



**HAL**  
open science

# Minimal Sparsity for Second-Order Moment-SOS Relaxations of the AC-OPF Problem

Adrien Le Franc, Victor Magron, Jean-Bernard Lasserre, Manuel Ruiz,  
Patrick Panciatici

► **To cite this version:**

Adrien Le Franc, Victor Magron, Jean-Bernard Lasserre, Manuel Ruiz, Patrick Panciatici. Minimal Sparsity for Second-Order Moment-SOS Relaxations of the AC-OPF Problem. IEEE Transactions on Power Systems, 2023, pp.1 - 10. 10.1109/TPWRS.2023.3333691 . hal-04110742v2

**HAL Id: hal-04110742**

**<https://hal.science/hal-04110742v2>**

Submitted on 7 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Minimal Sparsity for Second-Order Moment-SOS Relaxations of the AC-OPF Problem

Adrien Le Franc, Victor Magron, Jean-Bernard Lasserre, Manuel Ruiz, Patrick Panciatici \*

December 7, 2023

## Abstract

AC-OPF (Alternative Current Optimal Power Flow) aims at minimizing the operating costs of a power grid under physical constraints on voltages and power injections. Its mathematical formulation results in a nonconvex polynomial optimization problem which is hard to solve in general, but that can be tackled by a sequence of SDP (Semidefinite Programming) relaxations corresponding to the steps of the moment-SOS (Sum-Of-Squares) hierarchy. Unfortunately, the size of these SDPs grows drastically in the hierarchy, so that even second-order relaxations exploiting the correlative sparsity pattern of AC-OPF are hardly numerically tractable for large instances — with thousands of power buses. Our contribution lies in a new sparsity framework, termed minimal sparsity, inspired from the specific structure of power flow equations. Despite its heuristic nature, numerical examples show that minimal sparsity allows the computation of highly accurate second-order moment-SOS relaxations of AC-OPF, while requiring far less computing time and memory resources than the standard correlative sparsity pattern. Thus, we manage to compute second-order relaxations on test cases with thousands of power buses, which we believe to be unprecedented.

## Nomenclature

- For a finite set  $F$ , we write  $|F|$  for its cardinality.
- For a complex number  $z \in \mathbb{C}$ , we write  $\angle z$  for its angle;  $|z|$  for its magnitude;  $z^*$  for its complex conjugate;  $\Re(z)$  for its real part; and  $\Im(z)$  for its imaginary part.
- For a pair of integers  $(a, b) \in \mathbb{N}^2$  with  $a \leq b$ , we write  $\llbracket a, b \rrbracket$  for the sequence  $\{a, a + 1, \dots, b\}$ .
- For an  $N \times N$  real symmetric matrix  $M \in \mathbb{S}^N$ ,  $M \succeq 0$  means that  $M$  is positive semidefinite (PSD).
- For real matrices  $(A, B) \in (\mathbb{R}^{N \times N})^2$ , we write  $\langle A, B \rangle$  for the Frobenius inner product between  $A$  and  $B$ .
- For a polynomial function  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  that decomposes as  $\sum_{\alpha \in \mathbb{N}^N} f_{\alpha} x^{\alpha}$  in the standard monomial basis, we denote its support by  $\text{supp}(f) = \{\alpha \in \mathbb{N}^N \mid f_{\alpha} \neq 0\}$ , and the set of variables involved in  $f$  by  $\text{var}(f) = \{n \in \llbracket 1, N \rrbracket \mid \exists \alpha \in \text{supp}(f), \alpha_n \neq 0\}$ .

Whether  $|\cdot|$  denotes cardinality or magnitude is always clear from context.

---

\*Adrien Le Franc, Victor Magron and Jean-Bernard Lasserre are with LAAS CNRS, Toulouse, France (e-mail: adlefranc@laas.fr); Manuel Ruiz and Patrick Panciatici are with Réseau de Transport d'Électricité (RTE), Paris, France

# 1 Introduction

The AC-OPF (Alternative Current - Optimal Power Flow) problem plays a central role for the operational management and expansion planning of AC power grids. Originally introduced in 1962 [1], AC-OPF has now evolved to handle, e.g., discrete or random variables and N-1 security constraints. Nevertheless, even the original AC-OPF problem remains difficult to solve: it amounts to a nonconvex optimization program with thousands of decision variables, and solutions returned by a nonlinear solver might not be globally optimal [2]. The formulation of [1] is also the one adopted in the PGLib benchmark [3] to concentrate research efforts on the handling of nonlinear AC power flows. The latter is a critical technical challenge, as DC linear approximations — sometimes preferred to AC models to mitigate difficulties in advanced formulations — might lead to unrealistic voltage and power solutions [4].

This work positions in the abundant research stream of relaxations of AC-OPF — we refer to [5] for a recent survey. In particular, convex relaxations are motivated by at least two reasons. First, they provide tractable lower bounds allowing to estimate the global optimality gap of an AC-OPF solution. Second, they are used to combine mature methods from stochastic and mixed-integer optimization with advanced AC-OPF models: see e.g. [6] where a cutting-plane model for robust AC-OPF is developed.

In this paper, we follow the approach of [7, 8] to compute lower bounds for AC-OPF instances based on the moment-SOS (Sums-Of-Squares) hierarchy. We recall that the moment-SOS hierarchy introduces a sequence of SDP (Semidefinite Programming) relaxations which benefits from rich properties: it distinguishes from other convex relaxations by offering, e.g., convergence guarantees of the successive values to the global minimum; sufficient conditions for checking exactness at each step of the sequence; the possibility to extract a global minimizer when a stopping criterion is satisfied, and the possibility to incorporate discrete variables. The interested reader is referred to [9] for an introduction to the topic.

Experimentally, the second step of the hierarchy already achieves convergence for most of the AC-OPF test cases of [7, 8, 10]. However, the size of the SDPs involved in the hierarchy grows drastically with the number of AC-OPF variables and with the order of the relaxation, so that this method becomes rapidly intractable. Thus, it is crucial to leverage the sparsity pattern of AC-OPF to make the method competitive [11].

Among the collection of sparse moment-SOS relaxations — surveyed in [12] — the standard *correlative sparsity* pattern has shown some limitations: the corresponding sparse second-order relaxations yield out-of-memory errors for some instances with about a hundred of power buses on a computer with 125 GB of RAM [10]. The most scalable approach to our knowledge seems to be the recent *correlative-term sparsity* framework [13], which enables the computation of partial sparse second-order relaxations for instances with thousands of power buses [14].

Our contribution to such relaxations is threefold.

- First, we introduce a new sparsity framework, that we call *minimal sparsity*. This framework is inspired by correlative sparsity, but builds on the specific structure of the power flow equations to design sparse moment-SOS relaxations that have smaller matrix variables — which is generally preferred by SDP solvers based on interior-point methods [15].

- Second, we measure the accuracy of minimal sparsity on medium scale AC-OPF instances that display large optimality gaps for the first-order moment-SOS relaxation. This experiment lets us appreciate the empirical convergence of minimal sparsity, which is not theoretically grounded so far.

- Third, we investigate the numerical scalability of minimal sparsity on medium to large scale AC-OPF instances from PGLib [3]. We find that our method can handle second-order moment-SOS bounds for test cases with thousands of power buses, which drastically improves over standard correlative sparsity.

The paper is organized as follows. First, in §2, we recall background notions on sparse moment-SOS hierarchies and their application to AC-OPF. Second, in §3, we introduce our new minimal sparsity framework. Third, in §4, we illustrate the strengths of minimal sparsity by computing second-order moment-SOS relaxations of the AC-OPF problem on various numerical test cases.

## 2 Sparse moment-SOS relaxations for the AC-OPF problem

First, in §2.1, we recall the formulation of the AC-OPF problem. Second, in §2.2, we review basic concepts of moment-SOS hierarchies and their applications to AC-OPF. Third, in §2.3, we present background notions on sparse moment-SOS relaxations.

