52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

Stability and convergence of distributed algorithms for the OPF problem

Eoin Devane and Ioannis Lestas

Abstract— **Many modern power networks are partitioned in nature, with disjoint components of the overall network controlled by competing operators. The problem of solving the Optimal Power Flow (OPF) problem in a distributed manner is therefore of significant interest. For networks in which the high-level structure has tree topology, we analyze a dual decomposition approach to solving a recent convex relaxation of the OPF problem for the overall network in a distributed manner. Incorporating higher-order dynamics in terms of local auxiliary variables, we prove a result of guaranteed convergence to the solution set for sufficiently small values of the step size.**

I. Introduction

The Optimal Power Flow (OPF) problem for a power network involves minimizing the total cost of power generation over all of the network buses capable of generating power, subject to satisfying given power loads and a collection of physical inequality constraints on the system's operation [1], [2], [3]. Given that modern power networks are often partitioned in structure, distributed approaches to solving the OPF problem, whereby each operator may solve its own regional optimization problem locally without needing to disclose exact operational details to its competitors, are of significant interest. Several distributed approaches to the OPF problem have previously been studied, for example in [4] and [5].

In this paper we consider a decomposition based on consistency constraints for real and reactive power values and voltage magnitudes at a prescribed point on the tie-lines between regions. This allows us to split the problem into regional subproblems when these regions are connected over a tree structure. If the regional problems can be convexified by a recently-proposed LMI approach [6], [7], then distributed implementations can be achieved through this approach without loss of accuracy. Convergence properties of such algorithms are often improved by means of quadratic regularization associated with the consistency constraints, however, such an approach necessitates additional information sharing between competing regions. Instead, we propose an alternative means of regularization in terms of auxiliary variables that are local to each region, giving a distributed algorithm with augmented higher-order dynamics. We prove that, for sufficiently small step size, this algorithm is guaranteed to converge to a limiting set in which the collection of primal variables obtained is indeed a solution of our original relaxed global OPF problem. It should be noted that the convergence proof is more involved than classical convergence results associated with gradient-type dynamics, due to the fact that the regional subproblems are not strictly convex. Indeed, as noted in the seminal work in [8], for many classes of convex problems classical gradient methods can fail to converge to an optimal solution when the problem is not strictly convex.

The remainder of the paper is organized as follows. In Section II, we introduce the convex relaxation of the OPF problem that is to be considered and derive from this the iterative distributed approach that we shall investigate. In Section III we study the properties of the fixed points of our algorithm and present our main convergence result. Finally, conclusions are drawn in Section IV.

Eoin Devane is with the Cambridge Centre for Analysis, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom; esmd2@cam.ac.uk

Ioannis Lestas is with the Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom; icl20@cam.ac.uk

II. Formulation of the problem

Consider a simplified power system consisting of R dense regions, labeled by $\mathcal{I} \in \mathcal{E} = \{1, \ldots, R\}$, connected to each other by a net of tie-lines, each of which joins a particular bus within one region to a particular bus within another region. We allow at most one tie-line between any pair of regions.

Any region $\mathcal{I}$ may be considered as a set of $m_{\mathcal{I}}$ buses $\hat{N}^{\mathcal{I}}$, the subset of generator buses $\hat{G}^{\mathcal{I}} \subseteq \hat{N}^{\mathcal{I}}$ which are capable of supplying power to the network, and the set of flow lines that link these buses. Let us now define $\mathcal{C}(\mathcal{I})$ to be the set of other regions $\mathcal{J}$ that are connected to $\mathcal{I}$ by means of a tie-line. To model each such connection, we adjoin to $\mathcal{I}$ a single dummy generator bus by means of a flow line of twice the tie-line admittance connecting to the bus in $\hat{N}^{\mathcal{I}}$ from which the tie-line in question emanates. In so doing we adjoin to $\mathcal{I}$ exactly $n_{\mathcal{I}} = |\mathcal{C}(\mathcal{I})|$ dummy generator buses, each connecting to a pre-existing bus of $\mathcal{I}$ by exactly one flow line. We thus obtain an extension of the subsystem $\mathcal{I}$ which may be represented by the bus set $N^{\mathcal{I}}$ of cardinality $m_{\mathcal{I}} + n_{\mathcal{I}}$, the generator set $G^{\mathcal{I}} \subseteq N^{\mathcal{I}}$ of cardinality $|\hat{G}^{\mathcal{I}}| + n_{\mathcal{I}}$, and the expanded set of flow lines. We label the buses in $N^{\mathcal{I}}$ by the numbers $1, \ldots, m_{\mathcal{I}} + n_{\mathcal{I}}$ arbitrarily and let $\ell_{\mathcal{I}\mathcal{J}}$ denote the dummy bus corresponding to the tie-line joining region $\mathcal{I}$ to region $\mathcal{J} \in \mathcal{C}(\mathcal{I})$. Since $\mathcal{J} \in \mathcal{C}(\mathcal{I})$ if and only if $\mathcal{I} \in \mathcal{C}(\mathcal{J})$, it will also be prudent to define $\mathcal{D}(\mathcal{I}) = \{\mathcal{J} \in \mathcal{C}(\mathcal{I})\colon \mathcal{J} < \mathcal{I}\}$.

A. The OPF problem

Within each extended subsystem $\mathcal{I}$, we define:

- $P^{\mathcal{I}}_{d_k} + iQ^{\mathcal{I}}_{d_k}$ – the given total power of the load at bus $k \in N^{\mathcal{I}}$, set to the value 0 at the dummy buses,
- $P^{\mathcal{I}}_{g_k} + iQ^{\mathcal{I}}_{g_k}$ – the unknown total power generated at the at bus $k \in G^{\mathcal{I}}$,
- $V^{\mathcal{I}}_k$ – the unknown complex voltage at bus $k \in N^{\mathcal{I}}$,
- $f^{\mathcal{I}}_k\left(P^{\mathcal{I}}_{g_k}\right) = c^{\mathcal{I}}_{k2}\left(P^{\mathcal{I}}_{g_k}\right)^2 + c^{\mathcal{I}}_{k1}P^{\mathcal{I}}_{g_k} + c^{\mathcal{I}}_{k0}$ – the quadratic cost function of generation of real power at bus $k \in \hat{G}^{\mathcal{I}}$.

978-1-4673-5716-6/13/$31.00 ©2013 IEEE

The replacement of tie-lines by dummy generator buses introduces at each such dummy bus local copies of the complicating variables associated with the tie-line. For this model to be an accurate representation of the tie-line, these local copies must satisfy particular consistency constraints: namely continuity of the real and reactive powers and equality of the voltages. The classical overall OPF problem for our network can then be formulated as minimizing the total generation cost $\sum_{\mathcal{I}\in\mathcal{E}}\sum_{k\in\hat{G}^{\mathcal{I}}} f_k^{\mathcal{I}}\big(P_{g_k}^{\mathcal{I}}\big)$ over the vectors of decision variables $\mathbf{P}_G^{\mathcal{I}} = \big(P_{g_k}^{\mathcal{I}}\big)_{k\in G^{\mathcal{I}}}$, $\mathbf{Q}_G^{\mathcal{I}} = \big(Q_{g_k}^{\mathcal{I}}\big)_{k\in G^{\mathcal{I}}}$, and $\mathbf{V}^{\mathcal{I}} = \big(V_k^{\mathcal{I}}\big)_{k\in N^{\mathcal{I}}}$, for $\mathcal{I}\in\mathcal{E}$, subject to satisfying the power demand at each bus, the physical constraints[1] $\underline{P}_k^{\mathcal{I}} \le P_{g_k}^{\mathcal{I}} \le \overline{P}_k^{\mathcal{I}}$ and $\underline{Q}_k^{\mathcal{I}} \le Q_{g_k}^{\mathcal{I}} \le \overline{Q}_k^{\mathcal{I}}$ for all $k\in\hat{G}^{\mathcal{I}}$, and $\underline{V}_k^{\mathcal{I}} \le |V_k^{\mathcal{I}}| \le \overline{V}_k^{\mathcal{I}}$ for all $k\in\hat{N}^{\mathcal{I}}$, over all $\mathcal{I}\in\mathcal{E}$, and the consistency constraints $P_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} = -P_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}$, $Q_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} = -Q_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}$, and $V_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} = V_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}$, over all $\mathcal{I}\in\mathcal{E}$ and all $\mathcal{J}\in\mathcal{D}(\mathcal{I})$.

