RC25459 (IRE1404-010) April 14, 2014
Mathematics

IBM Research Report

Optimal Power Flow as a Polynomial Optimization Problem

Bissan Ghaddar, Jakub Marecek, Martin Mevissen
IBM Research
Smarter Cities Technology Centre
Mulhuddart
Dublin 15, Ireland

Optimal Power Flow as a Polynomial Optimization Problem

Bissan Ghaddar, Jakub Marecek, Martin Mevissen

Abstract—**Formulating the alternating current optimal power flow (ACOPF) as a polynomial optimization problem makes it possible to solve large instances in practice and to guarantee asymptotic convergence in theory. In this work, we utilize polynomial optimization to formulate the ACOPF as a degree two polynomial program and present strong convexifications of the ACOPF problem. This is done by adding violated but valid inequalities that tightens the first order relaxation of the quadratic polynomial program. We also exploit the polynomial program structure and use the sparse hierarchy to solve larger instances to global optimality. Computational results for both approaches are presented.**

Index Terms—**Power system management, Power system analysis computing, Optimization, Numerical Analysis (Mathematical programming, Method of moments, Sparse matrices)**

I. Introduction

Optimal Power Flow (OPF) in alternating current models (ACOPF) is one of the most important nonlinear optimization problems. Despite the fact that ACOPF has been studied since early 1960s [3] in thousands of papers, the progress has not been satisfactory. The problem is non-convex due to the non-linear power flow constraints, and hence difficult to solve.

There is a great variety [15] of relaxations and heuristic solution methods available. Such methods find locally optimal solutions or stationary points and are not guaranteed to find a global optimum. A particularly important line of research started with the semidefinite programming (SDP) relaxation of Bai et al. [1]. Lavaei et al. showed [10], [16] that the solution of the SDP is the global optimum, under some conditions. Several follow-up computational studies [5], [12] increased the dimension of SDP instances that can be solved, in practice.

Realising that OPF can be modeled as a polynomial program (PP) of degree at most four, we aim to tackle the problem using methods recently developed in polynomial optimization. Using this approach, one can capture the system behavior accurately, and obtain globally valid lower bounds and globally optimal solutions, under mild constraints. Notably, one can use a wide variety of objective functions, and incorporate further constraints easily, without affecting convergence properties. Computationally, one uses a hierarchy of SDP relaxations of Lasserre [8] or Waki et al. [18] to convexify the problem. Unfortunately, dimensions of the relaxations grow rapidly with the size of the power system, posing a major computational challenge.

In this paper, we present two techniques for tackling the computational challenge. The first technique uses "cutting surfaces" of Ghaddar et al. [4], which are valid inequalities, generated dynamically upon violation at each step of the algorithm. Instead of increasing the degree of the non-negative certificates in the hierarchy, the set of polynomial inequalities describing the feasible region of the polynomial program is changed in each iteration, while the degree of the polynomials is fixed. These valid inequalities yield stronger convexifications.

The second technique uses the sparse hierarchy of SDP relaxations of Waki et al. [18], which are computationally considerably more tractable than relaxations of the so called dense hierarchy of Lasserre studied previously [13], [6]. Specifically, the size of the relaxation does not grow exponentially in the number of buses, but only as a function of the density of the network. The relaxations are equivalent to the Lavaei-Low [10] SDP relaxation, where it is exact, and provide tighter relaxations, where it is not. Further, we employ matrix completion techniques to break down the largest SDP variable at the price of introducing additional equality constraints and several smaller matrix inequalities, to make the approach scale to power systems with thousands of buses.

Overall, the main contributions of the paper are:

IBM Research - Ireland, Damastown Industrial Estate, Mulhuddart, Dublin 15. email: bghaddar@ie.ibm.com, jakub.marecek@ie.ibm.com, martmevi@ie.ibm.com

- solve larger instances to global optimality than those published in the literature
- present stronger convexifications for the OPF problem than those presented in the literature
- exploit sparsity in existing SDP relaxations to tackle larger instances
- study the convergence of the sparse hierarchy.

Notably, either of the presented techniques improves upon the Lavaei-Low SDP relaxation, whenever the Lavaei-Low SDP relaxation does not provide the global optimum.

The rest of the paper is organized as follows. Section II describes the optimal power flow problem and presents the various formulations used to model this problem. Section III discusses the two proposed schemes. Section IV shows the performance of both schemes on two sets of instances. Finally, Section V provides concluding remarks and future research directions.