### 2.1 The AC-OPF problem

The AC-OPF problem aims at minimizing the operating costs of a power grid while satisfying power flow balance equations and infrastructure constraints. We model the grid by a directed graph  $(\mathcal{B}, \mathcal{L})$  where nodes  $\mathcal{B}$  represent buses and edges  $\mathcal{L}$  represent power lines. Line orientations model the asymmetry of power flow along transmission lines in AC power grids. The set of generating power units is denoted by  $\mathcal{G}$ . An illustrative  $(\mathcal{B}, \mathcal{L})$  example based on PGLib’s case 14 IEEE is given in Figure 1. For simplicity, we assume here that at most one single line can connect two buses  $(i, j) \in \mathcal{B}^2$ . Parallel

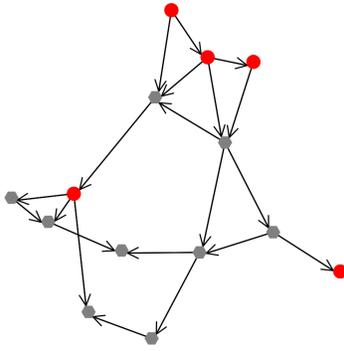


Figure 1: Example of  $(\mathcal{B}, \mathcal{L})$  graph model for PGLib’s case 14 IEEE. Red circle node markers highlight buses with power generators.

lines can be modeled by adequate edge labeling as in [3, Model 1].

Formally, AC-OPF amounts to solving the following optimization problem:

$$\min_{\substack{v \in \mathbb{C}^{|\mathcal{B}|} \\ s \in \mathbb{C}^{|\mathcal{G}|} \\ s^\ell \in \mathbb{C}^{2|\mathcal{L}|}}} \sum_{g \in \mathcal{G}} C_{2,g} \Re(s_g)^2 + C_{1,g} \Re(s_g) + C_{0,g} , \quad (1a)$$

s.t.

$$\forall i \in \mathcal{B}_{\text{ref}}, \quad \angle v_i = 0; \quad (1b)$$

$$\forall g \in \mathcal{G}, \quad \underline{s}_g \leq s_g \leq \overline{s}_g; \quad (1c)$$

$$\forall i \in \mathcal{B},$$

$$\underline{V}_i \leq |v_i| \leq \overline{V}_i, \quad (1d)$$

$$\sum_{g \in \mathcal{G}(i)} s_g - L_i - (Y_i^s)^* |v_i|^2 = \sum_{j \in \mathcal{N}(i)} s_{i,j}^\ell; \quad (1e)$$

$$\forall (i, j) \in \mathcal{L},$$

$$s_{i,j}^\ell = (Y_{i,j} + Y_{i,j}^c)^* \frac{|v_i|^2}{|T_{i,j}|^2} - Y_{i,j}^* \frac{v_i v_j^*}{T_{i,j}}, \quad (1f)$$

$$s_{j,i}^\ell = (Y_{i,j} + Y_{j,i}^c)^* |v_j|^2 - Y_{i,j}^* \frac{v_i^* v_j}{T_{i,j}^*}, \quad (1g)$$

$$|s_{i,j}^\ell| \leq \overline{S}_{i,j}, \quad |s_{j,i}^\ell| \leq \overline{S}_{i,j}, \quad (1h)$$

$$\underline{\Theta}_{i,j} \leq \angle(v_i v_j^*) \leq \overline{\Theta}_{i,j}. \quad (1i)$$

In this formulation, we use lower case letters for decision variables and capital letters for constant parameters. The original decision variables are the bus voltages  $\{v_i\}_{i \in \mathcal{B}}$  and the power generation values  $\{s_g\}_{g \in \mathcal{G}}$ . Additionally, for every edge  $(i, j) \in \mathcal{L}$ , we introduce  $s_{i,j}^\ell$  for the power flow from bus  $i$  to bus  $j$  and  $s_{j,i}^\ell$  for the power flow from bus  $j$  to bus  $i$ .

We now detail the components of Problem (1).

- We minimize power generation costs (1a), which are assumed to only depend on the real part of  $s_g$ , for  $g \in \mathcal{G}$  — that is, on active power generation — with parameters  $(C_{0,g}, C_{1,g}, C_{2,g}) \in \mathbb{R}^3$ .
- In constraint (1b), we set the voltage angle of some reference buses  $\mathcal{B}_{\text{ref}} \subseteq \mathcal{B}$  to zero to address the rotational invariance of voltage solutions.
- In constraints (1c)-(1d), we impose bounds  $(\underline{s}_g, \overline{s}_g) \in \mathbb{C}^2$  on the real and imaginary parts of the generated power  $s_g$ , for  $g \in \mathcal{G}$ , and bounds  $(\underline{V}_i, \overline{V}_i) \in \mathbb{R}_+^2$  on the magnitude of the bus voltage  $v_i$ , for  $i \in \mathcal{B}$ .
- In constraint (1e), we enforce the balance of power flows at every bus  $i \in \mathcal{B}$ . The balance equation involves power generations  $s_g$  for  $g$  in the (possibly empty) set  $\mathcal{G}(i) \subseteq \mathcal{G}$  of generators at bus  $i$ ; power flows  $s_{i,j}^\ell$  for  $j$  in the set  $\mathcal{N}(i) \subseteq \mathcal{B}$  of neighbors of bus  $i$ ; the load  $L_i \in \mathbb{C}$  and a shunt admittance term with  $Y_i^s \in \mathbb{C}$ .
- In constraints (1f)-(1g), we give the expression of power flows  $(s_{i,j}^\ell, s_{j,i}^\ell)$  along every line  $(i, j) \in \mathcal{L}$ , following the  $\Pi$ -circuit branch model with parameters  $(Y_{i,j}, Y_{i,j}^c, Y_{j,i}^c, T_{i,j}) \in \mathbb{C}^4$  detailed in [3, Appendix B].
- In constraints (1h)-(1i), we impose a thermal limit  $\overline{S}_{i,j} \in \mathbb{R}_+$  on power flows and voltage angle difference bounds  $(\underline{\Theta}_{i,j}, \overline{\Theta}_{i,j}) \in \mathbb{R}^2$  for every line  $(i, j) \in \mathcal{L}$ .

Due to constraints (1d)-(1g), Problem (1) is nonconvex and may have spurious local minima [2].

## 2.2 SDP lower bounds via moment-SOS hierarchies

Following the approach of [7, 8], AC-OPF can be cast as a POP (Polynomial Optimization Problem) to benefit from powerful results of the moment-SOS hierarchy. We introduce notations for such a reformulation of Problem (1) and recall some fundamental properties of moment-SOS relaxations.

**From AC-OPF to POP** by considering separately the real and imaginary parts of voltage and power generation variables of Problem (1), we obtain  $N = 2(|\mathcal{B}| + |\mathcal{G}|)$  real variables  $x \in \mathbb{R}^N$  (power flow variables are omitted by injecting (1f)-(1g) into (1e)). The correspondence between AC-OPF and POP variables is formalized by two bijective mappings

$$\mathfrak{b}^r : \mathcal{G} \cup \mathcal{B} \rightarrow \left[ \left[ 1, \frac{N}{2} \right] \right], \quad \mathfrak{b}^{\text{im}} : \mathcal{G} \cup \mathcal{B} \rightarrow \left[ \left[ \frac{N}{2} + 1, N \right] \right], \quad (2a)$$

so that

$$x_{\mathfrak{b}^r(g)} = \Re(s_g), \quad x_{\mathfrak{b}^{\text{im}}(g)} = \Im(s_g), \quad \forall g \in \mathcal{G}, \quad (2b)$$

$$x_{\mathfrak{b}^r(i)} = \Re(v_i), \quad x_{\mathfrak{b}^{\text{im}}(i)} = \Im(v_i), \quad \forall i \in \mathcal{B}. \quad (2c)$$

Then, we observe that all constraints in (1b)-(1i) can be equivalently formulated with  $K + 1$  multivariate polynomials  $\{f_k\}_{k \in \llbracket 0, K \rrbracket}$  of the variable  $x \in \mathbb{R}^N$  — we refer to [16, §5.1] for the explicit expression of the polynomials  $\{f_k\}_{k \in \llbracket 0, K \rrbracket}$ . Thus, Problem (1) can be written as a POP:

$$\rho = \min_{x \in \mathcal{X}} f_0(x), \quad \text{where} \quad (3a)$$

$$\mathcal{X} = \{x \in \mathbb{R}^N \mid f_k(x) \geq 0, \quad \forall k \in \llbracket 1, K \rrbracket\}. \quad (3b)$$