B. Convex relaxation and regional decomposition

Within each region $\mathcal{I}$, we can form the auxiliary matrices $\mathbf{Y}_k^{\mathcal{I}}$, $\bar{\mathbf{Y}}_k^{\mathcal{I}}$, and $\mathbf{M}_k^{\mathcal{I}}$ introduced in [7]. Then, following the method there, it is possible to reformulate the OPF problem in terms of rank-one matrix variables $W^{\mathcal{I}}$, in terms of which the key variables at all nodes $k\in N^{\mathcal{I}}$ can then be written as $P_{g_k}^{\mathcal{I}} = P_{d_k}^{\mathcal{I}} + \mathrm{Tr}\big(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, $Q_{g_k}^{\mathcal{I}} = Q_{d_k}^{\mathcal{I}} + \mathrm{Tr}\big(\bar{\mathbf{Y}}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, and $\big|\mathbf{V}_k^{\mathcal{I}}\big|^2 = \mathrm{Tr}\big(\mathbf{M}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, where $P_{g_k}^{\mathcal{I}} = Q_{g_k}^{\mathcal{I}} = 0$ whenever $k\in N^{\mathcal{I}}\setminus G^{\mathcal{I}}$. At these non-generating nodes, also define $\underline{P}_k^{\mathcal{I}} = \overline{P}_k^{\mathcal{I}} = \underline{Q}_k^{\mathcal{I}} = \overline{Q}_k^{\mathcal{I}} = 0$. The tie-line power equality constraints may then be written as

$$\mathrm{Tr}\big(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\big) = -\,\mathrm{Tr}\big(\mathbf{Y}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\big), \tag{1}$$

$$\mathrm{Tr}\big(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\big) = -\,\mathrm{Tr}\big(\bar{\mathbf{Y}}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\big), \tag{2}$$

respectively. For the voltage equalities, we shall in fact explicitly require the equality in the squared magnitudes, namely $\big|V_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}\big|^2 = \big|V_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}\big|^2$, which can in turn be written

$$\mathrm{Tr}\big(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\big) = \mathrm{Tr}\big(\mathbf{M}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\big). \tag{3}$$

As for the phase equalities, let us now suppose that the network of tie-lines connecting the subsystems $1,\ldots,R$ has tree topology[2]. If the voltage vectors $\mathbf{V}^{\mathcal{I}}$ correspond to an optimal solution of the OPF problem with tie-line constraints (1), (2), (3), then consider inductively forming the vectors $\mathbf{v}^{\mathcal{I}}$ as follows. For any subsystem $\mathcal{I}$ in the n^{th} level of the tree structure, let $\mathbf{v}^{\mathcal{I}} = e^{i\phi_{\mathcal{I}}}\mathbf{V}^{\mathcal{I}}$ with $\phi_{\mathcal{I}}$ chosen so that the phase equality constraint is satisfied along the tie-line connecting $\mathcal{I}$ to the $(n-1)^{\text{th}}$ level. The phase factor for the single top-level subsystem may be set arbitrarily. It can be seen that these phase factor multiplications leave the values of all the key variables $\mathrm{Tr}\big(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, $\mathrm{Tr}\big(\bar{\mathbf{Y}}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, $\mathrm{Tr}\big(\mathbf{M}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$ unchanged. It thus follows that the collection $\mathbf{v}^{\mathcal{I}}$ will also be optimal and moreover, by our construction, will satisfy the voltage phase constraints along the tie-lines. This justifies omitting the voltage phase constraints in what follows.

Furthermore, the work in [7] justifies relaxing the problem by removing the rank constraint on the matrix $W^{\mathcal{I}}$. Doing so gives the following convex relaxation of the OPF problem.

$$\begin{array}{ll}
\underset{(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}}{\text{minimize}} & \displaystyle\sum_{\mathcal{I}\in\mathcal{E}}\sum_{k\in\hat{G}^{\mathcal{I}}} f_k^{\mathcal{I}}\left(P_{d_k}^{\mathcal{I}} + \mathrm{Tr}\left(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\right)\right) \\
\underset{\forall\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{D}(\mathcal{I})}{\text{subject to}} & \mathrm{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right) + \mathrm{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\right) = 0, \\
& \mathrm{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right) + \mathrm{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\right) = 0, \\
& \mathrm{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right) - \mathrm{Tr}\left(\mathbf{M}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\right) = 0, \\
& W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}},
\end{array} \tag{4}$$

where $\mathcal{R}^{\mathcal{I}}$ denotes the set of all $2(m_{\mathcal{I}}+n_{\mathcal{I}})\times 2(m_{\mathcal{I}}+n_{\mathcal{I}})$ matrices for which the physical inequality constraints on $\mathrm{Tr}\big(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, $\mathrm{Tr}\big(\bar{\mathbf{Y}}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$, and $\mathrm{Tr}\big(\mathbf{M}_k^{\mathcal{I}}W^{\mathcal{I}}\big)$ all hold.

We consider now a Lagrangian of (4) by introducing dual variables for each of the equality constraints. In particular, for given values of the dual variables $\boldsymbol{\lambda} = \big(\lambda_{\mathcal{I}\mathcal{J}}^P, \lambda_{\mathcal{I}\mathcal{J}}^Q, \lambda_{\mathcal{I}\mathcal{J}}^V\big)^T_{\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{D}(\mathcal{I})}$, we have

$$\begin{array}{l}
\underset{(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}}{\text{minimize}}
\begin{cases}
\displaystyle\sum_{\mathcal{I}\in\mathcal{E}}\Bigg(\sum_{k\in\hat{G}^{\mathcal{I}}} f_k^{\mathcal{I}}\left(P_{d_k}^{\mathcal{I}} + \mathrm{Tr}\left(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\right)\right) \\
\displaystyle +\sum_{\mathcal{J}\in\mathcal{D}(\mathcal{I})}\Big[\lambda_{\mathcal{I}\mathcal{J}}^P\Big(\mathrm{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right) + \mathrm{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\right)\Big) \\
+\,\lambda_{\mathcal{I}\mathcal{J}}^Q\Big(\mathrm{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right) + \mathrm{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\right)\Big) \\
+\,\lambda_{\mathcal{I}\mathcal{J}}^V\Big(\mathrm{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right) - \mathrm{Tr}\left(\mathbf{M}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W^{\mathcal{J}}\right)\Big)\Big]\Bigg)
\end{cases} \\
\underset{\forall\mathcal{I}\in\mathcal{E}}{\text{subject to}}\ \ W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}}.
\end{array} \tag{5}$$

Writing $F\left((W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}},\boldsymbol{\lambda}\right)$ for the objective function in (5),

$$\max_{\boldsymbol{\lambda}\ge 0}\ \min_{(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}} F\left((W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}},\boldsymbol{\lambda}\right) \tag{6}$$

is then a dual problem of (4), which by the convexity of (4) has zero duality gap if Slater's condition also holds. Our aim within the paper will be to develop appropriate update schemes for the dual variables so as to solve (4) in a decentralized way. In order that our analysis be non-trivial, we make the following assumption.

