This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Scalable Economic Dispatch for Smart Distribution Networks

Elisabeth Kellerer and Florian Steinke

***Abstract*—We present a novel algorithm for economic dispatch in electric power grids. The method is inspired by statistical inference methods. Using discretized optimization variables, our algorithm finds the globally optimal, single time-step dispatch assignment for radial grids in linear time with respect to the number of network nodes. For such problems, the algorithm outperforms state-of-the-art mixed-integer scheduling, both in run-time and in the allowed complexity of component and line models. Moreover, the necessary computations can be performed in a distributed fashion, facilitating both practical implementation as well as information privacy. Our algorithm is thus optimally suited for the very large dispatch problems that will arise in future smart distribution grids with hosts of small, decentralized, and flexibly controllable prosumers, i.e., entities able to consume and produce electricity.**

***Index Terms*—Distributed algorithms, economic dispatch, graphical models.**

I. Introduction

ELECTRICITY generation is undergoing fundamental changes. In the past, a few large power stations have centrally supplied all demand. The wide-spread introduction of new technologies is changing this paradigm now: renewable energy, like photovoltaic and biogas, as well as decentral combined heat and power (CHP) generation, is typically small in size and distributed throughout electrical distribution grids. Moreover, intelligent prosumers, such as electric cars or stationary batteries, are flexible with regard to their consumption or production. The classic economic dispatch problem, i.e., determining a welfare maximizing assignment of who exactly should produce or consume how much electricity at each point in time (see e.g., the review [1]), thus receives renewed interest. Novel algorithmic requirements are the scalability to very large networks and the possibility to work in multi-owner environments where component models and cost functions comprise important business secrets and thus cannot be shared freely with a central coordinating dispatch controller.

Current commercial dispatch systems use almost exclusively the mixed integer linear programming (MILP) approach to economic dispatch, see, e.g., [2]. This formulation allows for moderate model complexity, while providing good solver performance control and scalability to medium-sized problems. However, it requires white box models, i.e., the dispatch system has to have full knowledge of all component models. This may be difficult in competitive, multi-owner environments. Moreover, even highly optimized solver packages like CPLEX[1] do not scale well to very large networks as our experiments show. For these reasons, market-based dispatch approaches have become more popular recently. For example, in the SO-EASY [3] and PowerMatcher [4] systems, intelligent local agents bid on a central market place whose market clearing then determines the dispatch result. While this approach is extremely scalable and suits a competitive environment well, it does not achieve a globally optimal dispatch result, even if the local agent models are very detailed and are solved in an optimal way, e.g., by employing dynamic programming approaches as in [5].

In this work we present a novel approach to economic dispatch that is scalable, well-suited for multi-owner environments, and finds the globally optimal solution for radial distribution networks and discrete optimization variables. The idea is to formulate the dispatch problem as an optimization problem on a graph and then to employ graphical model inference methods originating in the statistical learning community, see [6] for a good overview. These algorithms are designed to optimally exploit the graph structure of certain statistical estimation problems. We will show how to apply the same algorithms to a graph-oriented formulation of the economic dispatch problem.

Using probabilistic graphical model methods in power grids has been proposed before for state estimation applications [7], [8]. A one-to-one matching of customers and generators is computed using graphical model algorithms in [9]. Our novel contribution in this paper is to show how the graphical model approach can be used to solve economic dispatch problems. Furthermore, we discuss the properties of our approach in detail and describe several important steps facilitating an efficient implementation with standard tools from the probabilistic graphical model community. We demonstrate the algorithm's behavior for an exemplary future smart grid and show its optimal computational scaling for very large networks. In this performance test we beat the state of the art MILP solver CPLEX by orders of magnitude.

Manuscript received November 11, 2013; revised December 17, 2013, April 28, 2014, and July 17, 2014; accepted September 07, 2014. Paper no. TPWRS-01350-2013.

The authors are with Siemens Corporate Technology, 81739 München, Germany (e-mail: Anna.Kellerer.Ext@siemens.com; Florian.Steinke@siemens.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2014.2358375

[1]www.cplex.com.

Exploiting the graph structure for optimal power flow problem (OPF) has recently been discussed extensively in [10]–[12]. A convex conic-programming relaxation of the power flow equations [13], [14] is used for an OPF computation with convex costs in [10]. The results are then shown to be equivalent to the optimal solution of the unrelaxed problem for trees. Qualitatively similar steps but with different formulae are undertaken by [12] and [11], where [12] only requires monotonicity of cost functions instead of convexity. In contrast, our work does not make any restrictions on the shape of the cost functions, but requires discretization of the optimization variables. In [10], the resulting optimization is decomposed in line with the maximal cliques of a triangulated graph and linear run-time with respect to the number of such cliques is shown. Exactly same ideas underlie the Junction-Tree algorithm in the graphical model community, see, e.g., [15], that presents a generalization of the belief propagation algorithm employed in this paper. Whereas an iterative optimization routine exchanging simple Lagrange multipliers is used [10], our proposed approach exchanges more complex messages (i.e., functions of a few parameters), but only requires one iteration.

The remainder of the paper is structured as follows. In Section II we provide a formal description of the economic dispatch problem and describe how to formulate it as a graph problem. In Section III, we briefly review probabilistic graphical model theory, which is applied to the dispatch problem in Section IV. In this section, we also present an in-depth discussion of the proposed approach's properties. We describe details of our implementation and present the results of our experiments in Section V. We conclude in Section VI.

II. Economic Dispatch as a Graph Problem

The economic dispatch problem is to schedule a number of electricity generation or consumption units such that the overall welfare, i.e., the summed utility of all consumption minus the summed cost of all production, is maximized. This is to be done subject to generation and transportation constraints.

