Dynamic Simulation of Large-scale Power Systems Using a Parallel Schur-complement-based Decomposition Method

Petros Aristidou, *Student Member, IEEE,* Davide Fabozzi, and Thierry Van Cutsem, *Fellow Member, IEEE*

Abstract—Power system dynamic simulations are crucial for the operation of electric power systems as they provide important information on the dynamic evolution of the system after an occurring disturbance. This paper proposes a robust, accurate and efficient parallel algorithm based on the Schur complement domain decomposition method. The algorithm provides numerical and computational acceleration of the procedure. Based on the shared-memory parallel programming model, a parallel implementation of the proposed algorithm is presented. The implementation is general, portable and scalable on inexpensive, shared-memory, multi-core machines. Two realistic test systems, a medium-scale and a large-scale, are used for performance evaluation of the proposed method.

Index Terms—domain decomposition methods, Schur complement, power system dynamic simulation, OpenMP, shared-memory

1 Introduction

Dynamic simulations are frequently used in industry and academia to check the response of electric power systems to disturbances. Over the last decades, they have become indispensable to anyone involved in the planning, design, operation and security of power systems. Power system operation companies depend on fast and accurate dynamic simulations to train operators, analyze large sets of scenarios, assess the dynamic security of the network in real-time or schedule the day ahead operation. On the other hand, people designing future power systems depend on dynamic simulations to evaluate the proposed changes, whether these involve adding new transmission lines, increasing renewable energy sources, implementing new control schemes or decommissioning old power plants [1].

Complex electric components (like generators, motors, loads, wind turbines, etc.) used in dynamic simulations are represented by systems of stiff, non-linear Differential and Algebraic Equations (DAEs) [2]. Conversely, the network connecting the equipment together is represented by a system of linear algebraic equations. These equations try to describe the physical dynamic characteristics, the interactions and control schemes of the system. A large interconnected system may involve hundreds of thousands of such equations whose dynamics span over very different time scales and undergoing many discrete transitions imposed by limiters, switching devices, etc. Consequently, dynamic simulations are challenging to perform and computationally intensive [1].

In applications concerning the security of the system, for example Dynamic Security Assessment (DSA), the speed of simulation is a critical factor. In the remaining applications the speed of simulation is not critical but desired as it increases productivity and minimizes costs. This is the main reason why power system operators often resort to faster, static simulations. However, operation of non-expandable grids closer to their stability limits and unplanned generation patterns stemming from renewable energy sources require dynamic studies. Furthermore, under the pressure of electricity markets and with the support of active demand response, it is likely that system security will be more and more guaranteed by emergency controls responding to the disturbance. In this context, checking the sequence of events that take place after the initiating disturbance is crucial; a task for which the static calculation of the operating point in a guessed final configuration is inappropriate [11].

During the last decade the increase of computing power has led to an acceleration of dynamic simulations. Unfortunately, at the same time, the size and complexity of simulations has also grown. The increasing demand for more detailed and complex models (power electronic interfaces, distributed energy sources, active distribution networks, etc.) can easily push any given computer to its limits [1]. However, the emergence of parallel computing architectures resulted in a great boost of the simulation performance. The most prominent parallel algorithms developed are inspired from the field of Domain Decomposition Methods (DDMs).

DDM refers to a collection of techniques which revolve around the principle of "divide-and-conquer". Such methods have been used extensively in problems deriving from physics and structural engineering [3].

Petros Aristidou *is with the Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium, e-mail: p.aristidou@ieee.org.*
Davide Fabozzi *is with the Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium.*
Thierry Van Cutsem *is with the Fund for Scientific Research (FNRS) at the Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium, e-mail: t.vancutsem@ulg.ac.be.*

They were primarily developed for solving large boundary value problems of Partial Differential Equations (PDEs) by splitting them into smaller problems over sub-domains and iterating between them to coordinate the solution [4]. Following, because of their high efficiency and performance, DDMs became popular in DAE problems, such as those appearing in power system simulations, VLSI simulations, multi-domain mechanical design, etc. [5], [6], [7], [8], [9]. In these problems, the solution domain refers to the states describing the physics and mechanisms of the underlying process and not to the space domain as usually in PDE problems. A more detailed overview of DDMs is presented in App. A.

In this paper we propose a parallel algorithm for dynamic simulation of large-scale electric power systems based on the Schur complement DDM. A non-overlapping, topological-based, decomposition scheme is applied on large electric power systems revealing a star-shaped sub-domain layout. This decomposition is reflected to a separation of the DAEs describing the system. Following, the non-linear DAE system describing each sub-domain is solved independently by discretizing and using a Newton method with infrequent matrix update and factorization. The interface variables shared between sub-domains are updated using a Schur-complement-based method [8], [10].

The proposed algorithm augments the performance of the simulation in two ways: first, the independent calculations of the sub-systems (such as formulation of DAE system, discretization, formulation of linear systems for Newton methods, solution of linear system, check of convergence, etc.) are parallelized, thus providing computational acceleration. Second, it exploits the locality of the decomposed sub-systems to avoid many unnecessary computations and provide numerical acceleration.

Additionally, we present the implementation of the algorithm based on the shared-memory parallel programming model targeting common, inexpensive multi-core machines. For this, modern Fortran and the OpenMP Application Programming Interface (API) are used. The implementation is general, with no hand-crafted optimizations particular to the computer system, Operating System (OS), simulated electric power network or disturbance. The results were acquired using four different multi-core computers and show significant acceleration.