Assumption 1: There exists some $(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}$ in the relative interior of $\prod_{\mathcal{I}\in\mathcal{E}}\mathcal{R}^{\mathcal{I}}$ that is a feasible solution of (4).

The following theorem can then be stated.

Theorem 1: The collection of matrices $(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}$ is a solution of (4) if and only if there exists a choice of dual variables $\boldsymbol{\lambda}$ such that $(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}$ is also a solution of (5).

Proof: Noting that Assumption 1 expresses Slater's condition for both (4) and (5), the result follows immediately from strong duality. ■

Remark 1: It is clear that if, for some particular values of the dual variables, problem (5) admits a solution for which the equality constraints (1), (2), (3) all hold, then this solution will also be optimal for problem (4). Moreover, Theorem 1

[1]For simplicity, we have relaxed the apparent power constraints. These can easily be incorporated within our framework as shown in [7].

[2]Note that we only require the tree topology in the high-level network, consisting of the tie-lines connecting the subsystems indexed by $\mathcal{I}$. This imposes no restriction on the internal structures of the subsystems themselves.

guarantees that, whenever (4) has a solution, there must exist a choice of dual variables such that this occurs. Therefore, if our solution method for (5) in what follows produces a primal-dual pair $\big((W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}},\boldsymbol{\lambda}\big)$ such that the collection of primal variables $(W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}$ satisfies all of the equality constraints (1), (2), (3), then we can be assured that this choice of primal variables is indeed a solution of problem (4).

C. Distributed solution of the combined problem

We now consider splitting (5) into the regional problems

$$\underset{W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}}}{\text{minimize}}\quad \begin{cases} \sum_{k\in\hat{G}^{\mathcal{I}}} f_k^{\mathcal{I}}\left(P_{d_k}^{\mathcal{I}}+\operatorname{Tr}\left(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\right)\right)\\ +\sum_{\mathcal{J}\in\mathcal{C}(\mathcal{I})}\Big[\lambda_{\mathcal{I}\mathcal{J}}^{P}\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)\\ +\lambda_{\mathcal{I}\mathcal{J}}^{Q}\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)\\ +\delta_{\mathcal{I}\mathcal{J}}\lambda_{\mathcal{I}\mathcal{J}}^{V}\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)\Big], \end{cases} \tag{7}$$

where $\delta_{\mathcal{I}\mathcal{J}}$ takes the value 1 when $\mathcal{I}>\mathcal{J}$ and -1 when $\mathcal{I}<\mathcal{J}$, we have defined the additional dual variables $\lambda_{\mathcal{J}\mathcal{I}}^{P}=\lambda_{\mathcal{I}\mathcal{J}}^{P}$, $\lambda_{\mathcal{J}\mathcal{I}}^{Q}=\lambda_{\mathcal{I}\mathcal{J}}^{Q}$, $\lambda_{\mathcal{J}\mathcal{I}}^{V}=\lambda_{\mathcal{I}\mathcal{J}}^{V}$ for $\mathcal{J}\in\mathcal{C}(\mathcal{I})\backslash\mathcal{D}(\mathcal{I})$, and the values of the dual variables are regarded as given inputs. Using the Schur Complement Formula as in [7], these optimization problems can be expressed in the form of semidefinite programs and can thus be solved efficiently.

Theorem 2: The collection of problems given by (7) over all regions $\mathcal{I}\in\mathcal{E}$ together is equivalent to problem (5).

Proof: Denoting by $\operatorname{obj}_{\boldsymbol{\lambda}}^{\mathcal{I}}$ the objective function in (7) for given $\boldsymbol{\lambda}$, the minimization in (5) naturally splits as

$$\underset{(W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}}{\text{minimize}}\ F\left((W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}},\boldsymbol{\lambda}\right)=\sum_{\mathcal{I}\in\mathcal{E}}\underset{W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}}}{\text{minimize}}\ \operatorname{obj}_{\boldsymbol{\lambda}}^{\mathcal{I}}\left(W^{\mathcal{I}}\right)$$

and since this is now a sum of R independent minimization problems of the form (7), the choices of $W^{\mathcal{I}}$ that minimize this are exactly the optimal solutions of (7). ■

We thus have R local problems that can be considered separately, allowing the application of a classical dual decomposition algorithm here. Finally, since the convergence of a classical dual decomposition approach may not be guaranteed for problems that are non-strictly convex (e.g. [8], [9]), it can be of benefit to the system performance to augment the primal updates by means of quadratic regularization terms. A standard approach, as employed for example in the ADMM method [10], is to associate the regularization with the consistency constraints. This approach, however, requires the use of current primal variable values from neighbouring regions, which might not be feasible in cases of competing areas where a fully distributed implementation is sought. In order to avoid such additional information exchange, we instead use local auxiliary variables within each region, and regularize in terms of these local variables. The resulting iterative scheme is presented[3,4] in Algorithm 1.

[3]The scaling of the regularization is parameterized by ρ, while the positive constants α_n represent the iteration step size.

[4]In the cases when the algorithm is applicable, [7] provides a procedure by which each region $\mathcal{I}$ may obtain then a solution of the desired rank-one form from the output primal value obtained on termination of Algorithm 1.

Algorithm 1 Distributed dual decomposition algorithm for the relaxed OPF problem with quadratic regularization terms.

while $\Delta_n>\text{tolerance}$ **do**

1. Optimize regional problems to obtain $W_{n+1}^{\mathcal{I},\rho}$ as

$$\underset{W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}}}{\operatorname{argmin}}\begin{cases} \sum_{k\in\hat{G}^{\mathcal{I}}} f_k^{\mathcal{I}}\left(P_{d_k}^{\mathcal{I}}+\operatorname{Tr}\left(\mathbf{Y}_k^{\mathcal{I}}W^{\mathcal{I}}\right)\right)\\ +\sum_{\mathcal{J}\in\mathcal{C}(\mathcal{I})}\Big[\lambda_{\mathcal{I}\mathcal{J},n}^{P,\rho}\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)\\ +\lambda_{\mathcal{I}\mathcal{J},n}^{Q,\rho}\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)\\ +\delta_{\mathcal{I}\mathcal{J}}\lambda_{\mathcal{I}\mathcal{J},n}^{V,\rho}\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)\\ +\frac{1}{2}\rho\left(\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)-z_{\mathcal{I}\mathcal{J},n}^{P,\rho}\right)^2\\ +\frac{1}{2}\rho\left(\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)-z_{\mathcal{I}\mathcal{J},n}^{Q,\rho}\right)^2\\ +\frac{1}{2}\rho\left(\delta_{\mathcal{I}\mathcal{J}}\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W^{\mathcal{I}}\right)-z_{\mathcal{I}\mathcal{J},n}^{V,\rho}\right)^2\Big], \end{cases} \tag{8}$$

according to the consistency criterion that if given the same values of $\lambda_{\mathcal{I}\mathcal{J},n}^{P,\rho}$, $\lambda_{\mathcal{I}\mathcal{J},n}^{Q,\rho}$, $\lambda_{\mathcal{I}\mathcal{J},n}^{V,\rho}$, $z_{\mathcal{I}\mathcal{J},n}^{P,\rho}$, $z_{\mathcal{I}\mathcal{J},n}^{Q,\rho}$, and $z_{\mathcal{I}\mathcal{J},n}^{V,\rho}$ on multiple occasions the optimization method employed should on each occasion yield the same answer.

