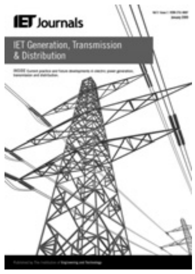
www.ietdl.org

Published in IET Generation, Transmission & Distribution
Received on 16th August 2012
Revised on 17th December 2012
Accepted on 22nd January 2013
doi: 10.1049/iet-gtd.2012.0473



ISSN 1751-8687

State estimation in power systems using linear model infinity norm-based trust region approach

Seshadri Sravan Kumar V., Thukaram Dhadbanjan

Department of Electrical Engineering, Indian Institute of Science, Bangalore 560012, India
E-mail: dtram@ee.iisc.ernet.in

Abstract: State estimation is one of the most important functions in an energy control centre. An computationally efficient state estimator which is free from numerical instability/ill-conditioning is essential for security assessment of electric power grid. Whereas approaches to successfully overcome the numerical ill-conditioning issues have been proposed, an efficient algorithm for addressing the convergence issues in the presence of topological errors is yet to be evolved. Trust region (TR) methods have been successfully employed to overcome the divergence problem to certain extent. In this study, case studies are presented where the conventional algorithms including the existing TR methods would fail to converge. A linearised model-based TR method for successfully overcoming the convergence issues is proposed. On the computational front, unlike the existing TR methods for state estimation which employ quadratic models, the proposed linear model-based estimator is computationally efficient because the model minimiser can be computed in a single step. The model minimiser at each step is computed by minimising the linearised model in the presence of TR and measurement mismatch constraints. The infinity norm is used to define the geometry of the TR. Measurement mismatch constraints are employed to improve the accuracy. The proposed algorithm is compared with the quadratic model-based TR algorithm with case studies on the IEEE 30-bus system, 205-bus and 514-bus equivalent systems of part of Indian grid.

1 Introduction

The necessity of having a robust state estimator in the energy control center was highlighted by the NERC report [1] on the 2003 blackout in USA and Canada. A working and reliable state estimation (SE) solution is a must before an online security assessment system can be implemented. The performance of SE depends mainly on the numerical techniques employed for the solution and also on the quality of measurements. Trust region (TR)-based SE algorithms [2–6] are introduced to enhance the convergence characteristics in the presence of topological errors/bad data. The present work proposes a new, computationally efficient SE algorithm based on a linear model and infinity norm TR method.

Research in the area of SE has seen the application of various numerical and optimisation methods [7, 8]. The Gauss–Newton (GN) approach [9] is one of the traditional methods employed for solving the least squares state estimation (LSSE) problem. Application of the GN approach to the SE problem gives rise to what are popularly known as normal equations (NE). From the numerical prospective, it is possible to classify the SE problems as either well-conditioned or ill-conditioned. Numerical ill-conditioning arising in NE approach because of large number of injection measurements, indifferent weighing factors can be tackled to certain extent using orthogonal transformation (QR-GN) approach [10–13]. The method of Peters and Wilkinson [14] is another solution to the least squares (LS) formulation which has a good trade-off between speed and stability. Bounded LS formulations [15] have also been proposed to overcome the problem of numerical ill-conditioning.

1.1 TR methods

Even QR-GN approach which can successfully overcome most ill-conditioning issues are found to be inefficient in providing a converged solution in the presence of topological errors. TR approach-based SE algorithms [2, 4] are proposed to overcome this problem. TR-based SE approaches were first proposed by Pajic and Clements [3, 4] with the motivation of enhancing the convergence properties of SE in the presence of topological errors/bad data. The authors show that, in the presence of topological errors, the conventional GN-based algorithm may fail to converge because of the presence of large measurement residuals. The TR formulation proposed by the authors to overcome this issue minimises a quadratic model in an l_2 norm TR to obtain a globally convergent solution. Costa *et al.* [5, 6] show that the basic TR equation can be interpreted as the solution to the linearised LSSE problem taking a priori information into account. Very limited literature [3–6] is available on the application of TR methods for SE.

Existing TR-based SE algorithms based on quadratic models need higher computational effort than the conventional LS estimators. Hence their application is limited to cases where convergence difficulties arise such as presence of topology errors.

1.2 Aim and contribution of the paper

In this paper, we present case studies on practical Indian systems where existing TR methods fail to provide a converged solution in the presence of topological errors.

The aim of this paper is three-fold. First, a new, computationally efficient SE algorithm based on linearised model TR method is presented. The second aim is to show that the convergence issues that occur in the orthogonal transformation-based GN approach in the presence of large measurement residuals can effectively overcome using the proposed approach as in the existing TR-based estimators, but at lower computational cost. The third aim is to show that the accuracy of the proposed algorithm is comparable to that of conventional LS and quadratic model-based TR estimators.

1.3 Organisation

This paper is organised as follows. Section 2 describes the basic mathematical formulation of the proposed approach. Section 3 outlines the proposed SE algorithm and discusses the implementation issues. Finally, case studies on IEEE 30-bus and practical Indian systems are presented in Section 4 followed by the conclusions in Section 5.

2 Mathematical formulation

In this section, the general framework of the proposed linear model infinity norm-based TR method is described. Solution to the SE problem involves solving a set of nonlinear equations. Let, the nonlinear model relating the states to measurements be given as

$$z = f(\boldsymbol{x}) + \boldsymbol{e} \tag{1}$$

where

z is the m-dimensional measurement vector,
$\boldsymbol{x}$ is the n-dimensional state vector,
f is the nonlinear function relating measurements to states,
$\boldsymbol{e}$ is the m-dimensional measurement error vector,
$\boldsymbol{R}$ is the $m \times m$ co-variance matrix,
$\boldsymbol{H}$ is the Jacobian ($\boldsymbol{H} = \nabla f$).

Usually, the error is modelled as a Gaussian distribution with zero mean and a known co-variance. If the errors are assumed to be independent, then the co-variance matrix ([$\boldsymbol{R}$]) is a diagonal matrix with diagonal elements set to σ_i^2 (variance corresponding to measurement i). Under these conditions, an estimate to state $\boldsymbol{x}$ can be obtained by using the least-squares criterion. The least-squares criterion for (1) yields

$$\min_x J(\boldsymbol{x}) = \left[z - f(\boldsymbol{x})\right]^{\mathrm{T}} [\boldsymbol{R}]^{-1} \left[z - f(\boldsymbol{x})\right] \tag{2}$$

2.1 TR algorithm

A TR reflects the region where the model approximates the objective function well. Algorithms based on TR methods employ a linear or quadratic model of the objective function and make use of the TR s to enhance convergence. These algorithms minimise a model m_k of the function $J(\boldsymbol{x})$ within the TR

$$\| \Delta \boldsymbol{x} \|_k \leq \boldsymbol{\Theta}_k \tag{3}$$

A TR is usually defined as

$$\beta = \left\{ \boldsymbol{x} \in R^n : \| \boldsymbol{x} - \boldsymbol{x}_k \|_k \leq \Theta_k \right\} \tag{4}$$

where

m is the model of the objective function,
β is the TR constraint,
Θ_k is the TR radius,
$\|.\|_k$ is the kth-norm which defines the geometry of the TR and
$\Delta \boldsymbol{x}$ is the model minimiser or correction step.

