# Observability analysis and bad data processing for state estimation using Hachtel's augmented matrix method†

FELIX F. WU‡, WEN-HSIUNG E. LIU‡, LARS HOLTEN§, ANDERS GJELSVIK¶, and SVERRE AAM¶

The triangular-factorization-based observability analysis and the normalized-residual-based bad data processing are extended to state estimation using Hachtel's augmented matrix method. This method is numerically robust, computationally efficient, and has a reasonable extra storage requirement. In this paper it is shown that the observability analysis can be carried out in the process of triangular factorization of the augmented coefficient matrix used in Hachtel's method. Moreover, the normalized residuals are shown to be obtainable using the sparse inverse of this augmented matrix. The algorithms have been successfully incorporated in the state estimation program developed at Norwegian State Power Board (Statkraft). Test results on the IEEE 14 bus system and a 99-bus system consisting of the main grid of southern Norway are presented. Hachtel's approach to state estimation provides an attractive alternative to the standard normal equations approach.

## 1. Introduction

State estimation in a power system is formulated as a nonlinear weight least square (WLS) problem. Traditionally this problem is solved iteratively by the normal equations approach (Schweppe and Handschin 1974).

This approach is known to exhibit poor numerical stability when the system is ill-conditioned (Golub and Van Loan 1983). In such cases, the nonlinear WLS problem may take more iterations to converge, or fail to converge at all. In power system state estimation, ill-conditioning does occur and the cause of ill-conditioning may be artificial or innate. It is experienced that artificially assigning very large and very small weights for different types of pseudo-measurements, for example, large weights for *zero-injections*, usually causes a convergence problem (Aschmoneit, Peterson and Adrian 1977). It is shown that the existence of a connection of long and short transmission lines in the power system may lead to ill-conditioning (Monticelli, Murari and Wu 1985). Alternative WLS solution algorithms exist in the literature. The formulation of WLS with equality constraints has been used to handle zero-injections (Aschmoneit, Peterson and Adrian 1977). Orthogonal transformation methods have been suggested (Monticelli *et al.* 1985; Simoes-Costa and Quintana 1981a, b; Gu, Clements, Krumpholz and Davis 1983; Wang and Quintana 1984; U.S. Department of Energy 1984) to ease the ill-conditioning problem.

Duff and Reid (1976) have conducted an extensive comparison of various means of solving the sparse linear WLS problem. The comparison considers numerical stability, preservation of sparsity, and computation time, based on numerical experiments on a number of random, as well as some structured, matrices. Their conclusion is that for equations with widely differing scales, the Peters–Wilkinson algorithm is recommended, otherwise Hachtel's algorithm is recommended. Gjelsvik, Aam and Holten (1985) have applied Hachtel's method with equality constraints to power system state estimation. Since zero-injections are treated as equality constraints, the remaining equations do not have widely differing scales. Moreover, since the equality constraints are included in the formulation, they are much easier to deal with. Gjelsvik, Aam and Holten have compared the performance of the proposed approach with the normal equations approach and the normal equations with equality constraints on a realistic 99-bus Norwegian power system and confirmed that for power system application, the Hachtel's method is numerically more stable, computationally efficient, reasonable in extra storage requirement, and simple to implement.

Estimating the state is only a part of the larger function of the state estimation software. Other related functions include observability analysis, bad data processing, and external network modeling (Wu and Monticelli 1986). Two approaches for observability analysis have been proposed; graph-theory based (Krumpholz, Clements and Davis 1980; Clements, Krumpholz and Davis 1983; Quintana, Simoes-Costa and Mandel 1982; VanCutsem 1985) and triangular-factorization-based (Monticelli and Wu 1985a, b; Monticelli and Wu 1986). The triangular-factorization-based observability analysis is derived for state estimation using the normal equations approach. It uses the same subroutines already in the state estimation and is simple and efficient. For bad data detection and identification, it is generally agreed that the methods based on the normalized residuals (Handschin, Schweppe, Kohlas and Fiechter 1975; Monticelli and Garcia 1983; Monticelli, Wu and Yen 1986) are effective. The computational burden of calculating the normalized residuals can be greatly reduced for state estimation using the normal equations approach by the use of the sparse inverse of the gain matrix (Broussolle 1978).

In state estimation using Hachtel's method with equality constraints, instead of the gain matrix, another matrix equation is solved at each iteration. Since the gain matrix is never formed here, the triangular-factorization-based observability analysis and the procedure for calculating normalized residuals for bad data processing derived for state estimation with normal equations are not directly applicable. In this paper, however, we extend the factorization-based observability analysis and the normalized-residual-based bad data processing to state estimation using Hachtel's method with equality constraints. Recently, the factorization-based observability analysis and normalized-residual-based bad data processing have been extended to state estimation using normal equations with equality constraints (Wu, Liu, and Lun 1987).

The application of Hachtel's method to state estimation is reviewed in § 2. The algorithm for observability analysis is discussed in § 3. The procedure for calculating the normalized residuals using the sparse inverse technique is discussed in § 4. The derivations of the results used in §§ 3 and 4 are included in the Appendices. Test results on the IEEE 14-bus system and the Norwegian 99-bus system are presented in § 5. The conclusion is presented in § 6. We have called the augmented matrix approach Hachtel's method following Duff and Reid (1976). Using the augmented

matrix for solving the WLS problem has been previously proposed by many researchers, including Siegel (1965) and Bjorck (1967). Since the use of the augmented matrix for the direct solution of the linear WLS problem for the sparse case was suggested to Duff and Reid (1976) by Hachtel, they have called it Hachtel's method.

## 2. Hachtel's method

The equations relating the telemetered measurements and the state variables in power systems are

$$z = h(x) + w \tag{1}$$

where $z$ is the $(m \times 1)$ measurement vector, $h(\cdot)$ is the $(m \times 1)$ vector of nonlinear functions, $x$ is the $(n \times 1)$ true state vector, $w$ is the $(m \times 1)$ measurement error vector and $m$ is the number of measurements.

The equality constraints of the zero injections are expressed by a set of nonlinear equations

$$c(x) = 0 \tag{2}$$

where $c(\cdot)$ is the $(r \times 1)$ vector of nonlinear functions. These zero injection constraints are to be satisfied.

The problem is to find an estimate of the state vector $\hat{x}$ which minimizes the weighted least square $J(x) = [z - h(x)]^T W[z - h(x)]$ while the equality constraints $c(x) = 0$ are satisfied, i.e.,

$$\min J(x) = [z - h(x)]^T W[z - h(x)]$$

*subject to*

$$c(x) = 0 \tag{3}$$

where $W$ is a diagonal matrix whose $i$th element $W_i$ is the weight for the $i$th measurement and is equal to the reciprocal of the variance of the $i$th measurement error.

The nonlinear WLS problem (3) may be solved for $\hat{x}$ by an iterative procedure, at each iteration the linearized problem of (3) is solved. Hachtel's method with equality constraints has been proposed to solve the linearized WLS problem, in which the following equations are solved:

$$\begin{bmatrix} 0 & 0 & C \\ 0 & \alpha W^{-1} & H \\ C^T & H^T & 0 \end{bmatrix} \begin{bmatrix} -\alpha^{-1}\lambda \\ \alpha^{-1}Wr \\ \Delta x \end{bmatrix} = \begin{bmatrix} \Delta c \\ \Delta z \\ 0 \end{bmatrix} \tag{4}$$

where

$$H := \frac{\partial h}{\partial x}(x) \quad \text{and} \quad C := \frac{\partial c}{\partial x}(x)$$

are the Jacobian matrices, $\Delta z := z - h(x)$, $\Delta c := -c(x)$, $\Delta r = \Delta z - H\Delta x$ and $x = x^k$ at the $k$th iteration, $\lambda$ is the Lagrange multiplier, and $\alpha$ is a parameter to control the numerical stability of the problem (Gjelsvik *et al.* 1985) that has been set to one in the tests. The 'optimal' selection of $\alpha$ to enhance stability is discussed in (Liu *et al.*

1987). It should also be pointed out that the weighting factors can be applied to $\mathbf{H}$ and $\mathbf{z}$ first as is done in (Gjelsvik *et al.* 1985).

