Title	Analytic loss minimization: Theoretical framework of a second order optimization method
Authors(s)	Dassios, Ioannis K.
Publication date	2019-01-26
Publication information	Symmetry, 11 (2): 1-11
Publisher	MDPI
Item record/more information	http://hdl.handle.net/10197/10577
Publisher's statement	This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).
Publisher's version (DOI)	10.3390/sym11020136

symmetry

MDPI

Article

Analytic Loss Minimization: Theoretical Framework of a Second Order Optimization Method

Ioannis K. Dassios

AMPSAS, University College Dublin, Dublin 4, Ireland; ioannis.dassios@ucd.ie

Received: 6 December 2018; Accepted: 22 January 2019; Published: 26 January 2019

check for updates

Abstract: In power engineering, the Y_{bus} is a symmetric $N \times N$ square matrix describing a power system network with N buses. By partitioning, manipulating and using its symmetry properties, it is possible to derive the K_{GL} and Y_{GGM} matrices, which are useful to define a loss minimisation dispatch for generators. This article focuses on the case of constant-current loads and studies the theoretical framework of a second order optimization method for analytic loss minimization by taking into account the symmetry properties of Y_{bus}. We define an appropriate matrix functional of several variables with complex elements and aim to obtain the minimum values of generator voltages.

Keywords: symmetric matrix; network; complex derivatives; system; loss minimization

1. Introduction

Electrical power system calculations rely heavily on Y_{bus}, a symmetric square $N \times N$ matrix, which describes a power system network with N buses. It represents the nodal admittance of the buses in a power system, see Figure 1.

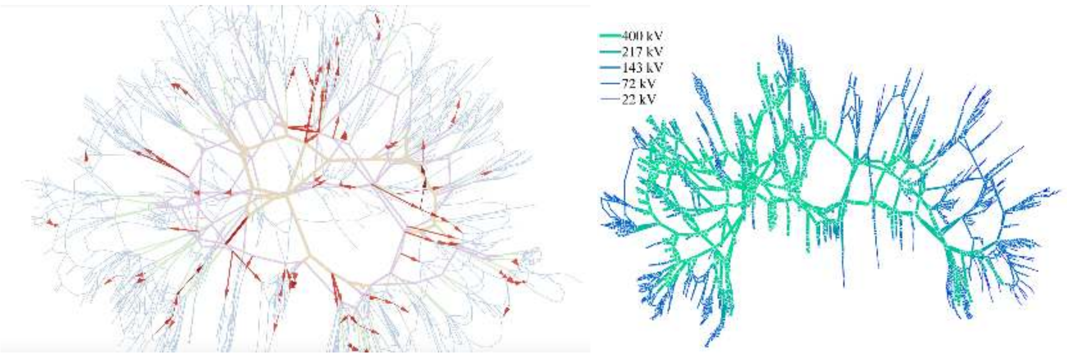

Figure 1. On the left: The red arrows in this network diagram indicate where overloading of power lines could occur in the case2382wp power system, due to short-term fluctuations in (notionally) renewable generator outputs. On the right: The diagram shows the nesta case2224 edin test power system, see [1,2].

By taking advantage of its symmetry properties, it is useful to split the Y_{bus} into sub-matrices and separately quantify the connectivity between load and generation nodes in the network; see [3]. This idea has been also further applied to several power engineering problems [4–7]. Currents (I) and voltages (V) in an electrical power system are related by the symmetric admittance matrix, Y_{bus}, in such a way as to group generator (G) and load (L) nodes separately (see [8,9]):

$$\begin{bmatrix} I_G \\ I_L \end{bmatrix} = Y_{bus} \begin{bmatrix} V_G \\ V_L \end{bmatrix}, \tag{1}$$

where $I_G, V_G \in \mathbb{C}^m$, $I_L, V_L \in \mathbb{C}^n$ and:

$$Y_{bus} = \begin{bmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{bmatrix}.$$

Since Y_{bus} is symmetric, we have $Y_{LG} = Y_{GL}^T$; where $Y_{GG} \in \mathbb{C}^{m\times m}$, $Y_{LL} \in \mathbb{C}^{n\times n}$, $Y_{LG} = Y_{GL}^T \in \mathbb{C}^{n\times m}$ and $n + m = N$ with $n \neq m$. The kernel of Y_{bus} is known and has dimension one, unless the grid graph is disconnected. From this, one can conclude that Y_{LL} is invertible. The use of a pseudo-inverse could be considered only if all nodes are loads. Let $Z_{LL} = Y_{LL}^{-1}$. Then, we can define ([10,11]) two useful sub-matrices:

$$Y_{GGM} = Y_{GG} + Y_{GL}F_{LG}, \tag{2}$$

and:

$$F_{LG} = -Z_{LL}Y_{LG} = -K_{GL}^T. \tag{3}$$

Using (2) and (3), we give (1) the form:

