



HAL
open science

A FAST ALGORITHM FOR THE SEMI-DEFINITE RELAXATION OF THE STATE ESTIMATION PROBLEM IN POWER GRIDS

Stephane Chretien, Paul Clarkson

► **To cite this version:**

Stephane Chretien, Paul Clarkson. A FAST ALGORITHM FOR THE SEMI-DEFINITE RELAXATION OF THE STATE ESTIMATION PROBLEM IN POWER GRIDS. *Journal of Industrial and Management Optimization*, 2019, 16, pp.431 - 443. 10.3934/jimo.2018161 . hal-02515906

HAL Id: hal-02515906

<https://hal.science/hal-02515906>

Submitted on 25 Mar 2020

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.

A FAST ALGORITHM FOR THE SEMI-DEFINITE RELAXATION OF THE STATE ESTIMATION PROBLEM IN POWER GRIDS

STEPHANE CHRETIEN AND PAUL CLARKSON

ABSTRACT. State estimation in power grids is a crucial step for monitoring and control tasks. It was shown that the state estimation problem can be solved using a convex relaxation based on semi-definite programming. In the present paper, we propose a fast algorithm for solving this relaxation. Our approach uses the Bürer Monteiro factorisation in a special way that solves the problem on the sphere and estimates the scale in a Gauss-Seidel fashion. Simulations results confirm the promising behavior of the method.

1. Introduction. Power networks have been a topic of extensive recent studies, both from an engineering and an applied mathematical viewpoint. The main problems addressed in Power Networks engineering are come from the need to control and monitor large and sometimes very large grids. The problem of estimating the state, i.e. the voltage at each bus, is one of the most basic problems in the field. One of the major difficulties with the state estimation problem is that it boils down to a non convex polynomial least-squares optimisation problem. Fortunately enough, several recent works have addressed this issue. In this paper, we leverage the structure of the problem in order to provide an efficient method based on the breakthrough results of Bürer and Monteiro [8] and the recent discoveries around matrix least-squares problems [29].

1.1. Mathematical background on power networks. The power system states are those parameters that can be used to determine all other parameters of the power system. These are the Node voltage phasor voltage magnitude V_j , phase angle θ_j , the complex power flow: (a) active power flow P_{ij}, P_{ji} , (b) reactive power flow Q_{ij}, Q_{ji} . Recall that the powers are quadratic functions of the voltages and it is sufficient to measure the powers in order to be able to estimate the voltages. Sometimes, one also has access to direct voltage measurements through PMU's. As in every systems, imperfections are often present in the measurement of current and voltage transformer transducers (A/D conversions and Tuning), RTU/IED Data storage, Rounding in calculations, Communication links.

1.2. Previous works on state estimation. Schweppe introduced the state estimation problem in the power systems research community [36]. He defined the state estimator as “a data processing algorithm for converting redundant meter readings and other available information into an estimate of the state of an electric power

2010 *Mathematics Subject Classification.* Primary: 93B, Secondary: 90-08.

Key words and phrases. Power grids, State estimation, Semi-Definite Relaxation, Bürer Monteiro, Optimisation on manifolds.

* Corresponding author: Stéphane Chrétien.

system”. State estimation is nowadays an essential part in almost every energy management system [40].

The state estimation problem is most often addressed via least-squares, which is justified in the case of Gaussian random measurement errors if one appeals to maximum likelihood theory. The main difficulty faced by the engineer in the process of state estimation is the fact that the least-squares problem is a non-convex optimization problem with potentially many stationary points and local optimisers.

Fortunately, in many cases, the State Estimation problem can be successfully approached using convex programming via standard relaxation techniques. The most remarkable result in this spirit is the breakthrough obtained in [28] stating that the power flow problem could be solved exactly via a natural positive semi-definite programming problem. This result has been extended in [32], [24], [30], [5], [20]. The case of State estimation was then studied in [43], [41], [25] and [31]. One of the main drawbacks of the convex relaxation is that it relies on a semi-definite programming which, although solvable in polynomial time, does not scale to large networks. Distributed approaches have been proposed in [43] in order to circumvent this problem. On the other hand, an interesting recent trend based on the Burer Monteiro factorisation [8] is currently extensively explored in the context of various low rank matrix estimation problems in statistics. Surprisingly, the Burer Monteiro was proved to be amenable to gradient-type algorithm with polynomial time complexity in several situations [7], [4], [34] and potential applicability of this approach deserves some attention.

1.3. Goal and organisation of the paper. The goal of the present paper is to propose a new factorisation based approach to the State Estimation problem and report on its efficiency via extensive computational experiments.

The paper is organised as follows. In Section 2, we present the mathematical details on power networks and the main notations used in the sequel. In Section 3, we introduce our factorisation approach to the problem and prove that it provides a solution to the original semi-definite programming relaxation of [43] under the assumption that the observation design satisfies a certain algebraic condition known as RIP [29]. Finally, in Section 4, we present some numerical experiments on some real networks.