We shall use $\mathbf{K}(\mathbf{x})$ to denote the augmented coefficient matrix in (4), i.e.,

$$\mathbf{K}(\mathbf{x}) := \begin{bmatrix} 0 & 0 & \mathbf{C} \\ 0 & \alpha\mathbf{W}^{-1} & \mathbf{H} \\ \mathbf{C}^{\mathrm{T}} & \mathbf{H}^{\mathrm{T}} & 0 \end{bmatrix} \tag{5}$$

Obviously a decoupled (P and Q) version of Hachtel's method can easily be derived using the decoupled $\mathbf{C}$ and $\mathbf{H}$.

$$\mathbf{H} \leftarrow \begin{bmatrix} \mathbf{H}_{P\theta} & 0 \\ 0 & \mathbf{H}_{QV} \end{bmatrix} \quad \mathbf{C} \leftarrow \begin{bmatrix} \mathbf{C}_P & 0 \\ 0 & \mathbf{C}_Q \end{bmatrix} \tag{6}$$

The $\mathbf{K}(\mathbf{x})$ will be decoupled into

$$\mathbf{K}(\mathbf{x}) \leftarrow \begin{bmatrix} \mathbf{K}_P(\mathbf{x}) & 0 \\ 0 & \mathbf{K}_Q(\mathbf{x}) \end{bmatrix} \tag{7}$$

where $\mathbf{K}_P(\mathbf{x})$ and $\mathbf{K}_Q(\mathbf{x})$ have the same form as in Eqn. (4).

Note that the coefficient matrix $\mathbf{K}(\mathbf{x})$ is sparse, symmetric, but indefinite. An ordinary sparsity-oriented pivoting scheme for symmetric matrices may lead to numerical stability problems. However a good symmetric factorization may be obtained using $2 \times 2$ pivots as well as simple $1 \times 1$ pivots as described in Appendix C.

It has been shown analytically (Liu *et al.* 1987) that Hachtel's method is numerically more robust than the method of state estimation with equality constraints (Aschmoneit *et al.* 1977; Wu 1987). More precisely, the error in the solution introduced due to the error in the input data $(\mathbf{H}, \mathbf{C}, \mathbf{z})$ can be much larger if the method of state estimation with equality constraints, rather than Hachtel's method, is used.

## 3.   Observability analysis

The theory of observability analysis (Krumpholz *et al.* 1980; Clements *et al.* 1983; Quintana *et al.* 1982; VanCutsem 1985; Monticelli and Wu 1985a, b; Monticelli and Wu 1986) was derived using the decoupled measurement Jacobian at the flat start (unity voltage at zero angle). In the following, the measurement (telemetered and zero-injections) Jacobian really refers to

$$\begin{bmatrix} \mathbf{C}_P^0 \\ \mathbf{H}_{P\theta}^0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \mathbf{C}_Q^0 \\ \mathbf{H}_{QV}^0 \end{bmatrix},$$

where superscript 0 denotes that the quantity is evaluated at the flat start.

A network is observable if for the given set of measurements the state estimation is solvable. The state estimation is solvable if and only if the measurement (including telemetered and zero-injections) Jacobian matrix $\begin{bmatrix} \mathbf{C} \\ \mathbf{H} \end{bmatrix}$ has full rank. It has been shown (Monticelli and Wu 1985a) that

—  If the Jacobian matrix has rank deficiency $p$, then the network needs $p$ angle (voltage) pseudo-measurements to make it observable.

—  The location of these angle (voltage) pseudo-measurements, together with state

estimation runs, can be used to identify the observable islands, the unobservable branches, and candidates for measurement (flow, injection, voltage) to make the network observable.

— If the rank of the measurement Jacobian matrix $\begin{bmatrix} \mathbf{C} \\ \mathbf{H} \end{bmatrix}$ is deficient, a zero pivot will be encountered in the triangular factorization of the gain matrix $[\mathbf{C}^T \mathbf{H}^T] \begin{bmatrix} \mathbf{C} \\ \mathbf{H} \end{bmatrix}$

— When the zero pivot occurs, the remaining elements in the column below it and the row beyond it are necessarily all zero. (We shall refer to this as a zero pivot with null row/column.)

In Appendix A, it is shown that there is a one-to-one correspondence in rank deficiency between $\begin{bmatrix} \mathbf{C} \\ \mathbf{H} \end{bmatrix}$ and the augmented matrix

$$\mathbf{K} := \begin{bmatrix} 0 & 0 & \mathbf{C} \\ 0 & 1 & \mathbf{H} \\ \mathbf{C}^T & \mathbf{H}^T & 0 \end{bmatrix} \tag{9}$$

(for observability analysis, we may take $\alpha = 1$, $W_i = 1$). In other words, the zero pivot with null row/column is encountered when the network is unobservable and it occurs only at a column corresponding to an angle (voltage) variable, and cannot occur at a column corresponding to $\lambda$, or $\mathbf{r}$. Note that because the matrix $\mathbf{K}$ is indefinite, using a pivoting strategy that preserves symmetry, a zero pivot may occur without the remaining row/column being zero. Therefore, unlike the case with the gain matrix (Monticelli and Wu 1985a, b), when a zero pivot is encountered one has to check further whether the remaining row/column are all zeros. However this does not really introduce any additional work in implementation since in the factorization process, the largest (in absolute value) element of the remaining row/column $R_{max}$ is indeed monitored to ensure numerical stability. ( A pivot $p_i$ is accepted only if $|p_i| > u \cdot R_{max}$.) Therefore, observability analysis can be carried out during the triangular factorization of the coefficient matrix $\mathbf{K}$ in Hachtel's method for state estimation.

The basic procedure in the observability analysis consists of the following.

1. Form the coefficient matrix $\mathbf{K}$ in (9).

2. Perform triangular factorization of $\mathbf{K}$. If a zero pivot is encountered, check whether the elements in the remaining row and column are zero. If so, introduce angle (voltage) pseudo-measurement at the corresponding bus.

3. Solve the state estimation equation

$$\begin{bmatrix} \mathbf{K} \end{bmatrix} \begin{bmatrix} -\lambda \\ \mathbf{r} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{z} \\ 0 \end{bmatrix} \tag{10}$$

with all the measurements set to zero, except the angle (voltage) pseudo-measurements which are set at $0, 1, 2, \ldots$ etc.

4. Remove injection measurements which are adjacent to a branch that has non-zero flows. These are irrelevant measurements.

5. Update the coefficient matrix **K**. Perform triangular factorization. Introduce angle (voltage) measurements when zero pivot with zero remaining row and column are encountered.

6. Solve the state estimation equation (10) with all measurements set to zero, except the angle (voltage) pseudo-measurements set at 0, 1, 2, etc. The sub-network whose nodes have the same angle (voltage) belongs to the same observable island. The branches crossing two observable islands are the unobservable branches.

7. The candidates for additional measurements to make the network observable are

   (i) the voltage measurement where a voltage pseudo-measurement was needed in step 5.

   (ii) line flow measurements on unobservable branches

   (iii) injection measurements adjacent to unobservable branches

The foregoing observability analysis can be used to place additional measurements until observability is ensured.