For an electric grid with buses/nodes V and lines/edges E, the one time step dispatch problem formally reads

$$\min_{P_v \in \mathcal{P}_v} \sum_{v \in \mathrm{V}} C_v(P_v), \tag{1}$$

$$\text{s.t.} \sum_{v \in \mathrm{V}} P_v = 0. \tag{2}$$

Here, P_v denotes the net power injection of node $v, C_v(P_v)$ the cost or negative utility function associated with it, and $\mathcal{P}_v$ the feasibility set for P_v. For a generator node a typical (non-convex) $\mathcal{P}_v$ would be $\mathcal{P}_v = \{0\} \cup [P_{\min}, P_{\max}]$, i.e., a generator with minimum and maximum load if switched on.

The power balance constraint (2) is *global* in the sense that it couples all decision variables P_v in one equation. The problem can therefore not trivially be decomposed. Equivalent to (2), however, is a set of power flow equations, e.g.,

$$\sum_{w \neq v} P_{v,w} = P_v \tag{3}$$

where $P_{v,w}$ is the power transport from v to w and $P_{w,v} = -P_{v,w}$. If this equation is satisfied at each node, no energy is lost in the network and the overall power balance constraint (2) is satisfied.

Note that using formulation (3), the optimization problem (1), (3) is *local* in the sense that each term of the objective and the constraints only contains variables from a neighborhood of v in the graph. Such a *graph problem* formulation of economic dispatch can be tackled by working locally on the graph, trying to find optimal local assignments while guaranteeing local consistency between neighboring sites that share common $P_{v,w}$ variables. In some sense, that is exactly what the proposed graphical model methods presented in the next section does. The rewriting of (2) into (3) thus is, while almost trivial, important for the proposed solution strategy.

The notation above, which we will also use for the remainder of the paper, hints at using the $P_{v,w}$ as optimization variables directly, the most simplified form of a power flow model with local active power conservation only. However, the full AC power flow equations with $S_{v,w} = U_v Y^*_{v,w}(U_v - U_w e^{-j\theta_{v,w}})$ could also be employed, using the voltage magnitudes U_v and phase angle differences $\theta_{v,w}$ as optimization variables. Here, complex valued power transports are denoted by $S_{v,w}$, line admittances by $Y_{v,w}$, and costs functions may be dependent on complex node power injections $S_v = \sum_w S_{v,w}$. Such a formulation would also preserve the locality property as defined above and thus would be amenable to the solution formalism proposed below. Only the number of real variables in each neighborhood would increase when using the three variables $\theta_{v,w}, U_v, U_w$ instead of one $P_{v,w}$ per line.

Using local power flow equations (3) in replacement of the global constraint (2) also easily allows us to define limits on power transport, e.g., $-T_{v,w} \leq P_{v,w} \leq T_{v,w}$ where $T_{v,w}$ is the line capacity.

III. Short Introduction to Probabilistic Graphical Models

A probabilistic graphical model describes a family of multivariate probability distributions that share a common (conditional) independence structure. The graphical representation is used to develop efficient graph-based algorithms for various statistical inference computations such as independence testing, marginalization or finding the most likely variable assignment. A recent introduction is given in [6].

Graphical models come in various flavors. We focus on undirected graphical models here. For these, the graph (V, E) represents all probability distributions $p(x_{\mathrm{V}})$ over random variables $x_v, v \in V$, that factor as

$$p(x_{\mathrm{V}}) = \frac{1}{Z} \prod_C \phi_C(x_C). \tag{4}$$

Here, x_A for set A denotes the set of all $x_v, v \in A$. C are cliques in the graph, i.e., fully connected subsets, and the factors ϕ_C are non-negative functions over the variables x_C. Z denotes the necessary normalization factor.

Finding the maximum probability assignment, also known as *decoding*, means determining

$$x^*_{\mathrm{V}} = \operatorname*{argmax}_{x_{\mathrm{V}}} \; p(x_{\mathrm{V}}).$$

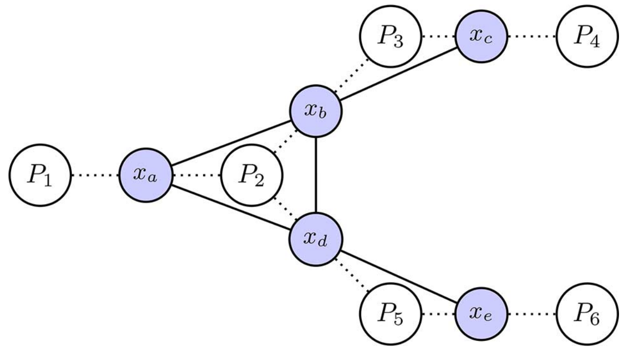


Fig. 1. Undirected graphical model with 5 nodes $x_a, \dots, x_e$ and three maximal cliques (solid lines, filled circles). Overlayed is shown an electrical grid with 6 power in-feeds $P_1, \dots, P_6$. Electric lines (dotted) are associated with the graphical model variables that correspond to power flows $P_{v,w}$.

For this inference task the graph structure can be exploited as the following example shows. The graphical model depicted in Fig. 1 contains three maximal cliques, namely $\{x_a, x_b, x_d\}$, $\{x_b, x_c\}$ and $\{x_d, x_e\}$. Computing the maximum probability assignment we obtain

$$\begin{aligned}\max_{x_a,\dots,x_e} & \frac{1}{Z}\phi(x_a,x_b,x_d)\phi(x_b,x_c)\phi(x_e,x_d) \\ &= \max_{\substack{x_a\\x_b\\x_d}} \frac{1}{Z}\phi(x_a,x_b,x_d) \\ &\quad\times \underbrace{\left(\max_{x_c}\phi(x_b,x_c)\right)}_{=m_{cb}(x_b)}\underbrace{\left(\max_{x_e}\phi(x_e,x_d)\right)}_{=m_{ed}(x_d)} \end{aligned} \tag{5}$$

where $m_{cb}(x_b), m_{ed}(x_d)$ are function-valued *messages* that can be computed independently.