The paper is organized as follows: in Section 2 we present the power system model formulation and some existing dynamic simulation algorithms. In Section 3, we explain the proposed algorithm and its application on dynamic simulations of electric power systems. In Section 4, some important parallel programming aspects and challenges of the implementation are discussed. Our simulation study is reported in Section 5, followed by a discussion in Section 6 and closing remarks in Section 7.

2 Power System Dynamic Simulation

Power system dynamic simulations fall in basically two categories: electromagnetic transients and quasi-sinusoidal approximation. In the former, fast electromagnetic transients are simulated; in steady state, voltages and currents evolve sinusoidally with time at a frequency close to the nominal value (50 or 60 Hz). The network itself is modeled through the differential equations relative to its inductances and capacitors. On the other hand, in the quasi-sinusoidal (or phasor) approximation, the network is represented through algebraic equations corresponding to sinusoidal regime. During transients, all phasors vary with time while in steady-state they take on constant values [2]. The dynamic model describes how the phasors evolve with time.

The dynamic simulations described in this paper fall in the second category. It is the one commonly used in global stability studies, where the simulated time interval can extend up to several minutes, if not more. Typical time step sizes range form a few milliseconds to seconds.

2.1 Model Overview

An electric power system, under the quasi-sinusoidal approximation, can be described in compact form by the following DAE Initial Value Problem (IVP):

$$\begin{aligned} \mathbf{0} &= \mathbf{\Psi}(\mathbf{x}, \mathbf{V}) \\ \mathbf{\Gamma}\dot{\mathbf{x}} &= \mathbf{\Phi}(\mathbf{x}, \mathbf{V}) \\ \mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{V}(t_0) &= \mathbf{V}_0 \end{aligned} \tag{1}$$

where $\mathbf{V}$ is the vector of voltages through the network, $\mathbf{x}$ is the expanded state vector containing the differential and algebraic variables (except the voltages) of the system and $\mathbf{\Gamma}$ is a diagonal matrix with

$$(\mathbf{\Gamma})_{\ell\ell} = \begin{cases} 0 & \text{if } \ell\text{-th equation is algebraic} \\ 1 & \text{if } \ell\text{-th equation is differential} \end{cases}$$

The first part of (1) corresponds to the purely algebraic network equations. The second part describes the remaining DAEs of the system. This system changes during the simulation due to the discrete dynamics of the system such as response to discrete controls or changes in continuous-time equations. For reasons of simplicity, the handling of these discrete events is not presented in this paper. More details concerning their handling can be found in [11] and its quoted references.

2.2 Review of Solution Algorithms

The existing sequential algorithms can be categorized as partitioned or integrated (simultaneous) [12]. Partitioned algorithms solve separately the differential equations of (1) for the state variables and the algebraic equations for the algebraic variables. Then, these solutions are alternated until convergence. On the other hand, integrated algorithms use an implicit integration method to convert (1) into a unique set of algebraic equations to be solved simultaneously. The most popular algorithm in this category [13] is the Very DisHonest Newton (VDHN) detailed in App. B.1.

As soon as parallel computers emerged, several algorithms trying to exploit the new computational potential

were proposed. These can be mainly classified into fine-grained [14], [15], [16], [17] and coarse-grained [6], [7], [18], [19], [20], [21], [22] parallel algorithms.

A more detailed overview of existing sequential and parallel algorithms in power system dynamic simulations can be found in App. B.

3 Proposed Algorithm

This section describes the proposed parallel Schur-complement-based algorithm.

3.1 Power System Decomposition

From topological perspective, power systems consist of electric components (e.g. synchronous machines, loads, motors, small distribution network equivalents, etc.) interfacing with each other through the network. For reasons of simplicity, all the aforementioned components connected to the network will be called *injectors* and refer to devices that either produce or consume power. The proposed algorithm employs a topological decomposition separating each injector to define one sub-domain and the remaining electric network as the central sub-domain of a star layout. The proposed decomposition is visualized in Fig. 1 with each shaded area representing one sub-domain.

This decomposition separates the problem (1) into several sub-problems, each defined over one sub-domain. Thus, the network sub-domain is described by the algebraic equations:

$$\begin{aligned} \mathbf{0} &= \mathbf{\Psi}(\mathbf{x}^{ext}, \mathbf{V}) \\ \mathbf{x}^{ext}(t_0) &= \mathbf{x}_0^{ext}, \mathbf{V}(t_0) = \mathbf{V}_0 \end{aligned} \tag{2}$$

while each injector sub-domain i is described by DAEs:

$$\begin{aligned} \mathbf{\Gamma}_i \dot{\mathbf{x}}_i &= \mathbf{\Phi}_i(\mathbf{x}_i, \mathbf{V}^{ext}) \\ \mathbf{x}_i(t_0) &= \mathbf{x}_{i0}, \mathbf{V}^{ext}(t_0) = \mathbf{V}_0^{ext} \end{aligned} \tag{3}$$

where the first $N-1$ sub-domains are injectors and N-th sub-domain the network. $\mathbf{x}_i$ and $\mathbf{\Gamma}_i$ are the projections of $\mathbf{x}$ and $\mathbf{\Gamma}$, defined in (1), on the i-th sub-domain. Furthermore, the variables of each injector $\mathbf{x}_i$ are separated into interior $\mathbf{x}_i^{int}$ and local interface $\mathbf{x}_i^{ext}$ variables and the network sub-domain variables $\mathbf{V}$ are separated into interior $\mathbf{V}^{int}$ and local interface $\mathbf{V}^{ext}$ variables.