II. Optimal Power Flow Problem

In this section we use the same notation used in [10] and [12]. The topology of the power system $P = (N, E)$ is represented as an undirected graph, where each vertex $n \in N$ is called a "bus" and each edge $e \in E$ is called a "branch".We use $|N|$ to denote the number of buses and $|E|$ to denote the number of branches. Let $G \subseteq N$ be the set of generators and $E \subseteq N \times N$ be the set of all branches modeled as Π-equivalent circuits. The matrix $y \in \mathbb{R}^{|N|\times|N|}$ represents the network admittance matrix, whose sparsity pattern is the same as that of the adjacency matrix of P. $\bar{b}_{lm}$ is the value of the shunt element at branch $(l, m) \in E$ and $g_{lm}+jb_{lm}$ is the series admittance on a branch (l, m). Let $S^d_k = P^d_k + jQ^d_k$ be the active and reactive load (demand) at each bus $k \in N$ and $P^g_k + jQ^g_k$ represent the apparent power of the generator at bus $k \in G$. Define $V_k = \Re V_k + j\Im V_k$ as the voltage at each bus $k \in N$ and $S_{lm} = P_{lm} + jQ_{lm}$ as the apparent power flow on the line $(l, m) \in E$. The edge set $L \subseteq E$ contains the branches (l, m) such that the apparent power flow limit is less than a certain given tolerance ε. In the next section, we describe one of the classical formulations of the optimal power flow problem in detail.

A. Formulation

In this work, we focus on the rectangular power-voltage formulation. To formulate the ACOPF model, we define the following parameters:

- $P^{\min}_k$ and $P^{\max}_k$ are the limits on active generation capacity at bus k, where $P^{\min}_k = P^{\max}_k = 0$ for all $k \in N/G$.
- $Q^{\min}_k$ and $Q^{\max}_k$ are the limits on reactive generation capacity at bus k, where $Q^{\min}_k = Q^{\max}_k = 0$ for all $k \in N/G$.
- $V^{\min}_k$ and $V^{\max}_k$ are the limits on the absolute value of the voltage at a given bus k.
- $S^{\max}_{lm}$ is the limit on the absolute value of the apparent power of a branch $(l, m) \in L$.

Let e_k be the k^{th} standard basis vector in $\mathbb{R}^{|N|}$, similar to [10], the following matrices are defined

$$
\begin{aligned}
y_k &= e_k e_k^T y,\\
y_{lm} &= (j\frac{\bar{b}_{lm}}{2} + g_{lm} + jb_{lm})e_l e_l^T - (g_{lm} + jb_{lm})e_l e_m^T,\\
Y_k &= \frac{1}{2}\begin{bmatrix} \Re(y_k + y_k^T) & \Im(y_k^T - y_k) \\ \Im(y_k - y_k^T)) & \Re(y_k + y_k^T) \end{bmatrix},\\
\bar{Y}_k &= -\frac{1}{2}\begin{bmatrix} \Im(y_k + y_k^T) & \Re(y_k - y_k^T) \\ \Re(y_k^T - y_k) & \Im(y_k + y_k^T) \end{bmatrix},\\
M_k &= \begin{bmatrix} e_k e_k^T & 0 \\ 0 & e_k e_k^T \end{bmatrix},\\
Y_{lm} &= \frac{1}{2}\begin{bmatrix} \Re(y_{lm} + y_{lm}^T) & \Im(y_{lm}^T - y_{lm}) \\ \Im(y_{lm} - y_{lm}^T) & \Re(y_{lm} + y_{lm}^T) \end{bmatrix},\\
\bar{Y}_{lm} &= -\frac{1}{2}\begin{bmatrix} \Im(y_{lm} + y_{lm}^T) & \Re(y_{lm}^T - y_{lm}) \\ \Re(y_{lm}^T - y_{lm}) & \Im(y_{lm} + y_{lm}^T) \end{bmatrix}.
\end{aligned}
$$

Let x be a vector of variables defined as $x := [\Re V_k \quad \Im V_k]^T$, and let the cost of power generation be $\sum_{k\in G} f_k(P^g_k)$ where $f_k(P^g_k) = c^2_k(P^g_k)^2 + c^1_k P^g_k + c^0_k$, with c^2_k, c^1_k, c^0_k non-negative. The classical OPF problem can be written as a polynomial optimization problem of degree 4,

[PP4]

$$
\begin{aligned}
\min \sum_{k\in G} & f_k(x) \\
\text{s.t. } & P^{\min}_k \le \operatorname{tr}(Y_k x x^T) + P^d_k \le P^{\max}_k && (1)\\
& Q^{\min}_k \le \operatorname{tr}(\bar{Y}_k x x^T) + Q^d_k \le Q^{\max}_k && (2)\\
& (V^{\min}_k)^2 \le \operatorname{tr}(M_k x x^T) \le (V^{\max}_k)^2 && (3)\\
& (\operatorname{tr}(Y_{lm} x x^T))^2 + (\operatorname{tr}(\bar{Y}_{lm} x x^T))^2 \le (S^{\max}_{lm})^2 && (4)
\end{aligned}
$$

The objective function often is the cost of power generation where $f_k(x) :=$

$\left(c_k^2(P_k^d + \mathrm{tr}(Y_k x x^T))^2 + c_k^1(P_k^d + \mathrm{tr}(Y_k x x^T)) + c_k^0\right)$. Constraints (1) and (2) impose a limitation on the active and reactive power. Constraints (3) restrict the voltage on a given bus. Constraints (4) limit the apparent power flow at each end of a given line. The problem can be modeled as a rank-constrained problem by defining the variable $W = xx^T$. Subsequently, one can drop the rank constraint to obtain the SDP relaxation [OPF-SDP] as has been done in [10]. Notice that:

Proposition 1. *For every $k \geq 1$, there exist an instance of the ACOPF optimization problem [PP4] with solution x^*, such that $z \leq kz \leq f(x^*)$, where z is the objective function of the [OPF-SDP] relaxation [10].*

Proof: Let us consider a system, where there is one slack bus and the Lavaei-Low relaxation [OPF-SDP] [10] does produce objective function value $z \geq 1$, which is not the global optimum $f(x^*)$. For example, this could be the two-bus instance of [2]. We can produce a variant of the instance with a new objective function given by $f'(\cdot) := C(k)f(\cdot)$, i.e. changing the c_i^0, c_i^1, c_i^2 for each generator i, to obtain the result when choosing $C(k)$ sufficiently large. ■

This motivates research for global optimization methods beyond SDP relaxation [OPF-SDP].

III. Polynomial Programming Approach

The OPF problem is a particular case of a polynomial optimization problem:

$$\textbf{[PP-P]} \quad \min \quad f(x) \qquad \text{s.t. } g_i(x) \geq 0 \qquad i = \{1, \ldots, m\}$$

Motivated by the seminal work of Lasserre [8], there are schemes applying representation theorems from algebraic geometry to characterize the set of polynomials that are non-negative on a given domain. The main idea of these schemes is based on applying representation theorems to characterize the set of polynomials that are non-negative on a given basic closed algebraic set S. We use the following notation: Define $\deg(f)$ to be the degree of the polynomial $f(x)$. Let $\mathbf{R}_d[x] := \mathbf{R}_d[x_1, \cdots, x_n]$ be the set of polynomials in n variables with real coefficients of degree at most d. Given $S \subseteq \mathbb{R}^n$, define $\mathcal{P}_d(S)$ to be the cone of polynomials of degree at most d that are non-negative over S. We use Σ_d to denote the cone of real polynomials of degree at most d that are sum-of-squares of polynomials. $\Sigma_d := \{\sum_{i=1}^N p_i(x)^2 : p_i(x) \in \mathbf{R}_{\lfloor \frac{d}{2} \rfloor}[x]\}$, with $N = \binom{n+d}{d}$. For any $\mathcal{G} \subseteq \mathbf{R}_d[x]$, we denote by $S_{\mathcal{G}} = \{x \in \mathbb{R}^n : g(x) \geq 0 \; \forall g \in \mathcal{G}\}$ the basic closed semialgebraic set defined by $\mathcal{G}$. Taking $\mathcal{G} = \{g_i(x) : i = 1, \ldots, m\}$, we can rephrase [PP-P] as

$$\textbf{[PP-D]} \quad \max \quad \varphi \qquad \text{s.t. } f(x) - \varphi \in \mathcal{P}_d(S_{\mathcal{G}}).$$

Although [PP-D] is a conic problem, it is unknown how to efficiently optimize over the cone $\mathcal{P}_d(S_{\mathcal{G}})$. Lasserre [8] introduced a hierarchy of SDP relaxations corresponding to liftings of polynomial problems into higher dimensions. Lasserre builds up a sequence of linear semidefinite relaxations of increasing size. Under assumptions slightly stronger than compactness, the optimal values of these problems converge to the global optimal value of the original problem, [PP-P]. The approximation of $\mathcal{P}_d(S_{\mathcal{G}})$ used in [8] is the cone $\mathcal{K}_{\mathcal{G}}^r$, where

$$\mathcal{K}_{\mathcal{G}}^r = \Sigma_r + \sum_{i=1}^m g_i(x)\Sigma_{r-\deg(g_i)}, \tag{5}$$

and $r \geq d$. The corresponding optimization problem over S can be written as:

$$\begin{aligned} &\max_{\varphi, \sigma_i(x)} \varphi \\ &\text{s.t. } f(x) - \varphi = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x) \\ &\sigma_0(x) \in \Sigma_r, \; \sigma_i(x) \in \Sigma_{r-\deg(g_i)}. \end{aligned} \tag{6}$$

Problem (6) can be reformulated as a semidefinite optimization problem. The computational cost of the problem clearly depends on both the degree of the polynomial and the dimension of the problem. The number of constraints can be large, especially when many variables and high-degree polynomials are used. As a result, the matrix inequalities grow with the value of r. Based on the described approach, Molzahn and Hiskens [13] and Josz et al. [6] used [PP4] and applied Lasserre's hierarchy to obtain global optimality on instances with up to

5 and 10 buses respectively, where Lavaei-Low is not globally optimal.