**The Moment-SOS hierarchy** despite its potential nonconvexity, the optimal value of Problem (3) can be approximated — and often exactly computed — by the moment-SOS hierarchy. In this framework, we consider two sequences of SDPs, starting from a minimal order  $r_0 = \max\{d_k\}_{k \in \llbracket 0, K \rrbracket}$  where  $d_k = \lceil \frac{\deg(f_k)}{2} \rceil$ . The moment hierarchy is defined by a sequence of SDPs indexed by the relaxation order  $r \in \llbracket r_0, +\infty \rrbracket$ :

$$\rho_r = \min_y \sum_{\alpha \in \text{supp}(f_0)} f_{0,\alpha} y_\alpha, \quad (4a)$$

$$\text{s.t. } M_r(y) \succeq 0, \quad (4b)$$

$$M_{r-d_k}(f_k y) \succeq 0, \quad \forall k \in \llbracket 1, K \rrbracket, \quad (4c)$$

$$y_{\mathbf{0}} = 1. \quad (4d)$$

The entries of the so-called *pseudo-moment* variable vector  $y$  in Problem (4) are indexed by elements of the truncated monomial basis  $\{x^\alpha\}_{\alpha \in \mathbb{N}_{2r}^N}$ , where  $\mathbb{N}_r^N = \{\alpha \in \mathbb{N}^N \mid \sum_{n \in \llbracket 1: N \rrbracket} \alpha_n \leq r\}$  for  $r \in \mathbb{N}$ . Subsequently, the moment matrix in (4b) and the localization matrices in (4c) are expressed as

$$M_r(y) = (y_{\alpha+\beta})_{\alpha, \beta \in \mathbb{N}_r^N}, \quad (5a)$$

$$M_{r-d_k}(f_k y) = \left( \sum_{\gamma \in \text{supp}(f_k)} f_{k,\gamma} y_{\alpha+\beta+\gamma} \right)_{\alpha, \beta \in \mathbb{N}_{r-d_k}^N}. \quad (5b)$$

These matrices have entries that are linear in the ones of  $y$ , so that we can write  $M_r(y) = \sum_{\alpha \in \mathbb{N}_{2r}^N} A_{0,\alpha} y_\alpha$  and  $M_{r-d_k}(f_k y) = \sum_{\alpha \in \mathbb{N}_{2r}^N} A_{k,\alpha} y_\alpha$  by introducing adequate matrices  $\{A_{k,\alpha}\}_{\alpha \in \mathbb{N}_{2r}^N}$  for all  $k \in \llbracket 0, K \rrbracket$ .

By taking the Lagrangian dual of (4), we obtain the SOS hierarchy of SDPs indexed by  $r \in \llbracket r_0, +\infty \rrbracket$ :

$$\theta_r = \max_{G,t} t, \quad (6a)$$

$$\text{s.t. } f_{0,\mathbf{0}} - t = \sum_{k \in \llbracket 0, K \rrbracket} \langle A_{k,\mathbf{0}}, G_k \rangle, \quad (6b)$$

$$f_{0,\alpha} = \sum_{k \in \llbracket 0, K \rrbracket} \langle A_{k,\alpha}, G_k \rangle, \quad \forall \alpha \in \mathbb{N}_{2r}^N \setminus \{\mathbf{0}\}, \quad (6c)$$

$$G_k \succeq 0, \quad \forall k \in \llbracket 0, K \rrbracket. \quad (6d)$$

In the context of AC-OPF, by adding redundant ball constraints to (3), we can enforce strong duality between Problems (4) and (6) and the convergence of the nondecreasing sequences of lower bounds  $\{\rho_r\}_{r \geq r_0}$  and  $\{\theta_r\}_{r \geq r_0}$  to the value  $\rho$  of (3) (see [9]). However, the sizes of the corresponding SDP relaxations grow drastically with the values of  $N$  and  $r$ , as the largest Gram matrix  $G_0$  in (6) and the moment matrix  $M_r(y)$  in (4) are of size  $|\mathbb{N}_r^N| = \binom{N+r}{r}$ .

### 2.3 Sparse relaxations

One way to bypass the curse of dimensionality mentioned hereabove is to exploit the sparsity of AC-OPF, as initially suggested in [11]. In the context of the moment hierarchy, sparsity consists in reducing the dimension of the search space of Problem (4) by selecting a subset of monomials in  $\{x^\alpha\}_{\alpha \in \mathbb{N}_{2r}^N}$  for indexing the pseudo-moment variable vector  $y$ . We concentrate on correlative sparsity [17], which introduces a hierarchy of sparse moment relaxations:

$$\rho_r(\mathcal{I}) = \min_y \sum_{\alpha \in \text{supp}(f_0)} f_{0,\alpha} y_\alpha, \quad (7a)$$

$$\text{s.t. } M_r(y; \mathcal{I}_p) \succeq 0, \quad \forall p \in \llbracket 1, P \rrbracket, \quad (7b)$$

$$M_{r-d_k}(f_k y; \mathcal{I}_p) \succeq 0, \quad \forall k \in \mathcal{K}_p, \quad (7c)$$

$$\forall p \in \llbracket 1, P \rrbracket,$$

$$y_0 = 1. \quad (7d)$$

Problem (7) is parameterized by a family of subsets of  $\llbracket 1, N \rrbracket$ , denoted  $\mathcal{I} = \{\mathcal{I}_p\}_{p \in \llbracket 1, P \rrbracket}$ , and satisfying  $\cup_{p \in \llbracket 1, P \rrbracket} \mathcal{I}_p = \llbracket 1, N \rrbracket$ . The constraints  $\{f_k\}_{k \in \llbracket 1, K \rrbracket}$  are distributed over a partition  $\{\mathcal{K}_p\}_{p \in \llbracket 1, P \rrbracket}$  of  $\llbracket 1, K \rrbracket$  such that for all  $p \in \llbracket 1, P \rrbracket$  and  $k \in \mathcal{K}_p$ ,  $\text{var}(f_k) \subseteq \mathcal{I}_p$ . Then, for  $p \in \llbracket 1, P \rrbracket$ , the sparse moment and localization matrices in (7b)-(7c) are defined after (5) by selecting only rows and columns indexed by monomials in  $\{x^\alpha\}_{\alpha \in \mathbb{N}_{2r}^N}$  satisfying  $\text{var}(x^\alpha) \subseteq \mathcal{I}_p$ .

As in the dense case, the Lagrangian dual of Problem (7) gives rise to a sparse SOS hierarchy, whose sequence of bounds is introduced as  $\{\theta_r(\mathcal{I})\}_{r \geq r_0}$ .

We remind that the choice of the subsets  $\mathcal{I}$  is of paramount importance. On the practical side, the cardinalities of these subsets control the sizes of the matrices in (7b)-(7c). In general, the smaller these matrices, the better the numerical performances of SDP solvers, especially for those based on interior-point methods [15]. On the theoretical side, the bounds  $\{\rho_r(\mathcal{I})\}_{r \geq r_0}$  are not guaranteed to converge to the value  $\rho$  of the POP (3) for any choice of  $\mathcal{I}$ . The most favorable case is when the subsets  $\mathcal{I}$  satisfy the RIP (Running Intersection Property) where asymptotic convergence is preserved [18]. These considerations on the design of  $\mathcal{I}$  are further investigated in the next section.

## 3 Minimal sparsity for scalable AC-OPF relaxations

We recall basic notions of clique-based sparsity and expose some of its limitations in §3.1. As an alternative, we introduce our minimal sparsity pattern in §3.2. We further detail a method to control the cardinalities of the subsets  $\mathcal{I}$  in §3.3.

### 3.1 Clique-based sparsity and its limitations

We recall how to compute clique-based subsets  $\mathcal{I}$  and discuss some limitations of this approach.

**Clique-based subsets** the design of subsets  $\mathcal{I}$  satisfying the RIP is usually based on the following algorithmic routine.

- (i) First, we define the *correlative sparsity pattern (csp) graph*  $(\mathcal{V}, \mathcal{E})$ . In this graph, the nodes  $\mathcal{V}$  represent the  $N$  variables of the POP (3) and undirected edges  $\mathcal{E}$  account for products between

variables in the polynomial functions  $\{f_k\}_{k \in \llbracket 1, K \rrbracket}$ : an edge  $(n_1, n_2) \in \mathcal{E}$  indicates that there exists  $\alpha \in \cup_{k \in \llbracket 0, K \rrbracket} \text{supp}(f_k)$  such that  $\{n_1, n_2\} \subseteq \text{var}(x^\alpha)$ .

