Multifrontal Solver for Online Power System Time-Domain Simulation

Siddhartha Kumar Khaitan, James D. McCalley, Fellow, IEEE, and Qiming Chen, Member, IEEE

Abstract—This paper proposes the application of unsymmetric multifrontal method to solve the differential algebraic equations (DAE) encountered in the power system dynamic simulation. The proposed method achieves great computational efficiency as compared to the conventional Gaussian elimination methods and other linear sparse solvers due to the inherent parallel hierarchy present in the multifrontal methods. Multifrontal methods transform or reorganize the task of factorizing a large sparse matrix into a sequence of partial factorization of smaller dense frontal matrices which utilize the efficient Basic Linear Algebra Subprograms 3 (BLAS 3) for dense matrix kernels. The proposed method is compared with the full Gaussian elimination methods and other direct sparse solvers on test systems and the results are reported.

Index Terms—Differential algebraic equations, dynamic simulation, frontal methods, linear solvers, multifrontal methods, UMFPACK.

I. INTRODUCTION

SIMULATION tools are an integral part in the design and operation of large interconnected power systems. The purpose of simulation is monitoring and tracking, and devising preventive or corrective action strategies for mitigating the frequency and impact of high consequence events. Dynamic simulation of power systems is important for secure power grid expansion, and it significantly impacts future design and operation of large interconnected power systems. The differential algebraic equations (DAEs) for the dynamic simulation are solved to get the transient response of the power system. The power system typically has thousands of components including generators and associated controls, loads, transformers, lines, and voltage control elements. Detailed modeling of the power system results in thousands of differential and algebraic equations forming the DAE.

There are two broad categories of numerical integration methods: explicit and implicit. Iterative methods like the Newton method are needed to solve the implicit nonlinear equations resulting from implicit numerical methods. The most attractive feature of implicit methods is that they allow very large time steps. Reference [1] reports the usage of 10-s time steps in EUROSTAG and [2] reports the usage of 20-s time steps in EXSTAB without losing numerical stability.

In the literature for the solution of dynamic algebraic equations, [3], there has been a significant effort to develop A-stable, accurate and fast numerical integration methods with variable time steps. Reference [1] developed a mixed Adams-BDF algorithm with variable step size and variable integration order to reliably discriminate between stable and unstable phenomena and efficiently tackle large stiff power system models. Reference [2] developed a variable time step implicit integration scheme based on a modified Trapezoidal method for extended term time-domain simulation of power systems. In reference [4], a new decoupled time-domain simulation algorithm is proposed that takes advantage of both explicit and implicit methods. It is based on decoupling the system into stiff and non-stiff parts via invariant subspace decomposition and using the implicit method for the stiff part and explicit method for the non-stiff part to gain computationally efficiency. All these approaches have focused on integration algorithm development to gain efficiency. Although the various numerical integration schemes differ in their convergence, order, stability, and other properties, they do not necessarily offer considerable improvement in computational gain. However, the core of the resulting nonlinear equations from any of the integration schemes is the solution of a sparse linear system, which is the most computationally intensive part of a DAE solver. This is exploited in the work described here, via implementation of the unsymmetric multifrontal algorithms for sparse linear systems, which, when combined with a robust integration scheme, achieve very fast time-domain simulation.

Section II provides background of the numerical methods used, namely, the integration scheme, the solution of nonlinear equations and the linear solvers. Section III introduces frontal methods, useful because they provide foundations on which multifrontal methods are based. Section IV describes fundamentals of multifrontal solvers and uses some simple examples to illustrate. Section V compares simulation results of a power system simulator deploying multifrontal solvers with the same simulator deploying other direct sparse solvers and the Gaussian elimination solver, using two test systems. Section VI presents a discussion and Section VII concludes.

II. NUMERICAL METHODS

Most current methods for performing power system dynamic simulation are developed for use on conventional sequential machines. This leads to the natural conclusion that there are two
viable options to reduce the wall-clock time to solve a computationally intensive problem like power system time-domain simulation. These are 1) advanced hardware technology in terms of speed, memory, I/O, and architecture, and 2) a more efficient algorithm. Although the emphasis is generally on the hardware, nevertheless efficient algorithms can offer great advantage in achieving the desired speed. There exists a symbiotic relationship between the two. In this paper, we focus on enhancing algorithm efficiency to achieve high computational gain.

We can safely divide our time-domain simulator software into three parts, namely 1) a user interface, 2) DAE solver kernel, and 3) the output assembler. Our studies indicate about 90%–95% of computation in time-domain simulation is spent on the DAE solver. Any DAE solver requires numerical analysis techniques which can be broadly classified into three categories namely 1) numerical integration, 2) solution of nonlinear equations, and 3) solution of linear equations. We review implementation of these numerical techniques in the following three subsections.