$$\begin{bmatrix} V_L \\ I_G \end{bmatrix} = \begin{bmatrix} Z_{LL} & F_{LG} \\ K_{GL} & Y_{GGM} \end{bmatrix} \begin{bmatrix} I_L \\ V_G \end{bmatrix},$$

from where more insightfully, one can use to derive an expression for this optimal generator dispatch:

$$I_G = K_{GL}I_L + Y_{GGM}V_G. \tag{4}$$

Furthermore, by using these expressions, we can arrive at $V_G^T I_G = V_G^T K_{GL} I_L + V_G^T Y_{GGM} V_G$ and $I_L^T V_L = I_L^T Z_{LL} I_L + I_L^T F_{LG} V_G$. Then (see [12–14]), given that generator powers will be positive and loads negative, the total system loss is given by the sum $V_G^T I_G + I_L^T V_L = V_G^T Y_{GGM} V_G + V_G^T (Y_{GL} Z_{LL} - Y_{LG}^T Z_{LL}^T) I_L + I_L^T Z_{LL}^T I_L$. From the three components in this sum, the circulating current loss is directly a consequence of generator voltage mismatches, i.e., it depends on the product $Y_{GGM}V_G$. How may in general generator voltage mismatches be avoided? It is trivial to achieve a consistent $\|Y_{GGM}V_G\|_1$ profile, as voltage magnitudes are directly controlled by arbitrary set points using automatic reactive power control. Under minimum loss condition, the work in [4] implies that V_G is homogeneous, and by obtaining the min $\|Y_{GGM}V_G\|_1$, the second term of (4), corresponding to current circulated between generators, reduces with the ideal conditions being: $I_G = K_{GL}I_L \approx S_G^{Opt}$. This is equivalent to the loss-minimizing formula presented in [4] and gives two terms for I_G, which the system operator can control by generator dispatch. We used $\|\cdot\|_1$ because of its properties in robustness and sparsity. However, in general, the proposed method could also work by replacing this norm with $\|\cdot\|_2$. Loss minimization in power systems is usually achieved using optimization techniques. There are several methods in the literature that avoid matrix factorizations and have low memory requirements; however, in most of these methods, which are usually first order methods, essential information is missing, and practical convergence is slow. In this article, we propose a second order method that aims to be memory efficient, provide effective computational results and have noticeable progress towards a solution. Note that when loads are constant-power, their current I_L becomes a function of their voltages, and therefore, the different terms in (4) cannot be treated independently. Hence, in a case like that, it would not be clear whether the proposed solution would minimize losses when constant-power loads are present. For this reason, in this article, we focus on and apply the results to the case of constant-current loads.

2. Mathematical Background

We are interested in the following optimization problem:

$$\text{minimize } \|Y_{GGM}V_G\|_1, \quad \text{subject to:} \quad AV_G = b,$$

where:

$$A = \begin{bmatrix} Y_{GG} \\ Y_{LG} \end{bmatrix}, \quad b = \begin{bmatrix} I_G \\ I_L \end{bmatrix} - \begin{bmatrix} Y_{GL} \\ Y_{LL} \end{bmatrix} V_L,$$

and $Y_{GGM}, Y_{GG} \in \mathbb{C}^{m \times m}$, $V_G, I_G - Y_{GL}V_L \in \mathbb{C}^m$, $Y_{LL} \in \mathbb{C}^{n \times n}$, $V_L, I_L - Y_{LL}V_L \in \mathbb{C}^n$. Let:

$$Y_{GGM} = W, \quad \text{and} \quad Y_{GGM}V_G = Y \in \mathbb{C}^m, \quad \text{with} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

In the next section with the main results, we will apply a second order optimization method, i.e., we will use derivatives of first and second order. However, the ℓ_1-norm is not differentiable. Many researchers in the literature use first order optimization methods and apply appropriate smoothing to the problem by using the Huber function. In our case, this is not possible since the Huber function is differentiable, but not twice differentiable. Hence, we propose to replace the ℓ_1-norm with the pseudo-Huber function [15,16], see Figure 2. The pseudo-Huber function parametrized with $\mu > 0$ is:

$$\Psi_\mu(WV_G) = \mu \sum_{i=1}^{m} \left(\sqrt{1 + \frac{y_i \bar{y}_i}{\mu^2}} - 1\right). \tag{5}$$

