# A Computational and Architectural Comparison of Benders Decomposition and ADMM for Solving Decentralized DC-OPF

Soner Candas[1*], Kai Zhang[2†], Thomas Hamacher [1‡]
[1]Technical University of Munich, Chair of Renewable and Sustainable Energy Systems,
Lichtenbergstr. 4a, 85748 Garching b. München, Germany
[2]TUMCREATE Ltd., 138602, Singapore
Email: *soner.candas@tum.de, †kai.zhang@tum-create.edu.sg ‡thomas.hamacher@tum.de

***Abstract*—To be used for decentrally solving the DC-optimal power flow (DC-OPF) problems, this work offers a comparative analysis between methods representing two distinct families of decomposition techniques: Benders decomposition (BD) from cutting-plane methods and consensus-based alternating direction method of multipliers (ADMM) from dual decomposition-like algorithms. Within the scope of the study, the computational performance and their communication requirements of these methods are compared. The results demonstrate a relative advantage of BD in terms of computational performance, as ADMM mostly requires a considerably larger number of iterations under the same convergence criterion. ADMM in turn has a completely distributed architecture, which allows regulators to withhold local information and realizes a higher parallelization potential.**

***Index Terms*—Decentralized optimization, DC-optimal power flow, Benders decomposition, ADMM**

## I. Introduction

In modern electricity supply, growing presence of distributed power generation units and increasing number of demand-side decision makers lead to impracticality of the conventional, central management of energy production and flows. Necessity arises, among other things, for distributed solution approaches to optimal power flow (OPF) problems, which have been serving as an effective tool for power system control, planning and market design [1,2]. Moreover, conventional operation of power systems brings further challenges when handling large-scale systems because of communication constraints and flexibility requirements, even more so with the liberalization of electricity markets. In this regard, distributed optimization techniques to solve OPF problems provide a framework that can accommodate large-scale system operation tasks under the presence of contingencies [2].

In order to enable decentralized optimization paradigms, partitioning the original optimization problem into smaller, minimally-interdependent subproblems via so-called *decomposition techniques* are used. In the literature, a wide family of these techniques have been applied in energy system modeling: Boyd's work about dual decomposition and alternating direction method of multipliers (ADMM) [3] has discussed the possibility of their deployment in distributed energy management systems, which has been intensively studied in recent years (see e.g. [4]–[7]). A few review studies [1,2] were conducted to compare ADMM to alternative methods (e.g. Karush-Kuhn-Tucker (KKT)-condition based decomposition [8], proximal message passing (PMP) [9] among others), which provides insights for their computational performance and organizational requirements. It turns out that ADMM has certain advantages due to its simplicity for implementation, robustness and scalability [2]. A decomposition technique of cutting-plane type, Benders decomposition (BD) has been initially introduced for mixed-integer linear optimization problems [10], in which the integer and continuous variables were solved for sequentially. Since its conception, various modifications to BD have been introduced; ranging from its generalization to any arbitrary convex problems [11], to methods aiming at improving convergence (refer to [12] for a comprehensive review). Despite having a mode of action fundamentally different from ADMM, recent works in literature demonstrate the suitability of BD also on decentralized OPF applications ([13,14]).

For each of these methods, computational efficiency, communications security and organizational aspects vary; hence these need to be carefully addressed. For example, BD needs a central coordinator to collect, process and redistribute the information, whereas only neighbor-to-neighbor information exchange is needed for ADMM's consensus protocols [2,7]. Thus depending on the regulatory framework, e.g. for applications where the sensitive information exchange is not allowed, BD may not be deployed despite its potentially faster convergence. Nevertheless, to the authors' knowledge, an assessment of BD at solving OPF problems within the domain of other decomposition methods has not been considered in previous works, which is the main focus of this work.

The contribution and main findings of this work are summarized as follows:

- We provide a comprehensive comparison of BD and ADMM in terms of organization, communication requirements and computational performance using standard networks. By solving DC-OPF problems with various number of network sizes, different network partition formations and levels of contingencies, the applicability of both algorithms is well understood,
- The study shows that BD generally outperforms ADMM computationally (despite limited parallelization due to the master problem), whereas the hierarchical information exchange makes BD less flexible compared to ADMM in the presence of regulatory restrictions.

## II. Decomposition of a multi-subsystem problem

A generic, decomposable optimization problem with $N$-subsystems has the following form:

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad \sum_{m=1}^{N} c_m(\boldsymbol{x}_m) \tag{1a}$$

$$\text{s.t.} \quad \mathbf{A}\boldsymbol{x} + \mathbf{E}\boldsymbol{y} \leq \boldsymbol{r}, \tag{1b}$$

$$\boldsymbol{y} \in \mathcal{Y} \tag{1c}$$

where the state of each subsystem $m \in \{1, \ldots, N\}$, is described by two types of variables: its local variables $\boldsymbol{x}_m \in \mathbb{R}^{n_m}$ and the so-called coupling (complicating) variables $\boldsymbol{y} \in \mathbb{R}^z$. The solution space is defined by two sets of constraints: (1b) where both types of variables appear, and (1c) which only concerns the complicating variables (see Figure 1 for a block representation). Note that these coupling variables are assumed to have no contribution to the cost function.