- (ii) Second, we perform a chordal extension of  $(\mathcal{V}, \mathcal{E})$ . We recall that a graph is chordal if each of its cycle of length four or greater has a chord. Therefore, chordal extension adds new edges, resulting in a new graph  $(\mathcal{V}, \overline{\mathcal{E}})$ , where  $\mathcal{E} \subseteq \overline{\mathcal{E}}$ .
- (iii) Third, we define the subsets  $\mathcal{I}^c$  as the nodes of the maximal cliques of the chordal graph  $(\mathcal{V}, \overline{\mathcal{E}})$ . We recall that a clique is a complete subgraph of  $(\mathcal{V}, \overline{\mathcal{E}})$ , and that it is maximal when it cannot be augmented by adding an adjacent node.

We illustrate the process (i) – (iii) for a POP example with  $f_0 = x_2x_5 + x_3x_6 - x_2x_3 - x_5x_6 + x_1(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6)$ ,  $f_1 = 1 - x_1^2 - x_4^2$  and  $N = 6$ . The csp graph (i) and its chordal extension (ii) are given in Figure 2. The maximal cliques (iii) give us the subsets  $\mathcal{I}^c = \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3\}$  with  $\mathcal{I}_1 = \{1, 4\}$ ,  $\mathcal{I}_2 = \{1, 2, 3, 5\}$  and  $\mathcal{I}_3 = \{1, 3, 5, 6\}$ .

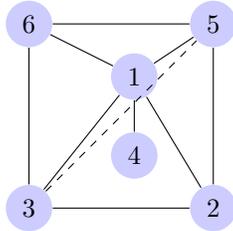


Figure 2: An example of csp graph together with its chordal extension (after adding the dashed edge).

By construction, the clique-based subsets  $\mathcal{I}^c$  automatically satisfy the RIP, and thus ensure the convergence of the correlative sparse moment-SOS hierarchy — we refer to [18] for technical details.

**Limitations of clique-based sparsity** the above routine for designing the subsets  $\mathcal{I}$  gives a systematic way to reduce the computing burden of the dense relaxation (4). However, for large AC-OPF instances, even the sparse relaxation (7) can be numerically challenging. Experimentally, [10] report that the second-order sparse moment relaxation trigger out-of-memory errors on a computer allowed with 125 GB of RAM for the instances 89 PEGASE and 162 IEEE from PGLib [3].

Therefore, some works concentrate on improving the algorithmic routine (i)–(iii) to reduce memory usage and computing time for solving (7). In particular, [19, 20] propose clique merging strategies as a post-processing of (iii). This line of work has allowed up to  $\times 3$  decreases in solving time for first-order relaxations [19]. However, extensions to second-order relaxations seem much less effective [21, §4.5]. We believe that it is due to the iteration complexity of interior-point SDP solvers, which typically perform operations that scale cubically with the size of the largest PSD matrix [15]. We recall that the largest matrix in (7) is of size

$$\binom{m+r}{r}, \quad \text{where } m = \max_{p \in \llbracket 1, P \rrbracket} |\mathcal{I}_p|, \quad (8)$$

hence the importance of moderating the cardinalities of the subsets in  $\mathcal{I}$  to alleviate memory requirements and computing time in second-order sparse relaxations.

### 3.2 Minimal sparsity

We introduce minimal subsets  $\mathcal{I}^m$  to address the principal limitations faced with clique-based sparsity in AC-OPF. Our definition builds on the specific structure of power flow equations: for each bus  $i \in \mathcal{B}$ ,

we select the minimal group of POP variables required to write the power flow balance equation at bus  $i$ . This results in  $P = |\mathcal{B}|$  subsets given by

$$\begin{aligned} \mathcal{I}_{\#i}^m = & \{\mathbf{b}^r(i), \mathbf{b}^{\text{im}}(i)\} \bigcup_{j \in \mathcal{N}(i)} \{\mathbf{b}^r(j), \mathbf{b}^{\text{im}}(j)\} \\ & \bigcup_{g \in \mathcal{G}(i)} \{\mathbf{b}^r(g), \mathbf{b}^{\text{im}}(g)\}, \end{aligned} \quad (9a)$$

where, assuming an arbitrary order on buses  $\mathcal{B}$ , we denote by  $\#i \in \llbracket 1, P \rrbracket$  the position of bus  $i \in \mathcal{B}$ . In term of correspondence between POP and AC-OPF formulations, we obtain the following relationship:

$$\begin{aligned} \{x_n\}_{n \in \mathcal{I}_{\#i}^m} = & \{\Re(v_i), \Im(v_i)\} \bigcup_{j \in \mathcal{N}(i)} \{\Re(v_j), \Im(v_j)\} \\ & \bigcup_{g \in \mathcal{G}(i)} \{\Re(s_g), \Im(s_g)\}. \end{aligned} \quad (9b)$$

The above expression highlights that in  $\mathcal{I}_{\#i}^m$ , we select the minimal amount of AC-OPF variables required to write constraints (1e)-(1g) at bus  $i \in \mathcal{B}$ . Considering any other constraint of Problem (1), it can always be matched to a power flow constraint (1e)-(1g) — hence a set  $\mathcal{I}_{\#i}^m$  — gathering all the variables in its support. As for the objective function (1a), each of its terms can also be dispatched to a minimal subset  $\mathcal{I}_{\#i}^m$ . Consequently, the minimal sets in (9a) are adapted to write the sparse relaxation (7) of Problem (1). If we were to consider another AC-OPF formulation with monomials not captured by (9a) in its objective or constraints — e.g. variable products between voltages  $v_i$  and  $v_j$  that are not neighbors in the power grid — then we would have to extend the minimal sets (9a) to incorporate them.

Minimal sparsity entails a trade-off between the number of subsets in  $\mathcal{I}$  and their cardinalities. We illustrate this trade-off by comparing clique-based subsets  $\mathcal{I}^c$  and minimal subsets  $\mathcal{I}^m$  for PGLib’s case 162 IEEE. We compute the chordal extension  $(\mathcal{V}, \overline{\mathcal{E}})$  and its maximal cliques using the *greedy fillin* heuristic implemented in the TSSOS package [22], as this heuristic yields smaller average clique numbers than other standard heuristics — see [23] for technical details. The histogram of the cardinalities of sets for both sparsity patterns is given in Figure 3. For case 162 IEEE,  $\mathcal{I}^c$  has  $P = 126$  sets, the largest of which has 70 variables, whereas  $\mathcal{I}^m$  has  $P = 162$  sets with at most 22 variables. Consequently, the sparse moment relaxations (7) written with  $\mathcal{I}^m$  have a larger amount of PSD matrices but their dimensions are much smaller: from (8), at order two,  $\mathcal{I}^c$  gives moment matrices of size up to  $2556 \times 2556$  whereas  $\mathcal{I}^m$  gives moment matrices of size at most  $276 \times 276$ . In general, the latter situation is preferred by SDP solvers based on interior-point methods [15].

Lastly, we recall that minimal sparsity is designed to improve the scalability of the sparse relaxation (7), but that it is not known so far whether its convergence to the global minimum of the POP (3) is preserved. Indeed, the definition of  $\mathcal{I}^m$  in (9a) might not enforce the RIP: we report that, for instance, the RIP is not satisfied by  $\mathcal{I}^m$  for cases 3 LMBD and case 5 PJM from PGLib. As a complete convergence analysis would go beyond the scope of this paper, we focus here on numerical evidences, presented in §4.

### 3.3 Finer control on the size of subsets

If the graph  $(\mathcal{B}, \mathcal{L})$  has nodes with a high number of neighbors, the minimal subsets  $\mathcal{I}^m$  defined by (9a) may still have large cardinalities. Assuming that we wish to impose a maximal cardinality threshold  $\bar{I}$  for the subsets  $\mathcal{I}^m$ , we propose a modification of the AC-OPF Problem (1) and of the minimal subsets  $\mathcal{I}^m$  to meet this requirement.