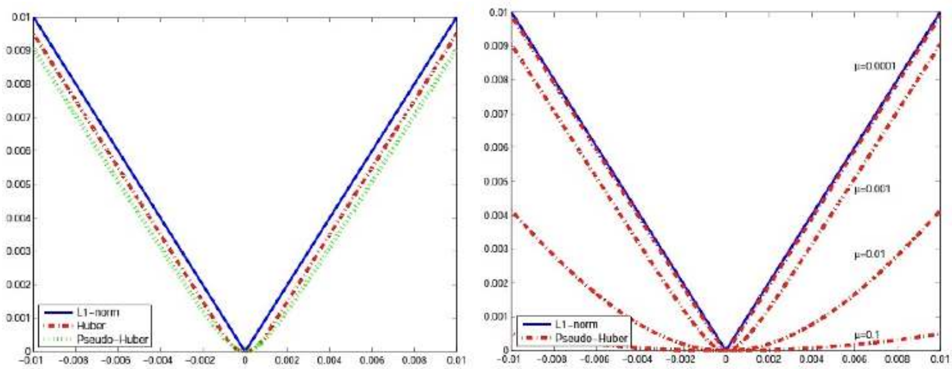

Figure 2. A comparison between the l_1 norm, the Huber & the pseudo–Huber function, and the l_1 norm & pseudo–Huber function for different values of μ; see [15,16].

The gradient of the pseudo-Huber function $\Psi_\mu(WV_G)$ is then given by:

$$\nabla\Psi_\mu(WV_G) = \frac{1}{2\mu} \begin{bmatrix} \frac{y_1}{\sqrt{1+\frac{y_1\bar{y}_1}{\mu^2}}}, & \ldots, & \frac{y_m}{\sqrt{1+\frac{y_m\bar{y}_m}{\mu^2}}}, & \frac{\bar{y}_1}{\sqrt{1+\frac{y_1\bar{y}_1}{\mu^2}}}, & \ldots, & \frac{\bar{y}_m}{\sqrt{1+\frac{y_m\bar{y}_m}{\mu^2}}} \end{bmatrix} \begin{bmatrix} \bar{W} \\ W \end{bmatrix} \tag{6}$$

and the Hessian is given by:

$$\nabla^2\Psi_\mu(WV_G) = \frac{1}{4\mu} \begin{bmatrix} \bar{W}^T & W^T \end{bmatrix} (\operatorname{diag}\begin{bmatrix} \hat{Y}_1, & \ldots, & \hat{Y}_m, & \hat{\bar{Y}}_1, & \ldots, & \hat{\bar{Y}}_m \end{bmatrix} \begin{bmatrix} W \\ \bar{W} \end{bmatrix} +$$

$$+ \operatorname{diag}\begin{bmatrix} Y_1^*, & \ldots, & Y_m^*, & \bar{Y}_1^*, & \ldots, & \bar{Y}_m^* \end{bmatrix} \begin{bmatrix} \bar{W} \\ W \end{bmatrix}), \tag{7}$$

where:

$$\hat{Y}_i = \frac{1}{\sqrt{(1 + \frac{y_1\bar{y}_1}{\mu^2})^3}} + \frac{1}{\sqrt{1 + \frac{y_1\bar{y}_1}{\mu^2}}}$$

and:

$$Y_i^* = -\frac{1}{\mu^2}\frac{y_i^2}{\sqrt{(1+\frac{y_1\bar{y}_1}{\mu^2})^3}}.$$

The following lemma shows that the gradient of the function $\Psi_\mu(WV_G)$ is bounded.

Lemma 1. *The gradient* $\nabla\Psi_\mu(WV_G)$ *satisfies:*

$$-2Km1_m \le \nabla\Psi_\mu(WV_G) \le 2Km1_m,$$

where 1_m *is a vector of ones of length* m*, and:*

$$K = max_{1\le i\le m, 1\le j\le n}\left[\begin{array}{cc} Rew_{ij}, & Imw_{ij}\end{array}\right].$$

Proof. Since:

$$(\mathrm{Re}y_i)^2 \le (\mu)^2 + (\mathrm{Re}y_i)^2 + (\mathrm{Im}y_i)^2, \quad \text{and} \quad (\mathrm{Im}y_i)^2 \le (\mu)^2 + (\mathrm{Re}y_i)^2 + (\mathrm{Im}y_i)^2$$

we get

$$\frac{1}{\mu}\frac{y_i}{\sqrt{1+\frac{y_1\bar{y}_1}{\mu^2}}} = \frac{\mathrm{Re}y_i}{\sqrt{\mu^2+y_1\bar{y}_1}} + i\frac{\mathrm{Im}y_i}{\sqrt{\mu^2+y_1\bar{y}_1}},$$

or equivalently,

$$\frac{1}{\mu}\frac{y_i}{\sqrt{1+\frac{y_1\bar{y}_1}{\mu^2}}} \le 1+i, \quad \text{and} \quad \frac{1}{\mu}\frac{\bar{y}_i}{\sqrt{1+\frac{y_1\bar{y}_1}{\mu^2}}} \le 1-i.$$