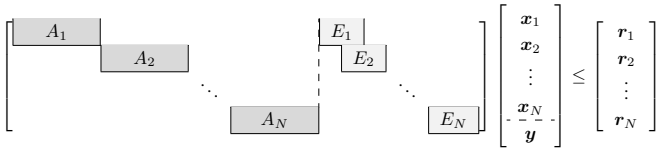


Fig. 1. Block structure of a multi-subsystem ($m \in \{1, \ldots, N\}$) optimization problem)

Presence of the constraints in form of (1b) prevents the direct separation of the problem into independent subproblems. Instead, decomposition methods reformulate the problem and are performed iteratively, through which the coupling information between the subsystems, e.g. the values of the complicating variables, is coordinated and inconsistencies in-between are cleared in convergence. Depending on the specific method that is employed, this coordination can take place by fixing the $\boldsymbol{y}$ values by solving a central master problem before the subproblems (as is the case for BD), or allowing each subproblem to solve for its local copy of $\boldsymbol{y}$ independently, which are then coordinated to calculate a global value (as is the case for ADMM).

## III. Cluster-based DC-OPF

OPF problems make a practical case for decomposition approaches as the original problem can be realistically considered as a combination of multiple subproblems, each of them representing bus clusters operated by a distinct regulatory body (e.g. Regional Transmission Organization (RTOs) in the North American electricity market). Decentralization reduces the required amount of cross-cluster information exchange, promoting confidentiality. Parallelized solvability of multiple smaller-scale problems, on the other hand, could offer a way to overcome computational hurdles accompanying large-scale models, e.g. for handling of the working memory or dampening a divergent computation time.

### A. Problem formulation

In this section, the cluster-based DC-OPF problem will be introduced. We consider a network that consists of a set of buses $\mathcal{N}$ and transmission lines $\mathcal{L}$ connecting them, and is clustered into $|\mathcal{M}|$ subsystems, $\mathcal{M}$ denoting the set of those. The set of buses contained in a given cluster $m \in \mathcal{M}$ is denoted by $\mathcal{N}_m$. For ease in formulation, these buses are grouped into two subsets: $\mathcal{N}_m^{int}$ being the internal buses of the cluster connected solely with buses within the same cluster, and the boundary buses $\mathcal{N}_m^{bo}$ that are neighboured at least by one boundary bus of another cluster (Fig. 2). The set of the external buses neighboring a cluster $m$ are denoted $\mathcal{N}_m^{nb}$. For each bus $i$, the set of its neighboring buses are denoted as $\mathcal{N}_i$.

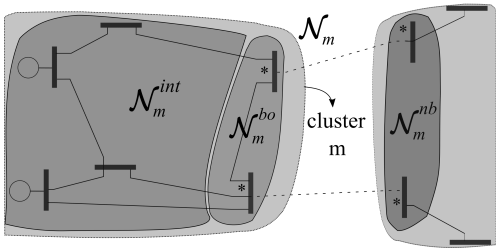


Fig. 2. Cluster topology.

Using these definitions, the cluster-based DC-OPF problem is formulated as follows:

$$\min_{\boldsymbol{p},\boldsymbol{\vartheta}} \left( \sum_{m \in \mathcal{M}} \left( \sum_{i \in \mathcal{N}_m} c_i p_i + q_i p_i^2 \right) \right) \quad \text{s.t.} \tag{2a}$$

Generator capacity constraints:

$$\forall m \in \mathcal{M},\ \forall i \in \mathcal{N}_m \qquad \underline{p_i} \leq p_i \leq \overline{p_i} \tag{2b}$$

Line capacity constraints:

$$\forall m \in \mathcal{M},\ \forall i \in \mathcal{N}_m,\ \forall j \in \mathcal{N}_i$$
$$-\overline{\tau}_{ij} \leq -b_{ij}(\vartheta_i - \vartheta_j) \leq \overline{\tau}_{ij} \tag{2c}$$

Nodal balance:

$$\forall m \in \mathcal{M},\ \forall i \in \mathcal{N}_m$$
$$\sum_{j \in \mathcal{N}_i \cap \mathcal{N}_m} -b_{ij}(\vartheta i - \vartheta j) + \sum_{j \in \mathcal{N}_i \cap \mathcal{N}_m^{nb}} -b_{ij}(\vartheta i - \vartheta j)$$
$$- p_i + d_i = 0 \tag{2d}$$

where the problem variables are the power generation at each bus $\boldsymbol{p} = \left[p_1, \ldots, p_{|\mathcal{N}|}\right]$ and the nodal phase angles $\boldsymbol{\vartheta} = \left[\vartheta_1, \ldots, \vartheta_{|\mathcal{N}|}\right]$. Capacity and susceptance of a given line connecting two buses $i$ and $j$ are denoted as $\overline{\tau}_{ij}$ and $b_{ij}$ respectively. At each node $i$, an active power demand of $d_i$ has to satisfied by local power generation and/or injection from neighboring buses. To each generator at a given bus $i$, linear and quadratic cost coefficients $p_i, q_i$ are assigned.