In our approach, when  $|\mathcal{I}_{\#i}^m| > \bar{I}$  at some bus  $i \in \mathcal{B}$ , we split neighboring buses  $\mathcal{N}(i)$  into a partition  $\{\mathcal{N}_a(i)\}_{a \in \mathcal{A}(i)}$ , where the set  $\mathcal{A}(i)$  is introduced to index additional complex variables  $\{z_{i,a}^\ell\}_{a \in \mathcal{A}(i)}$  for

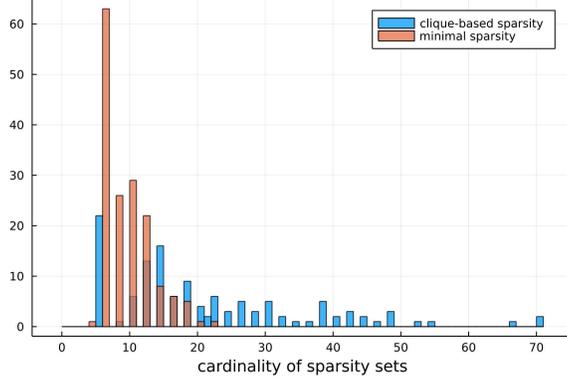


Figure 3: Histogram of the cardinalities of clique-based subsets  $\mathcal{I}^c$  (blue color) and minimal subsets  $\mathcal{I}^m$  (red color) for PGLib's case 162 IEEE

the AC-OPF Problem (1). Then, we rewrite the power flow equation (1e) at bus  $i$  as

$$\sum_{g \in \mathcal{G}(i)} s_g - L_i - (Y_i^s)^* |v_i|^2 = \sum_{a \in \mathcal{A}(i)} z_{i,a}^\ell, \quad (10a)$$

$$z_{i,a}^\ell = \sum_{j \in \mathcal{N}_a(i)} s_{i,j}^\ell, \quad \forall a \in \mathcal{A}(i), \quad (10b)$$

so that each constraint in (10a)-(10b) involves less variables than the original aggregated formulation (1e) — assuming that  $|\mathcal{A}(i)| < |\mathcal{N}(i)|$ . Next, we add  $2|\mathcal{A}(i)|$  real variables to the POP (3) and extend  $\{\mathbf{b}^r, \mathbf{b}^{\text{im}}\}$  so that

$$\begin{cases} \mathbf{b}^r(a) = \Re(z_{i,a}^\ell), \\ \mathbf{b}^{\text{im}}(a) = \Im(z_{i,a}^\ell), \end{cases} \quad \forall a \in \mathcal{A}(i). \quad (11)$$

Finally, we redefine minimal subsets as follows:

$$\begin{aligned} \mathcal{I}_{\#i}^m = & \{\mathbf{b}^r(i), \mathbf{b}^{\text{im}}(i)\} \cup_{a \in \mathcal{A}(i)} \{\mathbf{b}^r(a), \mathbf{b}^{\text{im}}(a)\} \\ & \cup_{g \in \mathcal{G}(i)} \{\mathbf{b}^r(g), \mathbf{b}^{\text{im}}(g)\}, \end{aligned} \quad (12a)$$

$$\begin{aligned} \mathcal{I}_{\#a}^m = & \{\mathbf{b}^r(i), \mathbf{b}^{\text{im}}(i)\} \cup_{j \in \mathcal{N}_a(i)} \{\mathbf{b}^r(j), \mathbf{b}^{\text{im}}(j)\} \\ & \cup \{\mathbf{b}^r(a), \mathbf{b}^{\text{im}}(a)\}, \quad \forall a \in \mathcal{A}(i). \end{aligned} \quad (12b)$$

In turn, the sets  $\mathcal{A}(i)$  and  $\{\mathcal{N}_a(i)\}_{a \in \mathcal{A}(i)}$  should be designed carefully to control the cardinalities of the subsets defined by (12a)-(12b). We suggest to use the solutions of the integer program

$$\min_{(n_{\mathcal{A}}, \bar{n}_a) \in \mathbb{N}^{*2}} n_{\mathcal{A}} \quad \text{s.t.} \quad \begin{cases} 2(\bar{n}_a + 2) \leq \bar{I}, \\ n_{\mathcal{A}} \times \bar{n}_a \geq |\mathcal{N}(i)|, \end{cases} \quad (13a)$$

which admits

$$n_{\mathcal{A}} = \left\lceil \frac{|\mathcal{N}(i)|}{\left\lfloor \frac{\bar{I}}{2} \right\rfloor - 2} \right\rceil \quad \text{and} \quad \bar{n}_a = \left\lceil \frac{|\mathcal{N}(i)|}{n_{\mathcal{A}}} \right\rceil \quad (13b)$$

as a solution, if  $\bar{I} \geq 6$ . The rationale behind the formulation of Problem (13a) is that we want to minimize  $n_{\mathcal{A}} = |\mathcal{A}(i)|$  so as to reduce the cardinality of  $\mathcal{I}_{\#i}^m$  in (12a). Meanwhile, we want to dispatch neighbors equally over the partition  $\{\mathcal{N}_a(i)\}_{a \in \mathcal{A}(i)}$ , which is composed of sets whose cardinalities are

at most  $\bar{n}_a$ . The constraints of Problem (13a) ensure that the subsets  $\{\mathcal{I}_{\#a}^m\}_{a \in \mathcal{A}(i)}$  in (12b) have cardinalities lower than  $\bar{I}$  (first inequality) and that the partition  $\{\mathcal{N}_a(i)\}_{a \in \mathcal{A}(i)}$  covers  $\mathcal{N}(i)$  (second inequality).

Applying the solution (13b), we obtain a reduction of the cardinality of  $\mathcal{I}_{\#i}^m$  provided that

$$\bar{I} \geq 4 + \frac{2|\mathcal{N}(i)|}{|\mathcal{N}(i)| - 1} . \quad (14)$$

However, we might still have that  $|\mathcal{I}_{\#i}^m| > \bar{I}$ . In this case, we can operate a similar partitioning of  $\mathcal{G}(i)$  to reduce the contribution of power generation variables  $\{s_g\}_{g \in \mathcal{G}(i)}$  to the cardinality of  $\mathcal{I}_{\#i}^m$  in (12a).

## 4 Numerical examples

We illustrate the success of minimal sparsity in computing second-order moment-SOS bounds in AC-OPF. In our experiments, we use Mosek 9.3 [24] to solve SDPs and IPOPT [25] for nonlinear programs. Both solvers are applied with their default parameters. We display the results of sparse SOS relaxations, i.e. the dual of (7), as they are usually better handled than moment relaxations by Mosek [24, §7.5]. The interface between data, models and solvers is implemented with JuMP [26] and PowerModels [27]. We run experiments on a 2.10 GHz Intel CPU with 150 GB of RAM. Our code is publicly available<sup>1</sup> and we use open data from [3, 28].

We measure the accuracy of a relaxation in term of its optimality gap

$$\gamma_r(\mathcal{I}) = \frac{\bar{\rho} - \theta_r(\mathcal{I})}{\bar{\rho}} \times 100 , \quad (15)$$

where  $\bar{\rho}$  is an upper bound computed with IPOPT. First, in §4.1, we measure the accuracy of minimal sparsity on modified case 57 IEEE instances that display large optimality gaps for first-order SOS relaxations. Second, in §4.2, we investigate the scalability of minimal sparsity on larger PGLib instances.

### 4.1 Case 57 IEEE modified

We consider the ten modified case 57 IEEE instances from [28, §5.4] displaying the largest optimality gaps at the first-order SOS relaxation. Following [28], we adopt a simplified AC-OPF model for this experiment: limits on power lines and angle differences in (1h)-(1i) are ignored. Since moreover case 57 IEEE has at most one generator per bus, we may consider a voltage-only formulation of Problem (1), and, for the sake of numerical stability, we scale all polynomial coefficients to  $f_{k,\alpha} \in [-1, 1]$ .

We present numerical results obtained with clique-based and minimal sparsity in Table 1.

**Bound accuracy** we observe that the second-order relaxation based on minimal sparsity always achieves zero optimality gap for all of the modified case 57 IEEE instances (Table 1, column 5). This suggests that, despite its heuristic nature, minimal sparsity is suitable to compute tight lower bounds for AC-OPF. In turn, clique-based sparsity performs equally well for the second-order relaxation (Table 1, column 4). Interestingly, for first-order sparse relaxations, the optimality gaps obtained with clique-based sparsity are smaller than the ones of minimal sparsity (Table (1), column 2-3).