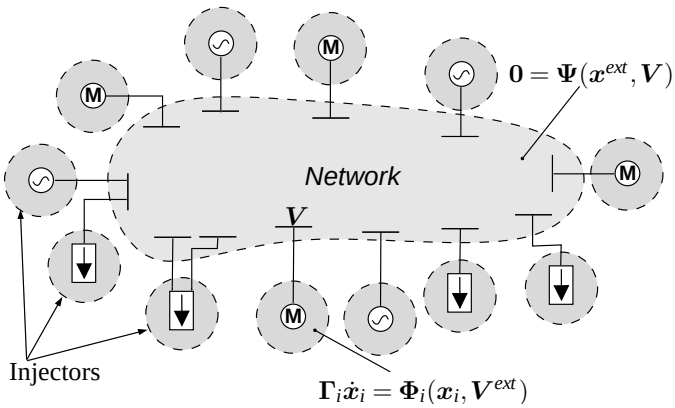


Figure 1. Proposed decomposition scheme

Although for decomposed DAE systems detecting the interface variables is a complicated task [8], for the proposed topological-based decomposition these are preselected based on the nature of the components. For the network sub-domain, the interface variables are the voltage variables of buses on which injectors are physically connected. For each injector sub-domain, the interface variable is the current injected in that bus.

The resulting star-shaped, non-overlapping, partition layout (see Fig. 1) is extremely useful when applying Schur-complement-based DDMs since it allows simplifying and further reducing the size of the global reduced system needed to be solved for the interface unknowns.

Moreover, the decomposition is based on topological inspection and doesn't demand complicated and computationally intensive partitioning algorithms. It increases modularity as the addition or removal of injectors does not affect the existing decomposition or other sub-domains besides the network. This feature can be exploited to numerically accelerate the solution.

3.2 Local System Formulation

The VDHN method is used to solve the algebraized injector DAE systems and the network equations. Thus, the resulting local linear systems take on the form of (4) for the injectors and (5) for the network.

$\mathbf{A}_{1i}$ (resp. $\mathbf{D}_1$) represents the coupling between interior variables. $\mathbf{A}_{4i}$ (resp. $\mathbf{D}_4$) represents the coupling between local interface variables. $\mathbf{A}_{2i}$ and $\mathbf{A}_{3i}$ (resp. $\mathbf{D}_2$ and $\mathbf{D}_3$) represent the coupling between the local interface and the interior variables. $\mathbf{B}_i$ (resp. $\mathbf{C}_j$) represent the coupling between the local interface variables and the external interface variables of the adjacent sub-domains.

3.3 Global Reduced System Formulation

Following, the interior variables of the injector sub-domains are eliminated (see App. A.2.2) which yields for the i-th injector:

$$\mathbf{S}_i \triangle \mathbf{x}_i^{ext} + \mathbf{B}_i \triangle \mathbf{V}^{ext} = \tilde{\mathbf{f}}_i \tag{6}$$

with $\mathbf{S}_i = \mathbf{A}_{4i} - \mathbf{A}_{3i}\mathbf{A}_{1i}^{-1}\mathbf{A}_{2i}$, the *local* Schur complement matrix and $\tilde{\mathbf{f}}_i = \mathbf{f}_i^{ext} - \mathbf{A}_{3i}\mathbf{A}_{1i}^{-1}\mathbf{f}_i^{int}$ the corresponding adjusted mismatch values.

The matrix $\mathbf{D}$ of the electric network is very sparse and structurally symmetric. Eliminating the interior variables of the network sub-domain requires building the local Schur complement matrix $\mathbf{S}_N$ which is in general not a sparse matrix. That matrix being large (since many buses have injectors connected to them) the computational efficiency would be impacted. On the other hand, not eliminating the interior variables of the network sub-domain increases the size of the reduced system but retains sparsity. The second option was chosen as it allows using fast sparse linear solvers.

$$\underbrace{\begin{pmatrix} \mathbf{A}_{1i} & \mathbf{A}_{2i} \\ \mathbf{A}_{3i} & \mathbf{A}_{4i} \end{pmatrix}}_{\mathbf{A}_i} \underbrace{\begin{pmatrix} \triangle \mathbf{x}_i^{int} \\ \triangle \mathbf{x}_i^{ext} \end{pmatrix}}_{\triangle \mathbf{x}_i} + \begin{pmatrix} 0 \\ \mathbf{B}_i \triangle \mathbf{V}^{ext} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{f}_i^{int}(\mathbf{x}_i^{int}, \mathbf{x}_i^{ext}) \\ \mathbf{f}_i^{ext}(\mathbf{x}_i^{int}, \mathbf{x}_i^{ext}, \mathbf{V}^{ext}) \end{pmatrix}}_{\mathbf{f}_i} \tag{4}$$

$$\underbrace{\begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \\ \mathbf{D}_3 & \mathbf{D}_4 \end{pmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} \triangle \mathbf{V}^{int} \\ \triangle \mathbf{V}^{ext} \end{pmatrix}}_{\triangle \mathbf{V}} + \begin{pmatrix} 0 \\ \sum_{j=1}^{N-1} \mathbf{C}_j \triangle \mathbf{x}_j^{ext} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{g}^{int}(\mathbf{V}^{int}, \mathbf{V}^{ext}) \\ \mathbf{g}^{ext}(\mathbf{V}^{int}, \mathbf{V}^{ext}, \mathbf{x}^{ext}) \end{pmatrix}}_{\mathbf{g}} \tag{5}$$

$$\underbrace{\begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \\ \mathbf{D}_3 & \mathbf{D}_4 - \sum_{i=1}^{N-1} \mathbf{C}_i \mathbf{S}_i^{-1} \mathbf{B}_i \end{pmatrix}}_{\widetilde{\mathbf{D}}} \underbrace{\begin{pmatrix} \Delta \mathbf{V}^{int} \\ \Delta \mathbf{V}^{ext} \end{pmatrix}}_{\Delta \mathbf{V}} = \underbrace{\begin{pmatrix} \mathbf{g}^{int} \\ \mathbf{g}^{ext} - \sum_{i=1}^{N-1} \mathbf{C}_i \mathbf{S}_i^{-1} \tilde{\mathbf{f}}_i \end{pmatrix}}_{\widetilde{\mathbf{g}}} \tag{8}$$