2. Problem formulation.

2.1. The estimation problem. In this section, we introduce the state estimation problem in power grids.

2.1.1. Notations. We consider a power network with N buses. Let $\mathcal{N} = \{1, \dots, N\}$ and $\mathcal{E} \subset \mathcal{N}^2$ we index the set of lines. For each bus indexed by n , its neighbours will have index set denoted by \mathcal{N}_n . We will denote by V_n the complex value of the voltage at bus n . We will work in rectangular coordinates as in [43].

2.1.2. Measurement model. We assume to measure the power at certain nodes and certain edges. These measurements are quadratic functions of the voltages at every nodes. More precisely we will observe the power injection at bus n , denoted by P_n (real) and Q_n (reactive), and flows from bus n to bus n' , denoted by $P_{n,n'}$ (real) and $Q_{n,n'}$ (reactive) as well as the squared magnitude of the voltages. The bus

admittance matrix Y is the matrix defined by

$$Y_{nn'} = \begin{cases} -y_{n,n'} & \text{if } (n, n') \in \mathcal{E} \\ \bar{y}_{n,n} + \sum_{\nu \in \mathcal{N}_n} y_{n,\nu} & \text{if } n = n' \\ 0 & \text{otherwise} \end{cases}$$

where $y_{n,n'}$ denotes the line admittance between buses n' and n , $\bar{y}_{n,n}$ is the shunt admittance at bus n . Let also $\bar{y}_{n,n'}$ denote the shunt admittance at bus n associated with line (n, n') . The current flowing from bus n to bus n' is denoted by $I_{n,n'}$ and the injected current at bus n is denoted by I_n . They satisfy

$$I = YV$$

and

$$I_{n,n'} = \bar{y}_{n,n'} V_n + y_{n,n'} (V_n - V_{n'}).$$

The AC power flow model asserts that the complex power injection at bus n is given by

$$P_n + jQ_n = V_n I_n^*,$$

while the complex power flow from bus n to bus n' is given by

$$P_{n,n'} + jQ_{n,n'} = V_n I_{n,n'}^*.$$

Let us collect the measurement values in a vector z as in [43], i.e. the *column* vector

$$z = [\{P_n\}_{n \in \mathcal{N}_P}, \{Q_n\}_{n \in \mathcal{N}_Q}, \{P_{n,n'}\}_{n,n' \in \mathcal{E}_P}, \{Q_{n,n'}\}_{n,n' \in \mathcal{E}_Q}, \{|V_n|^2\}_{n \in \mathcal{N}_V}]^* + \epsilon$$

where ϵ is a noise vector. Note that currently the z vector does not contain the voltage angles. The statistical problem of estimating the voltages from the observation vector z is called the State Estimation problem. Notice that all measurements are noisy quadratic functions of the voltage and this is the reason that makes the voltage estimation problem difficult.

2.2. Semi-Definite relaxation of State Estimation. After the breakthrough paper [28], a extensive effort has been devoted to the study of Semi-Definite relaxations of quadratic least-squares problems with rank one constraint in Power Flow estimation and State Estimation [43].

2.2.1. Notation. We adopt the same notations as [43]. Let e_n denote the vector of all one's in \mathbb{C}^n . Define

$$\begin{aligned} Y_n &= e_n e_n^T Y \\ Y_{n,n'} &= (\bar{y}_{n,n'} + y_{n,n'}) e_n e_n^T - y_{n,n'} e_n e_{n'}^T, \end{aligned}$$

and define $H_{P,n}, H_{Q,n}, \dots$ as

$$\begin{aligned} H_{P,n} &= \frac{1}{2} (Y_n + Y_n^*), & H_{Q,n} &= \frac{j}{2} (Y_n - Y_n^*) \\ H_{P,n,n'} &= \frac{1}{2} (Y_{n,n'} + Y_{n,n'}^*), & H_{Q,n,n'} &= \frac{j}{2} (Y_{n,n'} - Y_{n,n'}^*) \end{aligned}$$

and

$$H_{V,n} = e_n e_n^*.$$

Using these notations, we obtain that

$$\begin{aligned} P_n &= \text{trace}(H_{P,n}VV^*), & Q_n &= \text{trace}(H_{Q,n}VV^*) \\ P_{n,n'} &= \text{trace}(H_{P,n,n'}VV^*), & Q_{n,n'} &= \text{trace}(H_{Q,n,n'}VV^*) \end{aligned}$$

and

$$|V_n|^2 = \text{trace}(H_{V,n}VV^*).$$

2.2.2. The least squares problem and a first Semi-Definite Relaxation. After a small notational change allowing to enumerate the matrices H_l from 1 to L , independent of the subscripts P, Q, n, n' , the least squares estimation problem is thus given by

$$\min_{V \in \mathbb{C}^{N \times N}} \sum_{l=1}^L \left(z_l - \text{trace}(H_l V V^*) \right)^2 \quad (1)$$

Using the change of variable $W = VV^*$, we then have the equivalent rank constrained Semi-Definite Program

$$\min_{W \in \mathbb{C}^{N \times N}} \sum_{l=1}^L \left(z_l - \text{trace}(H_l W) \right)^2 \quad \text{s.t. } W \succeq 0 \quad \text{and} \quad \text{rank}(W) = 1.$$

A standard way to obtain a Semi-Definite relaxation is just to relax the rank one constraint. The resulting Semi-Definite Relaxation is given by

$$\min_{W \in \mathbb{C}^{N \times N}} f(W) \quad \text{s.t. } W \succeq 0 \quad (2)$$

with

$$f(W) = \sum_{l=1}^L \left(z_l - \text{trace}(H_l W) \right)^2. \quad (3)$$

3. The factorisation approach. Our goal in this section is to present a fast method for solving (2) based on the factorisation idea of [8].