**Computing time** the main improvement of minimal-sparsity over a clique-based approach lies in the reduction of computing time. Indeed, evaluating clique-based second-order sparse relaxation bounds  $\theta_2(\mathcal{I}^c)$  requires 3-6 hours of computation per instance (Table 1, column 8), whereas each of their minimal sparsity counterparts  $\theta_2(\mathcal{I}^m)$  can be computed within one minute (Table 1, column 9). We

<sup>1</sup><https://github.com/adrien-le-franc/MomentSOS.jl>

believe that this shrinkage of computing time is due to the reduction of the size of the largest subsets in  $\mathcal{I}$ : with clique-based sparsity, we have  $\max_p(|\mathcal{I}_p^c|) = 26$ , while minimal sparsity features smaller cardinalities with  $\max_p(|\mathcal{I}_p^m|) = 14$ . Lastly, we mention that this way of certifying optimality gaps also outperforms the branch-and-bound technique tested in [28], which achieves an average of 0.16% optimality gap after 120 hours of computation per instance.

| instances | optimality gap (%)        |                           |                           |                           | computing time (s)        |                           |                           |                           |
|-----------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
|           | $\gamma_1(\mathcal{I}^c)$ | $\gamma_1(\mathcal{I}^m)$ | $\gamma_2(\mathcal{I}^c)$ | $\gamma_2(\mathcal{I}^m)$ | $\theta_1(\mathcal{I}^c)$ | $\theta_1(\mathcal{I}^m)$ | $\theta_2(\mathcal{I}^c)$ | $\theta_2(\mathcal{I}^m)$ |
| 84        | 3.05                      | 3.30                      | *0.00                     | 0.00                      | $6.88 \cdot 10^{-1}$      | $1.91 \cdot 10^{-1}$      | $1.97 \cdot 10^4$         | $4.57 \cdot 10^1$         |
| 260       | 1.67                      | 1.85                      | 0.00                      | 0.00                      | $5.90 \cdot 10^{-1}$      | $2.58 \cdot 10^{-1}$      | $1.19 \cdot 10^4$         | $4.35 \cdot 10^1$         |
| 267       | 1.21                      | 1.42                      | 0.00                      | 0.00                      | $6.54 \cdot 10^{-1}$      | $1.86 \cdot 10^{-1}$      | $1.37 \cdot 10^4$         | $4.39 \cdot 10^1$         |
| 299       | 1.92                      | 2.06                      | 0.00                      | 0.00                      | $6.17 \cdot 10^{-1}$      | $1.96 \cdot 10^{-1}$      | $2.23 \cdot 10^4$         | $5.67 \cdot 10^1$         |
| 391       | 1.25                      | 1.54                      | 0.00                      | 0.00                      | $5.53 \cdot 10^{-1}$      | $1.83 \cdot 10^{-1}$      | $1.67 \cdot 10^4$         | $5.27 \cdot 10^1$         |
| 628       | 6.64                      | 6.89                      | 0.00                      | 0.00                      | $6.16 \cdot 10^{-1}$      | $2.00 \cdot 10^{-1}$      | $1.79 \cdot 10^4$         | $5.31 \cdot 10^1$         |
| 683       | 2.32                      | 2.50                      | 0.00                      | 0.00                      | $6.80 \cdot 10^{-1}$      | $1.95 \cdot 10^{-1}$      | $1.44 \cdot 10^4$         | $4.56 \cdot 10^1$         |
| 829       | 2.00                      | 2.21                      | *0.00                     | 0.00                      | $6.39 \cdot 10^{-1}$      | $1.84 \cdot 10^{-1}$      | $1.98 \cdot 10^4$         | $4.28 \cdot 10^1$         |
| 868       | 2.17                      | 2.33                      | 0.00                      | 0.00                      | $6.43 \cdot 10^{-1}$      | $1.92 \cdot 10^{-1}$      | $1.41 \cdot 10^4$         | $4.30 \cdot 10^1$         |
| 974       | 1.92                      | 2.08                      | 0.00                      | 0.00                      | $6.80 \cdot 10^{-1}$      | $2.08 \cdot 10^{-1}$      | $1.45 \cdot 10^4$         | $5.10 \cdot 10^1$         |

Table 1: Results for AC-OPF case 57 modified with clique-based subsets ( $|\mathcal{I}^c| = 38$ ,  $\max_p |\mathcal{I}_p^c| = 26$ ) and minimal sparsity subsets ( $|\mathcal{I}^m| = 57$ ,  $\max_p(|\mathcal{I}_p^m|) = 14$ ). In all cases, the solution returned is primal feasible. Instances for which Mosek terminated with the `SLOW_PROGRESS` status are marked with “\*”

## 4.2 Standard PGLib examples

We now present results on standard AC-OPF (1) for TYP, API and SAD PGLib cases [3]. We select instances with up to 1000 buses and RTE cases with thousands of buses. As larger instances are less numerically stable, we scale both polynomial coefficients to  $f_{k,\alpha} \in [-1, 1]$  and POP variables to  $x_n \in [0, 1]$ . We report the performance of second-order relaxations based on minimal sparsity in Table 2. Following §3.3, we apply a maximal cardinality threshold  $\bar{I} = 12$ , as we find empirically that this value of  $\bar{I}$  gives a good trade-off between the size of PSD matrices — at most  $91 \times 91$ , from (8) — and the number of constraints.

**Our results** second-order minimal sparsity relaxations successfully certify less than 1% of optimality gap for 47 of the 60 instances with up to 1000 buses (Table 2, columns 3, 5, 7). For 8 other instances (with gap values in bold font), Mosek stops at a feasible point with the `SLOW_PROGRESS` termination status, which suggests that the accuracy of the gap  $\gamma_2(\mathcal{I}^m)$  could be further reduced. Lastly, the solver returns an `UNKNOWN_RESULT_STATUS` for the 5 other instances.

Addressing larger AC-OPF instances appears numerically challenging, as Mosek stops with an `UNKNOWN_RESULT_STATUS` for 16 out of the 24 large RTE instances. Moreover, we obtain a negative gap value for case 1951 RTE TYP, which means that its bound  $\theta_2(\mathcal{I}^m)$  should be carefully certified.

Nevertheless, we manage to compute optimality gaps  $\gamma_2(\mathcal{I}^m)$  smaller than 1% for cases with thousands of buses. We believe that these results are unprecedented, and open new perspectives for second-order relaxations of large scale AC-OPF instances.

**Comparison with other approaches** we report that clique-based second-order relaxations  $\theta_2(\mathcal{I}^c)$  trigger out-of-memory errors for most cases of Table 2 — e.g. cases 89 PEGASE and 162 IEEE DTC, in line with the results of [10, Table II]. Comparatively, we report that instances with up to 1000 buses do not consume more than 10 GB of RAM for minimal sparsity bounds  $\theta_2(\mathcal{I}^m)$ .