A. Relationship with Lavaei-Low Formulation

In this work, instead of starting with [PP4] and growing the complexity of the SDP relaxations by applying the hierarchy, we reduce the OPF problem to a polynomial program of degree 2:

[PP2]

$$\min \sum_{k \in G} \left(c_k^2 (P_k^g)^2 + c_k^1 (P_k^d + \text{tr}(Y_k x x^T)) + c_k^0\right)$$
$$P_k^{\min} \leq \text{tr}(Y_k x x^T) + P_k^d \leq P_k^{\max}$$
$$Q_k^{\min} \leq \text{tr}(\bar{Y}_k x x^T) + Q_k^d \leq Q_k^{\max}$$
$$(V_k^{\min})^2 \leq \text{tr}(M_k x x^T) \leq (V_k^{\max})^2$$
$$P_{lm}^2 + Q_{lm}^2 \leq (S_{lm}^{\max})^2$$
$$P_k^g = \text{tr}(Y_k x x^T) + P_k^d$$
$$P_{lm} = \text{tr}(Y_{lm} x x^T)$$
$$Q_{lm} = \text{tr}(\bar{Y}_{lm} x x^T)$$

The number of variables added in this case is $|G| + 2|L|$, which can be small as typically the number of generators with non-zero c_k^2 value in the objective and the number of power flow constraints are relatively small.

Theorem 1. *The first level of the hierarchy approximation of [PP2], [D], is equivalent to the dual of [OPF-SDP], (i.e., optimization 4 in [10]).*

Hence, we apply two techniques to improve on the bounds obtained by [D]. The first approach is discussed in the next section where valid inequalities are added to [PP2] which are translated to more variables to [D]. The resulting scheme improves the [D] relaxation and hence the bound. In the second approach, we exploit the sparsity pattern of the two polynomial programs [PP2] and [PP4] to provide global optimal solutions for the OPF problem.

B. DIGS Approach

As seen in Section III-A, [PP2] is equivalent to the Lavaei-Low SDP relaxation. Hence, by adding valid inequalities to [PP2] one can improve on the Lavaei-Low bound and in some cases obtain the global optimum. In this section, we use the first level of the relaxation of [PP2], i.e. [D], and add valid quadratic inequalities of the form $p(x) \geq 0$. The approach is generic and can be applied to general polynomial programs as described in [4]. The idea is that the polynomial $p(x)$ needs to be a valid inequality and at the same time improve on the bound of the relaxation. This can be translated as $p \in \mathcal{P}_d(S) \setminus \mathcal{K}_{\mathcal{G}}^d$, where $d = 2$, the degree of [PP2], in this case. The iterative scheme can be summarized as follows:

- Start with $\mathcal{G}_0 = \mathcal{G}$
- Given $\mathcal{G}_i$ let $p_i \in \mathcal{P}_d(S_{\mathcal{G}}) \setminus \mathcal{K}_{\mathcal{G}_i}^d$, define $\mathcal{G}_{i+1} = \mathcal{G}_i \cup \{p_i\}$.

To be able to generate a polynomial $p(x)$, the scheme consists of a master problem and a subproblem. The master problem is of the following form

$$\textbf{[PP-M]} \quad \max \varphi \quad \text{s.t. } f(x) - \varphi \in \mathcal{K}_{\mathcal{G}}^d,$$

where $\mathcal{K}_{\mathcal{G}}^d$ is as defined in (5) with r fixed to d. [PP-M] provides lower bounds for OPF problem. The subproblem uses the optimal dual information from the master problem, Y, to generate polynomial inequalities that are valid on the feasible region. These valid inequalities are then incorporated into the master to construct new non-negativity certificates, obtaining better approximations of the OPF. The subproblem is also a semidefinite program and has the following form:

$$\textbf{[PP-S]} \quad \min_p \langle p, Y \rangle \quad \text{s.t. } p \in \mathcal{K}_{\mathcal{G}}^{d+2} \cap \mathbf{R}_d[x].$$

The iterative scheme terminates when the objective function of the subproblem is close to 0 [4]. As opposed to the approach proposed in [6], [13], where r is increased when using (6), in DIGS, r is fixed to d. Consequently, the growth in the size of the positive semidefinite matrices and the number of constraints can be significantly lower as compared to (6).

C. Exploiting OPF Structure

The current scalability of state-of-the-art SDP solvers limits the tractability of the Lasserre hierarchy even for medium-scale polynomial programs. One approach to improve the tractability of the Lasserre hierarchy is to exploit *correlative* sparsity of the underlying polynomial optimization problem due to Waki et al. [18]. Let $\{I_k\}_k$ the set of maximal cliques of the correlative sparsity pattern graph of [PP2] following the construction in [18]. The sparse approximation of $\mathcal{P}_d(S_\mathcal{G})$ is $\mathcal{K}^r_\mathcal{G}(I)$, given by

$$\mathcal{K}^r_\mathcal{G}(I) = \sum_{k=1}^{p} (\Sigma_r(I_k) + \sum_{j \in J_k} g_j \Sigma_{r-\deg(g_j)}(I_k)),$$

where $\Sigma_d(I_k)$ is the set of all sum-of-squares polynomials of degree up to d supported on I_k and $(J_1, \ldots, J_s)$ is a partitioning of the set of polynomials $\{g_j\}_j$ defining K such that for every j in J_k, the corresponding g_j supported on I_k. The sparse hierarchy of SDP relaxations is then given by

$$\begin{aligned} &\max_{\varphi, \sigma_{r,k}(x)} \varphi \\ &\text{s.t. } f(x) - \varphi = \sum_{k=1}^{p} \left(\sigma_k(x) + \sum_{j \in J_k} g_j(x) \sigma_{j,k}(x) \right) \\ &\qquad \sigma_k \in \Sigma_r((I_k)), \sigma_{j,k} \in \Sigma_{r-\deg(g_j)}(I_k). \end{aligned} \tag{7}$$