So, the global reduced system to be solved, with the previous considerations, takes on the form:

$$\begin{pmatrix} \mathbf{S}_1 & 0 & 0 & \cdots & 0 & \mathbf{B}_1 \\ 0 & \mathbf{S}_2 & 0 & \cdots & 0 & \mathbf{B}_2 \\ 0 & 0 & \mathbf{S}_3 & \cdots & 0 & \mathbf{B}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{D}_1 & \mathbf{D}_2 \\ \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 & \cdots & \mathbf{D}_3 & \mathbf{D}_4 \end{pmatrix} \begin{pmatrix} \triangle \mathbf{x}_1^{ext} \\ \triangle \mathbf{x}_2^{ext} \\ \triangle \mathbf{x}_3^{ext} \\ \vdots \\ \triangle \mathbf{V}^{int} \\ \triangle \mathbf{V}^{ext} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{f}}_1 \\ \tilde{\mathbf{f}}_2 \\ \tilde{\mathbf{f}}_3 \\ \vdots \\ \mathbf{g}^{int} \\ \mathbf{g}^{ext} \end{pmatrix} \tag{7}$$

Due to the star layout of the decomposed system, the resulting global Schur complement matrix is in the so called Block Bordered Diagonal Form (BBDF). Manipulating this structure, which is a unique characteristic of star-shaped layout decompositions, we can further eliminate from the global reduced system all the interface variables of the injector sub-domains and keep only the variables associated to the network sub-domain [23].

This results in the simplified, sparse, reduced system (8) where the elimination factors $\mathbf{C}_i \mathbf{S}_i^{-1} \mathbf{B}_i$ act only on the already non-zero block diagonal of sub-matrix $\mathbf{D}_4$ thus retaining the original sparsity pattern. System (8) is solved efficiently using a sparse linear solver to acquire $\mathbf{V}$ and the network sub-domain interface variables are backward substituted in (4), thus decoupling the solution of injector sub-domains. Then, the injector interior ($\mathbf{x}_i^{int}$) and local interface ($\mathbf{x}_i^{ext}$) variables are computed and the convergence of each sub-domain is checked using an infinite norm on its normalized state vector corrections. The solution stops when all sub-domains obey the convergence criteria, thus the non-linear systems (3) have been solved.

Other power system decomposition algorithms, like diakoptics [5], use a Schwartz-based approach (App. A) to update the interface variables. This is reflected into an equivalent system of (7) which, instead of BBDF, is in lower/upper triangular form. On the one hand, this form augments the algorithm's parallel performance by providing fewer synchronization points. On the other hand, it increases the number of iterations needed to converge as the influence of neighboring sub-domains is kept constant during a solution.

Although Schwartz-based algorithms are very useful when considering distributed memory architectures (where the cost of data exchange and synchronization is high), in modern shared-memory architectures the use of Schur-complement-based algorithms, as the one proposed, can compensate for the higher data exchange with less iterations because it retains the equivalent BBD structure of the final matrix.

3.4 Computational Acceleration

The proposed parallel algorithm -denoted *(P)*- is presented in Fig. 2. As described above, the VDHN method is used for the solution of the sub-domain systems, thus updating and factorizing the local matrices is done infrequently. If the decomposed algorithm does not converge after a given number of sub-domain VDHN iterations, then, all the local sub-domain matrices are updated and factorized, the local Schur complement and elimination factors are recomputed and the simplified reduced system (8) reformulated.

Thus, the main sources of parallelism of the proposed algorithm are the formulation and update of the sub-domain local systems, the elimination of the interior variables and calculation of local Schur complements, the calculation of the elimination factors, the formulation of the reduced system and the solution of the sub-domain systems. These can be seen in the shaded blocks in Fig. 2.

The proposed algorithm, as Schur-complement-based, suffers from the sequentiality introduced by the solution of the global reduced system (8). Because of the high sparsity, the linear nature of the network sub-domain and the infrequent system update, this bottleneck is bounded to $5-8\%$ of the overall computational cost (see App. E.2). Although this sequentiality can be tackled by the use of available parallel sparse linear solvers, the new synchronization points introduced would increase the overhead time and counteract any benefits arising from parallelization. Hence, a powerful but general solver [24] has been used in this work. However, in future development a customized advanced parallel solver (e.g. [25]) could be used to compute this remaining sequential portion in the most efficient manner.

3.5 Numerical Acceleration

The proposed algorithm can also provide numerical acceleration, if we exploit the locality of the sub-domains.