The sub-problems in the TR methods are (i) finding a model minimiser for m_k, (ii) evaluating the model minimiser $\Delta \boldsymbol{x}$ for acceptance and (iii) updating the TR radius at each iteration. The TR radius is updated based on the ratio of change in value of objective function to change in model value, ρ_k. If this ratio produced by $\Delta \boldsymbol{x}$ is large [i.e. $\rho_k \geq \eta_2$, signifying that the model has excellent agreement with $f(\boldsymbol{x})$], the TR radius is increased. If it is small, no change is made to the radius of TR. In both the cases, the value of $\boldsymbol{x}_{k+1}$ is updated to $\boldsymbol{x}_k + \Delta \boldsymbol{x}_k$. However, if this ratio is less than a minimum value (close to 0) or negative ($\rho_k \leq \eta_1$), the correction $\Delta \boldsymbol{x}$ is not accepted (i.e. $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k$), and the TR radius is reduced. A description of the constants η_1, η_2 and other details can be found in [2]. The basic TR algorithm is outlined in Algorithm 1 (see Fig. 1).

2.2 Proposed approach

The proposed algorithm differs from the existing TR-based SE algorithm in the following aspects:

- Choice of model m_k;
- Choice of TR β;
- Addition of measurement mismatch constraints.

Algorithm 1

$m_k \leftarrow$ Model for nonlinear function in β
Choose $\eta_1, \eta_2 \in [0, 1]$ ($\eta_1 = 0.1, \eta_2 = 0.9$, say).
Compute Δx.
Compute $\rho_k = \frac{J(x_k) - J(x_k + \Delta x)}{m_k(x_k) - m_k(x_k + \Delta x)}$
if $\rho_k \geq \eta_1$ **then**
    $x_{k+1} \leftarrow x_k + \Delta x_k$
else
    $x_{k+1} \leftarrow x_k$
end if
Update the trust region radius as

$$\Theta_{k+1} = \begin{cases} 2.0 * \Theta_k & \text{if } \rho_k \geq \eta_2 \\ \Theta_k & \text{if } \eta_1 \geq \rho_k \geq \eta_2 \\ 0.5 * \Theta_k & \text{if } \rho_k \leq \eta_1 \end{cases}$$

Fig. 1 *Algorithm 1: Updating solution and TR radius using basic TR algorithm*

2.2.1 Choice of model m_k: The models generally employed in the TR methods [16] are either quadratic models (5a) or linear models (5b) given by

$$m_k = J(\boldsymbol{x}_k) + \nabla J(\boldsymbol{x}_k)^{\mathrm{T}} \Delta \boldsymbol{x} + \frac{1}{2} \Delta \boldsymbol{x}^{\mathrm{T}} \nabla^2 J(\boldsymbol{x}) \Delta \boldsymbol{x} \tag{5a}$$

$$m_k = J(\boldsymbol{x}_k) + \nabla J(\boldsymbol{x}_k)^{\mathrm{T}} \Delta \boldsymbol{x} \tag{5b}$$

Most of the methods, including that adopted by Pajic and Clements [4] employ quadratic models. In this paper, linearised model-based TR algorithms are proposed for the SE problem. One of the issues with adopting linearised models is the accuracy of the estimated values. It is well-known that quadratic models are usually superior to linear models, but solution to quadratic models in the presence of the TR constraint involves higher computational burden. In this paper, measurement mismatches are employed as additional constraints to improve the accuracy of the proposed algorithm. The linearised model m_k is solved in the presence of the TR constraint and measurement mismatch constraints. It is shown that the accuracy of the proposed approach is comparable with the quadratic model-based SE algorithms.

2.2.2 Choice of TR β: The solution to the linearised model can be obtained using linear programming (LP) techniques, provided a proper choice is made in choosing the norm for defining the TR. The commonly employed norms for defining the TRs are the l_1, l_2 and l_∞ norms. For the proposed approach, the l_∞ norm (box-type constraints) is employed for defining TRs . The l_∞ norm is employed to define the TR because it can be handled by adding it as a simple variable bound in the linear programming problem (LPP). That is $\|\Delta \boldsymbol{x}\|_\infty \leq \boldsymbol{\Theta}_k \Rightarrow -\boldsymbol{\Theta}_k \leq \Delta \boldsymbol{x} \leq \boldsymbol{\Theta}_k$.

Existing TR-based SE approaches use a l_2 norm to define the TR. In such cases, computing the model minimiser at each iteration requires sub-iterations. However, in the proposed approach, since l_∞ norms are handled directly as variable bounds, the solution requires no sub-iterations. Consequently, there is a significant reduction in the computational effort required. In this paper, $-\Theta_k$ and $\Delta \boldsymbol{x}^{\min}$ are used interchangeably. Similarly, Θ_k and $\Delta \boldsymbol{x}^{\max}$ are used interchangeably.

2.2.3 Measurement mismatch constraints: Instead of merely solving the linearised model in the presence of TR constraints, measurement mismatch constraints (derived from the Taylor's series) are also employed as additional constraints to enhance the performance of the proposed approach. The measurement mismatches are derived from the Taylor's series expansion.

The truncated Taylor's series expansion of $f(\boldsymbol{x})$ around a nominal value $\boldsymbol{x}_k$ is given as

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_k) + (\boldsymbol{x} - \boldsymbol{x}_k) \nabla f(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_k} + \cdots \tag{6}$$

where $\nabla f(\boldsymbol{x})$ is the Jacobian $\boldsymbol{H}$.

From (6)

$$[\boldsymbol{H}][\Delta \boldsymbol{x}] \simeq f(\boldsymbol{x}) - f(\boldsymbol{x}_k) \tag{7}$$

The correction $\Delta \boldsymbol{x}$ should be such that $[\boldsymbol{H}][\Delta \boldsymbol{x}]$ should try to converge to $z - f(\boldsymbol{x}_k)$. That is

$$[\boldsymbol{H}][\Delta \boldsymbol{x}] \leq z - f(\boldsymbol{x}_k) \tag{8}$$

However, if an element of $z - f(\boldsymbol{x}_k) < 0$, then the inequality $[\boldsymbol{H}][\Delta \boldsymbol{x}] \leq z - f(\boldsymbol{x}_k)$ will force the corresponding element to be equally or even more negative, thereby not allowing the LPP to converge. Consequently, separate equations are required to handle cases where $z - f(\boldsymbol{x}_k) < 0$ and $z - f(\boldsymbol{x}_k) > 0$. For cases where $z - f(\boldsymbol{x}_k) < 0$, $[\boldsymbol{H}][\Delta \boldsymbol{x}] \geq z - f(\boldsymbol{x}_k)$ is employed to force the value of corresponding element close to zero or a positive value. That is

$$\begin{aligned} &[\boldsymbol{H}][\Delta \boldsymbol{x}] \leq z - f(\boldsymbol{x}_k) \quad \text{if} \quad z - f(\boldsymbol{x}_k) \geq 0 \\ &[\boldsymbol{H}][\Delta \boldsymbol{x}] \geq z - f(\boldsymbol{x}_k) \quad \text{if} \quad z - f(\boldsymbol{x}_k) < 0 \end{aligned} \tag{9}$$

The constraints in (9) are used for determining $\Delta \boldsymbol{x}$ such that the value of $|z - f(\boldsymbol{x}_k)|$ is reduced or made close to zero. This is similar to the objective function employed in the weighted least absolute value (WLAV) approaches [17–19]. However, here the same is employed as constraints. In the matrix form, the constraints can be represented as

$$[\tilde{\boldsymbol{H}}][\Delta \boldsymbol{x}] \leq \left[\tilde{\Delta z}\right] \tag{10}$$

where

$$[\tilde{\boldsymbol{H}}] = \begin{bmatrix} \boldsymbol{H}_1 \\ -1^* \boldsymbol{H}_2 \end{bmatrix} \tag{11a}$$

$$\left[\tilde{\Delta z}\right] = \begin{bmatrix} z - f(\boldsymbol{x}_k) \\ -1^*(z - f(\boldsymbol{x}_k)) \end{bmatrix} \tag{11b}$$

where $\boldsymbol{H}_1$ corresponds to the set of equations whose $z - f(\boldsymbol{x}_k) \geq 0$ and $\boldsymbol{H}_2$ corresponds to the set of equations whose $z - f(\boldsymbol{x}_k) < 0$.