3.1. Presentation of the method. We will use a Büler-Monteiro type approach to solving (2). In the Büler-Monteiro philosophy, one uses the factorisation $W = AA^*$ where $A \in \mathbb{C}^{n \times r}$. In this ideal case where the SDR relaxation is exact, one can take $r = 1$ and recover the voltage vector V directly as equal to A up to a 'rotation' R , i.e.

$$V = AR$$

with R a unitary matrix, i.e. $RR^* = I$. Finding the matrix R can be achieved by direct measurements of the voltage value at some particular buses using e.g. PMU's and using simple least squares estimation, a particular case of the Procrustes problem. This situation appears frequently in the State Estimation problems as was proved in the breakthrough paper [28]. In general however, it might happen that the rank at the solution of (2) is larger than one and one might account for this possible issue in the method. Therefore, one usually sets r to a value larger than one in general.

Let $g: \mathbb{C}^{n \times r} \mapsto \mathbb{C}^r$ denote the function

$$g(A) = \sum_{l=1}^L \left(z_l - \text{trace}(H_l AA^*) \right)^2. \quad (4)$$

This function is quartic in A and g has an infinite number of minimisers due to the fact that, for any unitary matrix R , $ARR^*A^* = AA^*$. The idea of Burer and Monteiro [8] is to find a relaxation of higher rank to the SDP leading to a solution of the Semi-Definite Relaxation while avoiding the use of standard SDP optimisation tools which might not scale to large real life problems. This was proved to be optimal for r sufficiently large in [7]. The approach is also known to achieve provable recovery if the matrices H_l , $l = 1, \dots, L$ satisfy a Restricted Isometry Property [4] [29].

Different approaches can be chosen for optimising g , based on first order information, i.e. an oracle which outputs the computation of the gradient $\nabla g(A)$ at a given A . One of the most relevant family of methods is the family of quasi-Newton methods such as BFGS algorithms with efficient line-search [6]. One of the main advantages is the guarantee that the method will reach a local minimizer instead of a saddle point. For a complete bibliography of the available techniques, we refer the reader to [33].

One of the main issues with our optimisation problem is that the gradient is a third order polynomial function of the variable and the Lipschitz constant of the gradient, which is fundamental for analysing the convergence of the method, can be very large even on reasonably small compact sets. As a result, numerical problems such as divergence of the iterates may be observed. A simple remedy to this issue is to incorporate some additional constraints into the problem in order to stabilise the algorithm. In this paper, we chose to use the unit ball of the operator norm. The reason for doing this is based on the following proposition.

Proposition 1. *Problem (1) is equivalent to the following problem:*

$$\min_{\alpha \in \mathbb{R}_+, \|A\| = \sqrt{\alpha}} \sum_{l=1}^L \left(z_l - \alpha \operatorname{trace}(H_l AA^*) \right)^2. \quad (5)$$

Proof. This is straightforward after making the change of variable $V = \sqrt{\alpha} A$ and observing that the feasible sets are the same in the two problems. \square

Based on the splitting approach given in Proposition 1, we now present a very simple and intuitive approach to the problem based on alternating optimisation on the unit ball for the operator norm and scaling the solution:

- The first stage corresponds to solving the optimisation problem over the set of unit norm matrices with k columns by simple projected gradient steps, i.e. we find a stationary point for the problem

$$\min_{\|A\|=1} \sum_{l=1}^L \left(z_l - \alpha \operatorname{trace}(H_l AA^*) \right)^2. \quad (6)$$

- The second stage corresponds to rescaling the matrix by tuning its norm via a least squares criterion, i.e.

$$\min_{\alpha \in \mathbb{R}} \sum_{l=1}^L \left(z_l - \alpha \operatorname{trace}(H_l AA^*) \right)^2 \quad (7)$$

This is done using a simple explicit formula.

The details are given in Algorithm 1 below. One of the main advantages of the method is its simplicity, scalability and speed, as demonstrated through computational experiments in Section 4.