A. Integration Scheme

The power system is modeled, in summary, by a set of differential and algebraic equations (DAE). These have inherent nonlinearities in them, and the resulting DAE is highly stiff. Switching events, contingencies, and forced outages introduce significant discontinuities in the system variables. The numerical scheme must converge quickly, give desired accuracy, and should be reliable and stable. The implicit integration scheme referred to as the theta method [4] satisfies all the above requirements and is used in our simulator. It does not have the infamous hyper-stability problem [2], which means that an algorithm will falsely report stability when the physical system is actually unstable. This method is also known as the weighted method [4]. Consider

$$\frac{dx}{dt} = f(t, y).$$

(1)

The theta method can be expressed in the general form as

$$y_{n+1} = y_n + h\left[\theta f(t_{n+1}, y_{n+1}) + (1 - \theta) f(t_n, y_n)\right]$$

$$n = 0, 1, \ldots$$

(2)

where $h$ is the time step of integration at time $t_n$, $n = 0, 1, 2, \ldots$. The integration scheme is explicit (or forward) Euler when $\theta = 0$, implicit (or backwards) Euler when $\theta = 1$, and trapezoidal when $\theta = 1/2$. The $\theta = 1$ case is a simple yet robust method for solving stiff ODEs.

Both the Euler and Trapezoidal integration scheme fit the equation of the above form. The choice of the theta method with $\theta \neq \{0, 0.5, 1\}$ is preferred over the Trapezoidal rule as it avoids the numerical oscillations following the occurrence of switching events, where such oscillations can occur when using the Trapezoidal rule [5]. The DAE for power systems can be summarized as

$$\frac{dx}{dt} = f(t, y)$$

$$0 = g(x, y)$$

(3)

(4)

Discretizing (3) and (4) using the theta-method results in

$$[x_{n+1} - x_n - h\frac{dx}{dt}] - (1 - \theta) h f(x_{n+1}, y_{n+1}) = 0$$

$$g(x_{n+1}, y_{n+1}) = 0.$$ 

(5)

(6)

In (5) and (6), only $x_{n+1}$ and $y_{n+1}$ are unknown variables and the rest are all known. Choosing $\theta = 0.47$, as suggested in [2], (5) and (6) constitute a set of nonlinear algebraic equations of the form

$$F(x_{n+1}, y_{n+1}) = 0.$$ 

(7)

B. Nonlinear Equation Solution

The set of nonlinear algebraic equations in (7) are solved at each time step using the Newton–Raphson method where, at the $i$th iteration, the unknowns are updated as follows:

$$x_{n+1}^{(i)} = x_{n+1}^{(i-1)} - \gamma \Delta x$$

$$y_{n+1}^{(i)} = y_{n+1}^{(i-1)} - \gamma \Delta y$$

(8)

(9)

where $\Delta x$ and $\Delta y$ are obtained by solving

$$J \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} R_x \\ R_y \end{bmatrix}$$

(10)

which can be represented by the set of linear equations

$$Ax = B$$

(11)

where

$$A = J = \text{Jacobian matrix}$$

$$R_x = \text{differential equation residual [Eq. (5)]}$$

$$R_y = \text{algebraic equation residual [Eq. (6)]}$$

$$x_{n+1}^{(i)} = \text{state variables at the } i\text{th iteration}$$

$$y_{n+1}^{(i)} = \text{additional variables at the } i\text{th iteration}$$

$$\gamma = \text{deceleration factor}$$

$$B = \begin{bmatrix} R_x \\ R_y \end{bmatrix}.$$ 

The Newton iterations terminate when the residual vectors are “smaller” than pre-specified tolerances based on norm and rate of convergence. Computation of the correction vector $[\Delta x, \Delta Y]^T$ requires the solution of the set of linear equations given by (11) which is discussed in the next section.

C. Formulation of Dynamic Algebraic Equations

The set of differential equations is in (12)–(18) at the bottom of the next page.

In the above formulations, the equations are for all the generations including the exciter, governor and AGC models. The notation for exciter, governor, and AGC are the same as we showed in [6]. For generators, the meanings of the notation are the same as what is in [7]. The limiter for each variable is implemented as logic in program and is not shown in the above equations. The program checks all the variable limits for each integration step and corrects them if necessary. We use the two-axis model [7] for generator dynamics.
The set of algebraic equations are as follows, where \( i \) means the \( i \)th generator and \( j \) means the \( j \)th generator bus:

\[
\begin{pmatrix}
V^i_q \\
V^i_d
\end{pmatrix}
- \begin{pmatrix} E^i_q^t \\
E^i_d^t
\end{pmatrix}
+ \begin{pmatrix}
X^i_q & -X^i_d \\
X^i_d & -X^i_d
\end{pmatrix}
\begin{pmatrix}
I^i_q \\
I^i_d
\end{pmatrix}
= 0
\tag{19}
\]

for each generator

\[
\begin{pmatrix}
V^j_q \\
V^j_d
\end{pmatrix}
- \begin{pmatrix} \cos \delta^j & - \sin \delta^j \\
\sin \delta^j & \cos \delta^j
\end{pmatrix}
\begin{pmatrix}
V^j_d \\
V^j_d
\end{pmatrix}
= 0
\tag{20}
\]

for each generator bus

\[
\begin{pmatrix}
I^j_x \\
I^j_y
\end{pmatrix}
- \begin{pmatrix} \cos \delta^j & - \sin \delta^j \\
\sin \delta^j & \cos \delta^j
\end{pmatrix}
\begin{pmatrix}
I^j_x \\
I^j_y
\end{pmatrix}
= 0
\tag{21}
\]