In this formulation, the second summation term on (2d) stands for the power injection at a boundary bus from neighboring clusters, and (2c) limits these flows by the respective line capacity. From the decomposition perspective, these two constraints create a coupling between these clusters; had the phase angles of the boundary buses been fixed, each cluster would obtain its independent subproblem. These thereby

constitute the *complicating variables* of the DC-OPF problem $\big(\boldsymbol{y} \equiv \boldsymbol{\vartheta}_{\Sigma}^{bo} = \big[\vartheta_i,\ i \in \bigcup_m \mathcal{N}_m^{bo}\big]\big)$.

In the following two sections, implementations of BD and ADMM for this problem will be introduced. Here, the phase angles of the internal buses of cluster $m$ are denoted as $\boldsymbol{\vartheta}_m^{int}$, the subset of $\boldsymbol{\vartheta}_{\Sigma}^{bo}$ belonging to a given cluster $m$ is denoted as $\boldsymbol{\vartheta}_m^{bo}$ and the variable space for the neighboring phase angles of a cluster $m$ is denoted as $\boldsymbol{\vartheta}_m^{nb}$.

### *B. Cluster-based solution of DC-OPF via Benders decomposition*

*1) Benders master problem:* The Benders master problem is formed by allocating the boundary bus angles as model variables as well as the estimation variables $\alpha_m$ corresponding to the costs arising from each cluster:

$$\min_{\boldsymbol{\vartheta}_{\Sigma}^{bo},\boldsymbol{\alpha}} \sum_{m\in\mathcal{M}} \alpha_m \qquad \text{s.t.} \tag{3a}$$

$$\forall m \in \mathcal{M},\ \forall i \in \mathcal{N}_m^{bo},\ \forall j \in \mathcal{N}_i \cap \mathcal{N}_m^{nb}$$
$$-\overline{\tau}_{ij} \le b_{ij}(\vartheta_i - \vartheta_j) \le \overline{\tau}_{ij} \tag{3b}$$

$$(\boldsymbol{\vartheta}_{\Sigma}^{bo}, \boldsymbol{\alpha}) \in \text{OC} \tag{3c}$$

where OC is the set of optimality cuts. The only set of physical constraints that are included in the master problem are the cross-cluster line capacity constraints (3b), as they do not include either of the internal variables $\boldsymbol{p}_m$ or $\boldsymbol{\vartheta}_m^{int}$. These, along with the optimality cuts (3c) generated from the subproblems, define the solution space of the master problem.

*2) Benders subproblems:* For each cluster $m \in \mathcal{M}$, subproblems of the following form are defined:

$$\min_{\boldsymbol{p}_m,\boldsymbol{\vartheta}_m^{int}} \left(\sum_{i\in\mathcal{N}_m} c_i p_i + q_i p_i^2\right) + M\sigma_m \qquad \text{s.t.} \tag{4a}$$

$$\forall i \in \mathcal{N}_m \qquad -p_i - \sigma_m \le \underline{p_i}, \qquad p_i - \sigma_m \le \overline{p_i} \tag{4b}$$

$$\forall i \in \mathcal{N}_m^{int},\ \forall j \in \mathcal{N}_i \cap \mathcal{N}_m^{int}$$
$$b_{ij}(\vartheta_i - \vartheta_j) - \sigma_m \le \overline{\tau}_{ij},\ -b_{ij}(\vartheta_i - \vartheta_j) - \sigma_m \le \overline{\tau}_{ij} \tag{4c}$$

$$\forall i \in \mathcal{N}_m^{bo},\ \forall j \in \mathcal{N}_i \cap \mathcal{N}_m^{int}$$
$$b_{ij}(\vartheta_i^* - \vartheta_j) - \sigma_m \le \overline{\tau}_{ij},\ -b_{ij}(\vartheta_i^* - \vartheta_j) - \sigma_m \le \overline{\tau}_{ij} \tag{4d}$$