In this paper we generally assume the variables x_v to take values in discrete sets. The messages are thus vectors of the size of these sets. The example then shows how one 5-D maximization problem is split into two sets of 2-D and one 3-D maximization problem by exploiting commutativity. Since the optimization effort for discrete variables is exponential in the dimension of the optimization problem (using full enumeration), this rewriting amounts to a tremendous reduction in computation time—especially if this trick is repeated again and again for larger graphs.

The same argument qualitatively also holds for continuous variables x_v. However, the messages then contain infinitely many values that can neither be fully computed nor be stored. *Approximate inference* algorithms for graphical models then try to approximate those messages with standard function classes, see [15] for an introduction.

A major limitation is the structure of the underlying graph: if it contains loops, the maximization operators cannot be shifted inside the formula, since variables to the right hand side of the formula have influence on variables on the left hand side. This is why the work presented in this paper is focused on tree-structured/radial networks, that do not contain loops.

The example shows the basis for deriving a general algorithm, known as *belief propagation*, for decoding in tree-structured graphs. In a first stage, the *messages* $m_{ij}(x_{S_{ij}})$ are computed for every pair of overlapping cliques C_i and C_j, i.e., cliques where the separator set $S_{ij} = C_i \cap C_j$ is not empty. The messages are computed as

$$m_{ji}(x_{S_{ij}}) = \max_{x_{C_j \setminus S_{ij}}} \phi(x_{C_j}) \prod_{k \neq i} m_{kj}\left(x_{S_{kj}}\right). \tag{6}$$

The optimal values for variables contained in one of the separator sets S_{ij} are then computed as

$$x^*_{S_{ij}} = \operatorname*{argmax}_{x_{S_{ij}}} m_{ij}\left(x_{S_{ij}}\right) m_{ji}(x_{S_{ij}}). \tag{7}$$

Computing the optimal values for variables that are not contained in one of the S_{ij} is described in [6, Ch. 10]. Message propagation starts at a leaf node, where (6) reduces to maximization over ϕ only, and then continues along the edges where all messages $m_{kj}(x_{S_{kj}})$, $k \neq i$ are known. A feasible ordering can always be found for trees, e.g., by depth-first-search. Examples and implementation tricks are explained in [6].

IV. Economic Dispatch With Graphical Models

We now show how to apply the graphical model framework to economic dispatch problems in electrical networks. Matching our introduction of graphical models we assume discrete variables $P_{v,w}$ in this paper.

We first note that due to monotonicity $\operatorname{argmax} \prod_c \phi_c = \operatorname{argmin} \sum_c - \log \phi_c$. Moreover, we can re-write problem (1), (3) as

$$\min_{P_{v,w} \in [-T_{v,w}, T_{v,w}]} \sum_{v \in V} C_v \left(\sum_{w \neq v} P_{v,w} \right) \tag{8}$$

where we have included the domain constraints $P_v \in \mathcal{P}_v$ into the cost function $C_v(P_v)$, by setting it to infinity if $P_v \notin \mathcal{P}_v$.

We then construct an undirected graphical model from the electric network as demonstrated in Fig. 1. We create a node for each transport variable $P_{v,w}$ and add edges between these nodes, iff the corresponding electric lines connect to the same electric bus. Buses in the electric network then match to cliques in the graphical model and we define the clique potentials ϕ_C by identifying the cost function of the electrical node, $C_v(\sum_{w \neq v} P_{v,w})$, with the negative logarithm of the clique's potential function in the graphical model $-\log \phi_C$.

With these ingredients the following statement holds.

Proposition 1: The decoding problem on the constructed undirected probabilistic graphical model is equivalent to the economic dispatch problem in the original electric network.

This equivalence implies that all methods for decoding of probabilistic graphical models can also be applied to economic dispatch. For belief propagation, specifically, the message computation (6) becomes

$$m_{vw}(P_{v,w}) = \min_{P_{u,v}, u \neq w} C_v\left(\sum P_{u,v}\right) + \sum_{u \neq w} m_{uv}(P_{u,v}) \tag{9}$$

and (7) then reads

$$P^*_{v,w} = \operatorname*{arg\,min}_{P_{v,w}} m_{vw}(P_{v,w}) + m_{wv}(P_{v,w}). \tag{10}$$

An implementation with standard graphical model inference codes is discussed in the experimental section. Using the full AC power flow equations leads to messages $m_{vw}(U_v, U_w, \theta_{v,w}) = \tilde{m}_{vw}(S_{v,w}(U_v, U_w, \theta_{v,w}))$. The optimization problems in (9) and (10) are then higher-dimensional, but still local.

A. Properties of the Proposed Approach

The equivalence of economic dispatch with a suitably constructed graphical model and the possibility to apply belief propagation has several important consequences for economic dispatch in smart distribution networks. They concern the optimality of results, the linear run-time of computations, and the information privacy principles inherent in the method.

Belief propagation is known to find maximum probability assignments on trees—for discrete variables. Applying this algorithm to economic dispatch will thus result in the **globally optimal** assignment, if the electrical network is radial, i.e., is a tree. Unlike most optimization algorithms, the cost functions and domain constraints may take any shape. Only the grid's structure is important. Low and medium voltage networks are often tree-shaped in practice. Low-voltage feeders tend to be composed of several strings connected to one transformer. Medium voltage lines are often built as rings, but during normal operation, which is of major interest to economic dispatch, the loops are typically split by an opened breaker.