If a decoupled version is used, P-observability and Q-observability may be analysed separately or if **P** and **Q** measurements always come in pairs, only one analysis is necessary. The observability analysis can also be implemented easily in an 'non-decoupled' state estimator. As a matter of fact, the P-observability and the Q-observability can be analysed simultaneously directly on the matrix **K** without explicit decoupling, provided that the fill-ins in the off-diagonal blocks of **C** and **H** in (6) are never entered. When a zero-pivot with null row/column is encountered, it is easy to determine whether it belongs to the **P** part or the **Q** part, thus whether an angle pseudo-measurement or a voltage pseudo-measurement needs to be added.

The proposed observability analysis algorithm has been implemented in the Norwegian National Control Center. The algorithm for the state estimation used happens to be in rectangular coordinates. It is interesting to note that the decoupled Jacobian (in polar coordinates) at the flat start can be obtained directly from the full Jacobian in rectangular coordinates at the flat start. Let the voltage at bus **k** be

$$E_k = e_k + f_k$$
$$= V_k e^{j\theta_k} \tag{11}$$

Let the Jacobian in rectangular coordinates be written

$$\begin{bmatrix} \dfrac{\partial \mathbf{P}}{\partial \mathbf{e}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{f}} \\ \dfrac{\partial \mathbf{Q}}{\partial \mathbf{e}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{f}} \end{bmatrix}$$

where **P** and **Q**, respectively, denote real power measurements and reactive power measurements (including voltage measurements.) Let the Jacobian in polar coordinates be written as

$$\begin{bmatrix} \dfrac{\partial \mathbf{P}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{V}} \\ \dfrac{\partial \mathbf{Q}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{V}} \end{bmatrix} \equiv \begin{bmatrix} \mathbf{H}_{P\theta} & \mathbf{H}_{PV} \\ \mathbf{H}_{Q\theta} & \mathbf{H}_{QV} \end{bmatrix} \tag{12}$$

By the chain rule we have

$$
\begin{bmatrix} \dfrac{\partial \mathbf{P}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{V}} \\[2mm] \dfrac{\partial \mathbf{Q}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{V}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \mathbf{P}}{\partial \mathbf{e}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{f}} \\[2mm] \dfrac{\partial \mathbf{Q}}{\partial \mathbf{e}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{f}} \end{bmatrix} \cdot \begin{bmatrix} \dfrac{\partial \mathbf{e}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{e}}{\partial \mathbf{V}} \\[2mm] \dfrac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{f}}{\partial \mathbf{V}} \end{bmatrix} \tag{13}
$$

When the Jacobian is evaluated at the flat start ($e_k = 1, f_k = 0$ or $\theta_k = 0, V_k = 1$) and under the assumption that the branch impedances are purely imaginary, it can be easily checked from the relevant mathematical equations that

$$
\begin{bmatrix} \dfrac{\partial \mathbf{P}}{\partial \mathbf{e}} & \dfrac{\partial \mathbf{P}}{\partial \mathbf{f}} \\[2mm] \dfrac{\partial \mathbf{Q}}{\partial \mathbf{e}} & \dfrac{\partial \mathbf{Q}}{\partial \mathbf{f}} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{H}_{Pf} \\ \mathbf{H}_{Qe} & 0 \end{bmatrix} \tag{14}
$$

$$
\begin{bmatrix} \dfrac{\partial \mathbf{e}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{e}}{\partial \mathbf{V}} \\[2mm] \dfrac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{f}}{\partial \mathbf{V}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{15}
$$

Combining eqs. (12–15) we have

$$
\mathbf{H}_{P\theta} = \mathbf{H}_{Pf}
$$
$$
\mathbf{H}_{Qv} = \mathbf{H}_{Qe} \tag{16}
$$

Thus the decoupled Jacobian matrix can be obtained from the full Jacobian matrix in rectangular coordinates. Similarly for **C**.

## 4. Bad data processing

Methods using the normalized residuals to detect and identify bad data have been found to be reliable for single or multiple bad data. It has been shown (Handshin *et al.* 1975; Monticelli and Garcia 1983) that for a single or multiple non-interacting bad data, the largest normalized residual corresponds to a bad data. For a general multiple bad data case, an efficient search scheme using normalized residuals has been developed (Monticelli *et al.* 1986).

The residual vector **r** is the difference between the measured quantities and the calculated quantities.

$$
\mathbf{r} = \mathbf{z} - \mathbf{h}(\hat{\mathbf{x}}) \tag{17}
$$

The normalized residuals are obtained from the residuals by the division of the square root of the diagonal elements of the covariance matrix of **r**, i.e.,

$$
\mathbf{r}^N = (\text{diag } \mathbf{R}_r)^{-1/2} \mathbf{r} \tag{18}
$$

where

$$
\mathbf{R}_r = \text{cov } (\mathbf{r}) \tag{19}
$$

In the normal equations approach to state estimation, the calculation of the diagonal elements of $\mathbf{R}_r$, as observed by Broussolle (1978), can be greatly facilitated by the use of the sparse inverse of the gain matrix ($\mathbf{H}^T\mathbf{W}\mathbf{H}$).

In Appendix B, it is shown that the covariance matrix $\mathbf{R}_r$ of the residual vector $\mathbf{r}$ is equal to

$$\mathbf{R}_r = \alpha\mathbf{W}^{-1}\mathbf{A}_4\mathbf{W}^{-1} \tag{20}$$

where

$$\begin{bmatrix} 0 & 0 & \mathbf{C} \\ 0 & \alpha\mathbf{W}^{-1} & \mathbf{H} \\ \mathbf{C}^\mathsf{T} & \mathbf{H}^\mathsf{T} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2^\mathsf{T} & \mathbf{A}_3^\mathsf{T} \\ \mathbf{A}_2 & \mathbf{A}_4 & \mathbf{A}_5^\mathsf{T} \\ \mathbf{A}_3 & \mathbf{A}_5 & \mathbf{A}_6 \end{bmatrix} \tag{21}$$

Hence, the diagonal elements of $\mathbf{R}_r$ which correspond to positions of non-zero element in $\mathbf{K}$ can be calculated using the sparse inverse of the coefficient matrix $\mathbf{K}$. The normalized residuals (18) can therefore easily be computed using the sparse inverse of $\mathbf{K}$. However, it should be pointed out that since $2 \times 2$ pivoting is employed in the triangular factorization, the sparse inverse technique needs to be extended to such a case, and this is described in Appendix C. Actually only certain diagonal elements of the inverse are required in computing the normalized residuals, the sparse vector technique (Tinney, Brandwain and Chan 1985) can be applied. For the tests performed in this paper, the sparse inverse technique is used. The detail implementation of sparse vector technique in calculating the normalized residuals will be reported elsewhere (Liu *et al*. 1987).