| PGLib cases  | TYP gap (%)                |                           | API gap (%)                |                           | SAD gap (%)                |                           | $\theta_2(\mathcal{I}^m)$ computing time (s) |                     |                     | $\theta'_1(\mathcal{I}^c)$ avg time (s) |
|--------------|----------------------------|---------------------------|----------------------------|---------------------------|----------------------------|---------------------------|--|---------------------|---------------------|---|
|              | $\gamma'_1(\mathcal{I}^c)$ | $\gamma_2(\mathcal{I}^m)$ | $\gamma'_1(\mathcal{I}^c)$ | $\gamma_2(\mathcal{I}^m)$ | $\gamma'_1(\mathcal{I}^c)$ | $\gamma_2(\mathcal{I}^m)$ | TYP  | API                 | SAD                 |   |
| 3 LMBD       | <b>1.16</b>                | 0.00                      | <b>4.90</b>                | 0.00                      | <b>3.58</b>                | 0.00                      | $5.3 \cdot 10^{-1}$                          | $9.8 \cdot 10^{-1}$ | $5.5 \cdot 10^{-1}$ | $9.0 \cdot 10^{-3}$                     |
| 5 PJM        | <b>14.55</b>               | 0.00                      | <b>4.07</b>                | *0.07                     | 0.00                       | 0.00                      | $7.2 \cdot 10^0$                             | $9.5 \cdot 10^0$    | $5.3 \cdot 10^0$    | $2.4 \cdot 10^{-2}$                     |
| 14 IEEE      | 0.00                       | 0.00                      | <b>5.18</b>                | 0.00                      | 0.09                       | 0.00                      | $2.0 \cdot 10^1$                             | $2.5 \cdot 10^1$    | $1.9 \cdot 10^1$    | $7.8 \cdot 10^{-2}$                     |
| 24 IEEE RTS  | 0.00                       | *0.00                     | <b>6.40</b>                | *0.00                     | * <b>4.37</b>              | *0.00                     | $5.2 \cdot 10^1$                             | $8.7 \cdot 10^1$    | $5.6 \cdot 10^1$    | $1.9 \cdot 10^{-1}$                     |
| 30 AS        | 0.00                       | 0.00                      | <b>42.74</b>               | ?                         | 0.23                       | 0.00                      | $3.0 \cdot 10^1$                             | $6.3 \cdot 10^1$    | $3.3 \cdot 10^1$    | $2.1 \cdot 10^{-1}$                     |
| 30 IEEE      | <b>8.05</b>                | *0.00                     | <b>4.38</b>                | *0.00                     | <b>8.05</b>                | *0.00                     | $3.9 \cdot 10^1$                             | $4.3 \cdot 10^1$    | $4.5 \cdot 10^1$    | $1.8 \cdot 10^{-1}$                     |
| 39 EPRI      | <b>2.00</b>                | 0.00                      | <b>1.76</b>                | *0.16                     | 0.29                       | 0.00                      | $2.5 \cdot 10^1$                             | $2.4 \cdot 10^1$    | $2.3 \cdot 10^1$    | $2.9 \cdot 10^{-1}$                     |
| 57 IEEE      | 0.00                       | 0.00                      | 0.00                       | 0.00                      | 0.05                       | 0.00                      | $4.7 \cdot 10^1$                             | $4.0 \cdot 10^1$    | $4.3 \cdot 10^1$    | $5.5 \cdot 10^{-1}$                     |
| 60 C         | 0.03                       | *0.00                     | <b>11.90</b>               | *0.09                     | <b>2.70</b>                | *0.03                     | $5.9 \cdot 10^1$                             | $1.2 \cdot 10^2$    | $8.9 \cdot 10^1$    | $5.0 \cdot 10^{-1}$                     |
| 73 IEEE RTS  | 0.00                       | *0.00                     | <b>5.55</b>                | * <b>4.67</b>             | <b>2.75</b>                | *0.05                     | $1.9 \cdot 10^2$                             | $2.8 \cdot 10^2$    | $1.9 \cdot 10^2$    | $7.5 \cdot 10^{-1}$                     |
| 89 PEGASE    | 0.02                       | *0.00                     | <b>20.23</b>               | ?                         | 0.00                       | *0.00                     | $8.6 \cdot 10^2$                             | $9.1 \cdot 10^2$    | $8.1 \cdot 10^2$    | $2.1 \cdot 10^0$                        |
| 118 IEEE     | 0.32                       | *0.00                     | <b>20.01</b>               | * <b>9.59</b>             | <b>3.09</b>                | *0.05                     | $2.8 \cdot 10^2$                             | $5.0 \cdot 10^2$    | $3.8 \cdot 10^2$    | $1.7 \cdot 10^0$                        |
| 162 IEEE DTC | <b>5.95</b>                | *0.70                     | <b>7.37</b>                | *0.44                     | <b>5.40</b>                | *0.41                     | $9.5 \cdot 10^2$                             | $1.2 \cdot 10^3$    | $1.1 \cdot 10^3$    | $4.2 \cdot 10^0$                        |
| 179 GOC      | 0.54                       | *0.04                     | <b>10.85</b>               | *0.35                     | <b>4.58</b>                | * <b>4.08</b>             | $2.6 \cdot 10^2$                             | $3.6 \cdot 10^2$    | $2.8 \cdot 10^2$    | $2.3 \cdot 10^0$                        |
| 200 ACTIV    | 0.00                       | *0.00                     | <b>10.05</b>               | *0.00                     | 0.00                       | *0.00                     | $3.3 \cdot 10^2$                             | $5.1 \cdot 10^2$    | $3.2 \cdot 10^2$    | $2.3 \cdot 10^0$                        |
| 240 PSERC    | * <b>2.36</b>              | * <b>1.61</b>             | *0.50                      | *0.24                     | * <b>3.76</b>              | * <b>2.94</b>             | $9.1 \cdot 10^2$                             | $1.4 \cdot 10^3$    | $1.0 \cdot 10^3$    | $4.8 \cdot 10^0$                        |
| 300 IEEE     | * <b>1.94</b>              | *0.00                     | *0.83                      | *0.29                     | *0.72                      | *0.00                     | $9.4 \cdot 10^2$                             | $8.7 \cdot 10^2$    | $1.0 \cdot 10^3$    | $5.7 \cdot 10^0$                        |
| 500 GOC      | *0.10                      | *0.00                     | * <b>4.66</b>              | * <b>2.21</b>             | * <b>5.53</b>              | * <b>3.66</b>             | $1.6 \cdot 10^3$                             | $1.9 \cdot 10^3$    | $1.5 \cdot 10^3$    | $7.5 \cdot 10^0$                        |
| 588 SDET     | * <b>1.35</b>              | *0.25                     | * <b>1.05</b>              | ?                         | * <b>5.14</b>              | *0.22                     | $1.1 \cdot 10^3$                             | $8.5 \cdot 10^2$    | $1.2 \cdot 10^3$    | $1.2 \cdot 10^1$                        |
| 793 GOC      | * <b>1.44</b>              | ?                         | * <b>6.48</b>              | ?                         | * <b>5.04</b>              | * <b>1.72</b>             | $1.1 \cdot 10^3$                             | $1.2 \cdot 10^3$    | $1.3 \cdot 10^3$    | $1.4 \cdot 10^1$                        |
| 1888 RTE     | * <b>6.56</b>              | ?                         | * <b>1.52</b>              | *0.05                     | * <b>3.34</b>              | * <b>2.58</b>             | $4.7 \cdot 10^3$                             | $5.0 \cdot 10^3$    | $4.5 \cdot 10^3$    | $5.3 \cdot 10^1$                        |
| 1951 RTE     | *0.11                      | *-0.01                    | ?                          | *0.18                     | *0.39                      | *0.25                     | $4.7 \cdot 10^3$                             | $5.7 \cdot 10^3$    | $4.5 \cdot 10^3$    | $5.9 \cdot 10^1$                        |
| 2848 RTE     | * <b>1.03</b>              | ?                         | ?                          | ?                         | * <b>1.16</b>              | ?                         | $6.1 \cdot 10^3$                             | $8.4 \cdot 10^3$    | $6.5 \cdot 10^3$    | $6.6 \cdot 10^1$                        |
| 2868 RTE     | *0.23                      | ?                         | * <b>1.08</b>              | ?                         | *0.67                      | *0.39                     | $7.0 \cdot 10^3$                             | $7.6 \cdot 10^3$    | $6.9 \cdot 10^3$    | $7.4 \cdot 10^1$                        |
| 6468 RTE     | *0.73                      | *0.27                     | *0.92                      | ?                         | *0.73                      | ?                         | $1.3 \cdot 10^4$                             | $1.9 \cdot 10^4$    | $1.5 \cdot 10^4$    | $3.0 \cdot 10^2$                        |
| 6470 RTE     | * <b>1.40</b>              | *0.74                     | * <b>1.00</b>              | ?                         | ?                          | ?                         | $1.6 \cdot 10^4$                             | $1.9 \cdot 10^4$    | $1.9 \cdot 10^4$    | $3.0 \cdot 10^2$                        |
| 6495 RTE     | * <b>14.09</b>             | ?                         | ?                          | ?                         | * <b>14.09</b>             | ?                         | $1.5 \cdot 10^4$                             | $1.7 \cdot 10^4$    | $1.7 \cdot 10^4$    | $3.2 \cdot 10^2$                        |
| 6515 RTE     | * <b>5.97</b>              | ?                         | ?                          | ?                         | ?                          | ?                         | $1.8 \cdot 10^4$                             | $1.9 \cdot 10^4$    | $1.5 \cdot 10^4$    | $3.4 \cdot 10^2$                        |

Table 2: Results for AC-OPF PGLib instances for first-order gaps  $\gamma'_1(\mathcal{I}^c)$  based on cliques and second-order gaps  $\gamma_2(\mathcal{I}^m)$  based on minimal sparsity. Instances for which Mosek terminated with the SLOW\_PROGRESS status are marked with “\*” and “?” indicates an UNKNOWN\_RESULT\_STATUS

Thus, as  $\gamma_2(\mathcal{I}^c)$  is often intractable, we restrict our study to first-order gaps. As constraints (1h) involve quartic polynomials, we need to perform a quadratic approximation to define first-order SDP bounds  $\theta'_1(\mathcal{I}^c)$  with gaps  $\gamma'_1(\mathcal{I}^c)$  (Table 2, columns 2, 4, 6). We use the same approximation as in [14], detailed in [16, §5.3], which is most often tighter than the second-order cone relaxation reported in [3].