for each generator bus

\[
Y_{bus} \times \begin{pmatrix}
V^1_q \\
V^1_d \\
\vdots \\
V^n_q \\
V^n_d
\end{pmatrix}
- \begin{pmatrix}
I^1_x \\
I^1_y \\
\vdots \\
I^n_x \\
I^n_y
\end{pmatrix}
= 0
\tag{22}
\]

for the whole linear impedance network with \( n \) voltage bus (\( Y_{bus} \) is the system admittance matrix)

\[
\begin{pmatrix}
P^i \\
Q^i
\end{pmatrix}
- \begin{pmatrix}
V^i_x \\
V^i_y
\end{pmatrix}
= 0
\tag{23}
\]

for each load bus with constant \( P \) and \( Q \). The loads in our test system are modeled as constant active and reactive power injection.

Equations (19)–(21) are for each individual generator. Equations (22) and (23) are for the whole network and each voltage bus respectively. The DAE developed in (12)–(23) are summarized as in (3) and (4).

\section*{D. Linear Solver}

As seen in the previous section the core of any iterative solver like Newton–Raphson is the solution of a system of equations represented by (11). For power systems, the Jacobian matrix \( A \) is highly sparse and the fill-in is very low. We use this fact to gain computational efficiency by employing a multifrontal based sparse linear solver.

In the solution of the DAE arising out of the dynamic modeling of the power system, the most computationally intensive steps are the Jacobian building and the solution of the sparse system of linear equations. However the purpose of the Jacobian is to provide adequate convergence and as long as it is achieved, one can minimize computation associated with Jacobian updating \([2]\). Since the terms of the Jacobian involve time step, there is a direct relation between the variation in time step, system condition, and frequency of Jacobian building. Effective strategy to rebuild the Jacobian has resulted in considerable time saving for each simulation. Thus the key computational step is the solution of the sparse linear system of equations.

In the solution of the linear equations, the Jacobian matrices do not have any of the desirable structural or numerical properties such as symmetry, positive definiteness, diagonal dominance, or bandedness, which are generally associated with sparse matrices, to exploit in developing of efficient algorithms for linear direct solvers. In general, the algorithms for sparse matrices are more complicated than for dense matrices. The complexity is mainly attributed to the need to efficiently handle fill-in in the factor matrices. A typical sparse solver consists of four distinct steps as opposed to two in the dense case.

1. The ordering step minimizes the fill-in and exploits special structures such as block triangular form.
2. An analysis step or symbolic factorization determines the nonzero structures of the factors and creates suitable data structures for the factors.
3. Numerical factorization computes the factor matrices.
4. The solve step performs forward and/or backward substitutions.

This paper describes multifrontal methods as a new approach for solving sparse linear systems arising in the power system dynamic simulation. Multifrontal methods are a generalization of
the frontal methods developed primarily for finite element problems [8] for symmetric positive definite systems which were later extended to unsymmetric systems [9]. These methods were then applied to a general class of problems in [10]. In the next two sections, we describe fundamentals of frontal and multifrontal methods.

III. FRONTAL METHODS

Frontal methods were originally developed for solving banded matrices from finite element problems [8]. The motivation was to limit computation on small matrices to solve problems on machines with small core memories. Presently frontal codes are widely used in finite element problems because very efficient dense matrix kernels, particularly Level 3 Basic Linear Algebra Subprograms (BLAS) [11], can be designed over a wide range of platforms. A frontal matrix is a small dense submatrix that holds one or more pivot rows and their corresponding pivot columns.

The frontal elimination scheme is summarized as follows.

1) Assemble a row into the frontal matrix.
2) Determine if any columns are fully summed in the frontal matrix. A column is fully summed if it has all of its nonzero elements in the frontal matrix.
3) If there are fully summed columns, then perform partial pivoting in those columns, eliminating the pivot rows and columns and doing an outer-product update on the remaining part of the frontal matrix.
4) Repeat until all the columns have been eliminated and matrix factorization is complete.

The basic idea in frontal methods is to restrict elimination operations to a frontal matrix, on which dense matrix operations are performed using Level 3 BLAS. In frontal scheme, the factorization proceeds as a sequence of partial factorization on frontal matrices, which can be represented as

\[
F = \begin{bmatrix}
F_{11} & F_{12} \\
F_{21} & F_{22}
\end{bmatrix}.
\]

(24)

Pivots can be chosen from the matrix \( F_{11} \) since there are no other entries in these rows and columns in the overall matrix. Subsequently, \( F_{11} \) is factorized, multipliers are stored over \( F_{12} \), and the Schur complement \( F_{22} - F_{21}^T F_{11}^{-1} F_{12} \) is formed, using full matrix kernels. At the next stage, further entries from the original matrix are assembled with this Schur complement to form another frontal matrix.