Then,

$$\nabla\Psi_\mu(WV_G) \le \frac{1}{2}\left[\begin{array}{cccccc} 1+i, & ..., & 1+i, & 1-i, & ..., & 1-i\end{array}\right]\left[\begin{array}{c}\bar{W}\\ W\end{array}\right],$$

or equivalently,

$$\nabla\Psi_\mu(WV_G) \le$$

$$\frac{1}{2}\left[\begin{array}{ccc} (1+i)\sum_{i=1}^m \bar{w}_{i1} + (1-i)\sum_{i=1}^m w_{i1}, & ..., & (1+i)\sum_{i=1}^m \bar{w}_{in} + (1-i)\sum_{i=1}^m w_{in}\end{array}\right],$$

or equivalently,

$$\nabla\Psi_\mu(WV_G) \le \left[\begin{array}{ccc}\sum_{i=1}^m(\bar{w}_{i1}+w_{i1}), & ..., & \sum_{i=1}^m(\bar{w}_{in}+w_{in})\end{array}\right] \le 2Km1_n.$$

Furthermore, in a similar way, it can be proven that $-2Km1_n \le \nabla\Psi_\mu(WV_G)$. The proof is completed. □

Lemma 2. *The Hessian matrix* $\nabla^2\Psi_\mu(WV_G)$ *satisfies:*

$$0I_n \le \nabla^2\Psi_\mu(WV_G) \le \frac{1}{\mu}LI_n,$$

where $L = \frac{1}{4}\left\|\left[\begin{array}{cc}\bar{W}^T & W^T\end{array}\right]\right\|(2\left\|\left[\begin{array}{c}W\\ \bar{W}\end{array}\right]\right\| + \left\|\left[\begin{array}{c}\bar{W}\\ W\end{array}\right]\right\|)$.

Proof. It is known that for every induced norm $\|\cdot\|$, we have:

$$\lambda \le \left\|\nabla^2\Psi_\mu(WV_G)\right\|,$$

where $\lambda \geq 0$ is a random eigenvalue of $\nabla^2\Psi_\mu(WV_G)$. Observe that:

$$|\hat{Y}_i| \leq \left| \frac{1}{\sqrt{(1+\frac{y_1\bar{y}_1}{\mu^2})^3}} \right| + \left| \frac{1}{\sqrt{1+\frac{y_1\bar{y}_1}{\mu^2}}} \right| \leq 2,$$

and:

$$|Y_i^*| = \left| \frac{\frac{|y_i|^2}{\mu^2}}{\left|1+\frac{y_1\bar{y}_1}{\mu^2}\right|^{\frac{3}{2}}} \right| \leq 1$$

because:

$$\frac{|y_i|^2}{\mu^2} \leq 1 + \frac{y_1\bar{y}_1}{\mu^2} \leq (1+\frac{y_1\bar{y}_1}{\mu^2})^{\frac{3}{2}}.$$

Thus:

$$\left\|\nabla^2\Psi_\mu(WV_G)\right\| \leq$$

$$\frac{1}{4\mu}\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\|(\left\|\text{diag}\begin{bmatrix} \hat{Y}_1, & ..., & \hat{Y}_m, & \hat{\bar{Y}}_1, & ..., & \hat{\bar{Y}}_m \end{bmatrix}\right\|\left\|\begin{bmatrix} W \\ \bar{W} \end{bmatrix}\right\| +$$

$$+\left\|\text{diag}\begin{bmatrix} Y_1^*, & ..., & Y_m^*, & \bar{Y}_1^*, & ..., & \bar{Y}_m^* \end{bmatrix}\right\|\left\|\begin{bmatrix} \bar{W} \\ W \end{bmatrix}\right\|),$$

or equivalently,

$$\left\|\nabla^2\Psi_\mu(WV_G)\right\| \leq \frac{1}{\mu}L.$$

The proof is completed. □

The next lemma shows that the Hessian matrix of the pseudo-Huber function is Lipschitz continuous.

Lemma 3. *The Hessian matrix $\nabla^2\Psi_\mu(WV_G)$ is Lipschitz continuous:*

$$\left\|\nabla^2\Psi_\mu(z) - \nabla^2\Psi_\mu(y)\right\| \leq \frac{1}{\mu^2}M\left\|z-y\right\|.$$

where:

$$M = \frac{1}{2}\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\|\left\|\begin{bmatrix} W \\ \bar{W} \end{bmatrix}\right\| + 2\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\|\left\|\begin{bmatrix} \bar{W} \\ W \end{bmatrix}\right\|.$$

Proof.