While (7) provides a weaker relaxation to a PP than (6) for a fixed relaxation order r in general, the asymptotic convergence result for the dense hierarchy extends to the sparse case:

Assumption 1. *Let K denote the feasible set of [PP2]. Let $\{I_k\}_k$ denote the p maximal cliques of a chordal extension of the sparsity pattern graph of the [PP2].*

(i) *Then, there is a $M > 0$ such that $\| x \|_\infty < M$ for all $x \in K$.*
(ii) *Ordering conditions for index sets as in Assumption 3.2 (i) and (ii) in [9].*
(iii) *Running-intersection-property, c.f. [9], holds for $\{I_k\}_k$.*

Remark 1. *Note, that K is compact. Therefore, it is easy to add up to p redundant quadratic inequality to the definition of K, s.t. Assumption 1 (i) is satisfied. (ii) can be satisfied by construction and re-ordering of the sets $\{I_k\}_k$. The running-intersection-property is satisfied for the maximal cliques of a choral graph, as pointed out in [7] . Thus, Assumption 1 is satisfied for [PP2], in case of the quadratic generation cost objective.*

Now, we can formulate the convergence result.

Proposition 2 (Asymptotic Convergence)**.** *If Assumption 1 holds and the interior of K nonempty, then there exists a sequence of SDP relaxations S_r such that*

(a) $\inf S_r \nearrow \min([PP2])$ *as* $r \to \infty$.
(b) *There is no duality gap between S_r and $S_r^\star$.*
(c) *Let $y_r^\star$ the optimal solution of S_r and $\hat{y}_r^\star := \{y_{r\,\alpha}^\star :| \alpha |= 1\}$. If [PP4] has a unique global minimizer $x^\star \in K$, then*

$$\hat{y}_r^\star \to x^\star \text{ as } r \to \infty.$$

Proof: Corollary of Theorem 3.6 of Lasserre [9]. ■

Assumption 1 is satisfied in the case of the generation costs objective of [PP4] and [PP2]. Moreover, for [PP2] the following proposition holds.

Proposition 3. *The sparse SDP relaxation (7) of order one is equivalent to the first order relaxation of the dense Lasserre hierarchy (6) and the Lavaei-Low relaxation [OPF-SDP].*

Proof: Follows from the fact that sparse and dense SDP relaxation of order 1 are equivalent for non-convex quadratic optimization problems, as proven in Section 4.5 of Waki et al. [18] and the Theorem 1. ■

Remark 2. *For a fixed order r, the sparse hierarchy (7) has $O(\kappa^{2r})$ variables, where κ the maximum number of of variables appearing in the objective or a inequality constraint of PP. The largest matrix inequality is of size $O(\kappa^r)$. This is in contrast to $O(n^{2r})$ variables and matrix variables of size $O(n^r)$ in the dense hierarchy (6). In case the PP is very sparse, i.e., $\kappa \ll n$, the size of the sparse hierarchy is vastly smaller then the dense one.*

IV. Numerical Results

In the computational results, we demonstrate the performance of both techniques, the digs approach and the sparsity exploitation approach, in improving the quality of the SDP relaxation bound. DIGS is used on the quadratic formulation and improves the relaxation iteratively and in some cases prove optimality. On the other hand, SparsePoP is used

TABLE I: WB2 computational results.

	DIGS			L1		L2		L3	
$V_2^{\max}$	Iter s	s	Time	Bound	Time	Bound	Time	Bound	Time
0.976	**905.76**	1	0.9	905.76	0.2	**905.76**	0.4		
0.983	**905.73**	6	5.1	903.12	0.2	**905.73**	1.8		
0.989	**905.73**	6	4.3	900.84	0.1	905.72	1.7	**905.73**	1.8
0.996	**905.73**	6	4.6	898.17	0.2	905.73	1.4	**905.73**	1.6
1.002	**905.73**	6	4.8	895.86	0.2	905.72	1.8	**905.73**	1.5
1.009	**905.73**	8	6.4	893.16	0.2	905.71	1.9	**905.73**	0.6
1.015	**905.73**	6	4.7	890.82	0.1	905.71	0.8	**905.73**	0.6
1.022	**905.73**	8	6.5	890.82	0.2	905.71	2.6	**905.73**	1.7
1.028	**905.73**	8	5.1	885.71	0.1	904.59	0.8	**905.73**	0.8

to take advantage of the sparse structure of the polynomial program to solve the higher order relaxations in the hierarchy and ensure global optimality for some instances.[1] The first level of the SDP hierarchy of [PP2] and [PP4] is referred to as L1 and L2 respectively. The second level of the hierarchy where r is set to 6 for [PP4] is referred to as L3.

A. Small-scale Instances

The first set of instances are small test cases widely used in the literature [2]. Two examples are considered, where in each example some of the parameters are modified according to [6]. DIGS and SparsePoP results are presented for these examples and bold entries indicate that the objective value is optimal.