$$\forall i \in \mathcal{N}_m$$
$$\sum_{\substack{j\in\mathcal{N}_i\cap\mathcal{N}_m^{int}\\ \text{if } i\in\mathcal{N}_m^{int}}} -b_{ij}(\vartheta i - \vartheta j) + \sum_{\substack{j\in\mathcal{N}_i\cap\mathcal{N}_m^{bo}\\ \text{if } i\in\mathcal{N}_m^{int}}} -b_{ij}(\vartheta i - \vartheta j^*)$$
$$+ \sum_{\substack{j\in\mathcal{N}_i\cap\mathcal{N}_m^{int}\\ \text{if } i\in\mathcal{N}_m^{bo}}} -b_{ij}(\vartheta i^* - \vartheta j) + \sum_{\substack{j\in\mathcal{N}_i\cap\mathcal{N}_m^{nb}\\ \text{if } i\in\mathcal{N}_m^{bo}}} -b_{ij}(\vartheta i^* - \vartheta j^*)$$
$$- p_i + d_i = 0 \tag{4e}$$

In each subproblem, the local generation costs within each cluster are allocated as the corresponding objective function to be minimized (6), and the phase angles at the boundary buses assume their fixed values obtained from the master problem (3) (denoted as $(\dots)^*$). Note that, in order to dispense with the generation of feasibility cuts (and thereby simplify the cut generation process), a *Big-M relaxation* is applied to the subproblems (as proposed in [15]). Herewith, the local objective functions and the inequality constraints are adjusted with a large parameter $M$ and the slack variables $\sigma_m$.

*3) Generation of the Benders cuts for DC-OPF:* For generating the optimality cuts, each constraint (4b)-(4e) is rewritten in the following form :

$$\mathbf{A}_m \boldsymbol{x}_m - (\mathbf{1})_m \sigma_m \le \underbrace{\boldsymbol{r}_m - \mathbf{E}_m \boldsymbol{y}_m^*}_{\text{RHS}_m^*}. \tag{5}$$

The right-hand-side vector ($\text{RHS}_m = \boldsymbol{r}_m - \mathbf{E}_m \boldsymbol{y}_m$) is then multiplied by the vector $\boldsymbol{\lambda}_m$ consisting of the dual variables of each constraint, and the optimality cut is given by the non-positivity condition of this product. These cuts are generated for all $m \in \mathcal{M}$ and added to the master problem before it is solved in the subsequent iteration.

### *C. Cluster-based solution of DC-OPF via Consensus ADMM*

ADMM in consensus form consists of solving a partially augmented Lagrangian using the dual ascent approach, where the primal and dual variables are sequentially calculated by fixing one another at each step.

*1) ADMM subproblems:* For each cluster $m \in \mathcal{M}$, the corresponding subproblems are solved using local copies of the boundarying and neighboring phase angles ($\boldsymbol{\vartheta}_m^{ex} := [\boldsymbol{\vartheta}_m^{bo}, \boldsymbol{\vartheta}_m^{nb}]$). These local copies are put into a consensus constraint to be equal to their global values $\vartheta_i^+$, and a partial augmented Lagrangian is set for the objective function using these consensus constraints $\vartheta_{m,i}^{ex} - \vartheta_i^+ = 0$:

$$\min_{\boldsymbol{p}_m,\boldsymbol{\vartheta}_m^{int},\boldsymbol{\vartheta}_m^{ex}} \left(\sum_{i\in\mathcal{N}_m} c_i p_i + q_i p_i^2\right) + \sum_{i\in\mathcal{N}_m^{bo}\cup\mathcal{N}_m^{nb}} \lambda_{m,i}(\vartheta_{m,i}^{ex} - \vartheta_i^+)$$
$$+ \sum_{i\in\mathcal{N}_m^{bo}\cup\mathcal{N}_m^{nb}} \frac{\rho_{m,i}}{2}(\vartheta_{m,i}^{ex} - \vartheta_i^+)^2 \ \text{ s.t. (2b) – (2d)} \tag{6}$$

where $\lambda_{m,i}$ and $\rho_{m,i}$ denote the dual variables and the penalty terms respectively, corresponding to the consensus constraint for subproblem $m$'s local copy of the phase angle at a bus $i$.

*2) Updating the global and dual variables, parameter tuning:* The primal optimizers $(\vartheta_{m,i}^{ex})^*$, after obtained by solving above problem, are used to update (a) the global variable $\vartheta_i^+$ ("averaging step"), followed by (b) the penalty terms $\rho_{m,i}$ and (c) the dual variables $\lambda_{m,i}$ associated with them:

$$(a)\ \forall i \in \mathcal{N}_m^{bo} \cup \mathcal{N}_m^{nb}$$
$$\vartheta_i^+(k+1) = \frac{1}{|\mathcal{M}_i|} \sum_{m\in\mathcal{M}_i} (\vartheta_{m,i}^{ex})^*(k), \tag{7}$$

$$(b)\ \forall m \in \mathcal{M},\ \forall i \in \mathcal{N}_m^{bo} \cup \mathcal{N}_m^{nb}$$
$$\rho_{m,i}(k+1) =$$
$$\begin{cases} \rho_{m,i}(k)\cdot(1+\tau) & \text{if } \mu \cdot r_{m,i}(k+1) > s_{m,i}(k+1), \\ \rho_{m,i}(k)\cdot(1+\tau)^{-1} & \text{if } \mu \cdot r_{m,i}(k+1) < s_{m,i}(k+1), \\ \rho_{m,i}(k) & \text{otherwise} \end{cases} \tag{8}$$