Although it appears intuitively that, the maximum values that can be obtained is z (because of the $<$ constraint), it will be shown that values less than and greater than z can be obtained (because of $\geq$ constraint). Hence both positive errors and negative errors can be smoothed out.

2.3 Computing the model minimiser Δx_k

As mentioned earlier, the model minimiser at each step has to be computed and checked for acceptance (using the basic TR algorithm). The solution to the generic model given by (12) gives the model minimiser $\Delta \boldsymbol{x}_k$ at each step. This generic form can be solved using any of the LP techniques. The solution $\Delta \boldsymbol{x}_k$ thus obtained satisfies both the measurement mismatch and TR constraints. The solution is obtained in an non-iterative approach, thereby reducing the computational time. The generic model is

$$\begin{aligned} &\min m_k(\boldsymbol{x}_k + \Delta \boldsymbol{x}_k) = J(\boldsymbol{x}_k) + [\boldsymbol{c}]^{\mathrm{T}} \Delta \boldsymbol{x} \\ &\text{s.t} \quad \widetilde{\boldsymbol{H}}[\Delta \boldsymbol{x}_k] \leq \left[\widetilde{\Delta z}\right] \\ &\| \Delta \boldsymbol{x}_k \|_\infty \leq \Theta_k \end{aligned} \tag{12}$$

$$\text{where} \quad [\boldsymbol{c}]^{\mathrm{T}} = \nabla J(\boldsymbol{x})^{\mathrm{T}} = \left[\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}_1} \frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}_2} \cdots \frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}_n} \right] \tag{13}$$

In essence, the gradient of $J(\boldsymbol{x})$ is minimised (minimising the gradient corresponds to maximising the direction of descent). That is the movement is in a direction that has maximum directional descent. The resulting direction $\Delta \boldsymbol{x}$ gives the maximum reduction in $J(\boldsymbol{x})$.

It is well known that as per the first order necessary condition (FONC), the value of $\nabla J(\boldsymbol{x}) = 0$ when $\boldsymbol{x} \to \hat{\boldsymbol{x}}$. The expression for $\nabla J(\boldsymbol{x})^{\mathrm{T}}$ is given as

$$\nabla J(\boldsymbol{x})^{\mathrm{T}} = (-2)[z - f(\boldsymbol{x})]^{\mathrm{T}}[\boldsymbol{H}] \tag{14}$$

In order to evaluate an acceptable $\Delta \boldsymbol{x}_k$, the predicted reduction in m_k is required. Using the Taylor's series, this can be obtained as

$$m_k - m_{k+1} = \left[-\nabla J(\boldsymbol{x}_k)\right]^{\mathrm{T}}[\Delta \boldsymbol{x}_k] \tag{15}$$

The (12) employed for determining the model minimiser is a bounded LPP. The infinity norm constraint here implies that the solution $\Delta \boldsymbol{x}$ is bounded between $-\Theta_k$ and Θ_k. Hence the solution Δx is allowed to take both positive and negative values. The LPP given by (12) is solved by an improved variant of the simplex method given in [20]. A simple implementation is to convert the bounded LP into the standard LP form [21], but this would require a two-phase simplex procedure. The improved variant of simplex method in [20] solves (12) in a single phase, thereby reducing the computational time.

2.4 On convergence, initial parameters

The algorithm is assumed to have converged if the gradient of the objective function is less than a specified tolerance (i.e. $\nabla(J(\boldsymbol{x})) < \varepsilon$). In most cases, the maximum number of iterations is also specified. Theoretically, for any TR algorithm, the choice of initial point x_0 and Θ_0 has little effect on final solution since TR methods give a globally convergent solution.

The other parameters of the basic TR algorithm which need to be set are (a) η_1, η_2 for checking the acceptance of solution and (b) the values of scaling employed whereas updating the TR radius. The typical values employed are mentioned in [2]. Since the TR algorithms are globally convergent, the effect of the choice of these parameters is minimal.

3 Proposed SE algorithm

3.1 Algorithm

Algorithm 2 summarises various steps involved in the proposed approach (see Fig. 2).

3.2 Initialisation of the estimator

In this paper, two possibilities for choosing the nominal point are explored:

1. The procedure adopted usually is to choose a flat start (that is $1\angle 0$) as the initial point.
2. In real-time operation, results of the SE is stored, This stored information (which represents the state of the system) can be used as an initial condition.

Algorithm 2

Process the measurements to form
  $z \leftarrow$ measurements, $R \leftarrow$ co-variance matrix
if the previous operating state is known **then**
  $x_0 \leftarrow$ previous state value
else
  $x_0 \leftarrow 1\angle 0$
end if
Evaluate $J(x_0)$ and $\| \nabla J(x_0) \|$.
$k \leftarrow$ iteration count
while $\| \nabla J(x_k) \| \leq \epsilon$ **do**
  Evaluate $J(x_k)$ and $\nabla J(x_k)$
  Compute the model minimiser Δx_k by solving the LPP

$$\begin{aligned} &\min\ m_k(x_k + \Delta x_k) = J(x_k) + [\nabla J(x_k)]^T \Delta x \\ &\text{s. t.} \quad \widetilde{H}\,[\Delta x_k] \leq \left[\widetilde{\Delta z}\right] \\ &\qquad \| \Delta x_k \|_\infty \leq \Theta_k \end{aligned}$$

  Evaluate $J(x_k)$, $J(x_k + \Delta x_k)$, $m_k - m_{k+1}$, and ρ_k.
  if $\rho_k \geq \eta_1$ **then**
    $x_{k+1} \leftarrow x_k + \Delta x_k$
  else
    $x_{k+1} \leftarrow x_k$
  end if

$$\Theta_{k+1} = \begin{cases} 2.0 * \Theta_k & \text{if } \rho_k \geq \eta_2 \\ \Theta_k & \text{if } \eta_1 \geq \rho_k \geq \eta_2 \\ 0.5 * \Theta_k & \text{if } \rho_k \leq \eta_1 \end{cases}$$

  $k \leftarrow k + 1$
end while

Fig. 2 *Algorithm 2 proposed linear model-based TR algorithm for SE*

3.3 Computational, numerical and implementation issues

The implementation of the proposed estimator, the orthogonal transformation-based NE approach and the quadratic model-based TR approach presented in [4] is developed in 'C' programming language. Sparsity features of the matrices are exploited for efficient implementation. The developed code also employs 'OpenMP" to execute the program in parallel.

For solving the LPP, simplex method with presolving [22] is employed. Presolving is a set of techniques applied on a LPP before the simplex method solves it. This set of techniques aim at reducing the size of the LPP by eliminating redundant constraints and variables. For the LP implementation carried out, only simple presolving techniques presented in [23] are employed. However, with advanced presolving techniques, the size of the problem can be further reduced, thereby reducing the computation time per iteration.

As the TR radius becomes smaller, the LPP would be completely solved in the presolving stage itself, thereby reducing the computational effort. Moreover, with present day efficient LP solvers such as CPLEX, the computation time to solve the proposed formulation is comparable to that of the other least squares solution.