Concerning the run-time of this algorithm, note that the number of messages, and thus the number of message computations, is exactly twice the number of overlaps/links between cliques. If the maximum number of values for any $P_{v,w}$ is k and there are n electrical buses in the system, then the total run-time of our algorithm is $O(2nk^d)$, using full search for the optimizations in (9) and denoting the maximal clique size of the graphical model (equal to the degree of the electrical network) by d. The **total run-time scales linearly** in the size of the network, which renders our proposed algorithm suitable for very large networks. Moreover, the run-time is precisely predictable and independent of the shape of cost and constraint functions.

A reduction of the exponential scaling factor k^d for high-degree nodes in the electrical network to at most k^3 can be achieved by re-writing the electrical network without changing its physical meaning—iteratively split nodes with a degree larger than three into two new nodes that are connected with an appropriate capacity, loss-less line. The neighborhood connections are distributed among the two new nodes, thus reducing the original degree of the node to roughly one half. This procedure is repeated until all nodes in the graph have maximum degree three. The price to pay for this rewriting is a few additional edges/message computations. However, the computational effort is constant per edge, whereas the reduction of the neighborhood dimension leads to an exponential cost decrease.

Solving economic dispatch via belief propagation also means that the component models do not have to be exchanged if the algorithm is implemented in a distributed fashion. The only exchanged information are the messages. Since these messages contain highly aggregated information from several users, individual cost functions typically cannot be reconstructed from them. This built-in **information hiding** principle is very helpful for competitive, multi-owner environments, where the true cost structure of a participant may comprise an important business secret.

To what degree it is possible to reconstruct the individual cost functions from the transmitted messages strongly depends on 1) the prior knowledge of the shape of the true cost functions and 2) the number and position of messages that are used for reconstruction. A complete characterization under which conditions which knowledge can be obtained goes beyond the scope of this paper. We only discuss a few important situations here. First, the message from a single leaf node is the cost function of that node. However, the second message up the hierarchy of the tree is already a mixture of at least two cost functions. The minimum operator in the message computation (9) is not one-to-one, meaning that only the lowest cost for each value $P_{v,w}$ is visible in the message and the information about the higher "bids" is lost. Knowing the potential shapes of cost functions with the only uncertainty being a scaling factor, one can potentially reconstruct the individual cost functions from this partial information. The assignment of each cost function to a particular node, however, would still not be clear. In a realistic scenario with limited prior knowledge one can thus assume that from the messages of a larger sub-graph no exact reconstruction of each node's contribution is possible.

The proposed algorithm is, in fact, equivalent to the *dynamic programming* approach to economic dispatch when it is applied to a string-like graph extending over several time steps for one generator, not different locations for one time step as discussed in this paper. The proposed framework thus forms an extension of the dynamic programming approach to economic dispatch for tree-like graphs. Apart from placing the novel framework within a well-known economic dispatch environment, this observation has an important practical consequence. Having computed all messages one can locally deduce **finite step-size marginal costs**, i.e., the global cost increase if the demand at the given node increases by a *finite* amount—all other things being equal. For (mixed integer) linear programming the optimal dual variables of the local power flow equation yield only the differential marginal costs, i.e., the total cost changes for *infinitesimal* changes in local demand.

This result can be understood via the similarity between belief propagation and dynamic programming. Our messages are nothing but the *value functions* transferred in dynamic programming. The values $m_{vw}(P_{v,w})$ thus express exactly the minimum cost that would be accumulated in the sub-tree behind edge vw, if $P_{v,w}$ was transported over that edge. This view point also explains formula (10), where the minimum cost of the left half-tree is added to the minimum cost of the right half-tree to obtain the globally optimal decision (note the similarity to the forward-backward pass in dynamic programming). Interpreting the messages as value functions, the finite step-size marginal costs can be computed by deciding locally at each node for each additional demanded unit from which source it should be taken, either via transport from one of the attached sub-networks or from a local production. For this computation only the messages of the local neighborhood and the node's own cost function have to be considered, while still obtaining the true global finite step-size cost-changes.

One final remark concerns the consistency of the solution if the minimizer of the optimization problem is not unique. Since the decisions for the individual $P^*_{v,w}$ in (10) are taken locally and independently from each other, a non-consistent selection of one of the solution sets of optimal $P_{v,w}$ may result in this case. This may then lead to non-optimal or even infeasible node injections P_v. A simple example is a chain of three nodes where each of the end nodes could supply a load in the middle at identical cost. Independent decisions on each of the two edges could then lead to the situation where the load in the middle is covered twice or not at all. To avoid such problems we propose to ensure solution uniqueness by breaking potential symmetries, e.g., via adding small random terms onto otherwise identical cost functions.

V. Experiments

We test our proposed algorithmic framework with two experiments. We first compare the computational performance to MILP for sampled distribution networks that can be scaled to very large size. Second, we examine an exemplary smart distribution network of the future in closer detail.

A. Implementation

In our first example, the line capacities are all equal and their size is on the same order of magnitude as the generators' and consumers' power intake. We thus use an equal, equidistant discretization on all lines. In the second example, the line capacities are very different from each other and relatively large in comparison to the loads and generators. In this case we proceed iteratively. We use a fixed number of discretization points for each line and start with a large line-capacity dependent discretization step-size at first. We then determine a provisional solution with one run of belief propagation where the power balance at each node is assumed to hold up to discretization errors only. Afterwards, we re-discretize around the initial solution with a finer grid and rerun the algorithm. This iterative procedure increases the run-time of the algorithm by a small factor and is not guaranteed anymore to find the globally optimal solution, as claimed in Section IV. However, we have found it to work well in practice as our experiments in Section V-C show. It allows us to tackle difficult optimization problems with variables with largely different value ranges.