2. Update the dual variables by

$$\begin{aligned} \lambda_{\mathcal{I}\mathcal{J},n+1}^{P,\rho}&=\lambda_{\mathcal{I}\mathcal{J},n}^{P,\rho}+\alpha_n\Big(\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)\\ &\qquad+\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W_{n+1}^{\mathcal{J}}\right)\Big),\\ \lambda_{\mathcal{I}\mathcal{J},n+1}^{Q,\rho}&=\lambda_{\mathcal{I}\mathcal{J},n}^{Q,\rho}+\alpha_n\Big(\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)\\ &\qquad+\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W_{n+1}^{\mathcal{J}}\right)\Big),\\ \lambda_{\mathcal{I}\mathcal{J},n+1}^{V,\rho}&=\lambda_{\mathcal{I}\mathcal{J},n}^{V,\rho}+\alpha_n\Big(\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)\\ &\qquad-\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W_{n+1}^{\mathcal{J}}\right)\Big), \end{aligned}$$

for $\mathcal{I}\in\mathcal{E}$, $\mathcal{J}\in\mathcal{D}(\mathcal{I})$. Update the remaining dual variables by the rule $\lambda_{\mathcal{J}\mathcal{I}}^{P,\rho}=\lambda_{\mathcal{I}\mathcal{J}}^{P,\rho}$, $\lambda_{\mathcal{J}\mathcal{I}}^{Q,\rho}=\lambda_{\mathcal{I}\mathcal{J}}^{Q,\rho}$, and $\lambda_{\mathcal{J}\mathcal{I}}^{V,\rho}=\lambda_{\mathcal{I}\mathcal{J}}^{V,\rho}$.

3. Update the local auxiliary variables by

$$\begin{aligned} z_{\mathcal{I}\mathcal{J},n+1}^{P,\rho}&=z_{\mathcal{I}\mathcal{J},n}^{P,\rho}+\rho\alpha_n\Big(\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)-z_{\mathcal{I}\mathcal{J},n}^{P,\rho}\Big),\\ z_{\mathcal{I}\mathcal{J},n+1}^{Q,\rho}&=z_{\mathcal{I}\mathcal{J},n}^{Q,\rho}+\rho\alpha_n\Big(\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)-z_{\mathcal{I}\mathcal{J},n}^{Q,\rho}\Big),\\ z_{\mathcal{I}\mathcal{J},n+1}^{V,\rho}&=z_{\mathcal{I}\mathcal{J},n}^{V,\rho}+\rho\alpha_n\Big(\delta_{\mathcal{I}\mathcal{J}}\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)-z_{\mathcal{I}\mathcal{J},n}^{V,\rho}\Big), \end{aligned}$$

for $\mathcal{I}\in\mathcal{E}$, $\mathcal{J}\in\mathcal{C}(\mathcal{I})$.

4. Update the residual error by setting Δ_{n+1} equal to

$$\begin{aligned} \max_{\mathcal{I},\mathcal{J}}\Big\{&\left\|\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)+\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W_{n+1}^{\mathcal{J}}\right)\right\|,\\ &\left\|\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)+\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W_{n+1}^{\mathcal{J}}\right)\right\|,\\ &\left\|\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)-\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{J}\mathcal{I}}}^{\mathcal{J}}W_{n+1}^{\mathcal{J}}\right)\right\|,\\ &\left\|\operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)-z_{\mathcal{I}\mathcal{J},n}^{P}\right\|,\left\|\operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)\right.\\ &\left.-z_{\mathcal{I}\mathcal{J},n}^{Q}\right\|,\left\|\delta_{\mathcal{I}\mathcal{J}}\operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}}W_{n+1}^{\mathcal{I},\rho}\right)-z_{\mathcal{I}\mathcal{J},n}^{V}\right\|\Big\}. \end{aligned}$$

end while

Remark 2: In step 1, each region calculates locally the value of $W_{n+1}^{\mathcal{I},\rho}$ for its own region using the values of the prices $\lambda_{\mathcal{I}\mathcal{J},n}^{P,\rho}, \lambda_{\mathcal{I}\mathcal{J},n}^{Q,\rho}, \lambda_{\mathcal{I}\mathcal{J},n}^{V,\rho}$ relayed to it at the end of the previous iteration and its own local auxiliary variables $z_{\mathcal{I}\mathcal{J},n}^{P,\rho}, z_{\mathcal{I}\mathcal{J},n}^{Q,\rho}, z_{\mathcal{I}\mathcal{J},n}^{V,\rho}$. It is then required to send each value $\mathrm{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} W_{n+1}^{\mathcal{I},\rho}\right)$, $\mathrm{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} W_{n+1}^{\mathcal{I},\rho}\right)$, $\mathrm{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} W_{n+1}^{\mathcal{I},\rho}\right)$ to the processing unit for the link between $\mathcal{I}$ and $\mathcal{J}$, which uses these values to update the relevant dual variables. Moreover, these processing units can be isolated from each other and can be local to the relevant link. The auxiliary variable updates are entirely localized within each region. Each link processor is then required to send only its own maximum local error magnitude to any central processor. Until this processor announces that the target error has been reached, the local processing units return the updated prices $\lambda_{\mathcal{I}\mathcal{J},n+1}^{P,\rho}, \lambda_{\mathcal{I}\mathcal{J},n+1}^{Q,\rho}, \lambda_{\mathcal{I}\mathcal{J},n+1}^{V,\rho}$ to their relevant regions $\mathcal{I}$ and $\mathcal{J}$ and the algorithm repeats. Consequently, operators are only required to share their local copies of the complicating variables with link processors and do not have to disclose any internal operational details to rivals, whence, as desired, the algorithm requires only very little information sharing.