3.4 Graphical interpretation of the proposed approach

The geometrical interpretation of the proposed approach is presented in this section. In order to make the interpretation simpler, an estimation problem with two variables is considered. $\|m(\boldsymbol{x})\| = \boldsymbol{c}$ represent the level sets of the linearised objective function. Since this is a linearised objective function, the level sets correspond to n-dimensional planes. In a two-dimensional case, the level sets correspond to a straight line as shown in Fig. 3. The solution of the proposed formulation lies at the boundary of the intersection of the l_∞ norm TR ($\|\Delta\boldsymbol{x}\|_\infty \le \Theta_k$) constraint (denoted by C_t) and the linearised measurement mismatch ($\widetilde{\boldsymbol{H}}[\Delta\boldsymbol{x}] \le \widetilde{\Delta z}$) constraints (denoted by C_m). This feasible region is given by the shaded polygon PQRST. For the l_∞ norm, the geometry of the TR is a 'box' [2], whose each boundary is given by a plane parallel to each of its axis and is located at distance Θ_k on the either side. For the two-dimensional case represented in Fig. 3, the TR has a square geometry, with the four sides at a distance Θ_k from the point $\boldsymbol{x}^k$. The solution $\boldsymbol{x}^{k+1}$ is the level set $\|m(\boldsymbol{x}^{k+1})\| = c$ which has the least value and lies on the boundary of the feasible region.

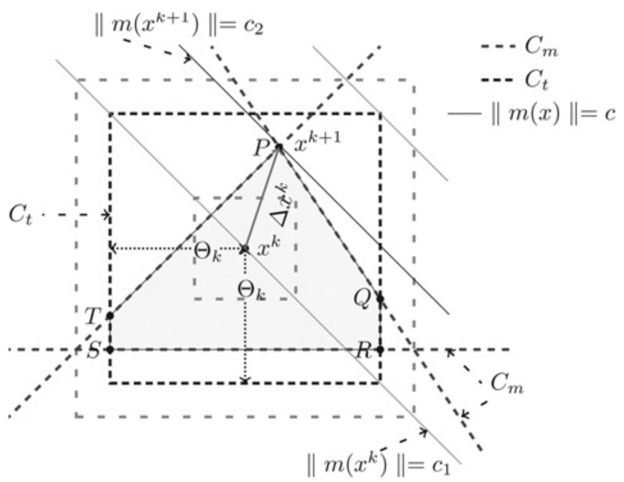


Fig. 3 *Graphical interpretation of the proposed approach*

3.5 Comparative evaluation

This section details a comparative evaluation of the proposed linear model-based TR approach with the quadratic model-based TR approach presented in [4]. Effective implementation of quadratic model TR method is detailed in [24]. In this paper, comparative evaluation of the proposed approach is carried out with the quadratic model algorithms that have been applied for SE in literature. The generalised formulation for the quadratic model TR approach presented is detailed in [25]. Existing TR methods for SE in power systems uses a quadratic model with a linear approximation to $\nabla^2 J(\boldsymbol{x})$. The model minimiser is computed by solving

$$\begin{aligned} &\min m_k = J(\boldsymbol{x}_k) + \nabla J(\boldsymbol{x}_k)^{\mathrm{T}}\Delta\boldsymbol{x} + \frac{1}{2}\Delta\boldsymbol{x}^{\mathrm{T}}\nabla^2 J(\boldsymbol{x})\Delta\boldsymbol{x} \\ &\text{s.t.} \quad \| \Delta\boldsymbol{x} \|_2 \le \Theta_k \end{aligned} \tag{16}$$

Applying the method of Lagrangian multipliers with the Gauss–Newton approximation for $\nabla^2 J(\boldsymbol{x})$, (16) will have the form

$$\begin{aligned} &(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H} + \mu I)\Delta\boldsymbol{x}(\mu) = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(z - f(\boldsymbol{x})) \\ &\text{s.t.} \quad \|\Delta\boldsymbol{x}(\mu)\|_2 = \Theta \end{aligned} \tag{17}$$

For this formulation given by (17), there is a practical difficulty in obtaining the solution $\Delta\boldsymbol{x}$ because of the l_2 norm constraint imposed. However, a reasonably accurate solution is obtained by using the 'hook step' method which is an approximate solution. The solution $\Delta\boldsymbol{x}(\mu)$ is given as

$$\begin{aligned} &\boldsymbol{U}_k^{\mathrm{T}}\boldsymbol{U}_k\Delta\boldsymbol{x}(\mu) = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{Q}r_w \\ &\mu_{k+1} = \mu_k + \frac{\| \Delta\boldsymbol{x}(\mu) \|_2}{\Theta_k}\frac{\phi(\mu_k)}{\phi'(\mu_k)} \end{aligned} \tag{18}$$

where

$$\begin{aligned} &\phi(\mu_k) = \| \Delta\boldsymbol{x}(\mu_k) \|_2 - \Theta_k \\ &\phi'(\mu_k) = \frac{\Delta\boldsymbol{x}(\mu_k)^T\left(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H} + \mu_k I\right)^{-1}\Delta\boldsymbol{x}(\mu_k)}{\| \Delta\boldsymbol{x}(\mu) \|_2} \end{aligned} \tag{19}$$

It can be observed that, for computing the $\Delta\boldsymbol{x}(\mu_k)$, additional computations such as factorisation/inverse of $(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H} + \mu_c I)$ are required. The enhanced solution comes with extra computational cost, because of the addition of the TR constraint.

In the proposed formulation, the choice of l_∞ norm fits the LPP as variable bounds and are taken care of while obtaining the solution. Hence no extra computations/approximations are needed. Moreover, as mentioned in Section 3.3, solving the proposed formulation is trivial with efficient presolving techniques.

IET Gener. Transm. Distrib., 2013, Vol. 7, Iss. 5, pp. 500–510
doi: 10.1049/iet-gtd.2012.0473

4 Case studies

In this section, test results for IEEE 30-bus system [26], 205-bus and 514-bus [20] equivalent systems which are a part of Indian grid are presented to validate the proposed approach. The 205-bus equivalent of Indian southern grid consists of 205 buses, 47 transformers, 235 transmission lines (includes 400 and 220 kV lines). The 514-bus equivalent of part of Indian grid consists of 514 buses, 296 transformers and 303 transmission lines (includes 400, 220, 132 and 33 kV lines).

The objective of the results presented in this section is to highlight the following two issues:

- In cases when numerical ill-conditioning and topological errors co-exist, the existing TR methods for SE fails to provide a converged solution. Results presented indicate how such cases can be handled using the proposed linear model TR approach.
- Even in case when only topological errors exist, the proposed approach is an computationally efficient alternative to existing TR methods for SE.

Potential advantages of the proposed approach in terms of the computational effort is discussed. The results obtained using the proposed approach are compared with those obtained by the quadratic model-based TR approach presented in [4].

The inputs to the SE are obtained from load flow studies. To simulate meter inaccuracies, errors are injected into all measurements. The errors in the measurements are modelled as Gaussian random variables with a known standard deviation. The error is presumed to be spread over a standard deviation of $\pm 3\%$ of its full scale. The algorithm is presumed to have converged when the value of $\|\nabla J(\boldsymbol{x})\| \leq 10^{-4}$. Zero injection measurements can be either handled as equality constraints or can be incorporated as two inequality constraints. In this paper, zero injection measurements are handled as two inequality constraints.

The results presented in the paper are for three categories.