Our implementation uses the UGM toolbox [16] with only slight modifications. This allows us to apply tested standard algorithms for all our purposes. Since we are only interested in costs, i.e., log probability functions, we replace the max-product implementation (6) with the min-sum equations (9) for improved numerical stability. Moreover, since the UGM toolbox only supports pairwise potentials, we rewrite the graphical model slightly as explained in [17, Appendix].

As a baseline we implement the MILP formulation of [2]. Our implementation is written using the GAMS programming system with either the CPLEX or the SCIP solver. CPLEX is a highly optimized commercial MILP solver whereas SCIP represents one of the best available open source implementations. We built piecewise linear cost functions for the MILP formulation using the same discretization values as for the graphical model approach. While convex generator cost functions with a minimum load require only one binary variable per generator, an additional binary variable is necessary for each discretization interval for non-convex functions. The additional binary variables indicate whether the previous interval is fully used. Moreover, since we consider problems with transmission constraints in this paper, we slightly extend [2] with additional linear constraints on the transported power.

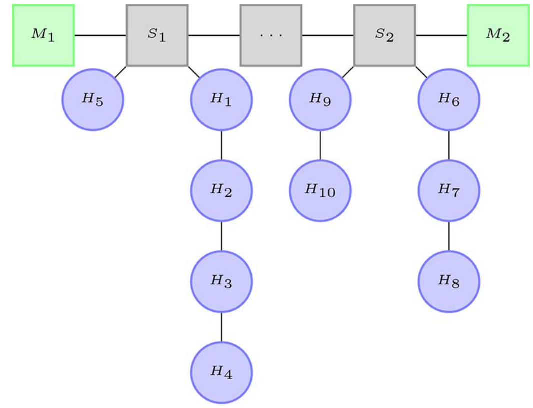


Fig. 2. Test system for scaling experiments: Two medium voltage transformers M_1 and M_2 feed a medium voltage ring of differing size. Radial low voltage lines with households H_i dissect at busbar nodes S_i.

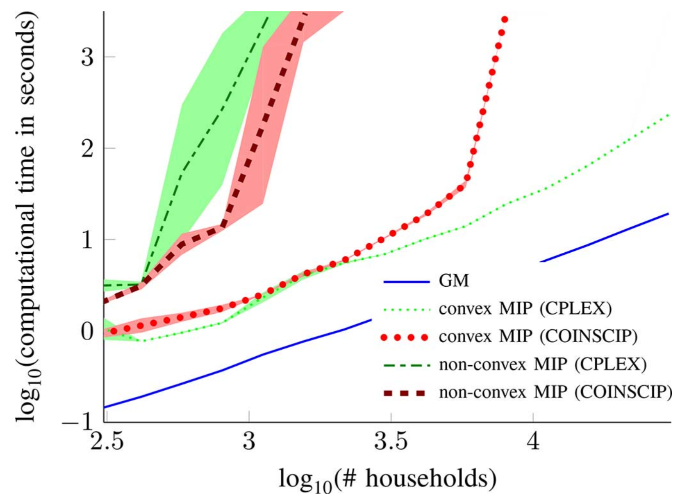
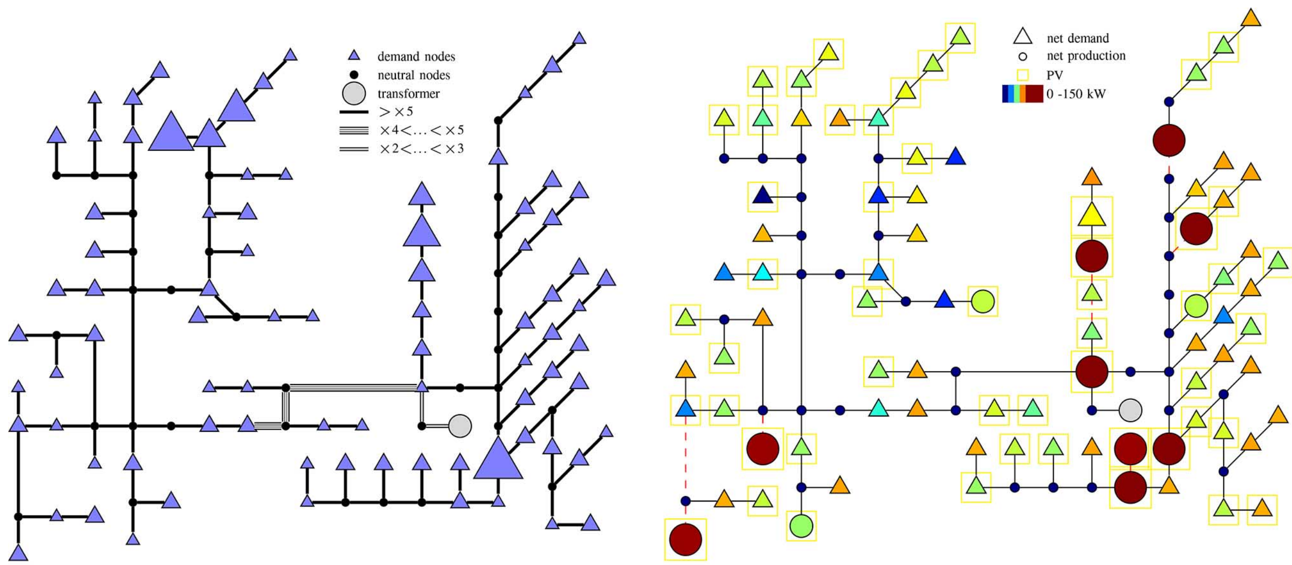


Fig. 3. Computational time of economic dispatch for the test distribution systems shown in Fig. 2, for convex (dotted) and non-convex (dashed) generator cost functions. Lines represent the mean computational time over 10 network random samples, shaded areas plus-minus one standard deviation. Our algorithm (solid line) is compared to MILP using the CPLEX solver (fine lines) and the SCIP solver (bold lines).