For example, consider the 6 by 6 matrix shown in Fig. 1(a). The non-zero entries in the matrix are represented by dots. The frontal method to factorize the matrix begins by assembling row 1 into an empty frontal matrix shown in Fig. 1(b). At this point, none of the variables are fully summed. Subsequently, we assemble row 2 to get the matrix in Fig. 1(c). Now variable 4 is fully summed, and hence column 4 can be eliminated. To eliminate a column, a pivot needs to be selected in that column. In this example, let the pivot be selected from row 2. Rearranging the matrix to bring the pivot element (2, 4) to the top left position, we obtain the matrix in Fig. 1(d). Here, \( u \) indicates an element of the upper triangular matrix, and \( f \) denotes an element of the lower triangular matrix. After elimination, the updated frontal matrix is as shown in Fig. 1(e).

IV. MULTIFRONTAL METHODS

Multifrontal methods have been reported in previous power system literature for solution of sparse linear systems arising in power flow studies. On serial platforms, the multifrontal methods were used for power flow in references [20] and [21]. Reference [20] implements an earlier version of multifrontal methods, and since then there has been a lot of research in the area of multifrontal methods. Currently, there is abundance of advanced algorithms for ordering schemes to reduce fill-in, for post-ordering and restructuring of elimination trees, assembly trees, and directed acyclic graphs (DAG). Efficient and optimized elimination trees, DAG, preordering strategies, reduced working storage and reduction in indirect memory access give higher speed and performance [22]–[29]. In [21],

with assembling rows. Again, when rows 3 and 4 are assembled, variable 1 is fully summed, and hence the column 1 can be eliminated. Choosing the pivot element to be (4, 1), the matrix with pivot element moved to the top left corner is shown in Fig. 1(f), and the updated frontal matrix after elimination is shown in Fig. 1(g). In this way, the frontal method continues until matrix factorization is complete.

Although frontal methods achieve large computational gain [9]–[15], there are many unnecessary operations on the frontal matrices which are often large and sparse, thus lowering overall performance. These deficiencies can be at least partially overcome through allowing the use of more than one front, resulting in the multifrontal method [16]–[19]. This permits pivot orderings that are better at preserving sparsity and also gives more possibility for exploitation of parallelism via simultaneous processing of different fronts.

Thus, in this paper we propose multifrontal methods as a viable solution methodology for large unsymmetric sparse matrices which are common in power system online dynamic simulation. The fundamentals of the multifrontal methods are discussed in the next section.

Fig. 1. Example for frontal method.
the main focus was to promote the FPGA technology for hardware implementation of sparse linear solver, as compared to the software solution for multifrontal solver UMFPACK [22]. However, all of these applications were for static power flow analysis, where matrices have symmetric zero pattern and nonzero diagonal elements. From the open literature, we have no indication that multifrontal methods have been applied for power system time-domain simulation, a particularly interesting application because the Jacobian is highly unsymmetric with unsymmetric zero pattern.

The multifrontal method, a generalization of the frontal method, was originally developed for symmetric systems [16]. Subsequently, an unsymmetric multifrontal algorithm UMFPACK [22] was developed for general sparse unsymmetric matrices. They make full use of the high performance computer architecture by invoking the level 3 Basic Linear Algebra Subprograms (BLAS) library. Thus memory requirement is heavily reduced, and computing speed is greatly enhanced.

In this section, we overview the multifrontal method for the solution systems characterized by large sparse matrices. Beginning with its development in 1983 by Duff and Reid [16], it has undergone many developments at different stages of its formulation, and different algorithms perform best for different classes of matrices. Broadly speaking, one can categorize them into six classes: 1) symmetric positive definite matrices [27]–[30]; 2) symmetric indefinite matrices [16], [31], [32]; 3) unsymmetric matrices with actual or implied symmetric nonzero pattern [18], [33]–[36]; 4) unsymmetric matrices where the unsymmetric nonzero pattern is partially preserved [37]; 5) unsymmetric matrices where the unsymmetric nonzero pattern is fully preserved [38]–[41]; and 6) QR factorization of rectangular matrices [42], [43]. There are significant differences among the various approaches. Here, we present fundamentals of multifrontal methods for symmetric positive definite linear systems because they are easier to understand, and they form the foundation for application to other classes of matrices. Reference [30] provides the basis for the concepts presented below on the theory of multifrontal methods.

Cholesky factorization of an \( n \) by \( n \) symmetric positive definite matrix is defined by \( A = LL^T \). Depending on the order in which the matrix entries are accessed and/or updated for factorization, the Cholesky factorization can be classified into row, column, or submatrix Cholesky schemes. Multifrontal methods perform Cholesky factorization by submatrices, where each factor column is formed and all of its updates to the submatrix remaining to be factored are computed. However, the novel feature of the multifrontal method is that the update contributions from a factor column to the remaining submatrix are computed, but not applied directly to the matrix entries. They are aggregated with contributions from other factor columns before updates are performed.

