the parameters scaling the ratio. In order to achieve convergence in the net and net-and-quantity specific PMP algorithms, these parameters would have to be insignificantly small. This is precisely why we use an update rule which depends on the sum rather than the ratio of the imbalance and the change in the imbalance relative to the previous iteration. We note that, unlike the exponential penalty updates, the use of our update rule does not require that we specify a limited number of iterations beyond which the penalty is no longer updated. Whereas this limitation is implemented in [2] to avoid problems of large penalty changes when the algorithm stops, our penalty update rule depends on the sum of the norms, rather than their ratio, assuring that the change in the penalty is close to zero when the stopping rule is satisfied. This ensures, that during many iterations before the stopping rule is satisfied, the penalty has been practically constant.

IV. Results

We first apply the PMP algorithms described above to a 47-bus realistic distribution feeder from Southern California Edison as published in [5]. We examine the following:

1) Constant penalty throughout the network with no iterative updates
2) Penalty updating after each iteration but constant throughout the network
3) Net-specific penalty with net-specific stopping criterion based on the sum of local stopping criterion flags. As can be seen in [5], the longest branch in this network is 12 busses long, and so the algorithm will stop when the observed sum of flags at the root node equals 47 for 12 sequential iterations.
4) Net-and-quantity specific penalties with net-specific stopping criterion as in the case above.

The required number of iterations till convergence for several penalty starting points are shown in Table I below.

TABLE I. Number of iterations until convergence

ρ(0)	**Constant Penalty**	**Adaptive Common Penalty**	**Net Specific Penalty**	**Net& Quantity Specific Penalty**
5	1205	395	340	204
10	2298	499	260	240
20	4595	505	354	330
50	>5000	498	440	341

We notice a significant drop in the required number of iterations when we allow the penalty factor to change. Net specific and Net-and-quantity specific PMP algorithms not only converge faster but also require much less information communication.

Tables II and III below show the average, minimum and maximum error in real and reactive power prices at convergence. These statistics are estimated relative to the exact values obtained in our benchmark solution of the market clearing problem.

TABLE II. Deviation of real prices from their optimal values at convergence

REAL PRICE DEVIATION (%)						
ρ0	Net specific ρ			Net-and-quantity specific ρ		
	Avg	Min	Max	Avg	Min	Max
5	0.066	0.001	0.542	0.025	0.001	0.107
10	0.059	0.001	0.541	0.020	0.000	0.058
20	0.088	0.000	0.530	0.019	0.003	0.078
50	0.022	0.000	0.187	0.027	0.000	0.140

TABLE III. Deviation of reactive prices from their optimal values at convergence

REACTIVE PRICE DEVIATION (%)						
ρ0	Net specific ρ			Net-and-quantity specific ρ		
	Avg	Min	Max	Avg	Min	Max
5	0.628	0.128	0.804	0.106	0.002	0.211
10	1.378	0.886	1.564	1.692	1.279	1.845
20	1.329	0.881	1.494	1.425	1.101	1.579
50	1.159	0.880	1.380	1.201	0.966	1.392

Figures 1 and 2 below show the convergence of the real and reactive prices as obtained by the proximal message passing algorithm for the case of net-and-quantity specific penalties, starting from a penalty of 20.

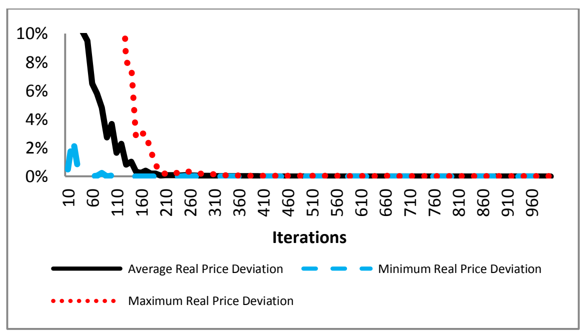

Figure 1. Average, Minimum and Maximum Deviation from Optimal Real Prices.

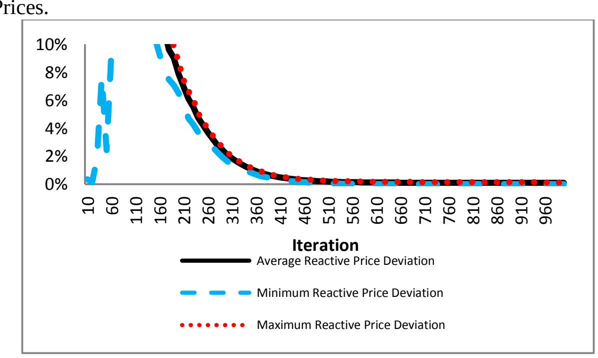

Figure 2. Average, Minimum and Maximum Deviation from Optimal Reactive Prices.

By examination of the values of $\|\mathbf{r}^k\|, \|\mathbf{s}^k\|$ we see that the decentralized stopping criterion is in fact tighter than the centralized stopping criterion proposed in [2]. This can also be verified by the fact that the real and reactive prices are closer to the optimal values obtained from the centralized AC OPF.

Figures 3 and 4 below show the evolution of the net specific and net-and-quantity specific penalty factors for starting penalty equal to 20 for a randomly chosen net.