B. Computational Scaling Performance

For testing the computational performance of our proposed algorithm in comparison to the MILP approach, we sample test distribution networks of different sizes n as illustrated in Fig. 2. From a medium voltage ring, $n/100$ radial low-voltage feeders dissect at busbar nodes S_i. The number of households H_i in a line is chosen randomly in such a manner that the total number of households equals n and each feeder connects to at least one household. Cost functions are designed as follows:

1) All households consume 1 power unit with infinite utility.
2) 70% of households additionally possess a decentralized power generation system able to produce between 3 and



(a) (b)

Fig. 4. Exemplary Distribution Grid. (a) Structure of the passive distribution network supplied through the transformer. Triangle size encodes household demand and the grey round node denotes the transformer. Understanding the network as a tree with the transformer as the root node, lines are drawn according to the ratio of their capacity and the accumulated power demand of the sub-tree below the line. Two lines denote a capacity more than twice the accumulated demand of the sub-tree, three lines more than threefold, four lines more than fourfold and solid lines more than the fivefold accumulated demand. (b) Structure of a future active distribution grid. Circles denote net production nodes, where large circles are decentral CHP plants of different sizes, yellow boxes PV installations and triangles net demand nodes. The color-coding denotes the result of our proposed dispatch approach. Note that when using MILP scheduling without line constraints, six line capacity violations result (red dashed lines). (a) Today's distribution grid. (b) Future Smart Grid.

6 units. The cost functions C_v are polynomials with maximal degree three and randomized parameters.

3) Busbar nodes do not produce or consume any power.
4) The medium voltage transformers act as slack nodes and supply any needed power at a constant-per-unit cost.

All low-voltage branches possess a capacity of 3 power units, medium voltage lines a capacity of 6. The discretization interval is uniformly 1 unit. Both optimization approaches are fed the same 10 independent network samples per network size $n \in [300, 30\,000]$. Timing tests are run on a virtual windows machine with 8 GB of RAM and a 3.20-GHz CPU. To ensure a comparison on equal terms the solvers called by GAMS are restricted to run on one processor only and to solve the problems to optimality, i.e., with no remaining optimality gap. The run-time per problem was limited to 3000 s.

When running these problems, all methods yielded the same optimal values up to machine precision. The different timings are shown in Fig. 3. In the first set of experiments, shown as dashed lines, the generator cost functions are chosen to be convex within the operating range. Convexity is a common assumption in economic dispatch. It allows MILP to run with one binary variable per generator node only. Nevertheless, the SCIP branch-and-bound algorithm still scales exponentially in the network size. The CPLEX algorithm performs much better due to highly optimized pre-solve routines. Our proposed approach, however, outperforms both methods in absolute numbers over the whole tested range of problem sizes. In the second set of experiments, shown as dashed lines, we examine non-convex generator cost functions. These are for example needed to describe the cost curves of real gas engines [18]. The MILP formulation then needs additional binary variables, one

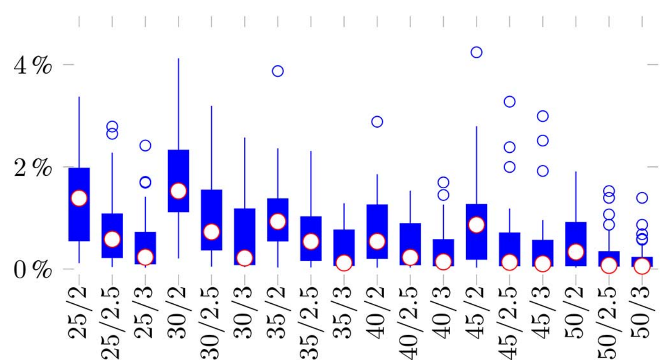
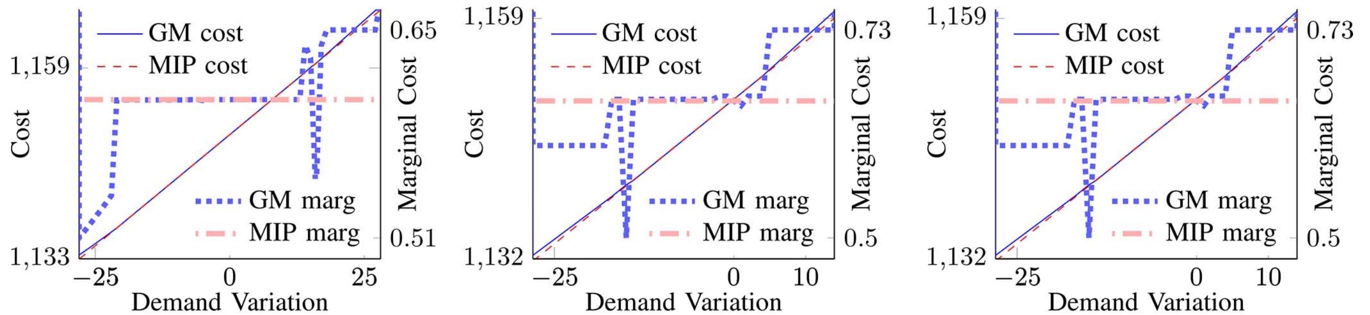


Fig. 5. Relative optimality gap for the iterative implementation of the proposed approach for 100 independent realizations of the exemplary smart distribution grid, see Fig. 4. Each interval is discretized with 25–50 points and the re-discretization in each iteration takes into account a band of 2, 2.5, or 3 times the previous discretization interval.

for each discretization interval per generator. In this case the run-time of the MILP solvers soars, while for our proposed approach no difference is observed.