III. CONVERGENCE ANALYSIS

Some additional notation will be helpful for analyzing the convergence of Algorithm 1. First, let us define the local dual variable and auxiliary variable vectors $\boldsymbol{\lambda}_{\mathcal{I}\mathcal{J}} = \left(\lambda_{\mathcal{I}\mathcal{J}}^P, \lambda_{\mathcal{I}\mathcal{J}}^Q, \lambda_{\mathcal{I}\mathcal{J}}^V\right)^T$ and $\mathbf{z}_{\mathcal{I}\mathcal{J}} = \left(z_{\mathcal{I}\mathcal{J}}^P, z_{\mathcal{I}\mathcal{J}}^Q, z_{\mathcal{I}\mathcal{J}}^V\right)^T$, and then stack these to give the global vectors $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_{\mathcal{I}\mathcal{J}}^T)_{\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{D}(\mathcal{I})}^T$ and $\mathbf{z} = (\mathbf{z}_{\mathcal{I}\mathcal{J}}^T)_{\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{C}(\mathcal{I})}^T$. We then introduce the auxiliary notation $\mathbf{T}_{\mathcal{I}\mathcal{J}}(W^{\mathcal{I}})$ to denote the vector of traces of a matrix $W^{\mathcal{I}}$, $\mathbf{T}_{\mathcal{I}\mathcal{J}}(W^{\mathcal{I}}) = \big(\mathrm{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} W^{\mathcal{I}}\right), \mathrm{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} W^{\mathcal{I}}\right), \delta_{\mathcal{I}\mathcal{J}}\,\mathrm{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}}^{\mathcal{I}} W^{\mathcal{I}}\right)\big)^T$. Using this notation, we then form $\mathbf{T}((W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}) = \left(\mathbf{T}_{\mathcal{I}\mathcal{J}}(W^{\mathcal{I}})^T\right)_{\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{C}(\mathcal{I})}^T$, stacked in corresponding order to the variables in $\mathbf{z}$, and $\mathbf{U}((W^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}}) = \left(\mathbf{T}_{\mathcal{I}\mathcal{J}}(W^{\mathcal{I}})^T + \mathbf{T}_{\mathcal{J}\mathcal{I}}(W^{\mathcal{J}})^T\right)_{\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{D}(\mathcal{I})}^T$, stacked in corresponding order to the variables in $\boldsymbol{\lambda}$. Finally, for given values of $\boldsymbol{\lambda}$ and $\mathbf{z}$, we denote the quadratic regularization term in (8) by $D_{\mathbf{z}}^{\mathcal{I},\rho}(W^{\mathcal{I}})$ and we write $\overline{\mathrm{obj}}_{\boldsymbol{\lambda},\mathbf{z}}^{\mathcal{I},\rho}$ for the augmented objective function, so that, recalling the definition of $\mathrm{obj}_{\boldsymbol{\lambda}}^{\mathcal{I}}$,

$$\overline{\mathrm{obj}}_{\boldsymbol{\lambda},\mathbf{z}}^{\mathcal{I},\rho}\left(W^{\mathcal{I}}\right) = \mathrm{obj}_{\boldsymbol{\lambda}}^{\mathcal{I}}\left(W^{\mathcal{I}}\right) + D_{\mathbf{z}}^{\mathcal{I},\rho}(W^{\mathcal{I}}). \tag{9}$$

In terms of this notation, the update rules in Algorithm 1 may be written more compactly as

$$W_{n+1}^{\mathcal{I},\rho} = \underset{W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}}}{\mathrm{argmin}}\ \overline{\mathrm{obj}}_{\boldsymbol{\lambda}_n^{\rho},\mathbf{z}_n^{\rho}}^{\mathcal{I},\rho}\left(W^{\mathcal{I}}\right), \tag{10}$$

$$\boldsymbol{\lambda}_{n+1}^{\rho} = \boldsymbol{\lambda}_n^{\rho} + \alpha_n \mathbf{U}\left((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\right), \tag{11}$$

$$\mathbf{z}_{n+1}^{\rho} = \mathbf{z}_n^{\rho} + \rho\alpha_n\left(\mathbf{T}\left((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\right) - \mathbf{z}_n^{\rho}\right). \tag{12}$$

For the convergence analysis, it will suffice to investigate the stability properties of the dual and auxiliary variable update schemes. Before we proceed, it is important to understand the significance of the fixed points of Algorithm 1.

Theorem 3: The collection of primal variables at any fixed point of Algorithm 1 solves the relaxed OPF problem (4).

Proof: Suppose first that $\left((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_*^{\rho}, \mathbf{z}_*^{\rho}\right)$ is a fixed point of Algorithm 1 for a general value of ρ. Then the fixed point conditions for (12) give $\mathbf{z}_{*,\mathcal{I}\mathcal{J}}^{\rho} = \mathbf{T}_{\mathcal{I}\mathcal{J}}(W_*^{\mathcal{I},\rho})$ for each pair of regions $\mathcal{I}$ and $\mathcal{J}$. Now let $(W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}}$ be optimal for the problems (7) with dual variables $\boldsymbol{\lambda}_*^{\rho}$ and consider the matrix $W_{\gamma}^{\mathcal{I}} = (1-\gamma)W_*^{\mathcal{I},0} + \gamma W_*^{\mathcal{I},\rho}$. Then $D_{\mathbf{z}_*^{\rho}}^{\mathcal{I},\rho}(W_{\gamma}^{\mathcal{I}}) = \frac{1}{2}\rho(1-\gamma)^2\sum_{\mathcal{J}\in\mathcal{C}(\mathcal{I})}\|T_{\mathcal{I}\mathcal{J}}(W_*^{\mathcal{I},\rho}) - T_{\mathcal{I}\mathcal{J}}(W_*^{\mathcal{I},0})\|_2^2$. Additionally, the convexity of the unregularized objective function $\mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}$ gives $\mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}(W_{\gamma}^{\mathcal{I}}) \le \mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},\rho}) + (1-\gamma)\big(\mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},0}) - \mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},\rho})\big)$. Invoking the decomposition (9) and the optimality of $W_*^{\mathcal{I},0}$ for $\mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}$, it is clear that the required optimality condition $\overline{\mathrm{obj}}_{\boldsymbol{\lambda}_*^{\rho},\mathbf{z}_*^{\rho}}^{\mathcal{I},\rho}(W_*^{\mathcal{I},\rho}) \le \overline{\mathrm{obj}}_{\boldsymbol{\lambda}_*,\mathbf{z}_*^{\rho}}^{\mathcal{I}}(W_{\gamma}^{\mathcal{I}})$ can only hold for all $\gamma\in[0,1]$ if $\mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},\rho}) = \mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},0})$. We thus conclude that, for each $\mathcal{I}$, $W_*^{\mathcal{I},\rho}$ is optimal for the problem (7) with dual variable values $\boldsymbol{\lambda}_*^{\rho}$, whence by Theorem 2 the collection $(W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}$ is a solution of (5). Moreover the fixed point conditions for (11) tell us that this collection satisfies all of the equality constraints (1), (2), (3). Therefore, according to Remark 1, $(W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}$ must solve (4). ■

Theorem 4: For any value of ρ, there exists a fixed point $\left((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_*^{\rho}, \mathbf{z}_*^{\rho}\right)$ of Algorithm 1.

Proof: First consider only the case $\rho = 0$. Recall from Theorems 1 and 2, together with Assumption 1, that there exists some choice of the dual variables $\lambda_{\mathcal{I}\mathcal{J}}^P$, $\lambda_{\mathcal{I}\mathcal{J}}^Q$, $\lambda_{\mathcal{I}\mathcal{J}}^V$ such that the solutions of systems (7) satisfy the equality constraints (1), (2), (3). Define $\lambda_{\mathcal{I}\mathcal{J},*}^{P,0}, \lambda_{\mathcal{I}\mathcal{J},*}^{Q,0}, \lambda_{\mathcal{I}\mathcal{J},*}^{V,0}$ to be equal to this choice and let $\boldsymbol{\lambda}_*^0 = \left(\lambda_{\mathcal{I}\mathcal{J},*}^{P,0}, \lambda_{\mathcal{I}\mathcal{J},*}^{Q,0}, \lambda_{\mathcal{I}\mathcal{J},*}^{V,0}\right)_{\mathcal{I}\in\mathcal{E},\mathcal{J}\in\mathcal{D}(\mathcal{I})}^T$. Then, since $\rho = 0$ means that the auxiliary variables $\mathbf{z}$ make no appearance in the primal optimization problems, we may calculate the corresponding primal optimal values $(W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}}$ from (8) and then let $\mathbf{z}_*^0 = \mathbf{T}((W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}})$. We then see that we see that $\left((W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_*^0, \mathbf{z}_*^0\right)$ is a fixed point of Algorithm 1.