Example 1. *The first example is an instance with two buses [2] where the maximum voltage on the second bus varies from 0.976 to 1.028 (see Table I). From Figure 1, it can be seen that after adding one inequality (i.e., one iteration of DIGS), the Lavaei-Low bound improves significantly. The optimal value is obtained in 8 iterations.*

Example 2. *The second example is LMBM3 with 3 buses [11]. The results from Table II indicate that both approaches are successful in providing the optimal solution of these problems when Lavaei-Low relaxation is not exact.*

[1]The two techniques are implemented in MATLAB running on a Red Hat Linux PC with a 3.5Ghz processor. The DIGS approach is implemented using APPS [4] and SparseColO was used for exploiting d- and r-space sparsity in the master and the subproblem. Exploiting sparsity of polynomial program [PP4] is done using SparsePoP [19]. To solve the resulting SDP relaxation for both approaches, SeDuMi [17] is used as the SDP solver. All test instances are taken from [2] and [14].

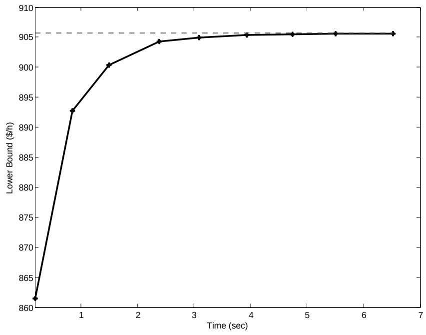

Fig. 1: WB2 Bounds for $V_2^{\max}$=1.022.

TABLE II: LMBM3 computational results.

	DIGS			L1		L2	
$S_{23}^{\max}$	Iter s	s	Time	Bound	Time	Bound	Time
28.35	**10294.88**	7	13.3	6307.97	0.3	**10294.88**	1.0
31.16	**8179.99**	6	11.2	6206.78	0.2	**8179.99**	0.7
33.96	**7414.94**	5	19.2	6119.71	0.2	**7414.94**	0.8
36.77	**6895.19**	5	19.5	6045.33	0.3	**6895.19**	0.7
39.57	**6516.17**	5	19.8	5979.38	0.3	**6516.17**	0.7
42.38	**6233.31**	5	18.1	5919.12	0.2	**6233.31**	0.7
45.18	**6027.07**	5	19.3	5866.68	0.3	**6027.07**	0.8
47.99	**5882.67**	3	12.1	5819.02	0.3	**5882.67**	0.7
50.79	**5792.02**	2	9.2	5779.34	0.3	**5792.02**	0.7
53.60	**5745.04**	1	0.7	5745.04	0.3	**5745.04**	0.8

B. Large-scale Instances

The second set of problems consists of larger instances. These instances are taken from [14]. Table III presents computational results for L1 and L2. L1 captures the Lavaei-Low dual relaxation and obtains the same bounds and has similar computational performance. For instances larger than 39 buses only L1 can be solved, as L2 becomes computationally expensive for SparsePoP. Using DIGS, optimality of case9mod is proven in 3 hours and 7 iterations while for case14mod, DIGS performed 2 iterations in a time limit of 5 hours and improved the Lavaei-Low bound. For instances larger than 30 buses DIGS was not able to improve on the SDP relaxation due to time limitations.

DIGS and exploiting sparsity are able to improve on the Lavaei-Low relaxation for small- to medium-scale instances and global optimality was proven for instances up to 39 buses. The main drawback of DIGS is that the subproblem is generic and hence is expensive to solve.

TABLE III: Computational results for IEEE and Polish network instances.

	L1			L2		
Instance	Bound	Dim	Time	Bound	Dim	Time
case9mod	2753.23	588×168	0.6	**3087.89**	1792×14847	17.5
case14mod	7792.72	888×94	0.9	7991.07	7508×66740	904.2
case30mod	576.89	4706×684	3.8	**578.56**	36258×49164	13740.0
case39	41862.08	7282×758	2.2	**41864.18**	26076×215772	4359.1
case57	41737.79	13366×356	3.2	*	*	
case118	129654.62	56620×816	6.1	*	*	
case300	719711.63	362076×1938	13.6	*	*	
case2383wp	1.814×10^6	22778705×47975	3731.5	*	*	
case2736sp	1.307×10^6	30019740×57408	3502.2	*	*	

V. Conclusion

In this work, we propose to formulate the optimal power flow as a polynomial programming problem, as well as two techniques which make the resulting SDP relaxations tractable for medium- and large-scale instances. This approach makes it possible to change objective functions and constraints rather freely, without any need to change the approach to solving the modified or extended variants. Additionally, binary variables can be included, e.g., to model discrete decisions in transmission switching.

We show that the proposed approach extends the Lavaei-Low SDP relaxation. For several instances, where the Lavaei-Low relaxation is not exact, we provide provide globally optimal solutions for the first time. Further research could focus on specializing the subproblem of DIGS to power flow problems, such that instances of larger sizes can be solved. Another research direction is to take advantage of the inequality generation and sparsity simultaneously, which could lead to further savings in terms of computational time.