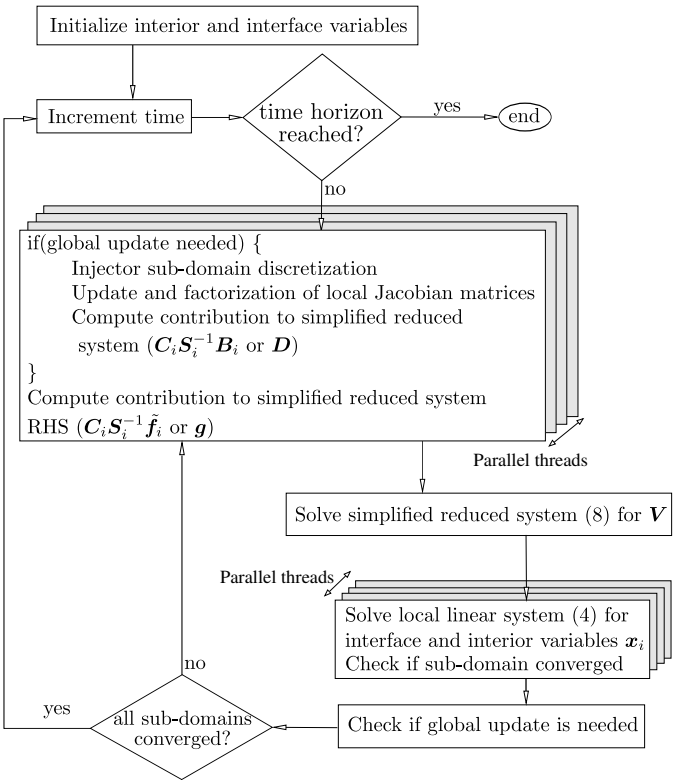


Figure 2. Proposed parallel algorithm (*P*)

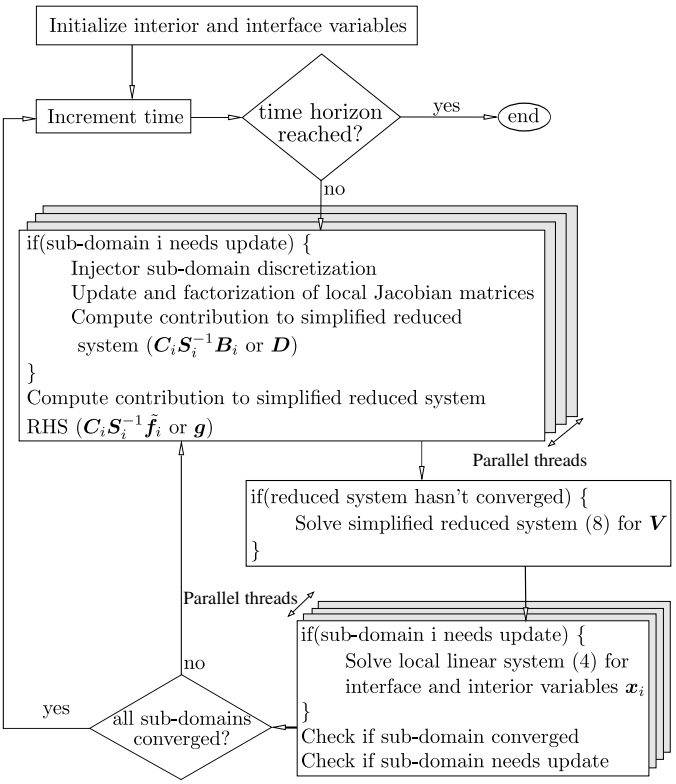


Figure 3. Proposed enhanced parallel algorithm (*EP*)

This is based on the observation that sub-domains described by strongly non-linear systems or with fast changing variables, converge slower and need more VDHN iterations, while sub-domains with "low dynamic activity" converge faster and need less VDHN iterations. We can exploit this in two ways.

First, in algorithm (*P*) the local matrices and the global Schur complement matrices are computed infrequently but in a synchronized way; when the criteria for global update is reached, all matrices are recomputed at the same time. Taking advantage of the fact that each sub-domain is solved by a separate VDHN method, we can decouple the matrix update criteria and allow local system matrices to be updated asynchronously. In this way, sub-domains which converge fast will keep the same local system matrices for many iterations and time-steps, while sub-domains which converge slower will update their local systems more frequently.

Second, in algorithm (*P*) all sub-domains need to be solved at every iteration up to global convergence. Conversely, if the convergence check is done locally on each separate sub-domain, then many sub-domain solutions can be saved. That is, at a given iteration not-converged sub-domains will continue being solved, while converged sub-domains will stop being solved.

Based on algorithm (*P*) and the two above enhancements, the enhanced parallel algorithm (*EP*) is presented in Fig. 3. Both algorithms have the same parallelization opportunities but the second algorithm (*EP*) offers some further, numerical acceleration [11].

4 Parallel Implementation

The main reason for applying parallelization techniques is to achieve higher application performance, or in our case, reduce simulation time. This performance depends on several factors, such as the percentage of parallel code, load balancing, communication overhead, etc. Other reasons, which are not considered in this paper, include the existence of proprietary data in some sub-domains or the usage of black-box models. In such cases, the method can work if the sub-domains are solved separately and only the interface values are disclosed.

Several parallel programming models were considered for the implementation of the algorithm (see App. C.1). The shared-memory model using OpenMP was selected as it is supported by most hardware and software vendors and it allows for portable, user-friendly programming. Shared-memory, multi-core computers are becoming more and more popular among low-end and high-end users due to their availability, variety and performance at low prices. OpenMP has the major advantage of being widely adopted on these platforms, thus allowing the execution of a parallel application, without any changes, on many different computers [26]. Moreover, OpenMP provides some easy to employ mechanisms for achieving good load balance among the working threads; these are detailed in App. C.2.

4.1 Parallel Work Scheduling

One of the most important tasks of parallel programming is to make sure that parallel threads receive equal

amounts of work [26]. Imbalanced load sharing among threads leads to delays, as some threads are still working while others have finished and remain idle.