Thus, there clearly exists a fixed point of in the case where $\rho = 0$. We now show that this is in fact gives rise to a fixed point of Algorithm 1 for any non-zero value of ρ by setting $\boldsymbol{\lambda}_*^{\rho} = \boldsymbol{\lambda}_*^0$, $\mathbf{z}_*^{\rho} = \mathbf{z}_*^0$, and defining $W_*^{\mathcal{I},\rho}$ to be the optimal solution to (10) with these dual and auxiliary variable values. According to the definitions $\mathbf{z}_*^{\rho} = \mathbf{z}_*^0 = \mathbf{T}((W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}})$, for any $\mathcal{I}$ we have $D_{\mathbf{z}_*^{\rho}}^{\mathcal{I},\rho}(W_*^{\mathcal{I},0}) = 0$. In addition, $W_*^{\mathcal{I},0}, W_*^{\mathcal{I},\rho} \in \mathcal{R}^{\mathcal{I}}$ and, since $\boldsymbol{\lambda}_*^{\rho} = \boldsymbol{\lambda}_*^0$, $W_*^{\mathcal{I},0}$ minimizes $\mathrm{obj}_{\boldsymbol{\lambda}_*^{\rho}}^{\mathcal{I}}$ over $\mathcal{R}^{\mathcal{I}}$. Thus, (9) gives $\overline{\mathrm{obj}}_{\boldsymbol{\lambda}_*,\mathbf{z}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},0}) \le \overline{\mathrm{obj}}_{\boldsymbol{\lambda}_*^{\rho},\mathbf{z}_*^{\rho}}^{\mathcal{I},\rho}(W_*^{\mathcal{I},\rho})$ with equality only possible here, according to the strict convexity of $D_{\mathbf{z}_*^{\rho}}^{\mathcal{I}}$, if $D_{\mathbf{z}_*^{\rho}}^{\mathcal{I}}(W_*^{\mathcal{I},\rho}) = 0$, which occurs if and only if $\mathbf{T}_{\mathcal{I}\mathcal{J}}(W_*^{\mathcal{I},\rho}) = \mathbf{z}_{*,\mathcal{I}\mathcal{J}}^{\rho}$ for all $\mathcal{J}\in\mathcal{C}(\mathcal{I})$. But $W_*^{\mathcal{I},\rho}$ is by definition a minimizer of $\overline{\mathrm{obj}}_{\boldsymbol{\lambda}_*^{\rho},\mathbf{z}_*^{\rho}}^{\mathcal{I},\rho}$, so equality must hold here, and thus we must have $\mathbf{T}_{\mathcal{I}\mathcal{J}}(W_*^{\mathcal{I},\rho}) = \mathbf{z}_{*,\mathcal{I}\mathcal{J}}^{\rho} = \mathbf{T}_{\mathcal{I}\mathcal{J}}(W_*^{\mathcal{I},0})$ for all $\mathcal{J}\in\mathcal{C}(\mathcal{I})$. Moreover, these trace equalities hold for all regions $\mathcal{I}$, and so we see clearly that the fixed point conditions for (12) hold. Furthermore, we then clearly have $\mathbf{U}((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}) = \mathbf{U}((W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}})$, which, according to the definition of $\left((W_*^{\mathcal{I},0})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_*^0, \mathbf{z}_*^0\right)$ as a fixed point of the $\rho = 0$ algorithm, means that the fixed point conditions for (11) must hold also. Finally, since $\boldsymbol{\lambda}_*^{\rho}$ and $\mathbf{z}_*^{\rho}$ are both fixed, the consistency constraint guarantees that the $W_*^{\mathcal{I},\rho}$ are then fixed under (10). Therefore, the triple $\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_*^\rho, \mathbf{z}_*^\rho\big)$ defines a fixed point for Algorithm 1. ■

We now prove our main convergence result.

Theorem 5: Suppose that $\rho \neq 0$. Then for sufficiently small positive values of α_n, given any starting value $\big((W_0^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_0^\rho, \mathbf{z}_0^\rho\big)$ in Algorithm 1, $\big((W_n^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_n^\rho, \mathbf{z}_n^\rho\big)$ converges to the set of fixed points as $n \to \infty$.

Proof: Let $\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_*^\rho, \mathbf{z}_*^\rho\big)$ be a fixed point. For a given value of n, consider an arbitrary point in the line segment $[(W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, (W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}]$, $\Omega_\gamma = (\Omega_\gamma^{\mathcal{I}})_{\mathcal{I}\in\mathcal{E}} = (1-\gamma)(W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}} + \gamma(W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}$ for $\gamma \in [0,1]$.

We now decompose the objective functions as in (9). Rearranging the trace sums into the equivalent form appearing in (5), invoking both the fixed point conditions for (11) and the linearity of the trace operator, and substituting for the appropriate trace sums from (11), we may express the sum of the unregularized objective values as

$$\sum_{\mathcal{I}\in\mathcal{E}} \mathrm{obj}_{\boldsymbol{\lambda}_n^\rho}^{\mathcal{I}}\left(\Omega_\gamma^{\mathcal{I}}\right) = \sum_{\mathcal{I}\in\mathcal{E}} \mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}\left(\Omega_\gamma^{\mathcal{I}}\right) + \frac{\gamma}{\alpha_n}\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right). \tag{13}$$

Additionally, invoking the linearity of the trace operator,

$$\sum_{\mathcal{I}\in\mathcal{E}} D_{\mathbf{z}_n^\rho}^{\mathcal{I}}(\Omega_\gamma^{\mathcal{I}}) = \frac{1}{2}\rho\Big\|(1-\gamma)\mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) + \gamma\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{z}_n^\rho\Big\|_2^2. \tag{14}$$

If we now sum (13) and (14), we obtain

$$\sum_{\mathcal{I}\in\mathcal{E}} \overline{\mathrm{obj}}_{\boldsymbol{\lambda}_n^\rho,\mathbf{z}_n^\rho}^{\mathcal{I}}\left(\Omega_\gamma^{\mathcal{I}}\right) = \sum_{\mathcal{I}\in\mathcal{E}} \mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}\left(\Omega_\gamma^{\mathcal{I}}\right) + \frac{\gamma}{\alpha_n}\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right) + \frac{1}{2}\rho\Big\|(1-\gamma)\mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) + \gamma\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{z}_n^\rho\Big\|_2^2. \tag{15}$$

Setting $\gamma = 1$ in (15) yields the corresponding expression

$$\sum_{\mathcal{I}\in\mathcal{E}} \overline{\mathrm{obj}}_{\boldsymbol{\lambda}_n^\rho,\mathbf{z}_n^\rho}^{\mathcal{I}}(W_{n+1}^{\mathcal{I},\rho}) = \sum_{\mathcal{I}\in\mathcal{E}} \mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}(W_{n+1}^{\mathcal{I},\rho}) + \frac{1}{\alpha_n}\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right) + \frac{1}{2}\rho\Big\|\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{z}_n^\rho\Big\|_2^2. \tag{16}$$