We explain the main concepts of the multifrontal method through an example. Consider a sparse symmetric positive definite \( n \) by \( n \) matrix \( A \) and its Cholesky factor as shown in Fig. 2. Each “\( * \)” represents an original nonzero in the matrix \( A \), and “\( o \)” represents a fill-in in the factor matrix \( L \). The elimination tree of the matrix \( A \) represented by \( T(A) \) is defined to be the structure with \( n \) nodes \( \{1, \ldots, n\} \) such that node \( p \) is the parent of \( j \) if and only if

\[
p = \min \{i > j | h_{ij} \neq 0\}.
\]

The elimination tree is a tree if \( A \) is irreducible, which we assume here. There are as many nodes in the tree as there are columns in the matrix (or the Cholesky factor). In this example, we have eight nodes.

From Fig. 2, we observe that for the fourth node \( d_4 \), the parent node is \( f_6 \), and for \( f_6 \) the parent node is \( h_8 \). Similarly for each node we can derive the parent node. Thus, traversing the path for all the nodes, we obtain the elimination tree shown in Fig. 3 for the example in Fig. 2.

Fundamental to understanding the multifrontal method are the concepts of descendents of a node in the elimination tree, subtree update matrix and frontal matrix, and update matrix, which we present in the following three subsections. A fourth subsection summarizes the method.

A. Descendents of A Node in the Elimination Tree

The descendents of the node \( j \) in the elimination tree \( T(A) \) contains \( j \) and the set of nodes in the subtree rooted at the node \( j \). The symbol \( T[j] \) is used to represent the set of descendents. For the above example the descendent of \( 1 \) is \( T[1] = \{1\} \), of \( 2 \) are \( T[2] = \{1, 2\} \), of \( 3 \) is \( T[3] = \{3\} \), of \( 4 \) are \( T[4] = \{1, 2, 4\} \) and of \( 6 \) are \( T[6] = \{1, 2, 3, 4, 5, 6\} \).
B. Subtree Update Matrix and Frontal Matrix

Let \((i_0, i_1, \ldots, i_r)\) be nonzero row subscripts in the \(j\)th column of the Cholesky factor, and \(i_0 = j\). Then, the subtree update matrix \(U_j\) at column \(j\) for the sparse matrix \(A\) is defined to be

\[
U_j = - \sum_{k \in \mathcal{I}[j] - \{j\}} \begin{pmatrix} l_{ij,k} \\ l_{i_1,j,k} \\ \vdots \\ l_{i_r,j,k} \end{pmatrix} \begin{pmatrix} l_{jk} & l_{i_1,j} & \cdots & l_{i_r,j} \end{pmatrix} + U_{j^*},
\]

and the frontal matrix \(F_j\) is defined to be

\[
F_j = \begin{pmatrix} a_{j,j} & a_{j,i_1} & \cdots & a_{j,i_r} \\ a_{i_1,j} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_r,j} & 0 & \cdots & 0 \end{pmatrix} + U_{j^*}.
\]

Both \(F_j\) and \(U_j\) have order \(r + 1\) which is equal to the number of nonzeros in the \(j\)th column of the Cholesky factor which includes the diagonal element. From the definition of \(F_j\) we see that when \(F_j\) is computed, the first row/column of \(F_j\) has been completely updated. Therefore, one step of elimination on \(F_j\) gives the nonzero entries of the factor column \(L_{n,j}\) of the Cholesky factor \(L\) of an \(n\) by \(n\) symmetric positive definite matrix is defined by \(A = LL^T\). So \(F_j\) can be factorized as

\[
F_j = \begin{pmatrix} l_{j,j} & 0 & \cdots & 0 \\ l_{i_1,j} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{i_r,j} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & U_j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix} + \begin{pmatrix} l_{i_1,j} & \cdots & l_{i_r,j} \end{pmatrix},
\]

(28)

This gives the update matrix \(U_j\) after one step elimination on \(F_j\). It is a full matrix and is derived from (16) to be

\[
U_j = - \sum_{k \in \mathcal{I}[j]} \begin{pmatrix} l_{i_1,j,k} \\ \vdots \\ l_{i_r,j,k} \end{pmatrix} \begin{pmatrix} l_{i_1,j} & \cdots & l_{i_r,j} \end{pmatrix}.
\]

(29)

The definition of \(U_j\) and \(U_j\) indicate that \(U_j\) is used to form the \(j\)th frontal matrix \(F_j\), whereas \(U_j\) is generated from \(F_j\) by an elimination step. \(U_j\) has one more row/column than \(U_j\).

From knowledge of the descendants given in Section IV-A, together with the definition of the subtree update matrix, we have \(U_1 = 0\), and so \(F_1\) is

\[
F_1 = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,7} & a_{1,8} \\ a_{2,1} & 0 & 0 & 0 \\ a_{7,1} & 0 & 0 & 0 \\ a_{8,1} & 0 & 0 & 0 \end{pmatrix} + U_{1^*}.
\]

(30)

\[
F_1 = \begin{pmatrix} a_{j,j} & a_{j,i_1} & \cdots & a_{j,i_r} \\ a_{i_1,j} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_r,j} & 0 & \cdots & 0 \end{pmatrix} + \begin{pmatrix} U_{c_1} \\ U_{c_2} \\ \vdots \\ U_{c_8} \end{pmatrix}.
\]

(35)