In the proposed algorithm, the injector sub-domains exhibit high imbalance based on the dynamic response of each injector to the disturbance. In such situations, the dynamic strategy (App. C.2) is to be preferred for better load balancing. Spatial locality can also be addressed with dynamic scheduling, by defining a minimum number of successive sub-domains to be assigned to each thread (*chunk*). Temporal locality, on the other hand, cannot be easily addressed with this strategy because the sub-domains treated by each thread, and thus the data accessed, are decided at run-time and can change from one parallel segment of the code to the next. When executing on Uniform Memory Access (UMA) architecture computers, where access time to a memory location is independent of which processor makes the request, this strategy proved to be the best.

On the other hand, when executing on Non-Uniform Memory Access (NUMA) architecture computers (App C.3), a combination of static scheduling (App. C.2) with a proper choice of successive iterations to be assigned to each thread (*chunk*), proved to provide the best performance. A small chunk number means that the assignment of sub-domains to threads is better randomized, thus practically providing good load balancing. A higher chunk number, as discussed previously, allows to better address spatial locality. So, a compromise chunk size is chosen at runtime based on the number sub-domains and available threads. The same static scheduling and chunk number is used on all the parallel loops, which means that each thread handles the same sub-domains and accesses the same data, at each parallel code segment thus addressing temporal locality.

4.2 Ratio of Sub-domains to Threads

In several decomposition schemes, the number of sub-domains is chosen to be the same as the number of available computing threads and of equal size in order to utilize in full the parallelization potential. Further partitioning is avoided as increasing the number of sub-domains leads to an increased number of interfaces and consequently higher synchronization cost. On the contrary, the proposed decomposition suggests an extremely large number of injector sub-domains.

First, due to the star-shaped layout, the injector sub-domains of the proposed scheme have no interface or data exchange with each other except with the network. Thus, the high number of injector sub-domains does not affect the synchronization cost. Second, groups (chunks) of sub-domains are assigned at run-time to the parallel threads for computation, thus better load-balancing is achieved as they can be dynamically reassigned. Last, this decomposition permits the individual injector treatment to numerically accelerate the procedure as seen in Section 3.5.

These benefits, that is dynamic load balancing and numerical acceleration, cannot be exploited when fewer but bigger sub-domains, aggregating many injectors, are used. In dynamic simulations it is not known which injectors will display high dynamic activity, thus high computational cost, before applying a disturbance. So, to achieve good load balancing the splitting of injectors to sub-domains would need to be changed for each system and applied disturbance. Further, to apply the numerical acceleration presented in Section 3.5, all the injectors of the big sub-domain should be treated at the same time. Hence, the whole big sub-domain should be solved even if only one injector is strongly active.

5 Results

In this section we present the results of the Schur Complement-based algorithm implemented in the academic simulation software RAMSES[1], developed at the University of Liège. The software is written in standard Fortran language with the use of OpenMP directives for the parallelization.

Three algorithms were implemented for comparison:

(I) Integrated sequential algorithm VDHN (see App. B.1) applied on the original system (1). The Jacobian matrix is updated and factorized only if the system hasn't converged after three Newton iterations at any discrete time instant. This update strategy gives the best performance for the proposed test-cases.

(P) Parallel decomposed algorithm of Fig. 2. The global update and global convergence criteria are chosen to be the same as of (*I*).

(EP) Parallel decomposed algorithm of Fig. 3. The local update and local convergence criteria, applied to each sub-domain individually, are chosen to be the same as of (*I*).

For all three algorithms the same models, algebraization method (second-order backward differentiation formula - BDF2) and way of handling the discrete events were used. For the solution of the sparse linear systems, the sparse linear solver HSL MA41 [24] was used in all three algorithms. For the solution of the dense injector linear systems in algorithms (*P*) and (*EP*), Intel MKL LAPACK library was used.

Keeping the aforementioned parameters of the simulation constant for all algorithms permits the better evaluation of the proposed algorithms' performance.

5.1 Performance Indices

Many different indices exist for assessing the performance of a parallel algorithm. The two indices most commonly used by the power system community, as proposed in [13], are *speedup* and *scalability*.

1. Acronym for "Relaxable Accuracy Multithreaded Simulator of Electric power Systems".

First, the *speedup* is computed by:

$$Speedup(M) = \frac{Wall\ time\,(I)\ (1\,core)}{Wall\ time\,(P/EP)\ (M\,cores)} \quad (9)$$

and shows how faster is the parallel implementation compared to the fast, well-known and widely used sequential algorithm presented in App. B.1.

The second index is the *scalability* of the parallel algorithm, computed by:

$$Scalability(M) = \frac{Wall\ time\,(P/EP)\ (1\,core)}{Wall\ time\,(P/EP)\ (M\,cores)} \quad (10)$$

and shows how the parallel implementation scales when the number of processors available are increased.

5.2 Computer Platforms Used

The following computer platforms were used to acquire the simulation results:

1) Intel Core2 Duo CPU T9400 @ 2.53GHz, 32KB private L1 and 6144KB shared L2 cache, 3.9GB RAM, Microsoft Windows 7
2) Intel Core i7 CPU 2630QM @ 2.90GHz, 64KB private L1, 256KB private L2 and 6144KB shared L3 cache, 7.7GB RAM, Microsoft Windows 7
3) Intel Xeon CPU L5420 @ 2.50GHz, 64KB private L1 and 12288KB shared L2 cache, 16GB RAM, Scientific Linux 5
4) AMD Opteron Interlagos CPU 6238 @ 2.60GHz, 16KB private L1, 2048KB shared per two cores L2 and 6144KB shared per six cores L3 cache, 64GB RAM, Debian Linux 6 (App. C.3, Fig. 1)

Machines (1) and (2) are ordinary laptop computers with UMA architecture, thus dynamic load balancing was used. Machines (3) and (4) are scientific computing equipment with NUMA architecture, hence static load balancing and the considerations of App. C.3 were used.