1. In the first category, results are presented for two cases with data sets having only topological errors. The objective is to illustrate the convergence behaviour and computational cost of the algorithms. Most of the case studies presented in this paper belong to this category because of two reasons. First, TR methods are specifically employed in cases when there are convergence issues in conventional approaches. Secondly, the proposed approach is a computationally efficient variant of the TR method and is intended to provide a converged solution at less computational cost in the presence of large measurement residuals.
2. In the second category, results are presented with data sets where topological errors and ill-conditioning co-exist. The objective is to indicate that the proposed approach provides a converged solution even in such scenarios where as existing quadratic model approaches for SE fail to do so.
3. The case studies presented in the third category are intended to show that the proposed approach is reasonably accurate as compared to those of the conventional LS estimators (LS techniques provide the optimal solution). The effect of initial TR radius and initial point on the convergence characteristics of the proposed algorithm is also presented.

4.1 Performance of the linear model estimator in the presence of topological errors

The presence of bad data or topological errors can cause the conventional LS approaches to diverge as shown by the Pajic and Clements [4]. A converged solution (although not exact) in such scenarios is required before any bad data or topological error detection technique can be applied. Application of TR methods can give enhanced convergence in such scenarios following which bad data/topological errors detection techniques can be applied. Case studies on the IEEE 30-bus system and 205-bus system of Indian grid are presented to compare the convergence characteristics and computational cost of the proposed approach with the quadratic model-based TR approach.

The first scenario considered in this set of results compares the convergence characteristics of various approaches in the presence of topology error for IEEE 30-bus system. The single line diagram of the IEEE 30-bus system and the measurement set is presented in Fig. 4.

The measurement set contains 72 measurements (15 injection (P,Q), 20 flow (P,Q) and two voltage magnitude measurements). Topological errors of exclusion type (erroneous exclusion of a branch) are created on three lines connecting the buses 12–14, 18–19 and 23–24. Fig. 5 shows the convergence characteristics of various approaches in the presence of topological errors. It can be observed that the orthogonal transformation-based GN approach fails to converge. This is due to the presence of extremely large residuals which are shown in Table 1.

The application of TR approaches in such scenarios results in good convergence. Fig. 5 indicates two mathematical aspects of the proposed linear model-based TR algorithm. First, it can be observed that during the initial stages, the value of $\log\|\nabla J(\boldsymbol{x})\|$ falls rapidly for the proposed approach than the quadratic model approach. This is because, for a given TR radius, the solution space (decided by the TR constraint) when defined using l_∞ norm is larger than when defined using the l_2 norm. In order to further explain this, the convergence characteristics of both approaches for the first few iterations of this case is presented in Table 1.

The initial TR radius Θ_0 for this case is 0.5. In the first iteration, the $\Delta\boldsymbol{x}^{\max}$ obtained using the proposed approach is 0.5000 and that obtained using the quadratic model approach is 0.2000. This is a direct consequence of the norm employed for defining the TR. With l_2 norm TR, the solution obtained should be such that $\|\Delta\boldsymbol{x}\| \leq 0.5$, whereas with the l_∞ norm TR, the solution $\Delta\boldsymbol{x}$ obtained would be such that $-0.5 \leq \Delta\boldsymbol{x}_i \leq 0.5$.

Table 1 also indicates other aspects of both the algorithms. First, the directional derivative $\nabla J(\boldsymbol{x})^{\mathrm{T}}\Delta\boldsymbol{x}$ at each iteration is presented. It is well known that for an converging algorithm, $\nabla J(\boldsymbol{x})^{\mathrm{T}}\Delta\boldsymbol{x} < 0$. The directional derivative for the TR algorithms is negative during all the iterations whereas the descent direction criterion is not satisfied for the orthogonal transformation-based GN approach (during the second iteration the value of $\nabla J(\boldsymbol{x})^{\mathrm{T}}\Delta\boldsymbol{x}$ is 3170.3). It is seen in column 5 of Table 1 that in the first two iterations, the directional derivative of the proposed approach is more in magnitude than that of the quadratic model approach and it is because of this reason that the convergence of the linear model approach is faster in initial iterations. Our observation is that this holds true for most of the cases. The number of internal iterations required to compute the model

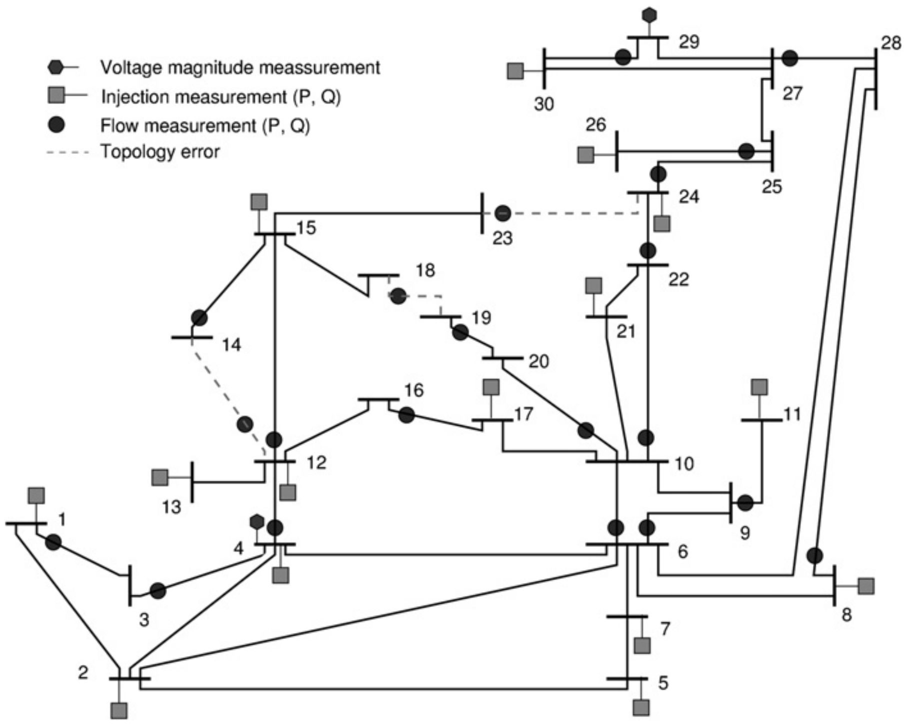


Fig. 4 *Single line diagram of IEEE 30-bus system with measurement sets and topology error considered*

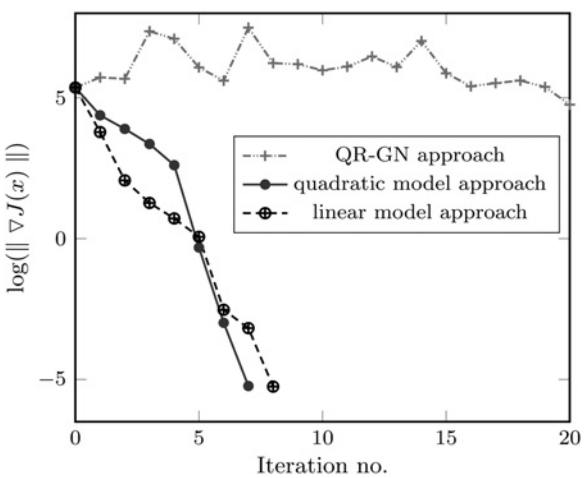


Fig. 5 *Comparison of convergence characteristics in the presence of topology error on IEEE 30-bus system*

minimiser for both the algorithms is shown in the last column of Table 1.