$$\left\|\nabla^2\Psi_\mu(z) - \nabla^2\Psi_\mu(y)\right\| = \left\|\int_0^1 \frac{\nabla^2\Psi_\mu(y+s(z-y))}{ds}ds\right\| \leq \left\|\int_0^1 \frac{\nabla^2\Psi_\mu(y+s(z-y))}{ds}\right\|ds,$$

or equivalently,

$$\left\|\nabla^2\Psi_\mu(z)-\nabla^2\Psi_\mu(y)\right\| \leq$$

$$\leq \tfrac{1}{4\mu}\left\|\begin{bmatrix}\bar{W}^T & W^T\end{bmatrix}\right\|\left\|\begin{bmatrix}W\\ \bar{W}\end{bmatrix}\right\|\int_0^1\left\|\tfrac{d}{ds}\mathrm{diag}\begin{bmatrix}\hat{Z}_1, & ..., & \hat{Z}_m, & \hat{\bar{Z}}_1, & ..., & \hat{\bar{Z}}_m\end{bmatrix}\right\|ds+ ,$$

$$+\tfrac{1}{4\mu}\left\|\begin{bmatrix}\bar{W}^T & W^T\end{bmatrix}\right\|\left\|\begin{bmatrix}\bar{W}\\ W\end{bmatrix}\right\|\int_0^1\left\|\tfrac{d}{ds}\mathrm{diag}\begin{bmatrix}Z_1^*, & ..., & Z_m^*, & \bar{Z}_1^*, & ..., & \bar{Z}_m^*\end{bmatrix}\right\|ds.$$

where:

$$\hat{Z}_i=\frac{1}{\sqrt{(1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^3}}+\frac{1}{\sqrt{1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}},$$

and:

$$Z_i^*=-\frac{1}{\mu^2}\frac{[y_i+s(z_i-y_i)]^2}{\sqrt{(1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^3}}.$$

Furthermore:

$$\left\|\frac{d}{ds}(\mathrm{diag}\begin{bmatrix}\hat{Z}_1, & ..., & \hat{Z}_m, & \hat{\bar{Z}}_1, & ..., & \hat{\bar{Z}}_m\end{bmatrix})\right\|=$$

$$=\left\|\mathrm{vec}(\frac{d}{ds}(\mathrm{diag}\begin{bmatrix}\hat{Z}_1, & ..., & \hat{Z}_m, & \hat{\bar{Z}}_1, & ..., & \hat{\bar{Z}}_m\end{bmatrix})\right\|_\infty=$$

$$=\max\left|[\frac{d}{ds}(\mathrm{diag}\begin{bmatrix}\hat{Z}_1, & ..., & \hat{Z}_m, & \hat{\bar{Z}}_1, & ..., & \hat{\bar{Z}}_m\end{bmatrix})]_{ii}\right|=$$

$$=\max\left|[\frac{d}{ds}\frac{1}{\sqrt{(1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^3}}+\frac{1}{\sqrt{1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}}]_{ii}\right|=$$

$$=\max\left|(-\frac{3}{2}\frac{1}{\sqrt{(1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^5}}-\frac{1}{2}(\frac{1}{\sqrt{1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}})^3)\right|\cdot$$

$$\cdot\left|(\frac{1}{\mu^2}([y_i+s(z_i-y_i)][\bar{y}_i-\bar{z}_i]+[y_i-z_i][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)])\right|\leq$$

$$\leq\frac{1}{\mu^2}\|z-y\|\max(\left|\frac{3}{2}\frac{[y_i+s(z_i-y_i)]}{\sqrt{(1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^5}}\right|+\left|\frac{1}{2}\frac{[y_i+s(z_i-y_i)]}{\sqrt{1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}})^3\right|+$$

$$+\left|[\frac{3}{2}\frac{[\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\sqrt{(1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^5}}\right|+\left|\frac{1}{2}(\frac{[\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\sqrt{1+\frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}})^3)\right|)$$

$$\leq\frac{1}{\mu^2}\|z-y\|\frac{\mu}{2}\frac{92}{25\sqrt{5}}<\frac{1}{\mu}\|z-y\|,$$

or equivalently,

$$\left\|\frac{d}{ds}(\mathrm{diag}\begin{bmatrix}\hat{Z}_1, & ..., & \hat{Z}_m, & \hat{\bar{Z}}_1, & ..., & \hat{\bar{Z}}_m\end{bmatrix})\right\|\leq\frac{1}{\mu}\|z-y\| \tag{8}$$

and:

$$\left\| \frac{d}{ds}(\text{diag}\begin{bmatrix} Z_1^*, & ..., & Z_m^*, & \bar{Z}_1^*, & ..., & \bar{Z}_m^* \end{bmatrix}) \right\| =$$

$$= \left\| \text{vec}(\frac{d}{ds}(\text{diag}\begin{bmatrix} Z_1^*, & ..., & Z_m^*, & \bar{Z}_1^*, & ..., & \bar{Z}_m^* \end{bmatrix}) \right\|_\infty =$$

$$= \max \left| [\frac{d}{ds}(\text{diag}\begin{bmatrix} Z_1^*, & ..., & Z_m^*, & \bar{Z}_1^*, & ..., & \bar{Z}_m^* \end{bmatrix})]_{ii} \right| =$$