5.3 Test-case A

The first test-case is based on a medium-size model of a real power system. This system includes 2204 buses and 2919 branches. 135 synchronous machines are represented in detail together with their excitation systems, voltage regulators, power system stabilizers, speed governors and turbines. 976 dynamically modeled loads and 22 Automatic Shunt Compensation Switching (ASCS) devices are also present. The resulting, undecomposed, DAE system has 11774 states. The models were developed in collaboration with the system operators.

The disturbance consists of a short circuit near a bus lasting seven cycles (116.7 ms at 60 Hz), that is cleared by opening two transmission lines (test-case A1). The same simulation is repeated with the difference that six of the ASCS devices located in the area of the disturbance are deactivated (test-case A2). Both test-cases are simulated over a period of 240 s with a time step of one cycle (16.667 ms) for the first 15 s after the disturbance (short-term) and 50 ms for the remaining time up to four minutes (long-term). The dynamic response of test-cases A1 and A2 as well as a comparison on the accuracy of the three algorithms can be found in App. D.1.

Tables 1 and 2 summarize the performance measured for test-cases A1 and A2, respectively. First, the average number of iterations needed for each discrete time step computation is presented. As expected, algorithms (*I*) and (*P*) need the same average number of iterations to compute the solution. Conversely, algorithm (*EP*) requires more iterations as converged sub-domains are not solved anymore and kept constant. However, each iteration of (*EP*) requires much smaller computational effort, as only non-converged sub-domains are solved, thus the overall procedure is accelerated.

Second, the performance indices achieved on each of the computational platforms. Both parallel algorithms (*P*) and (*EP*) provide significant acceleration when compared to the sequential algorithm (*I*). As expected, the best acceleration is achieved by algorithm (*EP*) which takes advantage of both the computational and numerical acceleration as described in Sections 3.4 and 3.5, respectively. It is noticeable that even on inexpensive portable computers (like UMA machines (1) and (2)) a significant acceleration is achieved with an up to 4.5 times speedup. On the bigger and more powerful computers used a speedup of up to 6.7 times is obtained.

Moreover, it is worth noticing that test-case A2, which is an unstable case (see App. D.1), requires on average more iterations per discrete time step computation. In fact, as the system collapses (App. D.1, Fig. 3) the injector sub-domains become more active, thus more iterations are needed for convergence. This also explains why test-case A2 shows slightly better scalability. That is, the additional number of sub-domain updates and solutions are computed in parallel, hence increasing the parallel percentage of the simulation (see App. E). At the same time, test-case A2 benefits less from the numerical speedup with algorithm (*EP*) as more active injector sub-domains delay to converge.

Table 1
Case A1: Performance summary

	average iterations per time step	platform	(1)	(2)	(3)	(4)
		cores	2	4	8	12
(*P*)	2.6	speedup	1.8	2.1	2.5	3.8
		scalability	1.1	1.5	1.7	2.8
(*EP*)	3.0	speedup	4.1	4.5	4.8	6.7
		scalability	1.1	1.4	1.4	2.3
(*I*)	2.6	-	-	-	-	-

Table 2
Case A2: Performance summary

	average iterations per time step	platform	(1)	(2)	(3)	(4)
		cores	2	4	8	12
(*P*)	3.5	speedup	1.9	2.3	2.6	3.9
		scalability	1.2	1.6	1.8	3.0
(*EP*)	4.2	speedup	3.8	4.2	4.5	6.2
		scalability	1.1	1.4	1.5	2.3
(*I*)	3.5	-	-	-	-	-

Table 3
Case B: Performance summary

	average iterations per time step	platform	(1)	(2)	(3)	(4)
		cores	2	4	8	24
(*P*)	2.2	speedup	1.4	2.0	2.7	3.9
		scalability	1.5	2.1	2.5	3.9
(*EP*)	2.6	speedup	2.9	3.3	4.1	4.7
		scalability	1.4	1.7	1.8	2.9
(*I*)	2.2	-	-	-	-	-

5.4 Test-case B

The second test-case is based on a large-size power system representative of the continental European main transmission grid. This system includes 15226 buses, 21765 branches and 3483 synchronous machines represented in detail together with their excitation systems, voltage regulators, power system stabilizers, speed governors and turbines. Additionally, 7211 user-defined models (equivalents of distribution systems, induction motors, impedance and dynamically modeled loads, etc.) and 2945 Load Tap Changing (LTC) devices are included. The resulting, undecomposed, DAE system has 146239 states. The models were developed in collaboration with industrial partners and operators of the power system.

The disturbance simulated on this system consists of a short circuit near a bus lasting five cycles (100 ms at 50 Hz), that is cleared by opening two important double-circuit transmission lines. The system is simulated over a period of 240 s with a time step of one cycle (20 ms) for the first 15 s after the disturbance (short-term) and 50 ms for the rest (long-term). The dynamic response of test-case B as well as a comparison on the accuracy of the three algorithms can be found in App. D.2.

Table 3 summarizes the performance measured for this test-case. As with test-cases A1 and A2, algorithms (*I*) and (*P*) need the same average number of iterations to compute the solution at each discrete time step, while (*EP*) slightly more. Furthermore, the performance indices show significant acceleration of the parallel algorithms compared to algorithm (*I*).