## 5.  Test results

### A. *IEEE 14-bus system observability analysis*

The measurement system of the IEEE 14-bus system used for the observability analysis is taken from Krumpholz, Clements, and Davis (1980) as shown in Fig. 1. Four cases are tested, and single precision is used. The branch admittances are set to $j1$ and the weighting factors are set to 1. The threshold for zero is set at 0·0001.
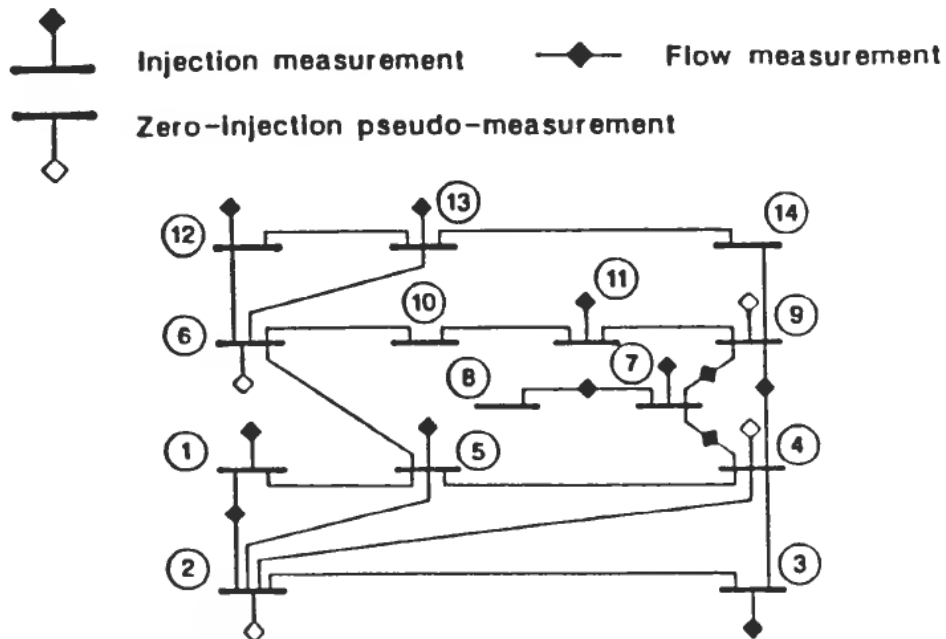


Figure 1.    IEEE 14-bus system for observability test.

Case 1. Complete measurement system, zero injections are treated as measurements.

Case 2. Injection measurement at bus 5 is removed, zero injections are treated as measurements.

Case 3. Complete measurement system, zero injections are treated as equality constraints.

Case 4. Injection measurement at bus 5 is removed, zero injections are treated as equality constraints.

Three methods for observability analysis have been employed for the test.

Method 1 Factorization-based observability analysis in state estimation using normal equations.

Method 2 Factorization-based observability analysis in state estimation using Hachtel's method as proposed in this paper.

Method 3 Graph-theory-based observability analysis (VanCutsem 1985).

All three methods give correct results, namely, Cases 1 and 3 are observable and Cases 2 and 4 are unobservable. In Cases 2 and 4, the network containing nodes (1, 2, 3, 4, 5, 7, 8, 9) forms an observable island, and the candidates for measurement placement are the injection measurements at buses 5, 10, and 14. However, when Method 1 is applied to Case 2, if the parameter $u$, which controls the trade-off between sparsity preservation and numerical stability in the triangular factorization routine of the software used (Harwell Subroutine Library routine MA27A), was set at $0.1$, then it mistakenly classifies the network as observable. The parameter $u$ is used for pivot testing. (The detail strategy is given in (Duff, Reid, Munksgaard and Nielsen 1979; Duff and Reid 1983) for both $1 \times 1$ and $2 \times 2$ cases.) A value of $u$ greater than $0.2$ is sufficient here to obtain correct results. This further confirms the fact that the formation of the gain matrix in the normal equation approach can make the system ill-conditioned (Monticelli *et al.* 1985). The misclassification is due to the round-off errors accumulated in an ill-conditioned system using an algorithm with unsatisfactory numerical stability. Another way of viewing this result is that the
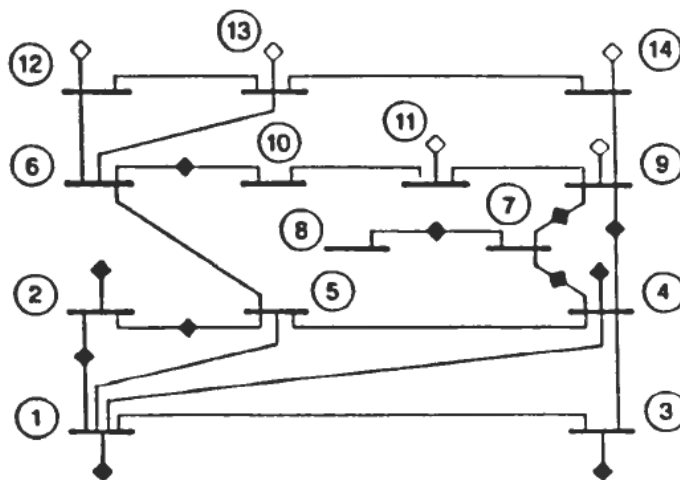


Figure 2. IEEE 14-bus system for bad data processing.

factorization-based observability analysis performs better with a more robust method such as that of Hachtel.


*Bad data processing*

The measurement system of the IEEE 14 bus-system used for testing bad data processing, shown in Fig. 2, is taken from Clements and Davis (1985). The impedances are all equal to $0 + j1$. The standard deviations are set at 0·01, and the weighting factors are set at 1. All the critical measurements ($\#4, 6, 8, 9, 10, 11, 13$) are treated as equality constraints. Two cases have been tested using the method for calculating the normalized residuals proposed in this paper.

Case 1.   Single bad data ($5\,\sigma$) in measurement $\#5$.

The normalized residuals are calculated using Eqn. (11). The three largest ones are shown below

| Measurement | $\#5$ | $\#3$ | $\#15$ |
|---|---|---|---|
| Normalized residual | 3·1160 | 2·9817 | 2·9763 |

The largest normalized residual corresponds to the bad data.

Case 2.   Multiple bad data in measurement $\#2\,(5\,\sigma)$ and $\#5\,(7\,\sigma)$

The three largest normalized residuals are

| Measurement | $\#2$ | $\#15$ | $\#7$ |
|---|---|---|---|
| Normalized residual | 4·2455 | 3·9030 | 3·2013 |

After removing the bad data $\#2$ and re-estimating the state, a set of new normalized residuals are obtained. The top three are

| Measurement | $\#5$ | $\#3$ | $\#15$ |
|---|---|---|---|
| Normalized residual | 3·1161 | 3·0559 | 3·0500 |

After removing measurement $\#5$, the normalized residuals pass the bad data detection test, hence, the two bad data are correctly identified.


B. *Norwegian 99-bus system*

Figure 3 is a simplified overview of the 99-bus system found in the main grid in southern Norway. The measurement system summarized below is the first metering configuration used in the state estimator of the National Control Center (Gjelsvik *et al.* 1987).
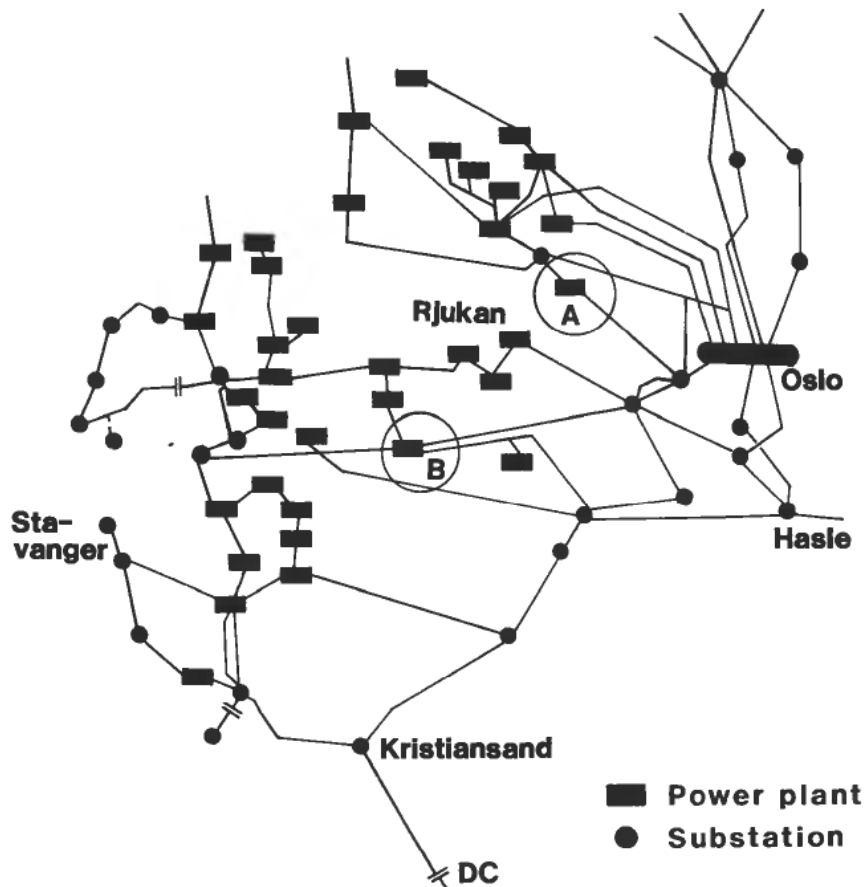
Figure 3.   Simplified Norwegian 99-bus network.