Another aspect of the algorithm is concerning the convergence in the final iterations. Since the linear model algorithm considers only the gradient information, situations could arise where the convergence closer to the solution could become slower, which is typical of any gradient-based algorithm. The consequence of which may result in the proposed approach requiring one or two additional iterations to converge than the quadratic model approach. Our observation is that during the final iterations, solution is obtained during the presolving stage itself. Hence even if the approach requires additional iterations the computational time is still less.

Whereas the first case compares both the TR-based algorithms, the next case is considered to study the performance of the proposed approach for large systems. In this case, a 205-bus equivalent system of Indian southern grid [20] is considered. Topological errors of exclusion type are created on the lines connecting the buses 179–91, 143–103, 106–171 and 107–144. The measurement redundancy employed for this case is 2.72 and the initial

Table 1 Performance of linear and quadratic model approaches for IEEE 30-bus system with topology error

	Iter. no.	$\Delta \boldsymbol{x}^{max}$	$J(\boldsymbol{x})$	$\nabla J(\boldsymbol{x})^T \Delta \boldsymbol{x}$	Sub-iter.
linear model approach	0	0.0000	1.482×10^5	0.00	0
	1	0.5000	1.45×10^3	-2.94×10^4	0
	2	0.0535	1.94×10^1	-2.90×10^3	0
	3	0.0010	1.49×10^1	-1.14×10^1	0
quadratic model approach	0	0.0000	1.48×10^5	0.00	0
	1	0.1997	2.062×10^3	-2.16×10^4	5
	2	0.1524	4.65×10^2	-2.45×10^3	3
	3	0.1432	7.44×10^1	-5.86×10^2	4

For the orthogonal transformation-based GN approach, the direction of descent ($\nabla J(\boldsymbol{x})^T \Delta \boldsymbol{x}$) for the first two iterations are −8108.4 and 3170.3.

doi: 10.1049/iet-gtd.2012.0473

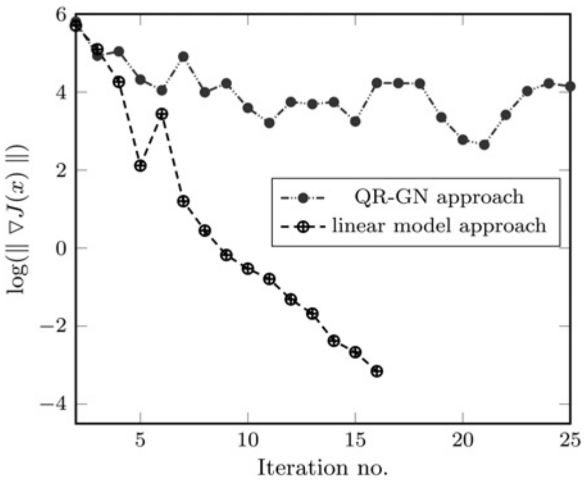


Fig. 6 *Comparison of convergence characteristics in the presence of topology error for 205-bus system*

TR radius is 1.00. Fig. 6 indicates the convergence characteristics of the proposed approach. Analysis of Fig. 6 leads to conclusions similar to that of previous case.

4.2 Performance in the presence of topological errors and numerical ill-conditioning

In the previous test cases, it is shown that the conventional Newton's algorithm may fail to converge in the presence of topology errors. In such case the failure of the algorithm is not because of the ill-conditioning and can be effectively tackled using the existing TR approaches. However, in certain cases the numerical ill-conditioning can occur in the presence of topological errors and such cases are difficult to handle even with existing TR methods. In order to illustrate the above observations, a case study on practical Indian systems is presented.

In this case, test results on 205 bus equivalent system of Indian southern grid is presented. The purpose of this case is to indicate that the TR method would fail to provide a converged solution in the presence of topological errors if the associated gain matrix is ill-conditioned. The measurement set considered for this set have injection measurement at all the buses, flow measurement at one end of the line and the voltage magnitude measurement. Topological error of erroneous exclusion is created on line 194–58. Even though the measurement redundancy considered is quite high, it was observed that neither the QR-GN approach nor the TR approach could provide a converged solution. The condition number of the gain matrix $\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H}$ for the QR-GN approach is 3.1924×10^{18}.

The reason for the TR method failing to provide a converged solution is also related with the ill-conditioning of the gain matrix. In the TR method, each iteration will have sub-iterations to determine the Lagrange multiplier μ such that the TR constraint is satisfied. The Lagrange multiplier μ is usually computed in an iterative way using (18) and (19). Observing (19), it can be observed that for computing $\phi'(\mu_k)$, the factorisation/inverse of $\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H} + \mu I$ is required. With such a high-condition number and with the initial value of μ_0 being 0, finding the factorisation/inverse becomes difficult.

The proposed linear model algorithm however is successful in providing a converged solution. The reason being that in the linear model approach, the solution is always determined by the basis which is always a full rank and well conditioned. The basis vector consists of few rows of Jacobin $\boldsymbol{H}$. The convergence characteristics of the proposed linear model approach for this case is presented in Fig. 7*a*.

As mentioned earlier, in order to test for successful convergence, it is better to look for the directional derivative and check that it always satisfies the descent criterion. Fig. 7*b* gives the directional derivative produced during each step of the linear model approach. It can be observed that the descent criterion is satisfied and the directional derivative smoothly reduces as it reaches closer to the solution.

4.3 Computation time

The ratio of the computation time per iteration of the linear model ($t_{\rm lm}$) and quadratic model ($t_{\rm qm}$)-based TR approaches is presented in Table 2.

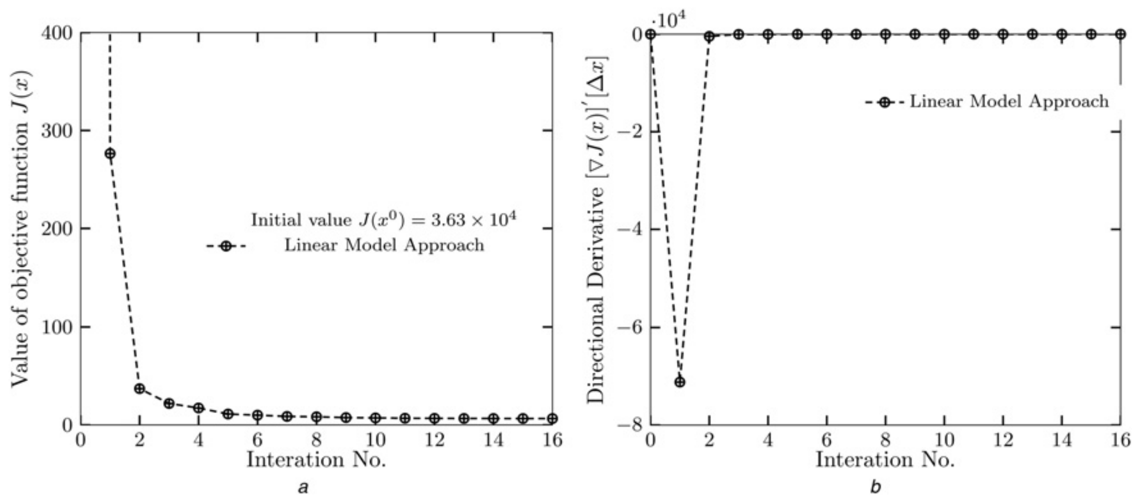


Fig. 7 *Results for 205 bus system with linear model approach when both topology error and ill-conditioning co-exist*

a Convergence characteristics of the linear model approach
b Directional derivative produced at each iteration of the linear model trust approach

Table 2 Computational time

System	Redundancy	Θ_0	t_{lm}/t_{qm} (per iter.)	Sub-iterations		
				1	2	3
30-bus	2.00	0.75	0.6098	6	4	3
	1.20		0.8577	5	3	4
205-bus	2.10	1.00	0.4093	6	5	5
		1.50	0.6302	6	4	4

It can be observed that the computational effort involved in the proposed approach is less than that of quadratic model approach and the reduction is significant in the cases when the measurement redundancy is high. Significant computational burden in the quadratic model approach lies in the computing of model minimiser. Quadratic model approach requires sub-iterations to compute the model minimiser at each step. The number of sub-iterations for computing the model minimiser in the first three iterations for the quadratic model approach is also indicated in Table 2.