References

[1] X. Bai, H. Wei, K. Fujisawa, and Y. Wang. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30(6):383–392, 2008.

[2] W. A. Bukhsh, A. Grothey, K. I. McKinnon, and P. Trodden. Local solutions of optimal power flow. *Power Systems, IEEE Transactions on*, 28(4):4780–4788, 2013.

[3] J. Carpentier. Contribution a l'etude du dispatching economique. *Bull. Soc. Francaise des Electriciens Ser. B*, 3:431–447, Aug 1962.

[4] B. Ghaddar. *New Conic Optimization Techniques for Solving Binary Polynomial Programming Problems*. PhD thesis, University of Waterloo, 2011.

[5] R. A. Jabr. Exploiting sparsity in sdp relaxations of the opf problem. *Power Systems, IEEE Transactions on*, 27(2):1138–1139, 2012.

[6] C. Josz, J. Maeght, P. Panciatici, and J. C. Gilbert. Application of the moment-sos approach to global optimization of the opf problem. arXiv:1311.6370v1, 2013.

[7] M. Kojima and M. Muramatsu. A note on sparse sos and sdp relaxations for polynomial optimization problems over symmetric cones. *Computational Optimization and Applications*, 42(1):31–41, 2009.

[8] J. Lasserre. Global optimization problems with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11:796–817, 2001.

[9] J. Lasserre. Convergent sdp-relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, 17(3):882–843, 2006.

[10] J. Lavaei and S. Low. Zero duality gap in optimal power flow problem. *Power Systems, IEEE Transactions on*, 27(1):92–107, 2012.

[11] B. Lesieutre, D. Molzahn, A. Borden, and C. DeMarco. Examining the limits of the application of semidefinite programming to power flow problems. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1492 –1499, 2011.

[12] D. Molzahn, J. Holzer, B. Lesieutre, and C. DeMarco. Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Transactions on Power Systems*, 28(4):3987–3998, 2013.

[13] D. K. Molzahn and I. A. Hiskens. Moment-based relaxation of the optimal power flow problem. arXiv:1312.1992v2, 2013.

[14] U. of Washington. Power systems test case archive.

[15] K. Pandya and S. Joshi. A survey of optimal power flow methods. *Journal of Theoretical & Applied Information Technology*, 4(5), 2008.

[16] S. Sojoudi and J. Lavaei. Semidefinite relaxation for nonlinear optimization over graphs with application to power systems. *Preprint*, 2013.

[17] J. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11-12, 1999.

[18] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1):218–242, 2006.

[19] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. SparsePOP : a sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Transactions on Mathematical Software*, 35(2):15, 2008.

APPENDIX A
PROOFS

Proof of Theorem 1: Notice that the variables in [PP2] are x, P_k^g, and P_{lm} and Q_{lm}. However, not all the monomials appear in the polynomial formulation and hence using the first level of the hierarchy, $\mathcal{K}_G^r$, where $r = 2$ one can approximate [PP2]:

$$
\begin{aligned}
\max \quad & \varphi \\
\text{s.t.} \quad & \sum_{k\in G}\left(c_k^2(P_k^g)^2 + c_k^1(P_k^d + \text{tr}(Y_k x x^T)) + c_k^0\right) - \varphi \\
&= A(x) + \sum_{k\in G} B_k(P_k^g) + \sum_{(l,m)\in L} C_{lm}(P_{lm}, Q_{lm}) \\
&+ \sum_{k\in N} \overline{\lambda}_k(P_k^{\max} - P_k^d - \text{tr}(Y_k x x^T)) \\
&+ \sum_{k\in N} \underline{\lambda}_k(-P_k^{\min} + P_k^d + \text{tr}(Y_k x x^T)) \\
&+ \sum_{k\in N} \overline{\gamma}_k(Q_k^{\max} - Q_k^d - \text{tr}(\bar{Y}_k x x^T)) \\
&+ \sum_{k\in N} \underline{\gamma}_k(-Q_k^{\min} + Q_k^d + \text{tr}(\bar{Y}_k x x^T)) \\
&+ \sum_{k\in N} \overline{\mu}_k((V_k^{\max})^2 - \text{tr}(M_k x x^T)) \\
&+ \sum_{k\in N} \underline{\mu}_k((-V_k^{\min})^2 + \text{tr}(M_k x x^T)) \\
&+ \sum_{(l,m)\in L} a_{lm}((S_{lm}^{\max})^2 - P_{lm}^2 - Q_{lm}^2) \\
&+ \sum_{k\in G} b_k(P_k^g - \text{tr}(Y_k x x^T) - P_k^d) \\
&+ \sum_{(l,m)\in L} c_{lm}(P_{lm} - \text{tr}(Y_{lm} x x^T)) \\
&+ \sum_{(l,m)\in L} d_{lm}(Q_{lm} - \text{tr}(\bar{Y}_{lm} x x^T))
\end{aligned}
$$