Result: W_{opt}
Choose $A^{(1,1)} \in \mathbb{C}^{n \times k}$

First stage

while $s \leq S - 1$ **do**

$$\nabla g(A) = 2 \sum_{l=1}^L \left(-z_l \alpha(H_l^* + H_l)A + 2 \alpha^2 \text{trace}(H_l A A^*) (H_l^* + H_l)A \right). \quad (8)$$

$$\tilde{A}^{(t,s+1)} = A^{(t,s)} - \eta \nabla g(A^{(t,s)}) \quad (9)$$

$$A^{(t,s+1)} = \frac{1}{\|\tilde{A}^{(t,s+1)}\|} \tilde{A}^{(t,s+1)}. \quad (10)$$

end

Set $A^{(t+1,1)} = A^{(t,S)}$.

Second stage

Set

$$W_{opt} = \alpha^{(t+1)} A^{(t+1,1)} A^{(t+1,1)*} \quad (11)$$

with

$$\alpha^{(t+1)} = \frac{\sum_{l=1}^L z_l \text{trace}(H_l A^{(t+1,1)} A^{(t+1,1)*})}{\sum_{l=1}^L (\text{trace}(H_l A^{(t+1,1)} A^{(t+1,1)*}))^2} \quad (12)$$

Algorithm 1: The two stage optimisation procedure

3.2. Theoretical analysis. In this section, we prove that the method recovers the true solution under standard assumptions on the measurement matrices H_l , $l = 1, \dots, L$.

Let r^* denote the smallest rank of an optimal solution to (2). Throughout this section, we will make the following assumption.

Assumption 1. The family H_l , $l = 1, \dots, L$ of measurement matrices satisfies the following Restricted Isometry Property

$$MI \succeq \nabla^2 f(W) \succeq mI \quad (13)$$

for all A with $\text{rank}(W) \leq 2r^*$

The Restricted Isometry Property played a key role in the field of Compressed Sensing [12] [10]. Compressed sensing is an elegant new paradigm for near optimal data acquisition which has gained paramount interest in the last fifteen years due to its wide applicability and various industrial and medical contexts [14], [2], [19], [27], [21]. Extension of this work to the case of unknown noise variance have been studied in [23], [16], [3], [35], [38]. Interpretation using Lagrange duality and improved reconstruction performance were studied in [15]. Extension of compressed

sensing ideas to low rank matrix and tensor estimation and completion was achieved rapidly after the vector case [11], [13], [9], [18], [42], [37], [17], [39] and the case of unknown noise variance was studied in [26]. The Lagrangian approach of [15] and also extended to the matrix setting in [1].

RIP was also found crucial in the analysis of fast non-convex factorisation methods for low rank estimation of Burer-Monteiro type for Matrix Sensing [29], [22], [44]. The RIP assumption may not hold in some practical settings but we think it is however useful as a conceptual guideline. In the sequel, we will use the main result from [29] in order to study a projected gradient type method on the sphere for the State Estimation problem.

The first step of our method produces a sequence of iterates on the unit sphere in the Frobenius norm. The following lemma characterises the limit behavior of this sequence.

In the sequel, we will assume that $S = +\infty$. The matrix $A^{(t,\infty)}$ will denote an arbitrary cluster point of the sequence $(A^{(t,s)})_{s \in \mathbb{N}}$.

Lemma 1. *For all $t \in \mathbb{N}$, $A^{(t,\infty)}$ is a Karush-Kuhn-Tucker point of the function g on the sphere.*

Proof. Using continuity of the function g , we obtain that the gradient of g is orthogonal to the unit sphere, which is equivalent to saying that $A^{(t,\infty)}$ is a stationary point of the minimisation problem. \square

Given Lemma 1, the following proposition is a standard result in optimisation theory.

Proposition 2. *Any cluster point of the sequence $((\alpha^{(t)}, A^{(t,\infty)}))_{t \in \mathbb{N}}$ is a stationary point of g .*

Proof. Straightforward given that g is differentiable. \square

Theorem 1. *If r^* is the smallest rank of a minimiser in (2), then any cluster point of the sequence $\alpha^{(t)^2} A^{(t,\infty)} A^{(t,\infty)*}$, $t = 1, \dots, +\infty$ is a global optimum of the original least-squares problem (2).*

Proof. Using Proposition 2 and Assumption 1, Theorem 1 in [29] concludes the proof. \square

4. Numerical experiments. We conducted some numerical experiments on various grids including IEEE - 30, a network called Malaga and the Polish network known as 'case2746wp'. In each experiment, we ran 100 Monte Carlo trials. Each trial corresponds to a different realisation of the observation noise, which was generated from a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with $\sigma = .1$. Other values of the noise level have been tested as well and it was observed that the behavior of the method was the same for the tested level values, with a Mean Squared Error increasing linearly as a function of the noise level.