$$= \max \left| \frac{d}{ds}(-\frac{1}{\mu^2}\frac{[y_i + s(z_i - y_i)]^2}{\sqrt{(1 + \frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2})^3}}) \right| \leq$$

$$\frac{1}{\mu^2}\|y - z\| \max \left| \frac{2[y_i + s(z_i - y_i)]}{\sqrt{(1 + \frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}^{\frac{3}{2}}} \right| +$$

$$+\frac{3}{\mu^2} \left| \frac{[y_i + s(z_i - y_i)]^3}{\sqrt{(1 + \frac{[y_i+s(z_i-y_i)][\bar{y}_i+s(\bar{z}_i-\bar{y}_i)]}{\mu^2}}^{\frac{5}{2}}} \right| \leq$$

$$\leq \frac{32}{25\mu}\|z - y\| < \frac{2}{\mu}\|z - y\|,$$

or equivalently,

$$\left\| \frac{d}{ds}(\text{diag}\begin{bmatrix} Z_1^*, & ..., & Z_m^*, & \bar{Z}_1^*, & ..., & \bar{Z}_m^* \end{bmatrix}) \right\| \leq \frac{2}{\mu}\|z - y\|. \tag{9}$$

From (8) and (9), we get:

$$\left\|\nabla^2\Psi_\mu(z) - \nabla^2\Psi_\mu(y)\right\| \leq \frac{1}{4\mu}\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\| \left\|\begin{bmatrix} W \\ \bar{W} \end{bmatrix}\right\| \frac{1}{\mu}\|z - y\| +$$

$$\frac{1}{4\mu}\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\| \left\|\begin{bmatrix} \bar{W} \\ W \end{bmatrix}\right\| \frac{2}{\mu}\|z - y\|,$$

or equivalently,

$$\left\|\nabla^2\Psi_\mu(z) - \nabla^2\Psi_\mu(y)\right\| \leq \frac{1}{\mu^2}M\|z - y\|,$$

where:

$$M = \frac{1}{2}\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\| \left\|\begin{bmatrix} W \\ \bar{W} \end{bmatrix}\right\| + 2\left\|\begin{bmatrix} \bar{W}^T & W^T \end{bmatrix}\right\| \left\|\begin{bmatrix} \bar{W} \\ W \end{bmatrix}\right\|.$$

The proof is completed. □

3. Main Results

In this section, we will present our main results and the proposed optimization method. We begin with a proposition that will be useful in the main Theorem.

Proposition 1. *The i^{th} row of the matrix Y_{GGM}, given in (2), sums to d_i, $\forall = 1, 2, ..., m$; where d_i is the sum of the i^{th} row of the matrix $Y_{GG} + Y_{GL}$ as defined in (1).*

Proof. From (3):

$$F_{LG} = -Z_{LL}Y_{LG},$$

whereby multiplying by 1_n, we get:

$$F_{LG}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = -Z_{LL}Y_{LG}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}.$$

Let c_i be the sum of the i^{th} row of the matrix $Y_{LL} + Y_{LG}$ as defined in (1). Then:

$$F_{LG}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = -Z_{LL}(\begin{bmatrix}c_1\\c_2\\\vdots\\c_m\end{bmatrix} - Y_{LL}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}),$$

or equivalently,

$$F_{LG}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = Y_{LL}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} - Z_{LL}\begin{bmatrix}c_1\\c_2\\\vdots\\c_m\end{bmatrix}.$$

For the sum of each row of Y_{GGM}, we have:

$$Y_{GGM}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = (Y_{GG} + Y_{GL}F_{LG})\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = Y_{GG}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} + Y_{GL}(\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} - Z_{LL}\begin{bmatrix}c_1\\c_2\\\vdots\\c_m\end{bmatrix}),$$

or equivalently,

$$Y_{GGM}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = (Y_{GG} + Y_{GL})\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} - Y_{GL}Z_{LL}\begin{bmatrix}c_1\\c_2\\\vdots\\c_m\end{bmatrix},$$

or equivalently, from [10]:

$$Y_{GGM}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \begin{bmatrix}d_1\\d_2\\\vdots\\d_m\end{bmatrix}.$$