Measurements

| | |
|---|---|
| Active power | 72 |
| Reactive power flows | 8 |
| Active power injections | 28 |
| Reactive power injections | 39 |
| Voltages | 38 |
| Zero injections | 80 |
| Total | 265 |
| Redundancy | 1·35 |

*Observability analysis*

Three cases are tested.

Case 1.   Complete measurement system.

Case 2.   Loss of measurements in station A (1 voltage, 1 $P$-injection, 1 $Q$-injection, 2 P-line flows).

Case 3.   Loss of measurements in station B. (1 voltage, 1 $P$-injection, 1 $Q$-injection, 2 P-line flows).

Again three methods for observability analysis have been employed in the test, as described in § A. All give correct results. Case 1 is both **P** and **Q** observable. Case

2 is P-observable, but not Q-observable. Bus 3019 is unobservable (forms an observable island by itself). The injections at buses 3011 and 3019 are candidates for measurement placement. Case 3 is neither P-observable nor Q-observable. Bus 1750 is unobservable and the candidates for measurement placement are the injections at buses 1750 and 1752.

The CPU-time required for the three methods in the test is shown in Table 1. The computer used was a PRIME 9955. The software for the graph-theory-based observability analysis was provided by (VanCutsem 1985).

|  | Normal equations w/equality constraints | Hachtel's method w/equality constraints | Graph-theory method |
|---|---|---|---|
| Case 1 | 1·930 | 0·767 | 0·245 |
| Case 2 | 2·000 | 0·918 | 4·163 |
| Case 3 | 1·518 | 0·973 | 89·042 |

Table 1. Comparison of CPU-time (in seconds).

One should ignore column 1 in Table 1 because we have an inefficient implementation of the normal-equation state estimation. The CPU-time for factorization-based observability analysis in state estimation using Hachtel's method does not vary very much from case to case. The large CPU-time required in cases 2 and 3 in the graph-theory method are for $Q$-observability analysis. Note that the Norwegian 99-bus system only has 8 reactive power flows. When there are few line flow measurements, the graph-theory method for observability analysis must try a large number of combinations of assigning the injection measurements, rendering the method inefficient.

*Bad data processing*

The section of the network where bad data occurs is shown in Fig. 4. Two cases are tested using the proposed procedure for calculating the normalized residuals.

Case 1. Bad data ($5\sigma$) in line flow measurement 3062–1122.

The three largest normalized residuals are

| Measurement | 3062–1122 | 3062–3052 | 3062 injection |
|---|---|---|---|
| Normalized residual | 4·63 | 2·63 | 1·43 |

The bad data is correctly identified.

Case 2. Bad data ($5\sigma$) in line flow measurements 3062–1122 and 1102–1092.

The three largest normalized residuals are

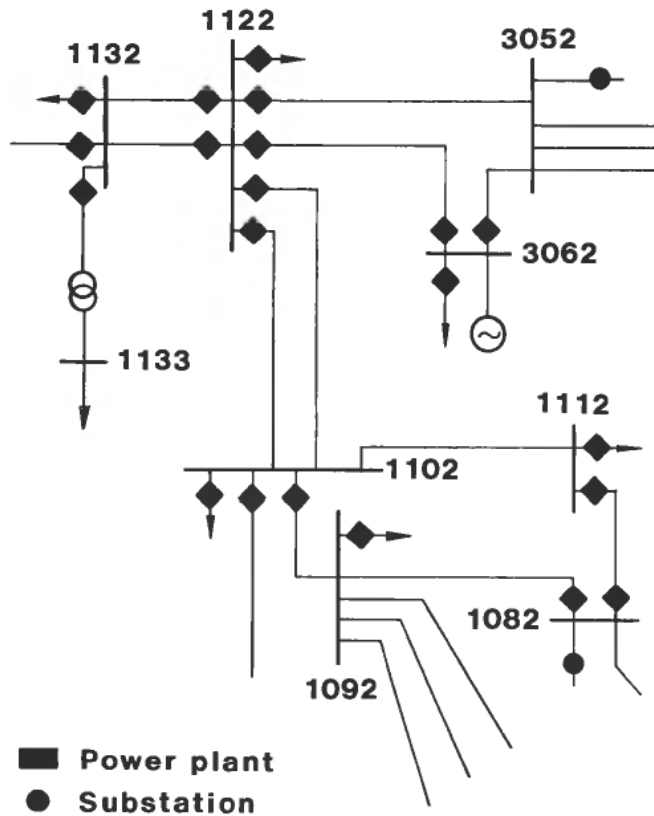| Measurement | 3062–1122 | 1102–1092 | 1092 injection |
|---|---|---|---|
| Normalized residual | 4·6589 | 3·8964 | 2·5546 |

Figure 4. Section of the 99-bus system for bad data.

After removing measurement 3062–1122 and re-estimating the state, the new normalized residuals have their maximum at measurement 1102–1092 with a value of 3·8592. After removing this measurement the bad data detection is passed, thus the two bad data are correctly identified.

The CPU-time required in various parts of the computation is shown in Table 2

| Optimal ordering | Triangular factorization | Back substitution | Sparse inverse |
|---|---|---|---|
| 0·25 | 0·57 | 0·07 | 0·55 |

Table 2. CPU-time (sec.) breakdown.

Note that the time required for computing the sparse inverse is about the same as that for triangular factorization.

## 6. Conclusion

Hachtel's augmented matrix method with equality constraints for the WLS problem has been applied to power system state estimation. It has previously been found that this method of estimating the state of a power system is numerically

more stable, computationally efficient, and simple to implement. In this paper we have extended the triangular-factorization-based observability analysis and the normalized-residual-based bad data processing to the state estimation algorithm using the Hachtel's method. In the Hachtel's method at each iteration the linearized equation has a different coefficient matrix $\mathbf{K}$ than the gain matrix in the normal equations approach. In this paper, we have shown that

— A network is observable if and only if in the process of triangular factorization of the coefficient matrix

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & \mathbf{C} \\ 0 & \mathbf{I} & \mathbf{H} \\ \mathbf{C}^{\mathrm{T}} & \mathbf{H}^{\mathrm{T}} & 0 \end{bmatrix}$$

no zero pivot with the remaining row and column being all zero, is encountered.

— When the network is unobservable, the result of the state estimation with all the measurements set to zero, except the angle (voltage) pseudo-measurements introduced during the triangular factorization process provides us with the observable islands, unobservable branches, relevant measurements, and candidates for additional measurements to make the network observable.