$(c)\ \forall m \in \mathcal{M},\ \forall i \in \mathcal{N}_m^{bo} \cup \mathcal{N}_m^{nb}$

$$\lambda_{m,i}(k+1) = \lambda_{m,i}(k) + \rho_{m,i}(k+1) \cdot (\vartheta_{m,i}^{ex} - \vartheta_i^{+}) \quad (9)$$

where $\mathcal{M}_i \subseteq \mathcal{M}$ are the set of clusters that either contain the bus $i$, or at least contain a bus that is connected with the bus $i$. $\mu$ and $\tau$ are the adaptive penalty parameters as introduced in [3]. $r_{m,i}$ and $s_{m,i}$ are the primal and dual residuals of $\vartheta_{m,i}$ respectively, which are calculated before step $(b)$ as follows:

$$r_{m,i}(k+1) = \left|\vartheta_{m,i}(k+1) - \vartheta_i^{+}(k+1)\right|, \quad (10)$$

$$s_{m,i}(k+1) = \rho_{m,i}(k) \cdot \left|\vartheta_{m,i}(k+1) - \vartheta_{m,i}(k)\right|. \quad (11)$$

### D. Stopping criterion

Since both decomposition methods are iterative, a uniform convergence criterion has to be defined, where the decentral solutions settle down to a point after a certain number of iterations and thereby are assumed to sufficiently approximate the original central solution. For ADMM, primal and dual residuals are typically used for checking convergence. Since Benders decomposition does not have a concept of local vs. global values which are involved by the primal residuals, the dual residuals (without the multiplied $\rho$ since it does not apply to BD) are selected as the indicator of the convergence:

$$\boldsymbol{s}^2(k) := \frac{\left\|(\boldsymbol{\vartheta}_{\Sigma}^{bo})^*(k) - (\boldsymbol{\vartheta}_{\Sigma}^{bo})^*(k-1)\right\|_2^2}{|\mathcal{N}|} \leq \epsilon, \quad (12)$$

where $\epsilon$ is a small positive real number. Note that the residual is normalized by the total number of the buses in the network, in order to make up for the accumulation of residuals for larger test cases.

## IV. Comparative study

### A. Organizational and communicational comparison

The two methods are clearly distinguished in their organization and communication; BD necessitates a hierarchical architecture whereas a flat organization can be deployed for ADMM. Figure 3 gives an overview on the problem hierarchy and information exchange between the subproblems for both methods. For BD, the presence of a master problem that requires a collection of the cuts from each subproblem prevent a completely distributed architecture. For consensus-based ADMM, the averaging step can be performed independently for each complicating variable, with the information exchange only taking place between the neighboring clusters where those variables appear. This allows for a distributed paradigm and does away with the requirement of a central entity.

### B. Computational evaluation using IEEE networks

In this section, convergence and computational properties of BD and ADMM is compared, where IEEE networks of 9, 14, 30, 39, 118 and a 62-bus network depicting the Indian utility system are used. The data for the IEEE cases are retrieved from MATPOWER [16], whereas the data for the 62-bus case is extracted from [17]. Furthermore, the following modifications were made: i) line capacities are added for IEEE 14-bus and IEEE 118-bus cases ([18] for the latter) as the original

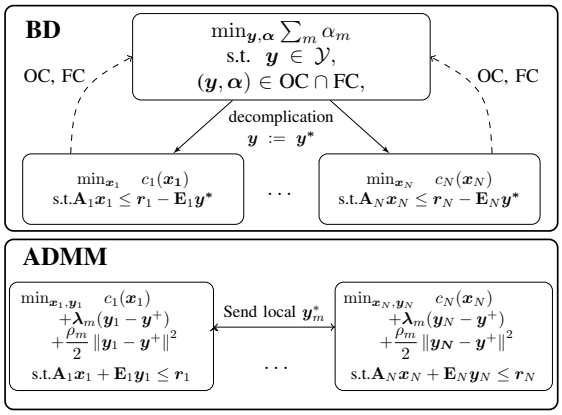


Fig. 3. Problem hierarchy and information flows for BD (above) and ADMM (below).