The proof is completed. □

Theorem 1. *We consider the following optimization problem:*

$$\textit{minimize}_{V_G}\ \|Y_{GGM}V_G\|_1, \quad \textit{subject to:} \quad AV_G = b, \tag{10}$$

where:

$$A = \begin{bmatrix}Y_{GG}\\Y_{LG}\end{bmatrix}, \quad b = \begin{bmatrix}I_G\\I_L\end{bmatrix} - \begin{bmatrix}Y_{GL}\\Y_{LL}\end{bmatrix}V_L,$$

and $Y_{GGM}, Y_{GG} \in \mathbb{C}^{m\times m}$, $I_G - Y_{GL}V_L \in \mathbb{C}^m$, $Y_{LG} \in \mathbb{C}^{n\times m}$, $I_L - Y_{LL}V_L \in \mathbb{C}^n$. Then, by using a second order optimization method, an approximate solution of (10) in respect of $V_G \in \mathbb{C}^m$ is given by the solution of the linear system:

$$\tilde{A}V_G = \gamma\nabla\psi_\mu(Y_{GGM}V_G^{(0)}) + A^*(AV_G^{(0)} - b) + \tilde{A}V_G^{(0)}, \tag{11}$$

where $\tilde{A} = [\gamma\nabla^2\psi_\mu(Y_{GGM}V_G^{(0)}) + A^*A]$, γ and μ are a priori-chosen scalars, $\psi_\mu(Y_{GGM}V_G^{(0)})$ is the pseudo-Huber function given by (5) and:

$$V_G^{(0)} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad Y_{GGM}V_G^{(0)} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix},$$

where d_i is the sum of the i^{th} row of the matrix $Y_{GG} + Y_{GL}$ as defined in (1).

Proof. We have the following optimization problem:

$$\text{minimize}_{V_G} \; \|Y_{GGM}V_G\|_1, \quad \text{subject to:} \quad AV_G = b.$$

In this case, the optimal solution of the following ℓ_1-analysis problem:

$$\text{minimize } f_\gamma(V_G) := \gamma\|Y_{GGM}V_G\|_1 + \frac{1}{2}\|AV_G - b\|_2^2,$$

is proven to be a good approximation to V_G; where γ is an a priori-chosen positive scalar and $\|\cdot\|_2$ is the Euclidean norm. Let:

$$Y_{GGM}V_G = Y \in \mathbb{C}^m, \quad \text{with} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Since we will apply a second order optimization method, we will use derivatives of first and second order. However, the ℓ_1-norm is not differentiable. As mentioned in Section 2, we can apply appropriate smoothing to the problem by using the pseudo-Huber function as defined in (5), which is a twice-differentiable function. Hence, we propose to replace the ℓ_1-norm with the pseudo-Huber function [15]. By using (5), the pseudo-Huber function parametrized with $\mu > 0$ is:

$$\psi_\mu(Y_{GGM}V_G) = \psi_\mu(Y) = \mu\sum_{i=1}^{m}\left(\sqrt{1 + \frac{y_i\bar{y}_i}{\mu^2}} - 1\right),$$

where $\bar{y}_i$ is the complex conjugate of y_i and μ controls the quality of approximation, i.e., for $\mu \to 0$, then $\psi_\mu(x)$ tends to the ℓ_1-norm. Our optimization problem is then approximated by:

$$\text{minimize } f_\gamma^\mu(V_G) := \gamma\psi_\mu(Y_{GGM}V_G) + \frac{1}{2}\|AV_G - b\|_2^2.$$

Note that $f_\gamma^\mu : \mathbb{C}^m \to \mathbb{C}$. From Lemmas 1 and 2, it can be observed that the Hessian of f_γ^μ is bounded and from Lemma 3 that it is Lipschitz continuous.

A second order approximation of f_γ^μ at a given state $V_G^{(0)}$ is:

$$\tilde{f}_\gamma^\mu(V_G) = f_\gamma^\mu(V_G^{(0)}) + \nabla f_\gamma^\mu(V_G^{(0)})^*(V_G - V_G^{(0)}) + \frac{1}{2}(V_G - V_G^{(0)})^*\nabla^2 f_\gamma^\mu(V_G^{(0)})(V_G - V_G^{(0)}).$$

With *, we denote the conjugate transpose. Note that $\nabla f_{\gamma}^{\mu}(Y_0)$ is $m \times 1$ and $\nabla^2 f_{\gamma}^{\mu}(Y_0)$ is $m \times m$. For the optimality condition at V_G^{opt}, we set:

$$\nabla \tilde{f}_{\gamma}^{\mu}(V_G^{opt})^* = 0_{1,m},$$

or equivalently,

$$\nabla f_{\gamma}^{\mu}(V_G^{(0)})^* + (V_G - V_G^{(0)})^* \nabla^2 f_{\gamma}^{\mu}(V_G^{(0)}) = 0_{1,m},$$

or equivalently,

$$\nabla f_{\gamma}^{\mu}(V_G^{(0)}) + \nabla^2 f_{\gamma}^{\mu}(V_G^{(0)})(V_G - V_G^{(0)}) = 0_{m,1}.$$