Observation of Table 2 also indicates that for smaller initial TR radius Θ_0 and higher measurement redundancy the quadratic model TR approach requires more number of sub-iterations to compute the model minimiser. In the case of the linear model approach, the model minimiser can be computed in one step irrespective of the initial TR radius. As a consequence, it can be observed that in such cases the proposed approach gives a significant reduction in computation burden.

The additional computational advantage in the proposed approach is with the employment of presolving techniques for solving the LPP. In cases with higher measurement redundancy, presolving helps in reducing the size of the problem. One more advantage of employing presolving is that during the final iterations, when the TR radius becomes

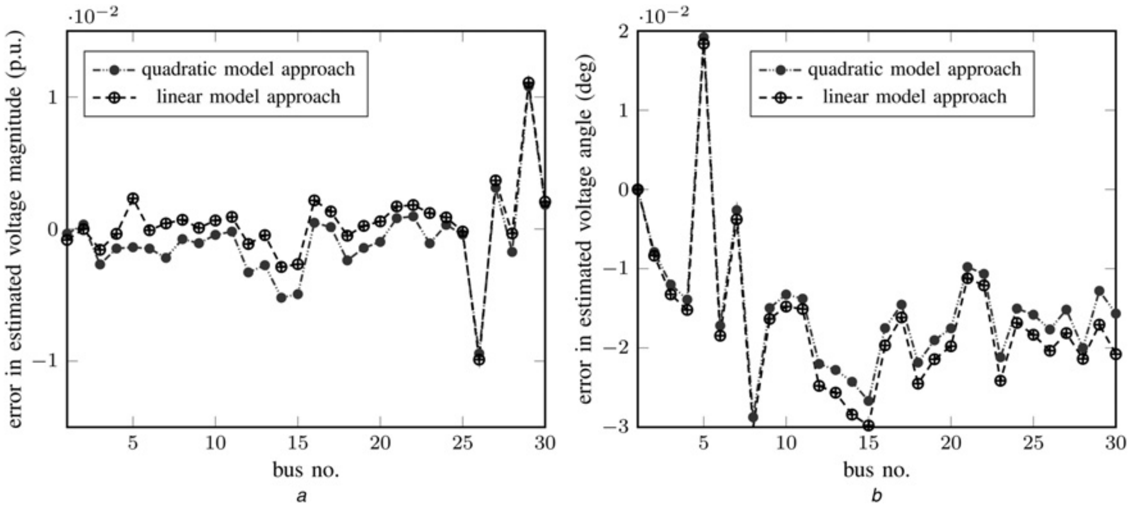


Fig. 8 *Error in estimated values (linear model and quadratic model approaches) for IEEE 30-bus system*

a Error in estimated voltage magnitudes
b Error in estimated voltage angles

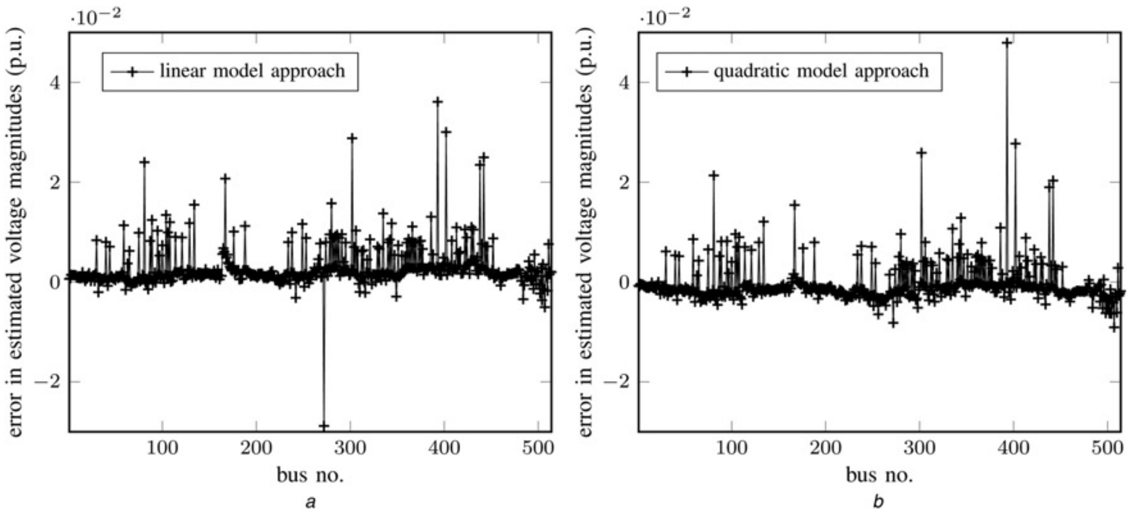


Fig. 9 *Error in estimated values (linear model and quadratic model approaches) for 514-bus system*

a Proposed linear model trust region approach
b Quadratic model TR approach]

doi: 10.1049/iet-gtd.2012.0473

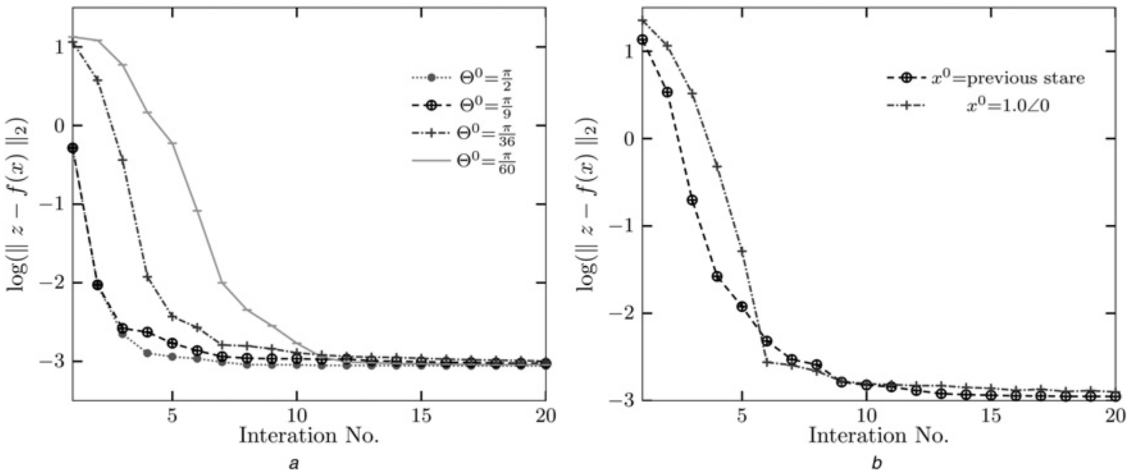


Fig. 10 *Effect of initial TR radius and initial conditions on the convergence of the proposed approach*

a Effect of initial TR radius

b Convergence characteristics of the proposed approach for two different initial points

low the solution is obtained in the presolving stage which involves very little computational effort.