— The normalized residuals can be calculated using the sparse inverse of the coefficient matrix $\mathbf{K}$.

The factorization-based observability analysis and the normalized-residual-based bad data processing have been implemented and the tests on the IEEE 14-bus system and the Norwegian 99-bus system show that

— The algorithms are easy to implement.

— The methods given consistently correct answers.

— The factorization-based observability analysis performs better with Hachtel's method because of its numerical robustness.

— The factorization-based observability analysis in Hachtel's state estimator has a predictable CPU-time requirement, whereas the graph-theory-based observability analysis routine used in the comparison requires considerably more CPU-time for difficult cases where there are only a few line flow measurements.

## REFERENCES

ASCHMONEIT, F. C., PETERSON, N. M., and ADRIAN, E. C. (1977). State estimation with equality constraints, *Tenth PICA Conference Proceedings*, May 1977, Toronto, pp. 427–430.

BJORCK, A. (1967). Iterative refinement of linear least square solutions, *BIT*, 7, 257–278.

BROUSSOLLE, F. (1978). State estimation in power systems: detecting bad data through the sparse inverse matrix method, *IEEE Trans. Power App. and Syst.*, 97, 678–682.

CLEMENTS, K. A. and DAVIS, P. W. (1985). Multiple bad data detectability and identifiability: a geometric approach, *Proc. 1985 PICA Conference*, San Francisco, May 6–10, 1985, pp. 461–466.

CLEMENTS, K. A., KRUMPHOLZ, G. R., and DAVIS, P. W. (1983). Power system state estimation with measurement deficiency: an observability measurement placement algorithm, *IEEE Trans. Power App. and Syst.*, **102**, 2012–2020.

DUFF, I. S. and REID, J. K. (1976). A comparison of some methods for the solution of sparse overdetermined systems of linear equations, *J. Inst. Math. Appl.*, **17**, 267–280.

DUFF, I. S. and REID, J. K. (1983). Multifrontal solution of indefinite sparse symmetric linear systems, *ACM Trans. on Mathematical Software*, **9**, No. 3., 302–325.

DUFF, I. S., REID, J. K., MUNKSGAARD, N., and NIELSEN, H. B. (1979). Direct solution of sets of linear equations whose matrix is sparse, symmetric, and indefinite, *J. Inst. Maths. Applics*, **23**, 1979, 235–250.

GJELSVIK, A., AAM, S., and HOLTEN, L. (1985). Hachtel's augmented matrix method—A rapid method improving numerical stability in power system static state estimation, *IEEE Trans. Power App. and Syst.*, **104**, 2987–2993.

GOLUB, G. H. and VAN LOAN, C. F. (1983). *Matrix Computations* (Johns Hopkins University Press, Baltimore, Maryland.)

GU, J. W., CLEMENTS, K. A., KRUMPHOLZ, G. R., and DAVIS, P. W. (1983). The solution of ill-conditioned power System state estimation problems via the method of Peters and Wilkinson, *PICA Conference Proceedings*, pp. 239–246, 1983.

HANDSCHIN, E., SCHWEPPE, F. C., KOHLAS, J., and FIECHTER, A. (1975). Bad data analysis for power systems state estimation, *IEEE Trans. Power App. and Syst.*, **94**, 329–337.

KRUMPHOLZ, G. R., CLEMENTS, K. A., and DAVIS, P. W. (1980). Power system observability: a practical algorithm using network topology, *IEEE Trans. Power App. and Syst.*, **99**, 1534–1542.

LIU, W.-H. E., WU, F. F., HOLTEN, L., GJELSVIK, A., and AAM, S. (1987). Computational issues in the Hachtel's augmented matrix method for power system state estimation, presented at the 1987 Power System Computation Conference, Cascais, Portugal, 30 August–4 September 1987.

MONTICELLI, A. and GARCIA, A. (1983). Reliable bad data processing for real-time state estimation, *IEEE Trans. Power App. and Syst.*, **102**, 1126–1139.

MONTICELLI, A., MURARI, C. A. F., and WU, F. F. (1985). A Hybrid State Estimator: Solving Normal Equations by Orthogonal Transformations, *IEEE Trans. Power App. and Syst.*, **105**, 3460–3468.

MONTICELLI, A. and WU, F. F. (1985a). Network observability: theory, *IEEE Trans. Power App. and Syst.*, **104**, 1042–1048.

MONTICELLI, A. and WU, F. F. (1985b). Network observability: identification of observable islands and measurement placement, *IEEE Trans. Power App. and Syst.*, **104**, 1035–1041.

MONTICELLI, A. and WU, F. F. (1986). Observability analysis for orthogonal transformation based state estimation, *IEEE Trans. Power Syst.*, **1**, 201–208.

MONTICELLI, A., WU, F. F., and YEN, M. (1986). Multiple bad data identification for state estimation by combinatorial optimization, *IEEE Trans. Power Delivery*, **1**, 361–369.

QUINTANA, V. H., SIMOES-COSTA, A., and MANDEL, A. (1982). Power system observability using a direct graph-theoretic approach, *IEEE Trans. Power App. and Syst.*, **101**, 617–626.

SCHWEPPE, F. C. and HANDSCHIN, E. J. (1974). Static state estimation in electric power systems, *Proc. IEEE*, **62**, 972–983.

SIEGEL, I. H. (1965). Deferment of computation in the method of least squares, *Math. Comp.*, **19**, 329–331.

SIMOES-COSTA, A. and QUINTANA, V. H. (1981a). A robust numerical technique for power system state estimation, *IEEE Trans. Power App. and Syst.*, **100**, 691–698.

SIMOES-COSTA, A. and QUINTANA, V. H. (1981b). An orthogonal row processing algorithm for power system sequential state estimation, *IEEE Trans. Power App. and Syst.*, **100**, 3791–3800.

TINNEY, W. F., BRANDWAJN, V., and CHAN, S. M. (1985). Sparse vector methods, *IEEE Trans. Power App. and Syst.*, **104**, 295–301.

U.S. DEPARTMENT OF ENERGY, (1984). Contribution to power system state estimation and transient stability analysis, prepared by ESCA Corporation, DOE/ET/29362-1, February 1984.

VANCUTSEM, Th. (1985). Power system observability and related functions—derivation of

WANG, J. W. and QUINTANA, V. H. (1984). A decoupled orthogonal row processing algorithm for power state estimation, *IEEE Trans. Power App. and Syst.*, **103**, 2337–2344.

WANG, J. W. and QUINTANA, V. H. (1984). A decoupled orthogonal row processing algorithm for power state estimation, *IEEE Trans. Power App. and Syst.*, 2337–2344.

WU, F. F., LIU, W.-H. E., and LUN, S.-M. (1987). Observability analysis and bad data processing for state estimation with equality constraints, *Paper No. 87WM103-5* presented at IEEE PES Winter Meeting, New Orleans LA, Feb. 1987.

WU, F. F. and MONTICELLI, A. (1986). Recent Progress in Real-Time Network Security Analysis, *Proc. IFAC Symp. on Power Systems and Power Plant Control*, August 12–15, 1986, Beijing, China, pp. 10–16.

## Appendices

### A. Observability

We wish to show that there is a one-to-one correspondence between the rank deficiency in the measurement Jacobian matrix and the coefficient matrix in Hachtel's method.