As expected, computing  $\gamma_2(\mathcal{I}^m)$  still represents a serious time overhead compared to  $\gamma'_1(\mathcal{I}^c)$  (Table 2, columns 8-11). Yet, we consequently reduce second-order computing times: even for cases with thousands of buses, the bounds  $\theta_2(\mathcal{I}^m)$  are computed within the same time as the clique-based  $\theta_2(\mathcal{I}^c)$  for case 57 IEEE (Table 1, column 8).

Regarding bound accuracy, we find that second-order optimality gaps  $\gamma_2(\mathcal{I}^m)$  can be much smaller than first-order gaps  $\gamma'_1(\mathcal{I}^c)$ , and are always tighter when  $\gamma'_1(\mathcal{I}^c)$  is not equal to zero. Minimal sparsity also proves competitive against the 1.5 CS-TSSOS hierarchy: on the 31 instances of [14] and Table 2 where both methods give reliable bounds — i.e. Mosek returns a feasible point —  $\gamma_2(\mathcal{I}^m)$  gives a strictly smaller (hence better) optimality gap in 14 cases, and a strictly larger (hence worse) optimality gap in 8 cases.

## 5 Conclusion

We have introduced minimal sparsity, designed to improve the scalability of second-order moment-SOS relaxations of AC-OPF. Our numerical test cases reveal that minimal sparsity gives very accurate lower bounds, while drastically reducing the computing times and memory requirements over standard clique-based sparse SDP relaxations. Our best achievement is to compute second-order relaxation bounds certifying less than 1% of optimality gaps for instances with thousands of buses. Yet, such large instances remain numerically challenging for state-of-the-art SDP solvers — in line with the conclusions of [14]. Regarding future improvements, we look forward to ongoing progresses in SDP solvers, and pre- or post-processing techniques enforcing numerical stability of SDP relaxations, as presented e.g. in [29].

## Acknowledgments

This work was supported by the PEPS2 FastOPF funded by RTE and AMIES, the EPOQCS grant funded by the LabEx CIMI (ANR-11-LABX-0040), the European Union’s Horizon 2020 program under the Marie Skłodowska-Curie Actions, grant agreement 813211 (POEMA), by ANITI through the French PIA3 program under the Grant agreement n° ANR-19-PI3A-0004 as well as by the National Research Foundation, Prime Minister’s Office, Singapore under its CREATE program. This work was performed using HPC resources from CALMIP (Grant 2023-P23035).

## 6 References

- [1] J. Carpentier, “Contribution à l’étude du dispatching économique,” *Bulletin de la Société Française des électriciens*, vol. 3, no. 8, pp. 431–447, 1962.
- [2] W. A. Bukhsh, A. Grothey, K. I. McKinnon, and P. A. Trodden, “Local solutions of the optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4780–4788, 2013.
- [3] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang *et al.*, “The Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms,” *arXiv preprint arXiv:1908.02788*, 2019.
- [4] K. Baker, “Solutions of dc opf are never ac feasible,” in *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, 2021, pp. 264–268.
- [5] D. K. Molzahn, I. A. Hiskens *et al.*, “A survey of relaxations and approximations of the power flow equations,” *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.
- [6] H. Yang, D. P. Morton, C. Bandi, and K. Dvijotham, “Robust optimization for electricity generation,” *INFORMS Journal on Computing*, vol. 33, no. 1, pp. 336–351, 2021.
- [7] C. Jozs, J. Maeght, P. Panciatici, and J.-C. Gilbert, “Application of the moment-SOS approach to global optimization of the OPF problem,” *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 463–470, 2014.
- [8] D. K. Molzahn and I. A. Hiskens, “Moment-based relaxation of the optimal power flow problem,” in *2014 Power Systems Computation Conference*. IEEE, 2014, pp. 1–7.
- [9] J. B. Lasserre, *An introduction to polynomial and semi-algebraic optimization*. Cambridge University Press, 2015, vol. 52.

- [10] S. Gopinath, H. L. Hijazi, T. Weisser, H. Nagarajan, M. Yetkin, K. Sundar, and R. W. Bent, “Proving global optimality of acopf solutions,” *Electric Power Systems Research*, vol. 189, p. 106688, 2020.
- [11] D. K. Molzahn and I. A. Hiskens, “Sparsity-exploiting moment-based relaxations of the optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3168–3180, 2014.
- [12] V. Magron and J. Wang, “Sparse polynomial optimization: theory and practice,” *Series on Optimization and Its Applications*, World Scientific Press, 2023, to appear.
- [13] J. Wang, V. Magron, J. B. Lasserre, and N. H. A. Mai, “CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization,” *ACM Transactions on Mathematical Software*, vol. 48, no. 4, pp. 1–26, 2022.
- [14] J. Wang, V. Magron, and J. B. Lasserre, “Certifying global optimality of AC-OPF solutions via sparse polynomial optimization,” *Electric Power Systems Research*, vol. 213, p. 108683, 2022.
- [15] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.
- [16] D. Bienstock, M. Escobar, C. Gentile, and L. Liberti, “Mathematical programming formulations for the alternating current optimal power flow problem,” *4OR*, vol. 18, no. 3, pp. 249–292, 2020.
- [17] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 218–242, 2006.
- [18] J.-B. Lasserre, “Convergent SDP-relaxations in polynomial optimization with sparsity,” *SIAM Journal on Optimization*, vol. 17, no. 3, pp. 822–843, 2006.
- [19] D. K. Molzahn, J. T. Holzer, B. C. Lesieutre, and C. L. DeMarco, “Implementation of a large-scale optimal power flow solver based on semidefinite programming,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3987–3998, 2013.
- [20] J. Sliwak, E. D. Andersen, M. F. Anjos, L. Létocart, and E. Traversi, “A clique merging algorithm to solve semidefinite relaxations of optimal power flow problems,” *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1641–1644, 2020.
- [21] J. Sliwak, “Résolution de problèmes d’optimisation pour les réseaux de transport d’électricité de grande taille avec des méthodes de programmation semi-définie positive,” Ph.D. dissertation, Polytechnique Montréal, 2021.
- [22] V. Magron and J. Wang, “TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization,” *Proceedings of MEGA: Effective Methods in Algebraic Geometry*, 2021. [Online]. Available: <https://github.com/wangjie212/TSSOS>
- [23] H. L. Bodlaender and A. M. Koster, “Treewidth computations I. Upper bounds,” *Information and Computation*, vol. 208, no. 3, pp. 259–275, 2010.
- [24] M. ApS, “Mosek modeling cookbook,” 2020.
- [25] A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Mathematical Programming*, vol. 106, pp. 25–57, 2006.
- [26] I. Dunning, J. Huchette, and M. Lubin, “JuMP: A modeling language for mathematical optimization,” *SIAM review*, vol. 59, no. 2, pp. 295–320, 2017.

- [27] C. Coffrin, R. Bent, K. Sundar, Y. Ng, and M. Lubin, “Powermodels. jl: An open-source framework for exploring power flow formulations,” in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, pp. 1–8.
- [28] H. Godard, “Résolution exacte du problème de l’optimisation des flux de puissance,” Ph.D. dissertation, Paris, CNAM, 2019.
- [29] A. Oustry, C. D’Ambrosio, L. Liberti, and M. Ruiz, “Certified and accurate SDP bounds for the ACOPF problem,” *Electric Power Systems Research*, vol. 212, p. 108278, 2022.