4.1. The IEEE - 30 network. In the experiments with the IEEE - 30 grid, we compared two approaches. The first approach is our new method for problem (2) and the second approach consists in using the YALMIP package in Matlab for solving the Semi-Definite Programming Relaxation. Figure 1 shows that a much

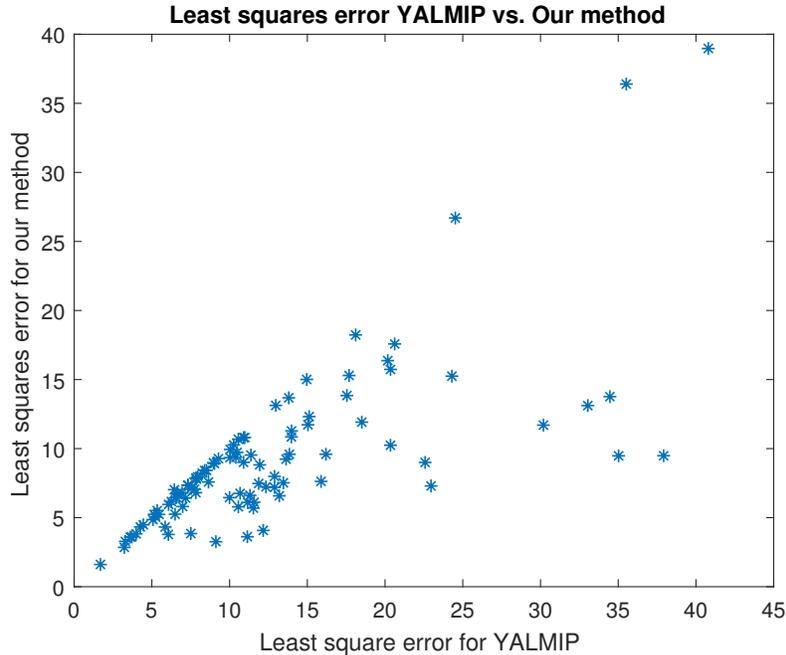


FIGURE 1. Comparison of Sum of Squared Errors for the IEEE-30 network: New method vs. SDP relaxation (using YALMIP) with noise standard deviation equal to .2 when power is observed at half the number of buses chosen at random.

better accuracy is obtained with our new approach. Figure 2 shows that the method converges much faster than YALMIP.

The evolution of the objective value function is presented in Figure 3 below. It shows that, while convergence is not monotonic, it is fast. The evolution of the distance between successive iterates is shown in Figure 4 and shows that a much more robust stopping criterion may be obtained based on this distance than based on the objective value.

4.2. A larger network. We then applied our method on a large network called 'case2746wp'. We quote the description given in [45]. The 'case2746wp' represents the Polish 400, 220 and 110 kV networks during winter 2003-04 evening peak conditions. Multiple centrally dispatchable generators at a bus have not been aggregated. Generators that are not centrally dispatchable in the Polish energy market are given a cost of zero.

Figure 5 shows the histogram of the Mean Squared Error and Figure 6 shows the computation time for the same set of Monte Carlo experiments as for the IEEE-30 network. Again, the method showed good performance despite the larger dimensionality of the network (2746 buses). The YALMIP method did not converge in reasonable computation time for this problem and only the results obtained with our method are shown below.

5. Conclusion. In this paper, we presented a factorisation approach for the State Estimation problem in power networks. Our approach is very efficient in practice

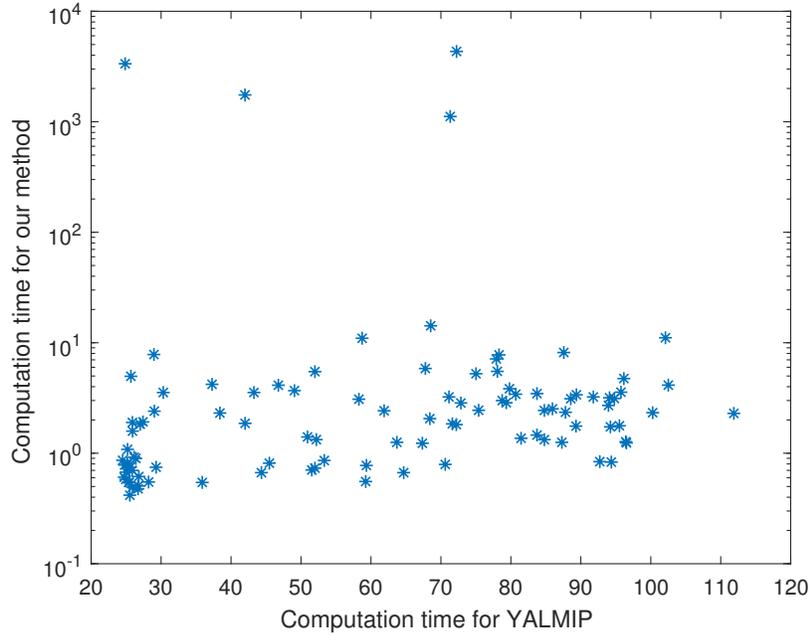


FIGURE 2. Comparison of computation times for the IEEE-30 network: New method vs. SDP relaxation (using YALMIP) with noise standard deviation equal to .2 when power is observed at half the number of buses chosen uniformly at random.

and outperforms e.g. YALMIP for the solution of the Semi-Definite Relaxation of the problem.