where $A(x), B_k(P_k^g), C_{lm}(P_{lm}, Q_{lm})$ are polynomials that are sum of squares as a function of x, P_k^g, and P_{lm} and Q_{lm} respectively. That is $A(x) = xAx^T$, $B_k(P_k^g) = \begin{bmatrix} 1 \\ P_k^g \end{bmatrix} B \begin{bmatrix} 1 \\ P_k^g \end{bmatrix}^T$, and $C_{lm}(P_{lm}, Q_{lm}) = \begin{bmatrix} 1 \\ P_{lm} \\ Q_{lm} \end{bmatrix} B \begin{bmatrix} 1 \\ P_{lm} \\ Q_{lm} \end{bmatrix}^T$, where A, B_k, and C_{lm} are positive semidefinite matrices of dimension $2|N| \times 2|N|$, 2×2 and 3×3 respectively. The variables $\overline{\lambda}_k, \underline{\lambda}_k, \overline{\gamma}_k, \underline{\gamma}_k, \overline{\mu}_k, \underline{\mu}_k$ and a_{lm} are non-negative variables. By equating the coefficients of the monomials of the above problem, we rewrite it as

$$
\begin{aligned}
\max \quad & \varphi \\
\text{s.t.} \quad & \sum_{k\in G} c_k^1 P_k^d + \sum_{k\in G} c_k^0 - \varphi \\
&= \sum_{k\in G} B_k^{00} + \sum_{(l,m)\in L} C_{lm}^{00} + \sum_{k\in N} \overline{\lambda}_k(P_k^{\max} - P_k^d) \\
&+ \sum_{k\in N} \underline{\lambda}_k(-P_k^{\min} + P_k^d) + \sum_{k\in N} \overline{\gamma}_k(Q_k^{\max} - Q_k^d) \\
&+ \sum_{k\in N} \underline{\gamma}_k(-Q_k^{\min} + Q_k^d) + \sum_{k\in N} \overline{\mu}_k(V_k^{\max})^2 \\
&- \sum_{k\in N} \underline{\mu}_k(V_k^{\min})^2 + \sum_{(l,m)\in L} a_{lm}(S_{lm}^{\max})^2 - \sum_{k\in G} b_k P_k^d \\
& \sum_{k\in N} c_1^k Y_k = A - \sum_{k\in N}\big(\overline{\lambda}_k Y_k + \underline{\lambda}_k Y_k - \overline{\gamma}_k \bar{Y}_k \\
&+ \underline{\gamma}_k \bar{Y}_k - \overline{\mu}_k(V_k^{\max})^2 + \underline{\mu}_k(V_k^{\min})^2 - b_k Y_k\big) \\
&- \sum_{(l,m)\in L}\big(c_{lm} Y_{lm} + d_{lm} \bar{Y}_{lm}\big) \\
& 0 = 2B_k^{12} + b_k \qquad c_k^2 = B_k^{22} \\
& 0 = c_{lm} + 2C_{lm}^{12} \qquad 0 = d_{lm} + 2C_{lm}^{13} \\
& 0 = 2C_{lm}^{23} \qquad 0 = -a_{lm} + C_{lm}^{22} \\
& 0 = -a_{lm} + C_{lm}^{33} \\
& A, B_k, C_{lm} \succeq 0.
\end{aligned}
$$

By substituting some of the variables:

$$
\begin{aligned}
\textbf{[D]} \ \max \quad & \sum_{k\in G} c_k^1 P_k^d + \sum_{k\in G} c_k^0 - \sum_{k\in G} B_k^{00} - \sum_{(l,m)\in L} C_{lm}^{00} \\
&- \sum_{k\in N} \overline{\lambda}_k(P_k^{\max} - P_k^d) - \sum_{k\in N} \underline{\lambda}_k(-P_k^{\min} + P_k^d) \\
&- \sum_{k\in N} \overline{\gamma}_k(Q_k^{\max} - Q_k^d) - \sum_{k\in N} \underline{\gamma}_k(-Q_k^{\min} + Q_k^d) \\
&- \sum_{k\in N} \overline{\mu}_k(V_k^{\max})^2 + \sum_{k\in N} \underline{\mu}_k(V_k^{\min})^2 \\
&- \sum_{(l,m)\in L} C_{lm}^{22}(S_{lm}^{\max})^2 - \sum_{k\in G} 2B_k^{12} P_k^d \\
\text{s.t.} \quad & A = \sum_{k\in N}\Big(c_1^k Y_k + \overline{\lambda}_k Y_k - \underline{\lambda}_k Y_k + \overline{\gamma}_k \bar{Y}_k \\
& -\underline{\gamma}_k \bar{Y}_k + \overline{\mu}_k(V_k^{\max})^2 - \underline{\mu}_k(V_k^{\min})^2 - 2B_k^{12} Y_k\Big) \\
&- \sum_{(l,m)\in L}\big(2C_{lm}^{12} Y_{lm} + 2C_{lm}^{13} \bar{Y}_{lm}\big) \\
& c_k^2 = B_k^{22} \qquad 0 = 2C_{lm}^{23} \\
& C_{lm}^{22} - C_{lm}^{33} = 0 \\
& A, B_k, C_{lm} \succeq 0.
\end{aligned}
$$

which is equivalent to optimization problem 4 described in [10] (i.e., the dual of [OPF-SDP]) and exhibits the same structure. ■