Therefore

\[
U_1 = - \begin{pmatrix} P_{2,1} & l_{2,1}l_{7,1} & l_{2,1}l_{8,1} \\ l_{7,1}l_{2,1} & P_{2,1} & l_{7,1}l_{8,1} \\ l_{8,1}l_{2,1} & l_{8,1}l_{7,1} & P_{8,1} \end{pmatrix}.
\]

(31)

Similarly, we can then calculate \(U_2\), and hence \(F_2\) and the update matrix. The process continues until the entire matrix is factorized. The advantage of the multifrontal method is that it is not sequential, rather it has a parallel hierarchy as can be observed from the elimination tree of the example. Formation of the subtree update matrix begins on multiple fronts, namely, \(U_{1^*}, U_{3^*}\), and \(U_{7^*}\). We can verify that for the above example the subtree update matrices \(U_1^*\) to \(U_4^*\) are as follows.

\[
U_1^* = 0
\]

\[
U_2^* = - \begin{pmatrix} (l_{2,1}) & 0 & 0 \end{pmatrix}
\]

\[
U_3^* = 0
\]

\[
U_4^* = - \begin{pmatrix} 0 & 0 & (l_{8,1}) \\ 0 & l_{8,1} & 0 \end{pmatrix}
\]

(32)

C. Matrix Extend Add Operator and Update Matrix

Let \(R\) be an \(r\) by \(r\) matrix with \(r \leq n\) and \(S\) be an \(s\) by \(s\) matrix with \(s \leq n\). Each row/column of \(R\) and \(S\) corresponds to a row/column of the given \(n\) by \(n\) matrix \(A\). Let \(i_1 \leq i_2 \leq \cdots \leq i_r\), be the subscripts of \(R\) in \(A\), and \(j_1 \leq j_2 \leq \cdots \leq j_s\) be those of \(S\). Let \(k_1, k_2, \ldots, k_l\) be the union of the two subscript sets. The matrix \(R\) can be extended to conform to the subscript set \(\{k_1, k_2, \ldots, k_l\}\), by introducing a number of zero rows and columns. In a similar way, the matrix \(S\) can be extended. Here, \(R \leftrightarrow S\) is defined to be the \(t\) by \(t\) matrix \(T\) formed by adding the two extended matrices of \(R\) and \(S\). The matrix operator \(\leftrightarrow\) is known as the matrix extend-add operator. For example, let

\[
R = \begin{pmatrix} p & q & 0 \\ u & v & 0 \end{pmatrix}, S = \begin{pmatrix} w & x \\ y & z \end{pmatrix}
\]

(33)

\[
R \leftrightarrow S = \begin{pmatrix} p & q & 0 \\ u & v & 0 \end{pmatrix} + \begin{pmatrix} w & 0 & 0 \\ 0 & 0 & y \end{pmatrix}
\]

(34)

In terms of the extend-add operator defined above, we see that the relationship between the frontal matrices \(\{F_j\}\) and the update matrices \(\{U_j\}\) is

\[
F_j = \begin{pmatrix} a_{j,j} & a_{j,i_1} & \cdots & a_{j,i_r} \\ a_{i_1,j} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_r,j} & 0 & \cdots & 0 \end{pmatrix} \leftrightarrow U_{c_1} \leftrightarrow \cdots \leftrightarrow U_{c_8}.
\]

(35)
where \( c_1, \ldots, c_s \) are the children of node \( j \) in the elimination tree. Thus \( \bar{U} = U_{c_1} + \cdots + U_{c_s} \), the aggregate of all outer-product updates from columns in \( T[j] \setminus \{j\} \). Since \( c_1, \ldots, c_s \) are the children of the node \( j \) in the elimination tree, \( T[j] \setminus \{j\} \) is the disjoint union of the nodes in the subtree \( T[C_1], \ldots, T[C_s] \), and all updates from columns in \( T[j] \setminus \{j\} \) are included in \( U_{c_1}, \ldots, U_{c_s} \). The process of forming \( J \)th frontal matrix \( F_j \) from \( A_{ij} \) and the update matrices of its tree children is the frontal matrix assembly operation, and the tree structure on which the assembly operations are based is called the assembly tree [14].

The operations described in the above three subsections, which are the essence of the multifrontal methods, are summarized in the form of an algorithm for multifrontal Cholesky factorization in Fig. 4.

D. Summary

The multifrontal method reorganizes the numerical computation, and the factorization is performed as a sequence of factorizations on multiple fronts. In practice, structural preprocessing [30] is done to reduce the working storage requirements by restructuring the tree and finding the optimal post-ordering of the tree. After the preprocessing, the computation of the Cholesky factor matrix by the multifrontal method is done as described in the above algorithm of Fig. 4.

E. Illustration

We illustrate the operations of the multifrontal method with the example of Fig. 5, using the same unsymmetrical matrix of Fig. 1 for which the operations of the frontal method were illustrated in Section III. The frontal matrices here are rectangular and not square.