Our future objective is to relax the Restricted Isometry Property which is the main ingredient in our analysis. The Restricted Isometry property is an essential theoretical tool which allows to understand if a problem can be efficiently approached by efficient procedures such as in Compressed Sensing or Matrix Sensing. However, one of the main drawbacks of this property is that it is NP-Hard to check and it is very difficult to know if the network at hand has this property.

Acknowledgements. The research leading to this paper is part of the European Metrology Research Program (EMRP), which is jointly funded by participating countries within EURAMET and the EU.

REFERENCES

- [1] Huang Bai, Gang Li, Sheng Li, Qiuwei Li, Qianru Jiang, and Liping Chang. Alternating optimization of sensing matrix and sparsifying dictionary for compressed sensing. *IEEE Transactions on Signal Processing*, 63(6):1581–1594, 2015.
- [2] Richard G Baraniuk. Compressive sensing [lecture notes]. *IEEE signal processing magazine*, 24(4):118–121, 2007.
- [3] Alexandre Belloni, Victor Chernozhukov, Lie Wang, et al. Pivotal estimation via square-root lasso in nonparametric regression. *The Annals of Statistics*, 42(2):757–788, 2014.
- [4] Srinadh Bhojanapalli, Behnam Neyshabur, and Nathan Srebro. Global optimality of local search for low rank matrix recovery. *arXiv preprint arXiv:1605.07221*, 2016.

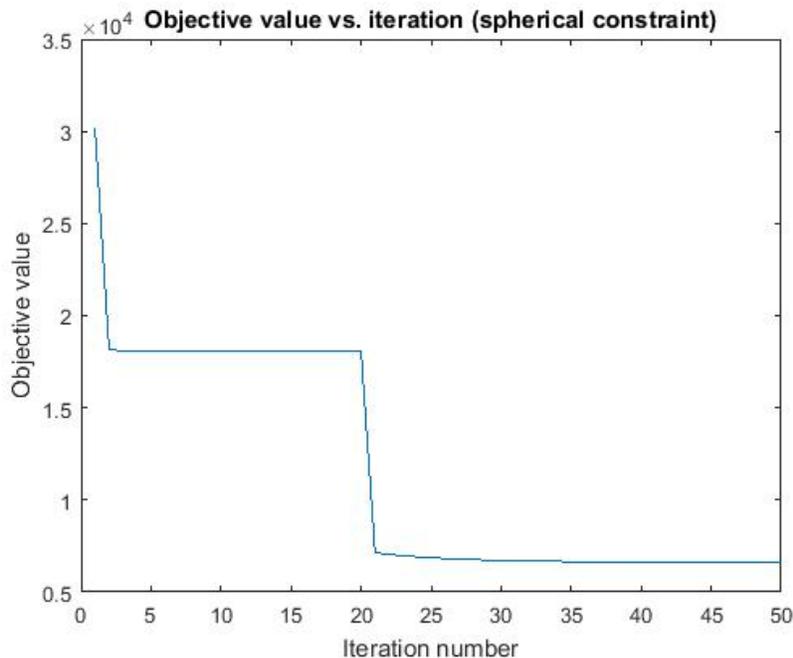


FIGURE 3. Example of evolution of the objective function as a function of iteration number for one realisation of a random noise for the IEEE-30 network.

- [5] D Bienstock and G Munoz. Lp formulations for mixed-integer polynomial optimization problems. *arXiv Preprint*, 2015.
- [6] Joseph-Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, and Claudia A Sagastizábal. *Numerical optimization: theoretical and practical aspects*. Springer Science & Business Media, 2006.
- [7] Nicolas Boumal, Vladislav Voroninski, and Afonso S Bandeira. The non-convex burer-monteiro approach works on smooth semidefinite programs. *arXiv preprint arXiv:1606.04970*, 2016.
- [8] Samuel Burer and Renato DC Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.
- [9] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.
- [10] Emmanuel J Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*, 346(9-10):589–592, 2008.
- [11] Emmanuel J Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717, 2009.
- [12] Emmanuel J Candès and Terence Tao. Decoding by linear programming. *IEEE transactions on information theory*, 51(12):4203–4215, 2005.
- [13] Emmanuel J Candès and Terence Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053–2080, 2010.
- [14] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30, 2008.
- [15] Stéphane Chrétien. An alternating l_1 approach to the compressed sensing problem. *IEEE Signal Processing Letters*, 17(2):181–184, 2010.
- [16] Stéphane Chrétien and Sébastien Darses. Sparse recovery with unknown variance: a lasso-type approach. *IEEE Transactions on Information Theory*, 60(7):3970–3988, 2014.