*Theorem*    Assume that the rows of the $r \times n$ matrix $\mathbf{C}$ are linearly independent.

$$\text{rank}\begin{bmatrix}\mathbf{H}\\\mathbf{C}\end{bmatrix} = n - k \Leftrightarrow \text{rank}\begin{bmatrix}0&0&\mathbf{C}\\0&\mathbf{I}&\mathbf{H}\\\mathbf{C}^T&\mathbf{H}^T&0\end{bmatrix} = r + m + n - k$$

where $\mathbf{H}$ is an $m \times n$ matrix, $m \geqslant n$, and $0 \leqslant k \leqslant n$.

*Proof* ($\Rightarrow$)

Let us partition the matrix

$$\begin{bmatrix}\mathbf{H}\\\mathbf{C}\end{bmatrix} = \begin{bmatrix}\mathbf{H}_1&\mathbf{H}_2\\\mathbf{C}_1&\mathbf{C}_2\end{bmatrix} \tag{A1}$$

where

(1) the $(m + r) \times (n - k)$ matrix $\begin{bmatrix}\mathbf{H}_1\\\mathbf{C}_1\end{bmatrix}$ has full rank.

(2) the columns of $\begin{bmatrix}\mathbf{H}_2\\\mathbf{C}_2\end{bmatrix}$ are the linear combinations of the columns of $\begin{bmatrix}\mathbf{H}_1\\\mathbf{C}_1\end{bmatrix}$, i.e.,

$$\mathbf{H}_2 = \mathbf{H}_1\mathbf{P} \quad \mathbf{C}_2 = \mathbf{C}_1\mathbf{P} \tag{A3}$$

$$\mathbf{K} = \begin{bmatrix}0&0&\mathbf{C}_1&\mathbf{C}_2\\0&\mathbf{I}&\mathbf{H}_1&\mathbf{H}_2\\\mathbf{C}_1^T&\mathbf{H}_1^T&0&0\\\mathbf{C}_2^T&\mathbf{H}_2^T&0&0\end{bmatrix} = \begin{bmatrix}\mathbf{I}&0&0&0\\0&\mathbf{I}&0&0\\0&0&\mathbf{I}&0\\0&0&\mathbf{P}^T&\mathbf{I}\end{bmatrix}\begin{bmatrix}0&0&\mathbf{C}_1&0\\0&\mathbf{I}&\mathbf{H}_1&0\\\mathbf{C}_1^T&\mathbf{H}_1^T&0&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}\mathbf{I}&0&0&0\\0&\mathbf{I}&0&0\\0&0&\mathbf{I}&\mathbf{P}\\0&0&0&\mathbf{I}\end{bmatrix}$$

$$\tag{A4}$$

Hence rank $\mathbf{K} \leqslant r + m + n - k$. Now we claim

$$\begin{bmatrix} 0 & 0 & \mathbf{C_1} \\ 0 & \mathbf{I} & \mathbf{H_1} \\ \mathbf{C_1^T} & \mathbf{H_1^T} & 0 \end{bmatrix} \quad \text{is nonsingular, i.e.,}$$

$$\begin{bmatrix} 0 & 0 & \mathbf{C_1} \\ 0 & \mathbf{I} & \mathbf{H_1} \\ \mathbf{C_1^T} & \mathbf{H_1^T} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Rightarrow \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 0 \\ 0 \end{bmatrix} \tag{A5}$$

The left-hand side of (A5) is

$$\mathbf{C_1}\boldsymbol{\alpha} = 0 \tag{A6}$$

$$\boldsymbol{\beta} + \mathbf{H_1}\boldsymbol{\gamma} = 0 \tag{A7}$$

$$\mathbf{C_1^T}\boldsymbol{\alpha} + \mathbf{H_1^T}\boldsymbol{\beta} = 0 \tag{A8}$$

Substituting (A8) into $\boldsymbol{\alpha}^T$(A6) $+ \boldsymbol{\beta}^T$(A7), we have $\boldsymbol{\beta}^T\boldsymbol{\beta} = 0$, hence $\boldsymbol{\beta} = 0$. Substituting $\boldsymbol{\beta} = 0$ into (A7) and combining it with (A6), because the matrix $\begin{bmatrix} \mathbf{C_1} \\ \mathbf{H_1} \end{bmatrix}$ has full rank, we have $\boldsymbol{\gamma} = 0$. Substituting $\boldsymbol{\beta} = 0$ into (A8), together with the fact that rows of $\mathbf{C_1}$ are linearly independent, we have $\boldsymbol{\alpha} = 0$. Hence the claim is proved, and rank $\mathbf{K} = r + m + n - k$.

($\Leftarrow$)  By contradiction. Suppose rank $\begin{bmatrix} \mathbf{H} \\ \mathbf{C} \end{bmatrix} \neq n - k$, but it is equal to $n - k + p$, $p > 0$ or $p < 0$. This implies that rank $\mathbf{K} = m + n + r - k + p$, which is not equal to $m + n + r - k$, as assumed. Thus we reach a contradiction.

## B. *Covariance matrix of the residuals*

Let $\mathbf{x}_t$ be the true state vector, $\hat{\mathbf{x}}$ be the estimated state, and $\delta\mathbf{x} = \mathbf{x}_t - \hat{\mathbf{x}}$. The estimated state $\hat{\mathbf{x}}$ satisfies the optimality conditions of the problem (3),

$$\mathbf{H^T W r} - \mathbf{C^T}\boldsymbol{\lambda} = 0 \tag{B1}$$

$$\mathbf{c}(\hat{\mathbf{x}}) = 0 \tag{B2}$$

where

$$\mathbf{r} = \mathbf{z} - \mathbf{h}(\hat{\mathbf{x}}) \approx \mathbf{w} + \mathbf{H}\delta\mathbf{x}. \tag{B3}$$

Combining (B1), (B3), and the linearized equation of (B2), we have

$$\begin{bmatrix} 0 & 0 & \mathbf{C} \\ 0 & \alpha\mathbf{W^{-1}} & \mathbf{H} \\ \mathbf{C^T} & \mathbf{H^T} & 0 \end{bmatrix} \begin{bmatrix} -\alpha^{-1}\boldsymbol{\lambda} \\ \alpha^{-1}\mathbf{Wr} \\ -\delta\mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{w} \\ 0 \end{bmatrix} \tag{B4}$$

Let

$$\begin{bmatrix} 0 & 0 & \mathbf{C} \\ 0 & \alpha\mathbf{W^{-1}} & \mathbf{H} \\ \mathbf{C^T} & \mathbf{H^T} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A_1} & \mathbf{A_2^T} & \mathbf{A_3^T} \\ \mathbf{A_2} & \mathbf{A_4} & \mathbf{A_5^T} \\ \mathbf{A_3} & \mathbf{A_5} & \mathbf{A_6} \end{bmatrix} \tag{B5}$$

We have

$$\alpha^{-1}\mathbf{Wr} = \mathbf{A}_4 \mathbf{w} \tag{B6}$$

Hence

$$\mathrm{cov}\,(\mathbf{r}) = (\alpha\mathbf{W}^{-1}\mathbf{A}_4)\mathbf{W}^{-1}(\alpha\mathbf{W}^{-1}\mathbf{A}_4)^{\mathrm{T}}$$

$$= \alpha\mathbf{W}^{-1}\mathbf{A}_4(\alpha\mathbf{W}^{-1})\mathbf{A}_4^{\mathrm{T}}\mathbf{W}^{-1} \tag{B7}$$

From (B5), we have

$$\mathbf{CA}_5 = 0 \tag{B8}$$

$$(\alpha\mathbf{W}^{-1})\mathbf{A}_4 + \mathbf{HA}_5 = \mathbf{I} \tag{B9}$$

$$\mathbf{C}^{\mathrm{T}}\mathbf{A}_2^{\mathrm{T}} + \mathbf{H}^{\mathrm{T}}\mathbf{A}_4 = 0 \tag{B10}$$