Consider the unsymmetrical matrix shown in Fig. 5(a). An initial pivot element is chosen, say element (1, 1). The corresponding first frontal matrix with this pivot row and column and all contributions to them is shown in Fig. 5(b). Subsequently, a pivot operation is performed to eliminate variable 1, which gives the resultant frontal matrix with \( u \) (upper triangular matrix), \( l \) (lower triangular matrix), and the nonzero entries in the non-pivot rows and columns corresponding to the contribution block (represented by dots). Further, after the elimination of variable 1, another pivot is selected, say (3, 2). A new frontal matrix is then constructed with row 3 and column 2, with all contributions to them from both the original matrix and the contribution block of the previous frontal matrix. The resulting frontal matrix is shown in Fig. 5(c). After performing a pivot operation to eliminate variable 2, we get the matrix as shown in Fig. 5(d), but here another pivot operation on element (4, 3) can be performed to eliminate variable 3 as well, since all contributions to row 4 and column 3 can also be assembled into the same matrix. A pivot operation on element (4, 3) reduces the frontal matrix further as shown in Fig. 5(e). In this way, assembly of the frontal matrices continues, and pivot operations are performed on them until the matrix is completely factorized.

In the present study, UMFPACK 4.4 [26] is used as the engine for the solution of (5) and (6) by the multifrontal method. UMFPACK consists of a set of ANSI/ISO C routines for solving unsymmetric sparse linear systems using the unsymmetric multifrontal method. It requires the unsymmetric, sparse matrix to be input in a sparse triplet (compressed sparse column) format.

The solver has within itself different fill reducing ordering schemes built in and it selects the best ordering scheme for the problem at hand to reduce fill-in and make it more memory efficient. The default ordering scheme [24], [44], [45] is approximate minimum degree (AMD) with suitable pivoting search during numerical factorization. It finds both a row and column pivot ordering as the matrix is factorized. No preordering or partial preordering is used. At the start of the factorization, no frontal matrix exists. It begins a new frontal matrix with a global Markowitz-style pivot search. All pivots with zero Markowitz cost are eliminated first and placed in the \( LU \) factors. The analyze phase then automatically selects one of three ordering and pivoting strategies (unsymmetric, 2-by-2, and symmetric). For symmetric matrices with a zero-free diagonal, the symmetric strategy is used. This computes a column ordering using AMD. It combines a column ordering strategy with a right-looking unsymmetric-pattern multifrontal numerical factorization. No modification of the column ordering is made during the numerical factorization. For symmetric indefinite problems with zeros on the diagonal, 2-by-2 strategy is chosen. This looks for a row permutation that puts nonzero entries onto the diagonal. The symmetric strategy is then applied to the permuted matrix.
The multifrontal method offers a significant performance advantage over more conventional factorization schemes by permitting efficient utilization of parallelism and memory hierarchy. We provide evidence of this by comparing the performance of Newton’s method when coupled with the multifrontal solver and when coupled with other sparse solvers or the Gaussian elimination solver for linear equations.

F. Implementation

Regarding the implementation of the solver, the major issues that we worked on were:

1) The UMFPACK is a set of ANSI/ISO C routines. Since we have developed our simulator in Visual C++, all the C routines of the solver were converted to their VC++ counterparts.

2) The solver takes input in the sparse triplet format (compressed sparse format). So the input matrices from the simulator were converted to the sparse triplet format, and the interface wrapper was written for the solver to be called from the simulator.

3) The interface wrapper was written for the output of the solver into the simulator for further analysis.

V. CASE STUDIES

The proposed method is tested on two systems 1) a Test system with six generators, 21 buses, 21 lines, nine transformers, and three tie lines as shown in Fig. 6 and 2) the IEEE RTS-96 [46], [47] with 33 generators as shown in Fig. 7. Contingencies for both systems are simulated for 3600 s on a Pentium 4, 2.8-GHz, and 1 GB of RAM, with the contingency at \( t = 0 \) and 20% load ramping from 900 to 2700 s. Performance comparison is performed against Gaussian elimination methods and other direct sparse solvers which include CHOLMOD [49], a set of ANSI C routines for sparse Cholesky factorization and update/downdate, GPLU [24] QR factorization [50] and a sparse LU factorization routine which utilizes routines from LAPACK. These sparse solvers are also present in Matlab. Fig. 8 below shows for the six-generator test system the comparison of simulation plots by the Multifrontal method and the other sparse solvers for an initial contingency of a generator and two line trips at \( t = 0 \) second. As seen in the figure, both the methods provided the same solution. This was also found to be true for all the other solvers for both test systems.

Figs. 9 and 10 show the structure of the sparse Jacobian matrix before and after reordering to improve computational efficiency. Table I shows the performance comparison for the multifrontal method with the full Gaussian elimination algorithm and the sparse solvers on the six-generator test system for six different critical initiating contingencies on the system. For the sparse solvers only the best results of all the sparse solvers are reported. Column 1 shows the contingency number, columns 2, 3, and 4 show the simulation time in seconds with the multifrontal algorithm, Gaussian elimination, and the sparse solvers, respectively. Column 5 shows the speed-up using the multifrontal method compared to Gaussian, which varies between 3.75 to 7 times; Column 6 shows the speed-up with other sparse solvers, varying between 3 to 5.4 times. For this system there are a total of around 300 contingencies including all N-1 contingencies, contingencies due to breaker failure, due to protection...
failure to trip, and due to inadvertent tripping. The time savings for 300 contingencies is provided in columns 7 and 8. For this number of contingencies, the multifrontal method would save around 16 to 22 h computing time compared to Gaussian elimination and 11 to 17 h compared to other sparse solvers.