data sets did not include any and ii) the quadratic and linear cost coefficients are adjusted for IEEE 39-bus and IEEE 118-bus cases respectively so that these cases have single global optimum. Comparison is done along two scaling aspects: i) increasing number of buses from the 9-bus case up to 118-bus, each divided in two cluster each and ii) increasing number of clusters using the 118-bus network as the base case. The optimization problems are built by MATLAB Problem-Based Optimization and the subproblems are solved in parallel using the Parallel Computing Toolbox on a workstation with Intel(R) Xeon(R) 2.4GHz, 16 cores processor (up to 6 of which were utilized in parallel) with 128 GB RAM. As solver, the built-in quadratic programming solver of MATLAB *quadprog* is used. For a fair benchmarking, all cases were run five times and the resultant average times are considered. For BD algorithm, a variable value of $(\$3) \times |\mathcal{N}|$ is set by inspection for the big-M term, which was observed to have good convergence rate while respecting all the inequality constraints at the same time. For the ADMM algorithm, adaptive penalty parameters of $\tau = 0.1$ and $\mu = 10$ are set [3]. For the stopping criterion, a threshold value of $\epsilon = 10^{-5}$ is selected, which corresponds to an mean absolute error of $10^{-5}$ in radians over each bus.

*1) Convergence properties for varying network sizes:* Figure 4 shows the convergence behavior of the nodal power generations for both methods, for the selected test cases ranging from 9 buses up to 118 buses, all with two clusters.

For each of these test cases, both methods converge to the solution of the original problem. To settle below the same error threshold, ADMM was observed to require more iterations than BD for all of the cases. This could be attributed to each BD iteration keeping the history of the algorithm through the Benders cuts that accumulate in the master problem, leading to overall more efficient iterations. On the other hand, ADMM's iteration steps are distinguished from one another only by the current global values of the coupling variables and the dual variables/penalty factors of their consensus constraints.

*2) Convergence properties with varying inter-cluster line contingencies:* Additional to the network size, the inter-cluster line capacities is expected to have an influence on the conver-

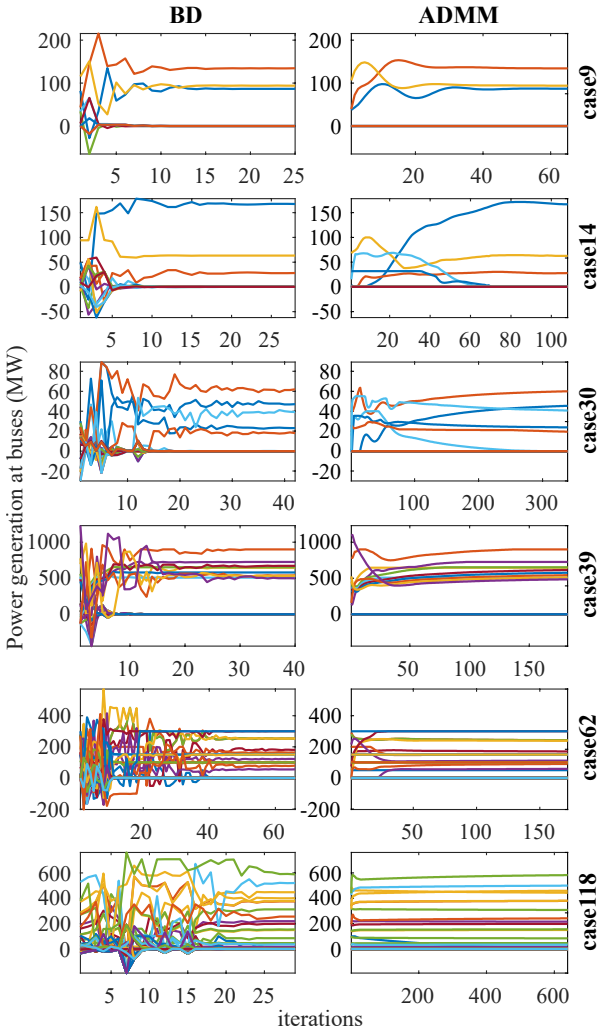


Fig. 4. Convergence profiles for 9- to 118-bus test cases, each network clustered into two.

gence. To analyse this, the IEEE-118 test case is divided in 2, 3, 4, and 6 clusters. Then, for each clustering formation, the line capacities between two neighboring clusters are adjusted by a scalar $\eta \in \{0.25, 0.5, 1, 2, 5\}$ in order to provide a wide range of instances between possible infeasibility and non-contingency. For ADMM, the four and six cluster cases with $\eta = 2$ exhibit non-convergence and are therefore not illustrated on the figure.