Now, $W_*^{\mathcal{I},\rho}$ and $W_{n+1}^{\mathcal{I},\rho}$ both lie in the convex feasible set $\mathcal{R}^{\mathcal{I}}$, so we must have $\Omega_\gamma^{\mathcal{I}} \in \mathcal{R}^{\mathcal{I}}$. The minimization property (8) thus tells us that, for each $\mathcal{I}$, $\overline{\mathrm{obj}}_{\boldsymbol{\lambda}_n^\rho,\mathbf{z}_n^\rho}^{\mathcal{I}}(\Omega_\gamma^{\mathcal{I}}) \geq \overline{\mathrm{obj}}_{\boldsymbol{\lambda}_n^\rho,\mathbf{z}_n^\rho}^{\mathcal{I}}(W_{n+1}^{\mathcal{I},\rho})$. Additionally, the proof of Theorem 3 tells us that $W_*^{\mathcal{I},\rho}$ is optimal for the unregularized minimization problem $\underset{W^{\mathcal{I}}\in\mathcal{R}^{\mathcal{I}}}{\mathrm{minimize}}\, \mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}\left(W^{\mathcal{I}}\right)$, whence the convexity of $\mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}$ and the definition of $\Omega_\gamma^{\mathcal{I}}$ imply that $\mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}(W_{n+1}^{\mathcal{I},\rho}) \geq \mathrm{obj}_{\boldsymbol{\lambda}_*^\rho}^{\mathcal{I}}(\Omega_\gamma^{\mathcal{I}})$. Then, writing $\boldsymbol{\omega}_\gamma = (1-\gamma)\mathbf{T}((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}) + \gamma\mathbf{T}((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}})$, subtracting (16) from (15) finally yields the key inequality that, for all $\gamma \in [0,1]$,

$$\frac{1}{2}\rho\left\|\boldsymbol{\omega}_\gamma - \mathbf{z}_n^\rho\right\|_2^2 - \frac{1}{2}\rho\left\|\mathbf{T}\left((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\right) - \mathbf{z}_n^\rho\right\|_2^2 \geq \frac{1}{\alpha_n}(1-\gamma)\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right). \tag{17}$$

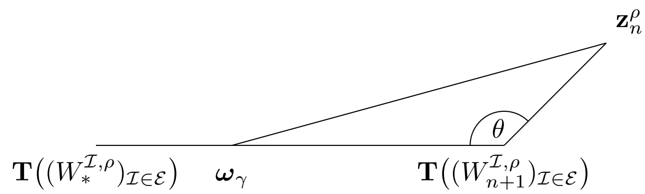

Fig. 1. Geometry of $\mathbf{z}_n^\rho$, $\mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)$, $\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)$, and $\boldsymbol{\omega}_\gamma$.

We now consider the plane passing through the points $\mathbf{z}_n^\rho$, $\mathbf{T}((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}})$, and $\mathbf{T}((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}})$, depicted in Fig. 1, where θ denotes the angle between the line segments $[\mathbf{T}((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}), \mathbf{T}((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}})]$ and $[\mathbf{T}((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}), \mathbf{z}_n^\rho]$. Substituting for $\|\boldsymbol{\omega}_\gamma - \mathbf{z}_n^\rho\|_2^2$ by the Cosine Rule applied to the triangle shown and dividing by $1-\gamma$ for $\gamma \neq 1$, (17) becomes

$$\frac{1}{2}\rho(1-\gamma)\Big\|\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)\Big\|_2^2 - \rho\Big\|\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)\Big\|_2\Big\|\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{z}_n^\rho\Big\|_2 \cos\theta \geq \frac{1}{\alpha_n}\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right), \tag{18}$$

which must be satisfied for all $\gamma \in [0,1)$. If we let $\gamma \uparrow 1$, then we see that (18) can only remain true if

$$\rho\left(\mathbf{z}_n^\rho - \mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)\right)\cdot\left(\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)\right) \geq \frac{1}{\alpha_n}\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right), \tag{19}$$

as the angle between the two vectors is $\pi - \theta$. Adding to this

$$\rho\Big\|\mathbf{z}_n^\rho - \mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)\Big\|_2^2 \geq 0, \tag{20}$$

noting that $\mathbf{z}_{n+1}^\rho - \mathbf{z}_n^\rho = \alpha_n\rho\left(\mathbf{T}\big((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) - \mathbf{z}_n^\rho\right)$ and $\mathbf{z}_*^\rho = \mathbf{T}\big((W_*^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)$, and finally rearranging, we thus get

$$\left(\boldsymbol{\lambda}_n^\rho - \boldsymbol{\lambda}_*^\rho\right)\cdot\left(\boldsymbol{\lambda}_{n+1}^\rho - \boldsymbol{\lambda}_n^\rho\right) + \left(\mathbf{z}_n^\rho - \mathbf{z}_*^\rho\right)\cdot\left(\mathbf{z}_{n+1}^\rho - \mathbf{z}_n^\rho\right) \leq 0. \tag{21}$$

Carefully considering the foregoing argument, we see that we get equality in (21) if and only if we have equality in both (19) and (20). Now let us suppose that we have equality in (21) for both $n = N$ and $n = N+1$. Then equality in (20) implies $\mathbf{z}_N^\rho = \mathbf{T}\big((W_{N+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)$ and $\mathbf{z}_{N+1}^\rho = \mathbf{T}\big((W_{N+2}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)$, whence (12) gives $\mathbf{z}_N^\rho = \mathbf{z}_{N+1}^\rho$ and so $\mathbf{T}\big((W_{N+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big) = \mathbf{T}\big((W_{N+2}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}\big)$. Then (11) can be seen to give $\boldsymbol{\lambda}_{N+1}^\rho - \boldsymbol{\lambda}_N^\rho = \boldsymbol{\lambda}_{N+2}^\rho - \boldsymbol{\lambda}_{N+1}^\rho$, in which case equality in (19) then implies $\boldsymbol{\lambda}_N^\rho = \boldsymbol{\lambda}_{N+1}^\rho$. The consistency criterion then tells us that also $W_{N+1}^{\mathcal{I},\rho} = W_{N+2}^{\mathcal{I},\rho}$ for all $\mathcal{I} \in \mathcal{E}$, so we see that $\big((W_{N+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_{N+1}^\rho, \mathbf{z}_{N+1}^\rho\big)$ is a fixed point. Consequently, if we define the combined variables $\mathbf{S}_n^\rho = \big((\boldsymbol{\lambda}_n^\rho)^T, (\mathbf{z}_n^\rho)^T, (\boldsymbol{\lambda}_{n+1}^\rho)^T, (\mathbf{z}_{n+1}^\rho)^T\big)^T$ and $\mathbf{S}_*^\rho = \big((\boldsymbol{\lambda}_*^\rho)^T, (\mathbf{z}_*^\rho)^T, (\boldsymbol{\lambda}_*^\rho)^T, (\mathbf{z}_*^\rho)^T\big)^T$, then, away from the set of fixed points, we are guaranteed to have the strict inequality

$$\left(\mathbf{S}_{n+1}^\rho - \mathbf{S}_n^\rho\right)\cdot\left(\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\right) < 0, \tag{22}$$

for all $n \in \mathbb{N}$. Geometrically, (22) says that the point $\mathbf{S}_{n+1}^\rho$ must lie within an open half-space defined by the tangent

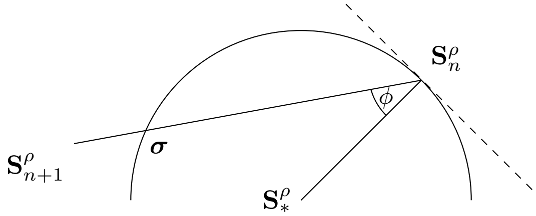

Fig. 2. The point $\mathbf{S}_{n+1}^\rho$ must lie to the left of the dashed tangent hyperplane. The point $\boldsymbol{\sigma}$ denotes the other intersection of the line through $\mathbf{S}_n^\rho$ and $\mathbf{S}_{n+1}^\rho$ with the hypersphere of radius $\|\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\|_2$ about $\mathbf{S}_*^\rho$.

hyperplane to the hypersphere of radius $\|\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\|_2$ about $\mathbf{S}_*^\rho$. The line segment $[\mathbf{S}_n^\rho, \mathbf{S}_{n+1}^\rho]$ must then have non-empty intersection with the interior of this hypersphere, as in Fig. 2.

