Algorithm Based Fault Tolerant State Estimation of Power Systems

Amitabh Mishra, Lamine Mili, and Arun G. Phadke

Abstract - Electric power industry deregulation has transformed state estimation from an important application into a critical one. State estimations in Power systems involve very tedious computations since large systems can consist of thousands of buses and lines. In this paper, we apply Algorithm Based Fault Tolerance techniques to the Gauss-Newton iterative algorithm that solves a weighted least-squares (WLS) state estimator of a power system. These techniques efficiently detect computational errors due to transient and permanent faults in the hardware. We show that the modified state estimation algorithm has low computational overhead and is free from false alarms.

Keywords- Algorithm based fault tolerance, state estimation, Gaussian elimination, error detection, power systems.

I. INTRODUCTION

Electric power industry deregulation has transformed state estimation from an important application into a critical one. State estimations in power systems involve very tedious computations since large systems can consist of thousands of buses and lines [1-3]. However, computations of that magnitude are always prone to errors caused by transient or permanent failures in hardware and also software failures to certain extents (buffer overflows, etc). Computationally efficient and accurate results are essential, since power system security monitoring and analysis hinge on them.

Algorithm Based Fault Tolerance (ABFT) is a cost effective technique to detect and correct errors caused by permanent or transient failures in the hardware. ABFT is based on encoding the data at the system level in the form of some error correcting or error detecting code and then designing algorithms to operate on encoded input data to produce encoded output data. One of the key advantages of ABFT which makes it very attractive is that it does not require any modifications to the underlying hardware to provide fault tolerance. This technique has already been applied to numerical algorithms e.g. matrix multiplications, LU factorizations etc. [4, 5]; the fast Fourier transforms [6, 7], QR factorization [8], singular value decomposition [9], Laplace equation [14], and the multigrid method [15], [16] etc. running on general-purpose multiprocessors. Later the same set of problems were resolved using parallel algorithms [10, 11] running on general-purpose multiprocessors.

The ABFT technique is known as ABED (algorithm based error detection) when applied only for the detection of errors. The ABED technique provides fault tolerance by preserving an invariant before the start and at the end of all computations in an algorithm. It can also preserve an invariant during the computations by slightly modifying the algorithm. Lack of preservation of the invariant during the course of the algorithm signals presence of errors in the computation due to faults. Since most of these computations are performed on computers that have finite precision, sometimes invariants are not preserved because of roundoff errors in the computation of the invariant and in the algorithm, and the ABFT technique can signal presence of faults even in the absence of any actual hardware faults. This situation is known as false alarms. Invariant computations, therefore should be corrected for roundoff errors to eliminate or minimize the false alarms by using some kind of tolerance.

It has been shown that methods which compute tolerance using error analysis of the algebraic manipulations in the algorithm are superior to other methods [12]. The error analysis provides upper bounds of roundoff errors which can occur in the computation of an expression on a machine with finite precision. The invariants when modified by adding the computed value of the roundoff error accumulated in an algorithm totally solves the problem of false alarms for any data set of any magnitude or dimension.

The solution of a large set of the order of several thousand of simultaneous algebraic equations is still an important issue in numerical linear algebra. Such large systems of equations often arise in load flow studies and the state estimation in the power system as well as when a partial differential equation modeling electromagnetic field, heat conduction or fluid flow is discretized using finite differences or finite elements. Typical methods for solving system of linear algebraic equations are either direct or iterative. Gaussian elimination and all its variants fall under the category of direct methods. A direct method computes the solution to a finite precision in a finite number of operations. Operation counts for Gaussian elimination vary from $O(N^3)$ to $O(N^2 \log N)$ for $N$ unknowns depending upon the properties of the coefficient matrix and how the sparsity structure of the matrix is exploited. The minimum operation count that can be achieved...
$O(N^2)$ is for the minimum degree algorithm, which is nearly optimal [13].

This paper presents an application of ABFT on a coefficient matrix that has been obtained as a result of state estimation algorithm of a power system that is solved using Gaussian elimination. We use error analysis to derive expressions for the upper bound on roundoff errors in the Gaussian elimination algorithm, which are then used to correct the invariants that are used in the checking step. We show that the modified Gaussian elimination algorithm has low computational overhead and is free from false alarms.

The paper is organized as follows. Section II contains a brief description of the estimation algorithm that we have considered for the decomposition using Gaussian elimination. Section III presents an ABFT encoded Gaussian elimination including the expressions for the row and column checksums that are used as invariants in the checking step and the derivations of expressions for the round off errors accumulated in the elements of the decomposed matrix.