Note also that the MILP approaches show a large variation in run-time, due to the fact that the pre-solving might, but is not guaranteed, to find a good starting point for the branch-and-bound algorithm that is applied afterward. In contrast, our proposed method has a deterministic run-time.

In sum, our method is the only one whose run-time scales linearly in the problem size. With the given time limit of 3000 s, it is the only method to be able to compute optimal solutions for non-convex networks with more than 10 k nodes. This may not be unrealistic in future smart distribution grids.



(a) (b) (c)

Fig. 6. Total cost change for a local demand variation—assuming that the rest of the network stays the same. For three nodes from the network in Fig. 4, the marginal cost approach derived from the MILP optimal dual variables is compared with the exact, locally computable result in our message based of approach (GM). (a) Node 25. (b) Node 72. (c) Node 36.

C. *Exemplary Smart Distribution Grid in Detail*

We now examine a smart distribution network of the future in closer detail. We construct such a grid based on experience with real projects in Southern Germany. Due to confidentially reasons, however, the original grid data could not be used.

We base our work on the 123-node test distribution feeder [19]. We compute the power line transmission limits using Europe's 230-V distribution grid voltage and scale the given consumer demands such that the smallest consumption of any node is 1.2 kW, a typical German household consumption. Note that the grid is tree-structured given the network switch states specified in [19].

The resulting grid model, see Fig. 4(a), shows features of a typical distribution grid of today. The line capacities are dimensioned to cover all demands from the transformer node alone and are rather large in comparison with individual household consumptions. Line capacities are thus unlikely to pose an active limitation for dispatch problems. This allows applying common algorithms such as MILP formulation without line constraints. The situation, however, changes in the future when additional generation and consumption units are deployed decentralized throughout the grid, as is demonstrated with our example here.

We model the following new technologies, see Fig. 4(b):

1) One electric vehicle per household chargeable with a maximal 20 kW and a utility function with marginal costs in the range 0.6–1 Euro per kW.
2) 60% of all households have an installed PV capacity of 14.9 $\mathrm{kW}_{p_{el}}$ and offer their power output at prices between 0–0.2 Euro per kW.
3) All PV owners additionally have a battery of equal size to the PV plant, generating a utility between 0.4–0.5 Euro per kW for charging.
4) Ten CHP plants with a power rating of 140 kW_{el} are present in the network, preferably located at nodes with large demand. The generator model follows [18] where the cost functions are neither convex nor concave and a minimum load of 35 kW_{el} applies in running state.
5) Households without CHP have an electric heater consuming maximally 10 $\mathrm{kW}_{P_{el}}$ while providing a utility of 0.2–0.3 Euro per kW.

The addition of these new prosumers to the distribution grid means that the network's line capacities can be violated for some dispatch assignments. For example in Fig. 4(b) it is shown that MILP algorithm without our additionally-introduced line constraints leads to 6 line violations for this network instance. In contrast, our approach always yields feasible dispatch results.

Due to the large difference between line and generator capacities, we had to resort in this example to an iterative re-discretization scheme as described above. The procedure is not guaranteed to find the global optimum, but it mostly finds a solution very close to the optimum as is demonstrated by the following test, see Fig. 5. We sampled 100 independent realizations of the exemplary smart distribution grid and computed the optimal dispatch with the MILP and with our proposed approach. Since the MILP approach with line constraints is guaranteed to find the optimal solution, we used it to benchmark the quality of our algorithm's results. With 50 discretization points for each interval and a re-discretization band of 2.5 times the discretization interval of the previous iteration, the sub-optimality of the iterative approach is less than 1% in more than 75% of the cases, and less than 4.2% in the worst case.

Next, we show that our proposed approach—despite hiding individual cost functions—delivers very broad information about the sensitivity of the solution. Remember that the optimal dual variables of a MILP approach denote the incremental total system cost for an infinitesimal demand change in one node. In contrast, the messages in our approach allow one to deduce the exact total system cost increments for all possible demand changes. The difference is especially large in the case where a small step in the local demand leads the MILP problem to change its set of active constraints and the local linearization as expressed in the dual variables is thus no longer valid.

In Fig. 6, we plot for three exemplary nodes the sum of the incoming messages plus the local node potential as a function of the local demand, minimized over all other local variables in the expression. With this local operation we then derive the curve of total system cost changes as well as its derivatives. For comparison we also plot the local linearization as derived from the optimal dual variables of the MILP approach. The difference between the local linearization of MILP and the holistic view derived from the local messages of our approach is especially obvious for nodes close to highly loaded lines, see, e.g., Fig. 6(b). In this case even a small change in local demand may lead to an infeasibility, and the linear continuation of the MILP costs is not valid any more.

VI. Conclusion

We have presented a novel economic dispatch algorithm for radial distribution networks. It is based on graphical model methods from the computational statistics community. The algorithm utilizes the grid topology to decompose the economic dispatch problem into decentralized computable parts. Only a finite number of messages are passed between the nodes of the network. The algorithm finds the globally optimal solution for discretized variables. It does so in linear time with respect to the network size. The approach protects the privacy of local cost functions while giving a comprehensive picture as to how the system cost would change for local, finite-size demand variations. The sum of all these features makes our algorithm optimally suited for economic dispatch in future large, multi-owner smart distribution grids.