Table II similarly shows the performance comparison results for the 32-generator RTS system for three different critical initiating contingencies on the system. We observe that the speed-up against the Gaussian elimination varies between 4 to 6.4 times and against the other sparse solvers it varies between 3.2 to 4.2 times. Again for this system if we are to analyze 300 contingencies as in the previous case, we can see from columns 7 and 8 that using the multifrontal method saves around 67 to 100 h computing time compared to Gaussian elimination and between 49 to 67 h compared to other commonly available sparse solvers.

VI. DISCUSSION

From the numerical results in the last section we observe that the time saving is almost linearly related to the increase in the number of generators in the system and also the speed up increases with the size of the system. In the test system we had six generators and in the RTS system there were 32 generators.
There was almost five times increase in the number of generators (and thus the number of dynamical states). From Tables I and II, we can see that there was a corresponding increase in time savings and speed up. This is because the size of the matrix linearly scales with the number of generators (assuming tenth-order ODE per generator) and the algebraic equations corresponding to the network. The gain offered by the multifrontal methods becomes significantly more important as system size increases.

To get an idea about the size of the problem which the large control centers deal with in practice, we can consider one large US ISO, where the real-time model represents approximately 32,000 buses and 42,000 branches [51]. It is likely this model contains over 3000 generators. The magnitude of transient simulation is therefore in terms of tens of thousands of differential algebraic equations. Thus computational speed is a main concern, together with data measurement, network processing, and modeling.

Modern power system operators are supervising one of the most complex engineering systems in existence. Under normal conditions they are able to control the power system with sufficient automatic control support, and they have computational support to help them respond quickly and effectively to disturbance conditions for likely contingencies (typically N-1). Yet, there is very little decision support available to them for responding to severe disturbances and complex unfolding of the post disturbance phenomena, when catastrophic conditions can sometimes, perhaps even often, be avoided if the operator has a decision aid. Such a decision-aid necessarily needs to be highly efficient to address extended-term phenomena for large systems. Thus, a fast time-domain simulator, capable of performing extended-term (several hours) of simulation, is highly desirable. Our particular interest is for online deployment of long-term time-domain simulation, and in this context, multifrontal methods may be viewed as a fundamentally important enabling technology.

VII. CONCLUSION

We observe from the two examples presented that the amount of time saving and speed-up increases with the size of the system. For systems having size orders of magnitude larger, the efficiency gain may also be orders of magnitude larger. We project for a 500 generator system, computation time for simulating 300 contingencies over 3600 s would be an order of magnitude faster than Gaussian elimination methods or other non-multifrontal direct sparse solvers. The computational gain in general increases with the size of the matrix [22], [44] due to inherent parallel hierarchy. However the complexity of the algorithm can vary depending on the problem. Thus the multifrontal methods are highly appealing for enhancing computational efficiency of power system time-domain simulation. Decreasing computation time for any heavily used tool is advantageous as it increases the engineer’s ability to complete work more quickly and/or to analyze more scenarios. Our particular interest is for online deployment of long-term time-domain simulation, and in this context, multifrontal methods may be viewed as a fundamentally important enabling technology.

REFERENCES


time step, implicit integration for extended-term power system dy-

[3] K. Brenan, S. Campbell, and L. Petzold, Numerical Solution of Initial-


[17] I. Duff and J. Scott, “The use of multiple fronts in Gaussian elimi-


KHAITAN et al.: MULTIFRONTAL SOLVER FOR ONLINE POWER SYSTEM TIME-DOMAIN SIMULATION


Siddharta Kumar Khaian received the B.E. degree in electrical engineering from Birla Institute of technology, Mesra, India, in 2003 and the M.Tech. degree from the Indian Institute of Technology, Delhi, India, in 2005.

He is currently a Post Doctoral Research Associate at Iowa State University, Ames. His current research interests are power system dynamic simulation, cascading, numerical analysis, linear algebra, and parallel computing.

Mr. Khaian was the topser and Gold Medalist in his undergraduate studies.

James D. McCalley (F’04) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Georgia Tech, Atlanta, in 1982, 1986, and 1992, respectively.

He was employed with Pacific Gas and Electric Company, San Francisco, CA, from 1985–1990 as a Transmission Planning Engineer. He is now a Professor of electrical and computer engineering at Iowa State University, Ames, where he has been employed since 1992.

Dr. McCalley is a registered Professional Engineer in California.

Qiming Chen (S’00–M’03) received the B.S. and M.S. degrees in science from Huazhong University of Science and Technology, Wuhan, China, in 1995 and 1998, respectively, and the Ph.D. degree from Iowa State University, Ames, in 2004.

He was a Planning Engineer with PJM Interconnection, Philadelphia, PA, from 2003 to 2008. He is currently a Manager for power system modeling with Macquarie Cook Power, Inc., Houston, TX.