II. WLS State Estimation

The role of static state estimator is to provide a complete, coherent, and reliable base-case power flow solution for contingency analysis and load forecasting function, among others [1, 2]. The state estimator processes a redundant collection of measurements based on a mathematical model that relates these measurements to the nodal voltage magnitudes and phase angles, which are taken as the state variables of the system. This model is derived from Kirchhoff’s and Ohm’s laws and hinges on several assumptions, which include: (1) random measurement errors that are Gaussian and uncorrelated with zero mean and known diagonal covariance matrix, $R = \text{diag} (\sigma^2)$; (2) a balanced three-phase system, (3) no time-skew between the metered values, (4) the exactness of the pi-equivalent models of the lines and transformers of their parameters, and (5) a known topology of the network. The latter is determined by a topology processor from the metered or given status of the circuit breakers of the system equipments such as lines, transformers, capacitors and inductors, FACTS devices, to name a few.

A general state estimation model is given by

$$z = h(x) + e,$$  

(1)

where $z$ is the measurement vector, $e$ is the measurement error vector, and $h(x)$ is a vector-real value function. The latter is derived from the foregoing assumptions; in particular, it depends on the given values for the line and transformer parameters and the given network topology. Many commercial software packages seek a solution to the WLS normal equation,

$$H^T R^{-1} (z - h(x)) = 0,$$  

(2)

through the Gauss-Newton iterative algorithm [3] expressed as

$$(H^T R^{-1} H) \Delta^{(k)} = H^T R^{-1} (z - h(x^{(k)}))$$  

(3)

where $H = \partial h(x)/\partial x$ is the Jacobian matrix, $\Delta^{(k)} = z - h(x^{(k)})$ is the residual vector, and $\Delta^{(k)} = x^{(k+1)} - x^{(k)}$ is the increment of the state vector at the $k$th iteration step. The error covariance matrix of the state estimates, Cov(δ$x$) = $\Sigma_x$, and of the estimates associated with the measurements, Cov(δ$z$) = $\Sigma_z$, are respectively given by

$$\text{Cov}(\delta x) = \Sigma_x = (H^T R^{-1} H)^{-1}$$  

(4)

$$\text{Cov}(\delta z) = \Sigma_z = H (H^T R^{-1} H)^{-1} H^T$$  

(5)

The diagonal elements of $\Sigma_x$ and $\Sigma_z$ are the variances of the estimate errors. Under the assumptions stated earlier and under the validity of the linear approximation of the model given by (2) around the solution, these variances provide a measure of the accuracy of the estimates; the smaller they are, the more accurate the state estimator solution is.

III. ABFT Encoded Gaussian Elimination

The solution of a set of linear algebraic equations of the form $Ax = b$ for a single right hand side $b$ is usually obtained by performing Gaussian elimination to create an upper triangular matrix. The system is then easily solved by back substitution from the last unknown upwards. Here matrix $A = (H^T R^{-1} H)$ given by equation (3). The basic algorithm (incorporating no error detection features) proceeds as follows: For a system of size $m$ there are $m - 1$ iterations in the algorithm from 1 to $m - 1$, during each of which one more column of 0’s is introduced into matrix $A$. Iteration $k$ of the algorithm results in the update of an $(n-k) \times (n-k)$ sub-matrix. The update step of the algorithm can be given as

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - a_{ik}^{(k)} \times a_{jk}^{(k)}/a_{kk}^{(k)}, \quad k \leq i \leq m, \quad k < j \leq m$$  

(6)

where $a_{ij}^{(k)}$ represents the element in the $i$th row and $j$th column of $A$ at the start of the $k$th iteration. Error detection capability is introduced into the algorithm by the introduction of row and column checksums which are described below. Let us denote the $i$th row of $A$ by $\text{row}$. We also denote the column checksum of the $j$th column of first $k-1$ rows at the start of the $k$th iteration by $CCA_{ij}^{(k)}$ and the column checksum of the remaining rows by $CCB_{ij}^{(k)}$. We denote the row checksum of row $i$ at the start of the $k$th iteration by $RC_{ij}^{(k)}$. At the start of the $k$th iteration of the algorithm, assuming error-free operation of the algorithm for the first $k-1$ steps, the following invariants are maintained:

1. $I_1 = RC_{ij}^{(k)}$ for all rows;
2. $I_2 = CCA_{ij}^{(k)}$ and $CCB_{ij}^{(k)}$ for all columns.
The algorithm computes the row checksum of all the rows and partial column checksums. At the start of the first iteration invariants $I_1$ and $I_2$ are correctly maintained for each row and each column. Then at the start of the $k^{th}$ iteration we first update the tolerances for the row and column checksums for the current iteration using the expressions derived in the next section. We next perform row and column checksum checks on the $k^{th}$ row and column of the results of the previous iteration. Since $a_k^{(k)}$ is known for all rows the column checksum computation for column $k$ can be easily computed. Then a row checksum check on row $k$ is performed. If both the row checksum and the column checksums pass, it means there is no error in $k^{th}$ row and $k^{th}$ column. The row checksums are updated in the following manner:

$$ R_c^{(k+1)} = R_c^{(k)} - a_k^{(k)} \times R_c^{(k)}/a_k^{(k)}, \quad \forall \ j > k $$  \hspace{1cm} (7)

and the column checksums $CCA$ and $CCB$ as:

$$ CCA^{(k+1)} = CCA^{(k)} + \delta_k^{(k)}, \quad \forall j \leq k $$  \hspace{1cm} (8)

$$ CCB^{(k+1)} = CCB^{(k)} - a_k^{(k)} \times CCB^{(k)}/a_k^{(k)}, \forall j \leq k $$  \hspace{1cm} (9)

We then proceed to update its rows of $(n-k) \times (n-k)$ sub-matrix using equation (6). If no errors are introduced in the $k^{th}$ iteration, the start of $(k+1)^{th}$ iteration results in the preservation of $I_1$ and $I_2$.

**A. Error Analysis**

In the following, we shall provide the derivations of the error expressions for the Gaussian elimination algorithm which is described by equation (6). We denote by notation $|\text{err}(a)|$, the upper bound on the absolute error associated with a variable $a$ that results due to a floating point operation on $a$. We can write therefore for a variable $v$,

$$ |\overline{v}| \leq |v| + |\text{err}(v)| $$  \hspace{1cm} (10)

knowing that any floating point operation $z = \text{fl}(x \oplus y)$ results in an error $\text{err}(z) = (x+y)\delta$, where $|\delta| \leq \varepsilon = 2^{-t}$ for a $t$ digit mantissa [11].

In these derivations the variables $LE(x)$ and $GE(x)$ denote the local error resulting from the last computation of a quantity $x$, and the global error or the accumulated error over all steps of computations respectively. The variable $EB(x)$ is used to store an upper bound on the absolute value of the global error in $x$. So we have $EB(x) \geq GE(x)$. In the derivations that follow, $O(\varepsilon^2)$ terms are dropped from the inequalities wherever they arise.

Equation (6) leads us to the algebraic expression which we need to analyze for round off errors for the Gaussian elimination and has the form:

$$ z^* = \text{fl}(v^* - w^* \times x^*/y^*) $$  \hspace{1cm} (11)

where, $v^* = v + GE(v), \ldots, z^* = z + GE(z)$. Then we have

$$ |GE(z)| \leq |xGE(w)/y| + |wGE(x)/y| + |v| \varepsilon + |wGE(y)/y^2| + |GE(v)| $$  \hspace{1cm} (12)

**Proof:** Let’s denote:

(i) $t_1 = \text{fl}(w^* x^*)$

(ii) $t_2 = \text{fl}(t_1/y^*)$

(iii) $t_3 = \text{fl}(v^* - t_2)$

Let’s simplify case (i) first.

$$ t_1 = \text{fl}(w^* x^*) = \text{fl}((w + GE(w))(x + GE(x))) $$

$$ = (w + GE(w))(x + GE(x))(1 + \delta_1) $$

$$ = wx + xGE(w) + wGE(x) + GE(w)GE(x) + \delta_1(wx + xGE(w) + GE(x)) $$

$$ = wx + xGE(w) + wGE(x) + \delta_1 wx $$

upon ignoring the second order terms. Similar simplification of (ii) and (iii) yields:

$$ t_2 = \left(\frac{1}{y}\right)(wx+ xGE(w) + wGE(x) + \delta_1 wx - \frac{wx}{y} GE(y) + \delta_2 wx) $$

$$ z^* = z + GE(z) = \text{fl}(v^* - t_2) = (v + GE(v) - t_2)(1 + \delta_1) $$

$$ = v + GE(v) - \frac{1}{y}(wx + xGE(w) + wGE(x) + \delta_1 wx - \frac{wx}{y} GE(y) + \delta_2 wx) $$

$$ + \delta_2 wx $$

Since $z = v - wx/y$, we can further simplify the expression for $z^*$ and write it as

$$ GE(z) = -x \frac{GE(w)}{y} - w \frac{GE(x)}{y} + \delta_1 - \frac{wx}{y} (\delta_1 + \delta_2 + \delta_3) + \frac{wxGE(y)}{y^2} + GE(v) $$