Premultiply (B9) by $\mathbf{A}_4$, we have

$$\mathbf{A}_4(\alpha\mathbf{W}^{-1})\mathbf{A}_4 = \mathbf{A}_4 - \mathbf{A}_4\,\mathbf{HA}_5$$

$$= \mathbf{A}_4 + \mathbf{A}_2\,\mathbf{CA}_5$$

$$= \mathbf{A}_4 \tag{B11}$$

Hence

$$\mathrm{cov}\,(\mathbf{r}) = \alpha\mathbf{W}^{-1}\mathbf{A}_4\mathbf{W}^{-1} \tag{B12}$$

### C. Sparse inverse with $2 \times 2$ pivoting

Consider a symmetric sparse matrix $\mathbf{Y}$, $\mathbf{Y}$ is to be factorized by Gaussian elimination. If $\mathbf{Y}$ is not positive definite, $2 \times 2$ matrices may be used as pivots so that symmetry is still preserved and numerical stability is maintained (Duff *et al.* 1979; Duff and Reid 1983). This means that the triangular factorization of $\mathbf{Y}$ takes the form

$$\mathbf{Y} = \mathbf{U}^{\mathrm{T}}\mathbf{DU} \tag{C1}$$

where $\mathbf{D}$ is a symmetric block diagonal matrix with a mixture of blocks of size $1 \times 1$ or $2 \times 2$. U is upper triangular with 1's along the diagonal. Furthermore, if rows $i$ and $i + 1$ belong to a $2 \times 2$ pivot block, then $u_{i,\,i+1} = 0$. In this Appendix we shall consider the issues of fill-ins and the calculation of the sparse inverse when $2 \times 2$ pivoting is used.

Let us consider the symmetric matrix

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_3 \\ \mathbf{Y}_3^{\mathrm{T}} & \mathbf{Y}_2 \end{bmatrix} \tag{C2}$$

Suppose $\mathbf{Y}_1 = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$

and $y_{11}$ may be zero, but the $2 \times 2$ matrix $\mathbf{Y}_1$ is nonsingular. We want to use $\mathbf{Y}_1$ as the pivot. The triangular factorization (C1) in block form can be written as

$$\begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_3 \\ \mathbf{Y}_3^{\mathrm{T}} & \mathbf{Y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1^{\mathrm{T}} & 0 \\ \mathbf{U}_3^{\mathrm{T}} & \mathbf{U}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 & 0 \\ 0 & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_3 \\ 0 & \mathbf{U}_2 \end{bmatrix} \tag{C3}$$

The first step of triangular factorization may be expressed as

$$\begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_3 \\ \mathbf{Y}_3^T & \mathbf{Y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1^T & 0 \\ \mathbf{U}_3^T & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 & 0 \\ 0 & \mathbf{Y}_2^1 \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_3 \\ 0 & \mathbf{I} \end{bmatrix} \tag{C4}$$

where

$$\mathbf{U}_1 = \mathbf{I} \tag{C5}$$

$$\mathbf{U}_3 = \mathbf{Y}_1^{-1} \mathbf{Y}_3 \tag{C6}$$

$$\mathbf{D}_1 = \mathbf{Y}_1 \tag{C7}$$

$$\mathbf{Y}_2^1 = \mathbf{Y}_2 - \mathbf{Y}_3^T \mathbf{Y}_1^{-1} \mathbf{Y}_3 \tag{C8}$$

The matrix $\mathbf{Y}_2^1$ will be further factorized and eventually we have

$$\mathbf{Y}_2^1 = \mathbf{U}_2^T \mathbf{D}_2 \mathbf{U}_2 \tag{C9}$$

Thus we may view $\mathbf{Y}$ in (C2) as the submatrix after triangular factorization has performed up to the $i$-th row/column, i.e., $y_{11}$ is actually occupying the $i$-th row and the $i$-th column of the partially factorized matrix.

Consider the case $y_{ij} \neq 0$, and $y_{i+1,j} = 0$, the corresponding elements in the triangular factor, i.e. $u_{ij}$ and $u_{i+1,j}$, will all be non-zero because of the multiplication (see (C6)) by a generally full matrix $\mathbf{Y}_1^{-1}$. Therefore, $u_{i+1,j}$ is a fill-in. The fill-ins during this step of using a $2 \times 2$ pivot in the triangular factorization is illustrated in Fig. 5. It is thus clear that in $\mathbf{U}$, the $i$-th row and the $(i + 1)$-th row have the same sparsity structure.

$$\begin{array}{ccccc}
& i & i+1 & j & k \\
i & x & x & \cdots \; x & \otimes \\
i+1 & x & x & \otimes & x \\
& & & & \\
j & x & 0 & \cdots \; x & \otimes \\
k & 0 & x & \otimes & x
\end{array}$$

Figure 5. The fill-ins introduced during triangular factorization with $2 \times 2$ pivot. $x$ represents a nonzero and $\otimes$ represents a fill-in.

Let

$$\mathbf{Y}^{-1} = \mathbf{Z} \tag{C10}$$

or

$$\begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_3 \\ \mathbf{Y}_3^T & \mathbf{Y}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{Z}_3 \\ \mathbf{Z}_3^T & \mathbf{Z}_2 \end{bmatrix} \tag{C11}$$

It can easily be verified that

$$\mathbf{Z} = (\mathbf{I} - \mathbf{U})\mathbf{Z} + (\mathbf{U}^T \mathbf{D})^{-1} \tag{C12}$$

Using the block form of the matrices, we have

$$\begin{bmatrix} \mathbf{Z}_1 & \mathbf{Z}_3 \\ \mathbf{Z}_3^T & \mathbf{Z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} - \mathbf{U}_1 & -\mathbf{U}_3 \\ 0 & \mathbf{I} - \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1 & \mathbf{Z}_3 \\ \mathbf{Z}_3^T & \mathbf{Z}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{U}_1^T \mathbf{D}_1 & 0 \\ \mathbf{U}_3^T \mathbf{D}_1 & \mathbf{U}_2^T \mathbf{D}_2 \end{bmatrix}^{-1} \tag{C13}$$

Since $U_1 = I$, from (C13) we obtain

$$Z_3 = -U_3 Z_2 \tag{C14}$$

$$Z_1 = -U_3 Z_3^T + (D_1)^{-1} \tag{C15}$$

In term of its elements, (C14) and (C15) can be expressed as

$$Z_{rj} = -\sum_{k=i+2}^{n} u_{rk} Z_{kj} \quad r = i, i+1; \quad j \geqslant i+2 \tag{C16}$$

$$Z_{rr} = -\sum_{k=i+2}^{n} u_{rk} Z_{rk} + (D_1^{-1})_{rr}, \quad r = i, \quad i+1 \tag{C17}$$

$$Z_{i,i+1} = -\sum_{k=i+2}^{n} u_{ik} Z_{i+1,k} + (D_1^{-1})_{i,i+1} \tag{C18}$$

Suppose we want to find $Z_{ij}$ at a location $u_{ij} \neq 0$. From (C16) it is seen that we need $Z_{kj}$ as long as $u_{ik} \neq 0$, therefore $Z_{kj}$ also occupies a location for which the corresponding element in $U$ is nonzero. Similar arguments, together with the fact that row $i$ and row $j$ of $U$ have the same sparsity structure, lead us to the result that $Z_{ii}$, $Z_{i+1,i+1}$, $Z_{i,i+1}$ can all be computed with the elements in $Z$ that occupy nonzero elements in $U$. Thus we conclude that the sparse inverse can be similarly computed with $2 \times 2$ pivoting.