Hence:

$$\gamma \nabla \psi_{\mu}(Y_{GGM} V_G^{(0)}) + Y_{GG}^*(A V_G^{(0)} - b) + [\gamma \nabla^2 \psi_{\mu}(Y_{GGM} V_G^{(0)}) + A^* A](V_G - V_G^{(0)}) = 0_{m,1},$$

and consequently:

$$[\gamma \nabla^2 \psi_{\mu}(Y_{GGM} V_G^{(0)}) + A^* A](V_G - V_G^{(0)}) = \gamma \nabla \psi_{\mu}(Y_{GGM} V_G^{(0)}) + A^*(A V_G^{(0)} - b).$$

We may choose $V_G^{(0)} = 1_m$. Then, from Proposition 1:

$$Y_{GGM} V_G^{(0)} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}.$$

The proof is completed. □

4. Conclusions

This work has derived a new optimization method for loss minimization of power systems. We defined the function of several variables with complex coefficients that describes this optimization problem and obtained the minimum values of generator voltages, so that the active power losses reduce the irreducible component, which arises from serving load currents. For this purpose, we proposed a second order method and provided all the theoretical framework needed. This new result on the optimization of network topology may bring insights from the established literature on graph analysis to bear on electrical engineering problems. In particular, several parts in the method proposed are written generally and in a way that it can be applied to a much bigger class of problems, including sparsity-promoting fitting problems.

Funding: This material is supported by the Science Foundation Ireland (SFI), by funding Ioannis Dassios under Investigator Programme Grant No. SFI/15/IA/3074.

Conflicts of Interest: The author declares no conflict of interest.

References

1. MD1: Structural and Dynamic Modelling of Integrated Energy Systems. Available online: https://esipp.ie/research/modelling (accessed on 26 January 2019).
2. Dassios, I.; Keane, A.; Cuffe, P. Calculating Nodal Voltages Using the Admittance Matrix Spectrum of an Electrical Network. *Mathematics* **2019**, *7*, 106. [CrossRef]
3. Kessel, P.; Glavitsch, H. Estimating the voltage stability of a power system. *IEEE Trans. Power Deliv.* **1986**, *1*, 346–354. [CrossRef]
4. Thukaram, D.; Vyjayanthi, C. Relative electrical distance concept for evaluation of network reactive power and loss contributions in a deregulated system. *IET Gener. Transm. Distrib.* **2009**, *3*, 1000–1019. [CrossRef]

5. Visakha, K.; Thukaram, D.; Jenkins, L. Transmission charges of power contracts based on relative electrical distances in open access. *Electr. Power Syst. Res.* **2004**, *70*, 153–161. [CrossRef]
6. Vyjayanthi, C.; Thukaram, D. Evaluation and improvement of generators reactive power margins in interconnected power systems. *IET Gener. Transm. Distrib.* **2011**, *5*, 504–518. [CrossRef]
7. Yesuratnam, G.; Thukaram, D. Congestion management in open access based on relative electrical distances using voltage stability criteria. *Electr. Power Syst. Res.* **2007**, *77*, 1608–1618. [CrossRef]
8. Merris, R. Laplacian matrices of graphs: A survey. *Linear Algebra Appl.* **1994**, *197*, 143–176. [CrossRef]
9. Sanchez-Garcia, R.J.; Fennelly, M.; Norris, S.; Wright, N.; Niblo, G.; Brodzki, J.; Bialek, J.W. Hierarchical spectral clustering of power grids. *IEEE Trans. Power Syst.* **2014**, *29*, 2229–2237. [CrossRef]
10. Cuffe, P.; Dassios, I.; Keane, A. Analytic Loss Minimization: A Proof. *IEEE Trans. Power Syst.* **2016**, *31*, 3322–3323. [CrossRef]
11. Dassios, I.K.; Cuffe, P.; Keane, A. Visualizing voltage relationships using the unity row summation and real valued properties of the FLG matrix. *Electr. Power Syst. Res.* **2016**, *140*, 611–618. [CrossRef]
12. Abdelkader, S.M.; Morrow, D.J.; Conejo, A.J. Network usage determination using a transformer analogy. *IET Gener. Transm. Distrib.* **2014**, *8*, 81–90. [CrossRef]
13. Abdelkader, S.M.; Flynn, D. A new method for transmission loss allocation considering the circulating currents between generator. *Eur. Trans. Electr. Power* **2010**, *20*, 1177–1189. [CrossRef]
14. Abdelkader, S.M. Characterization of transmission losses. *IEEE Trans. Power Syst.* **2011**, *26*, 392–400. [CrossRef]
15. Dassios, I.; Fountoulakis, K.; Gondzio, J. A preconditioner for a primal-dual newton conjugate gradient method for compressed sensing problems. *SIAM J. Sci. Comput.* **2015**, *37*, A2783–A2812. [CrossRef]
16. Dassios, I.; Baleanu, D. Optimal solutions for singular linear systems of Caputo fractional differential equations. *Math. Methods Appl. Sci.* **2018**. [CrossRef]