Since $|\delta_i| \leq \varepsilon$, the last expression can be simplified to

$$ |GE(z)| = \left|\frac{xGE(w)}{y}\right| + \left|\frac{wGE(x)}{y}\right| + |v| \varepsilon + \left|\frac{3wx\varepsilon}{y^2}\right| + \left|\frac{wxGE(y)}{y^2}\right| + |GE(v)| $$

which is identical to (12).
Using equation (12) and (6) we can show that the recurrence for the upper bound for error in an individual element holds and is given by (13) below.

\[ EB(a^{(i+1)}) \leftarrow \left( |a^{(i)}_v| +3 |a^{(i)}_v a^{(i)}_w / a^{(i)}_w| \right) \varepsilon + EB(a^{(i)}) + |a^{(i)}_v EB(a^{(i)}) / a^{(i)}_w| + |a^{(i)}_v a^{(i)}_w EB(a^{(i)}) / (a^{(i)}_w)^2 \]  

Equation (13) can be further simplified to

\[ EB(a^{(i+1)}) \leftarrow 4 |a^{(i)}_v| \varepsilon + EB(a^{(i)}) \]  

which is arrived at by using the simplification \( |a^{(i)}_v| \geq |a^{(i)}_v| \) in equation (13). Using equation (14) and the fact that \( n-k \) elements are summed in the \( k^{th} \) iteration the following approximate error expression for the column checksums is obtained

\[ EB(CCB^{(i+1)}) \leftarrow (n-k)EB(a^{(i+1)}) \]  

Similarly one can derive the error bound expression for the row checksum using the procedure described above and it’s given by equation (16).

\[ EB(RCB^{(i+1)}) \leftarrow \sum_{j=1}^{n} (n+1-i) |a^{(i)}_v| \varepsilon, \forall i \]  

These error expressions are used to compute error bounds which are then added to the invariants to protect the algorithm from reporting false alarms.

### IV. SIMULATION RESULTS

The error coverage results for the ABFT encoded Gaussian elimination method described in the previous section are obtained by fixing the range and the size of the data sets and then varying one parameter while fixing the others as constants. The results that are reported here are the error coverage, significant error coverage for significance levels of 2 and 10 and the error acceptance level to display the robustness of algorithm. Here significance level \( \gamma \) is defined as the computation error in which the norm of deviation of the faulty result from the exact result is bounded by \( \gamma \) times the norm of deviation of fault–free results from exact results. The error acceptance level (EAL) of a test is defined as the smallest value of the significance \( \gamma \) which results in 100% error detection by the test. These comparisons are performed for transient bit-level floating addition errors, which are expected to have the lowest coverage of all classes of errors. We report results on timing overhead of this scheme as compared to the basic Gaussian elimination algorithm with no checks. We finally report the error latency for this algorithm for transient bit-level floating-point errors, since these are expected to be the hardest to detect.

#### A. Error Coverage

Tables 1 and 2 contain the error coverage results for the Gaussian elimination algorithm. Both tables contain the following columns representing the percentage computed values: The FA (False Alarm), EC (Error Coverage), SEC (2) (Significant Error Coverage – Significance Level 2) and SEC (10) (Significant Error Coverage – Significance Level – 10). The tolerance values used in these results came from the error analysis that was presented earlier. Table 1 shows the error coverage results with respect to the data range. The error coverage varies from 80-91% for a data range of 0.1 to 10000 respectively. This low variation suggests that the scheme is very robust against the data range variation.

<table>
<thead>
<tr>
<th>Data Range</th>
<th>FA</th>
<th>EC</th>
<th>SEC (2)</th>
<th>SEC (10)</th>
<th>EAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0</td>
<td>80</td>
<td>87</td>
<td>97</td>
<td>17.05</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>79</td>
<td>90</td>
<td>98</td>
<td>8.17</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>84</td>
<td>91</td>
<td>99</td>
<td>5.37</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>86</td>
<td>93</td>
<td>100</td>
<td>3.73</td>
</tr>
<tr>
<td>1000</td>
<td>0</td>
<td>90</td>
<td>94</td>
<td>99</td>
<td>5.60</td>
</tr>
<tr>
<td>10000</td>
<td>0</td>
<td>91</td>
<td>96</td>
<td>100</td>
<td>2.38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Range</th>
<th>FA</th>
<th>EC</th>
<th>SEC (2)</th>
<th>SEC (10)</th>
<th>EAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0</td>
<td>84</td>
<td>91</td>
<td>100</td>
<td>3.58</td>
</tr>
<tr>
<td>60</td>
<td>0</td>
<td>86</td>
<td>93</td>
<td>100</td>
<td>3.73</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>86</td>
<td>97</td>
<td>100</td>
<td>2.51</td>
</tr>
<tr>
<td>125</td>
<td>0</td>
<td>84</td>
<td>94</td>
<td>99</td>
<td>2.50</td>
</tr>
<tr>
<td>250</td>
<td>0</td>
<td>84</td>
<td>90</td>
<td>99</td>
<td>2.30</td>
</tr>
<tr>
<td>500</td>
<td>0</td>
<td>84</td>
<td>90</td>
<td>99</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Table 2 shows the variation of error coverage with the change in the size of the data set, other parameters being held constant. The error coverage and error acceptance levels do not show much variation with data set size either.