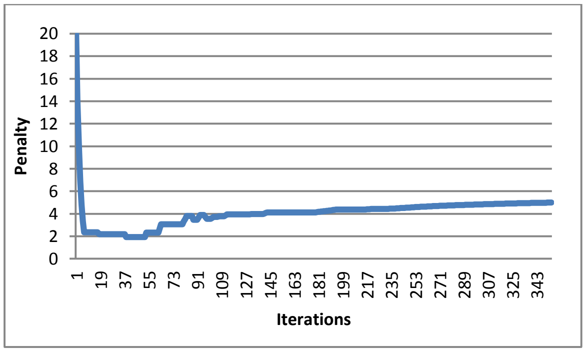

Figure 3. Net Specific Penalty, starting from a penalty of 20, Net 37.

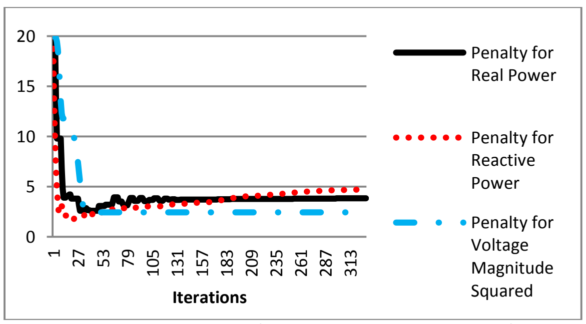

Figure 4. Net-and-quantity Specific Penalty, starting penalty of 20, Net 37.

PMP algorithms are superior in computational efficiency relative to the centralized market clearing algorithm. In a real implementation environment, the PMP algorithm may be applied on an hourly or five minute basis, in which case it can benefit from a hot start that determines the primal variable values from the previous hour or five minute solution. Numerical experience shows that this is indeed the case.

We finally consider the scalability of the proposed PMP algorithms by applying them to the 253 bus distribution network investigated in our benchmark [1]. Measured by the number of devices, the 253 bus network is 6 times bigger. Table IV reports the number of iterations needed for convergence. Unlike the claim in [2] that problem size does not affect computational requirements, we observe an increase in computation requirements which we measure in terms of the number of iterations required to converge. As such, the increases appears to evolve with the square root of the network size. It is possible that an asymptotic result may be in the works, but we could not test it for the time being as we only had benchmark results for the 253 network.

TABLE IV. Number of iterations needed for 253 nets versus 47 nets.

ρ(0)	NET SPECIFIC PENALTY		NET & QUANTITY SPECIFIC PENALTY	
	253 Bus Net Specific Penalty	Increase In # of iterations wrt 47 bus	253 Bus Net & Quantity Specific Penalty	Increase in # of iterations wrt 47 bus
5	619	1.821	480	2.353
10	666	2.562	565	2.354
20	717	2.025	742	2.248
50	745	1.693	804	2.358

We can use the results of Table IV to project the computational times needed for the solution of the 253 bus system using peer-to-peer implementation. Given the computational time of the bottleneck device, the proximal message passing algorithms needs approximately five times less computational effort than the centralized AC OPF required 51.42 seconds.

V. Conclusion

We have presented and tested several extensions of the PMP algorithms applied to the solution of complex distribution network power markets. Computational results have shown computational efficiency and accuracy, scalability and lower communication requirement benefits. In future work we will focus on asynchronous PMP algorithms.

References

[1] E. Ntakou, M. Caramanis, Price Discovery in Dynamic Power Markets with Low-Voltage Distribution Network Participants, *IEEE Transmission and Distribution 2014*, to be published.

[2] M. Kranning, E. Chu, J. Lavaei, S. Boyd, Dynamic Network Energy Management via Proximal Message Passing, in *Foundations and Trends in Optimization*, Vol.1, No. 2, 2013, pp.70-122.

[3] Y. Dai, Y. X. Ni, F. S. Wen, Z. X. Han, "Analysis of reactive power pricing under deregulation", *IEEE Power Engineering Society Summer Meeting*, 2000, Vol. 4, pp. 2162-2197

[4] M. L. Baughman, S. N. Siddiqi, "Real-time pricing of reactive power: theory and case study results", *IEEE Transactions on Power Systems*, Vol. 12, No. 1, February 1997, pp. 95-104.

[5] M. Farivar, C. R. Clark, S. H. Low, K. M. Chandy, "Inverter VAR control for distribution systems with renewables", 2011 International Conference on Smart Grid Communications, pp. 457-462.

[6] J. Lavaei, D. Tse, and B. Zhang, Geometry of power flows in tree networks, in *IEEE Power & Energy Society General Meeting, 2012*.

[7] N. Li, L. Chen, S. Low, Demand Response in Radial Distribution Networks: Distributed Algorithm, in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers*, 4-7 November 2012, pp.1549-1553.

[8] M. Caramanis, Is It Time For Power Market Reform To Allow for Retail Customer Participation and Distribution Network Marginal Pricing, *IEEE Smart Grid*, March 2012.

[9] M. Baran, F. Wu, "Optimal Capacitor Placement in Radial Distribution Systems", IEEE Transactions on Power Delivery, Vol. 4, No.1, January 1989, pp. 725-734.

[10] M. J. Wainwright, T. S. Jaakkola, A. S. Willsky, "Tree-based reparametrization framework for analysis of sum-product and related algorithms", *IEEE Transactions on Information Theory*, Vol. 49, Issue 5, pp. 1120-1146.

[11] S. Boyd, K. Parikh, E. Chu, B. Peleato, J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers", *Foundations and Trends in Machine Learning*, vol. 3, pp. 1-122, 2011.