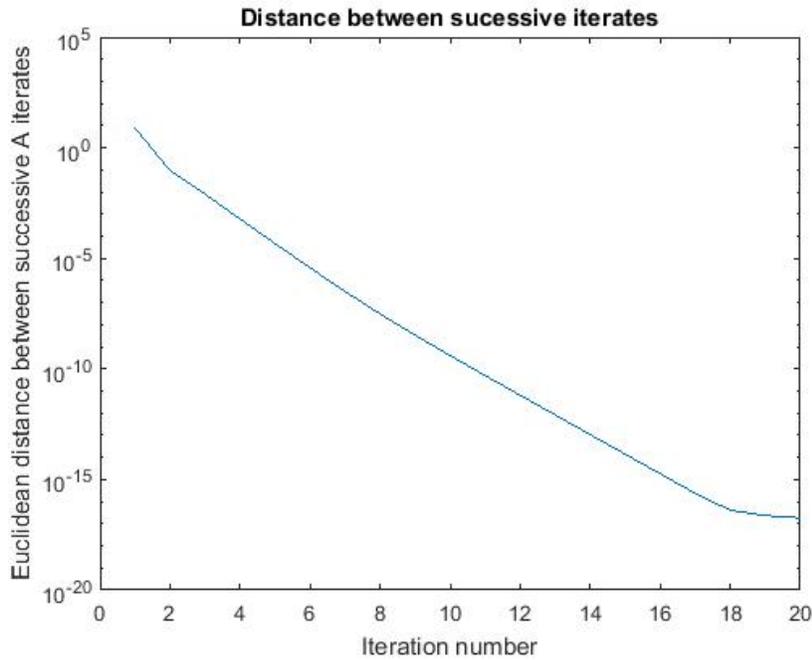


FIGURE 4. Example of evolution of the euclidean distance between successive A -iterates as a function of iteration number for one realisation of a random noise for the IEEE-30 network.

- [17] Stéphane Chrétien and Tianwen Wei. Sensing tensors with gaussian filters. *IEEE Transactions on Information Theory*, 63(2):843–852, 2017.
- [18] Mark A Davenport and Justin Romberg. An overview of low-rank matrix recovery from incomplete observations. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):608–622, 2016.
- [19] Yonina C Eldar and Gitta Kutyniok. *Compressed sensing: theory and applications*. Cambridge University Press, 2012.
- [20] Ghazal Fazelnia, Ramtin Madani, and Javad Lavaei. Convex relaxation for optimal distributed control problem. In *53rd IEEE Conference on Decision and Control*, pages 896–903. IEEE, 2014.
- [21] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*. Birkhäuser Basel, 2013.
- [22] Rong Ge, Chi Jin, and Yi Zheng. No spurious local minima in nonconvex low rank problems: A unified geometric analysis. *arXiv preprint arXiv:1704.00708*, 2017.
- [23] Christophe Giraud, Sylvie Huet, and Nicolas Verzelen. High-dimensional regression with unknown variance. *Statistical Science*, pages 500–518, 2012.
- [24] Rabih A Jabr. Exploiting sparsity in sdp relaxations of the opf problem. *IEEE Transactions on Power Systems*, 2(27):1138–1139, 2012.
- [25] Cecilia Klauber and Hao Zhu. Distribution system state estimation using semidefinite programming. In *North American Power Symposium (NAPS), 2015*, pages 1–6. IEEE, 2015.
- [26] Olga Klopp and Stéphane Gaïffas. High dimensional matrix estimation with unknown variance of the noise. *arXiv preprint arXiv:1112.3055*, 2011.
- [27] Gitta Kutyniok. Theory and applications of compressed sensing. *GAMM-Mitteilungen*, 36(1):79–101, 2013.
- [28] Javad Lavaei and Steven H Low. Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2012.

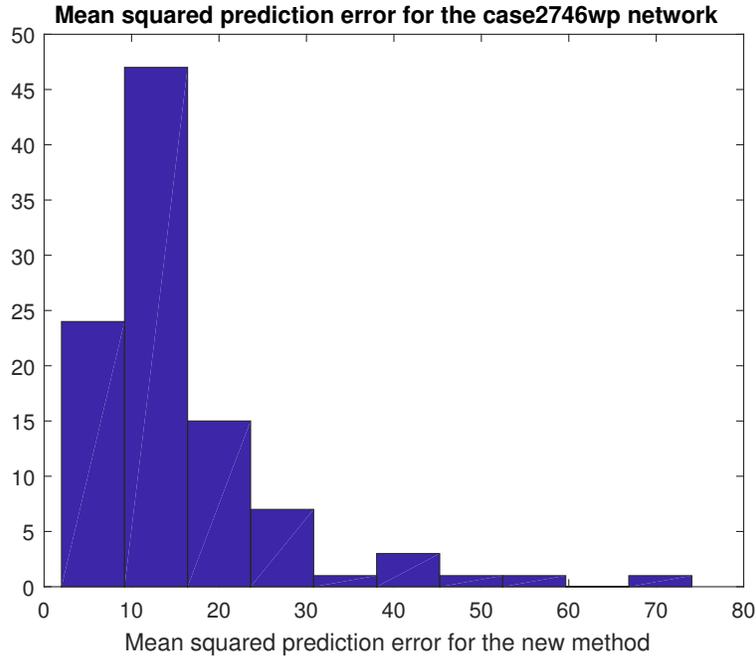


FIGURE 5. Mean Squared Error obtained using the estimator based on the new method with noise standard deviation equal to .2 when power is observed at half the buses. The buses selected for observation were selected uniformly at random.