Figure 5 shows the dual residuals for all test scenarios. Following observations are made: first, increasing number of clusters has led to a lower number of required iteration steps for ADMM, whereas this is not the case for BD. It is observed that BD leads to long-term linear residual profiles on the log-scale, indicating an exponential decay. In ADMM, larger line capacities between the clusters allow to higher oscillation amplitudes of the power injections across these lines. This has an adverse effect on the number of iterations needed for these oscillation to settle, and hence the time for the algorithm to converge. This behavior is also partially present for BD; the case with $\eta = 0.25$ mostly led to the least number of iterations. However, the effect is not as direct as for ADMM, since the subproblems assume fixed values for the boundary phase angles (and hence constant power injections over the inter-cluster lines). A notable case has been ADMM with two clusters, where the given range of line capacities had no influence on the convergence profile whatsoever; as no contingency was present across the adjusted lines.

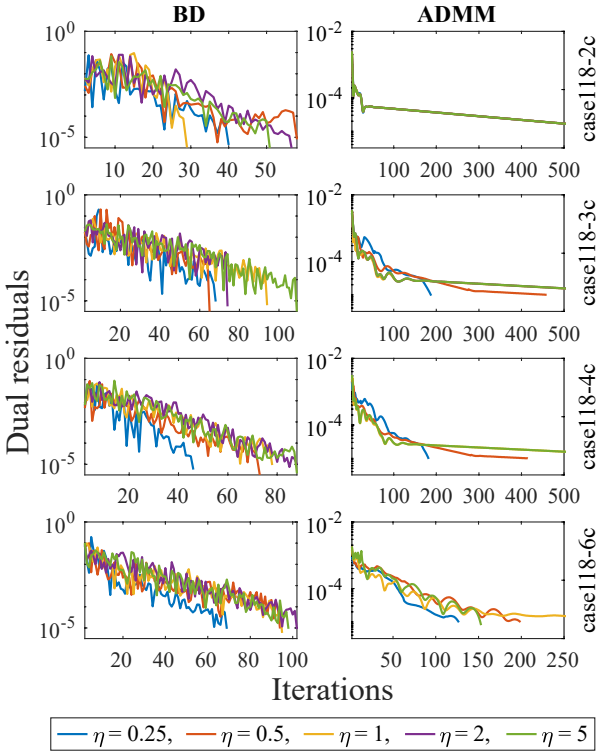


Fig. 5. Dual residuals of 118-bus test case with varying number of clusters and inter-cluster line capacity multipliers $\eta$.

*3) Runtime considerations:* Besides the required number of iterations, the actual computation time elapsed until convergence holds high practical importance. Figure 6 shows a comparison of the runtimes for both algorithms, across all network cases and line capacity multipliers. Reflecting the overall larger number of iterations for ADMM (as demonstrated at the previous section), BD shows overall better computational performance. Performance improvement by increasing the number of clusters was not achieved, which demonstrates the significance of the parallel overhead in these implementations. In addition to the total runtime, Figure 7 shows the evolution of the time elapsed between each iteration until convergence. Here, a qualitative difference is present: the ADMM iterations take a relatively constant amount of time in every step, whereas the time intensity of the iteration steps for BD increase roughly linearly. This is largely caused by the increasing size of the master problem in BD via the addition of Benders cuts at each step. This indicates that the time intensity of solving the master problem starts to dominate over the subproblems, which reduces the overall benefit of parallelizing the algorithm, as

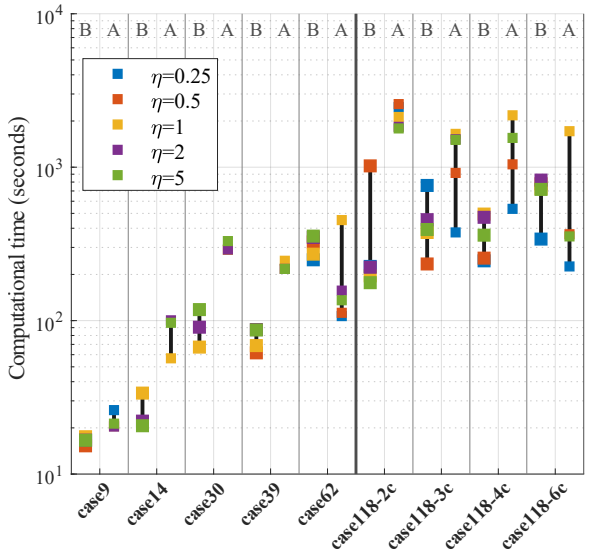


Fig. 6. Computation times resulting from parallelized Benders decomposition (B) and ADMM (A) algorithms, run over each test case with varying inter-cluster line capacity multipliers $\eta$.

the master problem has to be solved sequentially.