Define composite vectors $\mathbf{R}_n^\rho = \rho\left(\frac{1}{\rho}\mathbf{U}((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}})^T, \left[\mathbf{T}((W_{n+1}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}) - \mathbf{z}_n^\rho\right]^T, \frac{1}{\rho}\mathbf{U}((W_{n+2}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}})^T, \left[\mathbf{T}((W_{n+2}^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}) - \mathbf{z}_{n+1}^\rho\right]^T\right)^T$. Then suppose that, for all $n \in \mathbb{N}$, consecutive step sizes in Algorithm 1 satisfy the strict inequality

$$\alpha_n < \min\left\{\frac{-2\left(\mathbf{S}_{n-1}^\rho - \mathbf{S}_*^\rho\right)\cdot\mathbf{R}_{n-1}^\rho}{\|\mathbf{R}_{n-1}^\rho\|_2^2}, \frac{-2\left(\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\right)\cdot\mathbf{R}_n^\rho}{\|\mathbf{R}_n^\rho\|_2^2}\right\}. \tag{23}$$

It is important to note that (22) together with the update rules (11), (12) imply that the right-hand side of (23) is always strictly positive. Therefore, it is always possible to make α_n sufficiently small, yet still positive, such that (23) is satisfied.

Considering the geometry of Fig. 2, this implies $\|\mathbf{S}_{n+1}^\rho - \mathbf{S}_n^\rho\|_2 \leq \max\{\alpha_n, \alpha_{n+1}\}\|\mathbf{R}_n^\rho\|_2 < 2\|\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\|_2 \cos\phi = \|\sigma - \mathbf{S}_n^\rho\|_2$, which shows that the point $\mathbf{S}_{n+1}^\rho$ must lie within the interior of the hypersphere. Therefore, we conclude that, under the assumption (23), we are guaranteed to have $\|\mathbf{S}_{n+1}^\rho - \mathbf{S}_*^\rho\|_2 < \|\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\|_2$. Thus, if we define $\mathcal{V}_n = \|\mathbf{S}_n^\rho - \mathbf{S}_*^\rho\|_2^2 = \|\boldsymbol{\lambda}_n^\rho\|_2^2 + \|\mathbf{z}_n^\rho\|_2^2 + \|\boldsymbol{\lambda}_{n+1}^\rho\|_2^2 + \|\mathbf{z}_{n+1}^\rho\|_2^2$, then we see that, away from the set of fixed points of Algorithm 1, $\mathcal{V}_n$ is a positive definite function that is strictly monotone decreasing as n increases. That is, $\mathcal{V}_n$ defines a Lyapunov function for Algorithm 1. Therefore, Lasalle's Theorem guarantees that, given any initial condition, $\left((W_n^{\mathcal{I},\rho})_{\mathcal{I}\in\mathcal{E}}, \boldsymbol{\lambda}_n^\rho, \mathbf{z}_n^\rho\right)$ must converge to the set of fixed points as $n \to \infty$. ■

Remark 3: Theorem 5 provides a guarantee that Algorithm 1 will always converge, from any starting value, provided that the step sizes are small enough. However, since the bound (23) depends on the current and last state values, the required step sizes are not known a priori. Nevertheless, this result does allow one to repeatedly run the system with decreasing step sizes with confidence that eventually convergent behavior will be observed. In this way, Theorems 3–5 guarantee that Algorithm 1 can be used to obtain a solution to our original relaxed OPF problem (4). For particular problems it may be possible to obtain a priori bounds on the terms in (23) in terms of objective function differences.

We illustrate in Fig. 3 an application of Algorithm 1 to a composite network formed from a single top-level IEEE 30-bus test system and three separate IEEE 14-bus test systems, interconnected over a star topology (which has the tree structure required for our analysis). We plot the average relative distances of the dual variable values $\boldsymbol{\lambda}_n^1$ from their values at the fixed point $\boldsymbol{\lambda}_*^1$ starting from 25 randomly-chosen starting points, and we observe rapid decay of this error.

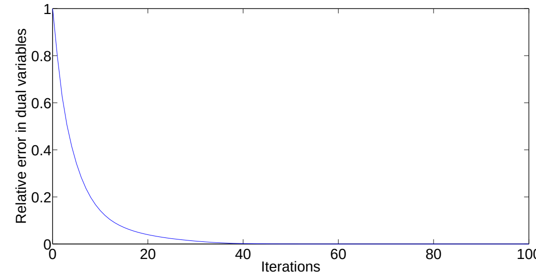

Fig. 3. Average relative dual variable error against iteration count for Algorithm 1 applied with unit step size to 14/30-bus star interconnection. Convergence to within a 0.1% tolerance level is obtained in just 44 iterations.

IV. Conclusions

We have considered a decomposition of an important convex relaxation of the OPF problem into a collection of local optimization problems and have justified its applicability for networks whose high-level structures have tree topology. This was facilitated by the introduction of dual variables for the complicating equality constraints along the tie-lines joining the network's dense regions. In order to avoid the requirement of additional information sharing associated with classical quadratic regularization, we then considered a modification of this approach in terms of local auxiliary variables. This leads to schemes with higher-order dynamics which can be of benefit to the overall system performance. The fixed points of the modified system were shown to correspond to solutions of the original relaxed global OPF problem. We then proved that, for sufficiently small values of the step size, this augmented algorithm is guaranteed to yield convergence to the set of fixed points. The analysis demonstrates that the iterative approach considered does indeed provide a distributed method for obtaining a solution of the original global OPF problem.

References

[1] J. Carpentier, "Contribution to the economic dispatch problem," *Bull. Soc. Francaise Elect.*, vol. 3, no. 8, pp. 431–447, 1962.

[2] J. A. Momoh, M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. I: Nonlinear and quadratic programming approaches," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 96–104, Feb. 1999.

[3] J. A. Momoh, M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. II: Newton, linear programming and interior point methods," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 105–111, Feb. 1999.

[4] B. H. Kim and R. Baldick, "Coarse-grained distributed optimal power flow," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 932–939, Aug. 1997.

[5] A. Y. S. Lam, B. Zhang, and D. N. Tse, "Distributed algorithms for optimal power flow problem," *Proc. 51st IEEE Conf. Decision Control*, pp. 430–437, Dec. 2012.

[6] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int. J. Elect. Power Energy Syst.*, vol. 30, no. 6–7, pp. 383–392, Jul.–Sep. 2008.

[7] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[8] K. J. Arrow, L. Hurwicz, and H. Uzawa, *Studies in linear and non-linear programming*. Stanford University Press, 1958.

[9] T. Holding and I. Lestas, "On the emergence of oscillations in distributed resource allocations," *Proc. 52nd IEEE Conf. Decision Control*, Dec. 2013.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers", *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp.1–122, Jan. 2011.