- [29] Qiuwei Li and Gongguo Tang. The nonconvex geometry of low-rank matrix optimizations with general objective functions. *arXiv preprint arXiv:1611.03060*, 2016.
- [30] Steven H Low. Convex relaxation of optimal power flow, part ii: Exactness. *arXiv preprint arXiv:1405.0814*, 2014.
- [31] Ramtin Madani, Javad Lavaei, and Ross Baldick. Convexification of power flow equations for power systems in presence of noisy measurements. *Preprint*, 2016.
- [32] Daniel K Molzahn, Jesse T Holzer, Bernard C Lesieutre, and Christopher L DeMarco. Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Transactions on Power Systems*, 28(4):3987–3998, 2013.
- [33] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.
- [34] Dohyung Park, Anastasios Kyrillidis, Constantine Caramanis, and Sujay Sanghavi. Non-square matrix sensing without spurious local minima via the burer-monteiro approach. *arXiv preprint arXiv:1609.03240*, 2016.
- [35] Joseph Salmon. *On high dimensional regression: computational and statistical perspectives*. PhD thesis, HDR, École normale supérieure Paris-Saclay, 2017.
- [36] Fred Schweppe. Recursive state estimation: unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13(1):22–28, 1968.
- [37] Qingquan Song, Hancheng Ge, James Caverlee, and Xia Hu. Tensor completion algorithms in big data analytics. *arXiv preprint arXiv:1711.10105*, 2017.
- [38] Alain Viroulet, Agathe Guilloux, Stéphane Gaïffas, and Malgorzata Bogdan. High-dimensional robust regression and outliers detection with slope. *arXiv preprint arXiv:1712.02640*, 2017.
- [39] Andong Wang and Zhong Jin. Near-optimal noisy low-tubal-rank tensor completion via singular tube thresholding. In *Data Mining Workshops (ICDMW), 2017 IEEE International Conference on*, pages 553–560. IEEE, 2017.

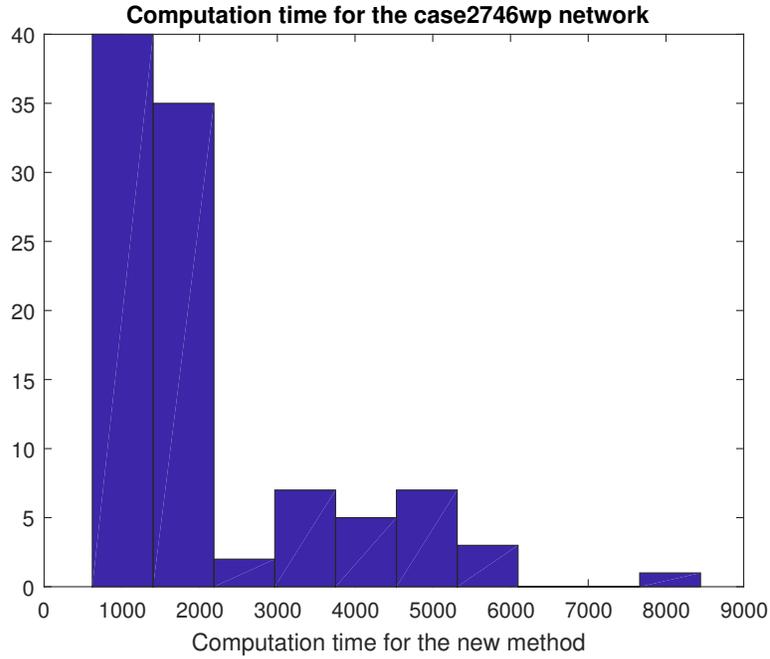


FIGURE 6. Computation times using the new method with noise standard deviation equal to .2 when power is observed at half the buses. The buses selected for observation were selected uniformly at random.

- [40] Felix F Wu. Power system state estimation: a survey. *International Journal of Electrical Power & Energy Systems*, 12(2):80–87, 1990.
- [41] Yu Zhang, Ramtin Madani, and Javad Lavaei. Power system state estimation with line measurements, 2016.
- [42] Zemin Zhang and Shuchin Aeron. Exact tensor completion using t-svd. *IEEE Transactions on Signal Processing*, 65(6):1511–1526, 2017.
- [43] Hao Zhu and Georgios B Giannakis. Power system nonlinear state estimation using distributed semidefinite programming. *IEEE Journal of Selected Topics in Signal Processing*, 8(6):1039–1050, 2014.
- [44] Zhihui Zhu, Qiuwei Li, Gongguo Tang, and Michael B Wakin. The global optimization geometry of nonsymmetric matrix factorization and sensing. *arXiv preprint arXiv:1703.01256*, 2017.
- [45] Ray D Zimmerman, Carlos E Murillo-Sánchez, and Deqiang Gan. Matpower. *PSERC.[Online]. Software Available at: <http://www.pserc.cornell.edu/matpower>*, 1997.
E-mail address: stephane.chretien@npl.co.uk
E-mail address: paul.clarkson@npl.co.uk