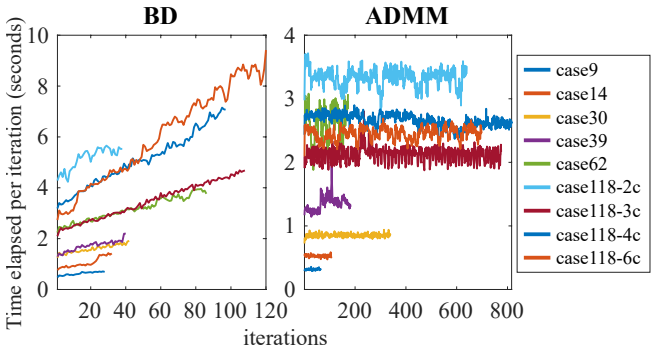


Fig. 7. Comparison of the time elapsed per iteration.

## V. Conclusions

In this work, a qualitative and computational comparison was done for decentrally solving DC-OPF problems using Benders decomposition and ADMM. While BD had a relative advantage in terms of computational performance, ADMM has been shown to be the truly distributed method that does not require a central optimizer. Furthermore, it was shown that the runtime inferiority of ADMM can be remedied if partitioning is done between weakly coupled clusters characterized by low interconnector capacities. To this end, optimal partitioning algorithms may prove useful for accelerating these algorithms. If performance improvement is desired, measures for optimizing the parallelization architecture also need to be taken, so that the parallel and communicational overheads are minimized.

## VI. Acknowledgments

This work was supported by the Federal Ministry for Economic Affairs and Energy of Germany (BMWi) [project eXtremOS, grant number 03ET4062C]. This work was also partially supported by the Singapore National Research Foundation under its Campus for Research Excellence And Technological Enterprise (CREATE) programme.

## References

[1] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar, and R. Baldick, "Toward distributed/decentralized DC optimal power flow implementation in future electric power systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2574–2594, July 2018.

[2] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov 2017.

[3] S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers*, ser. Foundations and Trends in Machine Learning Series. Now Publishers Incorporated, 2011.

[4] T. Erseghe, "Distributed optimal power flow using ADMM," *IEEE Transactions on Power Systems*, vol. 2015, no. 5, pp. 2370–2380, Sept 2014.

[5] K. Zhang, S. Hanif, and D. Recalde, "Clustering-based decentralized optimization approaches for DC optimal power flow," in *2017 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, Dec 2017, pp. 1–6.

[6] A. Engelmann, T. Mühlpfordt, Y. Jiang, B. Houska, and T. Faulwasser, "Distributed ac optimal power flow using ALADIN," *20th IFAC World Congress*, vol. 50, no. 1, pp. 5536 – 5541, 2017.

[7] K. Zhang, S. Hanif, C. M. Hackl, and T. Hamacher, "A framework for multi-regional real-time pricing in distribution grids," *IEEE Transactions on Smart Grid*, pp. 1–1, 2019.

[8] J. Mohammadi, G. Hug, and S. Kar, "Fully distributed DC-OPF approach for power flow control," in *2015 IEEE Power Energy Society General Meeting*, July 2015, pp. 1–5.

[9] M. Caramanis, E. Ntakou, W. W. Hogan, A. Chakrabortty, and J. Schoene, "Co-Optimization of Power and Reserves in Dynamic T&D Power Markets With Nondispatchable Renewable Generation and Distributed Energy Resources," *Proceedings of the IEEE*, vol. 104, no. 4, pp. 807–836, April 2016.

[10] J. F. Benders, "Partitioning procedures for solving mixed-variables programming problems," *Numer. Math.*, vol. 4, no. 1, pp. 238–252, Dec. 1962.

[11] A. M. Geoffrion, "Generalized Benders decomposition," *Journal of Optimization Theory and Applications*, vol. 10, no. 4, pp. 237–260, Oct 1972.

[12] R. Rahmaniani, T. G. Crainic, M. Gendreau, and W. Rei, "The Benders decomposition algorithm: A literature review," *European Journal of Operational Research*, vol. 259, no. 3, pp. 801 – 817, 2017.

[13] C. Lin, W. Wu, B. Zhang, B. Wang, W. Zheng, and Z. Li, "Decentralized reactive power optimization method for transmission and distribution networks accommodating large-scale dg integration," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 1, pp. 363–373, Jan 2017.

[14] Z. Li, W. Wu, B. Zhang, and B. Wang, "Decentralized multi-area dynamic economic dispatch using modified generalized Benders decomposition," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 526–538, Jan 2016.

[15] H. D. Sherali and B. J. Lunday, "On generating maximal nondominated Benders cuts," *Annals of Operations Research*, vol. 210, no. 1, pp. 57–72, Nov 2013.

[16] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb 2011.

[17] B. Krishnakumar and R. Arthi, "Contingency assessment of Indian utility 62 bus system," in *2012 IEEE International Conference on Engineering Education: Innovative Practices and Future Trends (AICERA)*, July 2012, pp. 1–8.

[18] Electrical and Computer Engineering Department, Illinois Institute of Technology. [Online]. Available: http://motor.ece.iit.edu/data/JEAS_IEEE118.doc