Power dispatch problems are naturally formulated with continuous dispatch variables. Whether the discretization approach of this paper is suitable depends on the structure of the individual problem. If the sizes of producers and consumers are relatively similar and the line capacities are not too large in comparison, such an approach can work well as shown in the first example of this paper. However, the second example already shows the arising problems with strongly differing line and generator/consumer sizes. For such problems as well as computations with the AC power flow model with higher dimensional messages a truly continuous approach would be needed. The proposed divide and conquer trick (5) could still be used in this case. However, the messages, which are then functions of continuous parameters, can and will typically become arbitrarily complex for growing graph sizes. Thus, neither their exact representation nor their exact computation is feasible in general. Only for a special class of functions, namely quadratic functions, can one show that the complexity of the messages stays limited and exact updates can be computed, see Gaussian belief propagation [6]. In all other cases—and we think useful dispatch problems would mostly fall into this category—one will generally approximate the messages with a finite-dimensional representation and approximate the update computations. Many algorithms exist for this *approximate inference* task, see [15]. Global optimality can typically not be proven but the locality and scaling properties of belief propagation persists and state-of-the-art optimization results are obtained with these methods in many application areas, see [15], [6], and references therein. A detailed examination of this field of work with respect to its application for power dispatch problems will be a major focus of our future work.

Another open point is the generalization of our method to networks with loops. Loopy belief propagation has been successfully used in the graphical model community [15], [6], but our initial tests with dispatch in loopy electrical networks showed that defining a reasonable initial solution is non-trivial. A similar problem holds for scheduling over multiple time steps. This problem can also be understood as optimization on a loopy graph, since each node's dispatch decision at one time is dependent on the decision at the previous time as well as on the decisions of its neighbors at the same time step.

References

[1] T. Logenthiran and D. Srinivasan, "Formulation of unit commitment (UC) problems and analysis of available methodologies used for solving the problems," in *Proc. IEEE Int. Conf. Sustainable Energy Technologies (ICSET)*, 2010, pp. 1–6.
[2] M. Carrion and J. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1371–1378, Aug. 2006.
[3] M. Metzger, A. Szabo, and J. Bamberger, "Control as a key technology for the integration of renewables," in *Proc. IFAC World Congr.*, 2011.
[4] K. Kok, B. Roossien, P. MacDougall, O. van Pruissen, G. Venekamp, R. Kamphuis, J. Laarakkers, and C. Warmer, "Dynamic pricing by scalable energy management systems—Field experiences and simulation results using powermatcher," in *Proc. IEEE PES General Meeting*, 2012.
[5] P. Wolfrum, M. Kautz, and J. Schäfer, "Smart operation of CHP units," *Power Plants Power Syst. Control*, vol. 8, pp. 61–66, 2012.
[6] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: MIT Press, 2009.
[7] Y. Hu, A. Kuh, T. Yang, and A. Kavcic, "A belief propagation based power distribution system state estimator," *IEEE Computat. Intell. Mag.*, vol. 6, no. 3, pp. 36–46, 2011.
[8] Y. Li, "Fully distributed state estimation of smart grids," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2012, pp. 6580–6585.
[9] L. Zdeborová, A. Decelle, and M. Chertkov, "Message passing for optimization and control of a power grid: Model of a distribution system with redundancy," *Phys. Rev. E*, pp. 1–10, 2009.
[10] A. Y. Lam, B. Zhang, and D. N. Tse, "Distributed algorithms for optimal power flow problem," in *Proc. IEEE Conf. Decision and Control (CDC)*, Dec. 2012, pp. 430–437.
[11] J. Lavaei, D. Tse, and B. Zhang, "Geometry of power flows and optimization in distribution networks," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 572–583, Mar. 2014.
[12] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—part I," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554–2564, Aug. 2013.
[13] R. Jabr, "Radial distribution load flow using conic programming," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1458–1459, Aug. 2006.
[14] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int. J. Elect. Power Energy Syst.*, vol. 30, pp. 383–392, 2008.
[15] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Found. Trends Mach. Learn.*, vol. 1, no. 1–2, pp. 1–305, 2008.
[16] M. Schmidt, UGM: Matlab code for Undirected Graphical Models, 2012 [Online]. Available: http://www.di.ens.fr/~mschmidt/Software/UGM.html
[17] Y. Weiss, "Correctness of local probability propagation in graphical models with loops," *Neural Computat.*, vol. 12, no. 1, pp. 1–41, 2000.
[18] C. Hernandez-Aramburo, T. Green, and N. Mugniot, "Fuel consumption minimization of a microgrid," *IEEE Trans. Ind. Applicat.*, vol. 41, no. 3, pp. 673–681, 2005.
[19] W. IEEE Distribtuion Test Feeder (2013), Distribution test feeders, IEEE PES Distribution System Analysis Subcommittee.

Elisabeth Kellerer received the Dipl. degree in mathematics from the Technische Universitt München (TUM), Munich, Germany, in 2011. She is currently pursuing the Ph.D. degree at Siemens Corporate Technology, Munich, Germany.

She is currently the Chair for Energy Economy and Application Technology, TUM, Munich, Germany. Her current research focuses on novel algorithms for Smart Grid dispatch, probabilistic graphical models, and optimization. She is member of the TUM International Graduate School of Science and Engineering.

Florian Steinke received the Ph.D. degree in machine learning from the Max Planck Institute for Intelligent Systems, Tübingen, Germany, in 2009.

He is now a Senior Key Expert Engineer at Siemens Corporate Technology in Munich, Germany. His major focus is on smart power systems that can integrate large amounts of renewable energy. He leads and contributes to projects, ranging from economic optimization of designs and operation to technical grid stability analyses. He has published several papers and patents on these topics.

Dr. Steinke is member of the VDE Association for Electrical, Electronic and Information Technologies e.V. working group on power generation.