4.4 Performance in the absence of topological errors/bad data

This category of results aims at comparing the accuracy of the proposed linear model approach with the quadratic model approach. In the absence of topological error, the results obtained using the QR-GN approach and the quadratic model TR approach are the same. As a first case in this category, the IEEE 30-bus system is considered. The measurements are presumed to be the real and reactive powers at selected locations and the voltage magnitude at the slack bus. The errors injected in the power measurements have a $\sigma = 0.02$ p.u. and that injected in voltage magnitude measurements have a $\sigma = 0.01$ p.u. The errors in the estimated magnitudes and angles for linear model and quadratic model-based TR approaches are given in Figs. 8*a* and *b*, respectively.

It can be observed from Figs. 8*a* and *b* that the estimates obtained using the linear model-based approach are comparable in terms of accuracy with that obtained using the quadratic model-based approach.

In the last case, test results for a 514-bus equivalent system of part of Indian grid is presented. The measurement set considered for this system includes injection and flow measurements along with voltage measurements at bus 1 (which is not the slack bus). The measurement redundancy for this case is 2.15. Figs. 9*a* and *b* give the errors in estimated voltage magnitudes for linear model and quadratic model approach, respectively. It can be observed that in this case, the proposed approach could get a better estimate than the quadratic model approach.

In the next scenario, the effect of initial TR radius on the convergence of the proposed algorithm is tested. Fig. 10*a* represents the convergence of proposed algorithm for various choice of initial TR radius. Fig. 10*b* gives the convergence characteristics of the proposed approach with two different initial conditions namely flat start and the previous state estimate. Whereas using the previous state as initial condition, first the estimate is obtained for 80% of the system load and the resulting estimate is used as the initial condition for case presented (100% load) in Fig. 10*b*.

The value of $\log\|z - f(x)\|$ is plotted for each iteration. The results indicate that the algorithm reaches global minima irrespective of the initial conditions and initial TR radius.

5 Conclusions

A new method for SE-based on linear model TR method has been presented. The proposed approach can effectively provide a converged solution in cases where numerical ill-conditioning and topology errors co-exist. Another advantage of the proposed approach lies in the fact that computing the model minimiser does not require any sub-iterations, thereby bringing down the computation time. Measurement mismatches are handled as constraints to make the accuracy of the linear model mismatches comparable to that of the quadratic model TR approaches and QR-based GN methods. In view of the above advantages, the proposed approach is a computationally efficient alternative for cases in which orthogonal transformation-based GN algorithm fails to converge. Case studies on practical Indian systems are presented to clearly illustrate the effectiveness of the proposed approach.

6 References

1 'Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations,' NERC, Tech. Rep., Apr. 2004. [Online]. Available: http://www.nerc.com/_lez/blackout.html

2 Conn, A., Gould, N., Toint, P.: 'Trust-region methods' (Society for industrial and applied mathematics, 2000)

3 Pajic, S., Clements, K.: 'Globally convergent state estimation via the trust region method'. Power Tech. Conf. Proc., 2003 IEEE Bologna, IEEE, 2003, vol. 1, pp. 6–pp

4 Pajic, S., Clements, K.A.: 'Power system state estimation via globally convergent methods', *Power Syst., IEEE Trans.*, 2005, **20**, (4), pp. 1683–1689

5 Costa, A., Salgado, R., Haas, P.: 'Globally convergent state estimation based on givens rotations'. Bulk Power System Dynamics and Control-VII. Revitalizing Operational Reliability, 2007 iREP Symp. IEEE, 2007, pp. 1–9

6 Costa, A., Salgado, R., Haas, P.: 'Trust region optimization methods via Givens rotations applied to power system state estimation' (NOVA Publishers, 2009), pp. 27–51, chapter 2

7 Monticelli, A.: 'Electric power system state estimation', *Proc. IEEE*, 2000, **88**, (2), pp. 262–282

8 Conejo, A.J., de la Torre, S., Canas, M.: 'An optimization approach to multiarea state estimation', *IEEE Trans. Power Syst.*, 2007, **22**, (1), pp. 213–221

9 Larson, R.E., Tinney, W.F., Peschon, J.: 'State estimation in power systems-parts i, ii', *Power Appar. Syst., IEEE Trans.*, 1970, **PAS-89**, (3), pp. 345–363

10 Strang, G.: 'Linear algebra and its applications' (Brooks/Cole, 2006, 4th edn.)

11 Simoes-Costa, A., Quintana, V.H.: 'A robust numerical technique for power system state estimation', *Power Appar. Syst., IEEE Trans.* 1981, **PAS-100**, (2), pp. 691–698

12 Simoes-Costa, A., Quintana, V.H.: 'An orthogonal row processing algorithm for power system sequential state estimation', *Power Appar. Syst., IEEE Trans.*, 1981, **PAS-100**, (8), pp. 3791–3800

13 Monticelli, A., Murari, C.A.F., Wu, F.F.: 'A hybrid state estimator: solving normal equations by orthogonal transformations', *Power Appar. Syst., IEEE Trans.*, 1985, **PAS-104**, (12), pp. 3460–3468

14 Gu, J.W., Clements, K.A., Krumpholz, G.R., Davis, P.W.: 'The solution of ill-conditioned power system state estimation problems via the method of Peter and Wilkinson', *Power Appar. Syst., IEEE Trans.*, 1983, **102**, (10), pp. 3473–3480

15 Gjelsvik, A., Aam, S., Holten, l.: 'Hatchel's augmented matrix method - a rapid method improving numerical stability in power system state estimation', *Power Appar. Syst., IEEE Trans.* , 1985, **PAS-104**, (11), pp. 2987–2993

16 Marazzi, M., Nocedal, J.: 'Wedge trust region methods for derivative free optimization', *Math. Program.*, 2002, **91**, (2), pp. 289–305

17 Singh, H., Alvarado, F.: 'Weighted least absolute value state estimation using interior point methods', *Power Syst., IEEE Trans.*, 1994, **9**, (3), pp. 1478–1484

18 Celik, M., Abur, A.: 'A robust WLAV state estimator using transformations', *Power Syst., IEEE Trans.*, 1992, **7**, (1), pp. 106–113

19 Falcao, D., De Assis, S.: 'Linear programming state estimation: error analysis and gross error identification', *Power Syst., IEEE Trans.*, 1988, **3**, (3), pp. 809–815

20 Kumar, V.S.S.: 'Development of algorithms for power system state estimation incorporating synchronized phasor measurements'. Master's thesis, Indian Institute of Science, Bangalore, India, January 2011

21 Luenberger, D.G.: 'Linear and nonlinear programming' (Springer Science, 2005)

22 Andersen, K.D., Andersen, A.E.D.: 'Presolving in linear programming', *Math. Program.*, 1995, **71**, pp. 221–245

23 Brearley, A., Mitra, G., Williams, H.: 'Analysis of mathematical programming problems prior to applying the simplex algorithm', *Math. Program.*, 1975, **8**, (1), pp. 54–83

24 Steihaug, T.: 'The conjugate gradient method and trust regions in large scale optimization', *SIAM J. Numer. Anal.*, 1983, **20**, pp. 626–637

25 Dennis, J., Schnabel, R.: 'Numerical methods for unconstrained optimization and nonlinear equations' (Society for Industrial and Applied Mathematics, 1996)

26 'Power system test cases archive'. University of Washington, Tech. Rep. [Online]. Available at http://www.ee.washington.edu/research/pstca

IET Gener. Transm. Distrib., 2013, Vol. 7, Iss. 5, pp. 500–510
doi: 10.1049/iet-gtd.2012.0473