From the results contained in Tables 1 and 2 it’s clear that the scheme is robust which achieves high error coverage and low error acceptance levels across data sets of widely varying ranges and sizes. Significant error coverage levels are in the range of 87-96. Error acceptance levels are around 20 in the worst case, while error acceptance levels of around 10 are common. Please note that an error acceptance level of 100 would correspond to an error of a fraction of 1% in the final results for almost all applications. The scheme presented in the paper is not observed to give any false alarms as seen from column 2 of Tables 1 and 2. Note that the coverage results for memory errors as well as truly gross errors such as word level errors and control flow errors were found to be always 100% but these are not reported here due to the length of the paper considerations. These will be presented in a paper in the near future.
B. Timing Results

For any proposed error detection algorithm, it’s desired that the error detection latency be as small as possible. The error latency is defined as the time taken between the appearance of an error and its detection. One is interested in devising error detection schemes with low error latency in order to minimize data corruption.

The error latency is calculated in two ways. First, error latency is reported as the number of checks which are passed without the error being detected after the injection of the error. This is achieved by setting a flag when the error is first injected in a run and incrementing a counter every time a check is passed thereafter with the error not being detected. An error may be detected in subsequent checks even if it was not detected in the first check following its insertion due to error magnification, propagation, or due to the occurrence of additional errors. Second, error latency is also reported as a percentage of a total runtime.

From Figure 1 we note that overhead incurred in introducing checks into basic algorithm can be amortized by going to larger matrix sizes. One can see from these figures that the overhead decreases with the increase in the problem size. This is not unexpected since the check steps involve $O(N^2)$ operations while the basic algorithm required $O(N^3)$ operations for a system of size N. Thus, the additional overhead caused by introduction of the check steps is not of concern for large problem sizes.

![Figure 1: Time Overhead of the current scheme over algorithm with no checks.](image)

C. Error Latency

Figures 2 and 3 indicate the error latencies for the Gaussian elimination application for transient bit-level floating-point errors for various bits in error starting with the first bit whose error was consistently detected. As the erroneous bit position becomes more and more significant, the error latency decreases until beyond a certain bit, all bit level errors are detected in the first check following the error. For larger problem sizes, error latency for errors in more significant bits corresponds to a smaller percentage of the total run-time since the error is detected at the first check following its occurrence. Since there is a check following each iteration, and for larger problem sizes the execution time for each iteration is a smaller fraction of the total runtime, the error latency also decreases with matrix size for errors in more significant bits. But, this need not be true for errors in less significant bits since it was often observed that the latency in terms of number of checks passed following error injection tended to decrease with problem size. However, the figures indicate that errors in more significant bits were detected in the first check following error injection and the latency amounted to no more than 5% of the total runtime. Error latency studies for word level floating point and all other classes of errors injected were also conducted and in all of our simulation runs, we found that these errors were detected at the very next check following their injection.

![Figure 2: Error latency as number of checks passed before detection](image)

![Figure 3: Error Latency as percentage of total runtime](image)
V. CONCLUSIONS

In this paper the algorithm based error detection scheme has been applied to the Gaussian elimination method for solving a WLS power system state estimator. We have derived the error expressions for the round off errors in the algebraic processes within the elimination procedure and modified the invariants for checking by taking into account the accumulated round off errors. The Gaussian elimination algorithm with checking was implemented on a SUN Enterprise server for experimental evaluation. Our results suggest that the algorithm provides error detection at low costs (less than 5% for 500 unknowns), excellent error coverage (98-100%) for floating point arithmetic in the presence of permanent bit and word errors. Moreover, the implemented algorithm is insensitive to the data range and the data size and reports no false alarms.

VI. REFERENCES