The scalability of test-case B is higher than of A1 and A2 as the significantly larger number of injectors

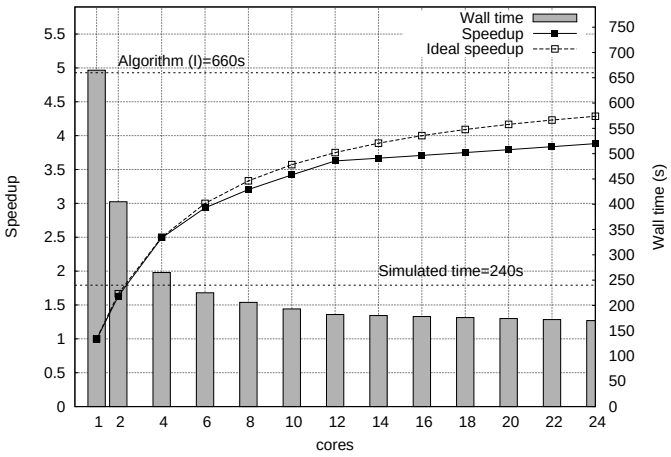


Figure 4. Case B: Speedup of algorithm (*P*)

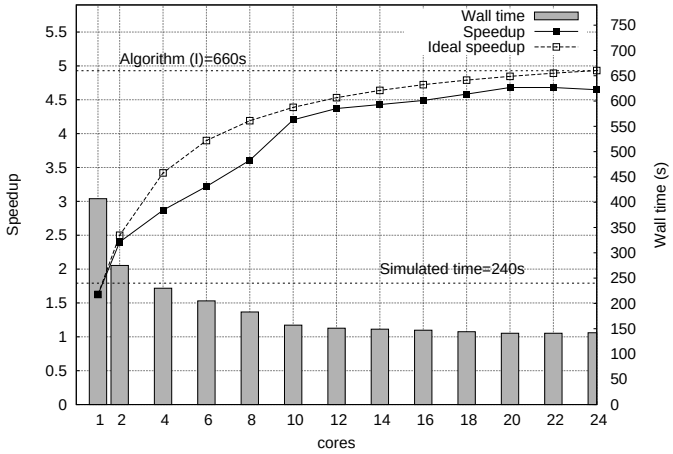


Figure 5. Case B: Speedup of algorithm (*EP*)

in the second system provides higher percentage of computational work in the parallelized portion. On the other hand, test-case B exhibits smaller overall speedup (4.7) compared to A1 and A2 (6.7 and 6.2 respectively). This can be attributed to the numerical acceleration of algorithm (*EP*), better benefiting test-cases A1 and A2. That is, many injectors in test-cases A1 and A2 converge early during the iterations and stop being solved (see Section 3.5), thus the amount of computational work inside each iteration is decreased and the overall simulation further accelerated.

Finally, Figs. 4 and 5 show the speedup and wall time of the algorithms (*P*) and (*EP*) when the number of threads on machine (4) are varied between one and 24. The ideal speedup is calculated using the profiling results of the test-case in sequential execution to acquire the parallel and sequential portion (App. E.2) and Amdahl's law (App. E.1).

6 Discussion

This section presents a discussion concerning the sequential, parallel and real-time performance of the proposed algorithms executed on NUMA machine (4).

6.1 Sequential Performance

In sequential execution, algorithms (*I*) and (*P*) perform almost the same (see Fig. 4, one core). On the other hand, due to the numerical accelerations presented in Section 3.5, algorithm (*EP*) is faster by a factor of $1.5-2.0$ (see Fig. 5, one core). Therefore, the proposed algorithm (*EP*) can provide significant acceleration even when executed on single-core machines. Moreover, it can be used on multi-core machines to simulate several disturbances concurrently using one core for each.

6.2 Parallel Performance

Equations 9 and 10 provide the actual speedup and scalability achieved by the parallel implementation. These indices take into consideration the OverHead Cost (OHC) needed for creating and managing the parallel

threads, as well as any other delays occurring in the implementation (e.g. due to load imbalances). To assess the performance of a parallel implementation Amdahl's law can be used [27] to provide its *theoretical scalability*, that is the theoretical maximum performance of the implementation if no OHC and perfect load balancing is considered. This theoretical maximum can be used to evaluate the quality of an implementation (see App. E).

Figure 6 shows how algorithms (*P*) and (*EP*) scale over a changing number of cores for test-case B. The theoretic scalability curves are plotted using the profiling results and Amdahl's law. The difference between the theoretic scalability and the measured scalability is due to costs of managing the threads and communication between them. This is better explained by the modified Amdahl's law (see App. E.1) which includes the OHC. The spread among theoretic and measured curves can be minimized if the costs of synchronization and communication, and thus OHC, are minimized.

To achieve this, further study on the algorithmic part (merging some parallel sections of the code to avoid synchronizations, parallelizing other time consuming tasks, etc.) and on the memory access patterns and management (avoiding false sharing, arranging better memory distribution, etc.) is needed. Especially for the latter, particularly interesting is the effect of LAPACK solvers used for the parallel solution of sub-domains. These popular solvers do not use optimized data structures to minimize wasting memory bandwidth or minimize conditionals to avoid pipeline flushing in multi-core execution. Techniques have been proposed for the development of customized solvers [25] to help alleviate this issue and need to be considered in future development.

Finally, a slight decrease of the scalability of (*EP*) is observed when the number of threads exceeds 22 (see Fig. 6). That is, when using more than 22 threads, the scalability degrades compared to the maximum achieved. The degradation occurs as the execution time gained $(\frac{P}{22} - \frac{P}{24})$, when increasing from 22 to 24 threads, is smaller than the overhead cost $OHC(24) - OHC(22)$ of creating and managing the two extra threads (see App. E.1).

6.3 Real-time Performance

Fast dynamic simulations of large-scale systems can be used for operator training and testing global control schemes implemented in Supervisory Control and Data Acquisition (SCADA) systems. In brief, measurements (such as the open/closed status from a switch or a valve, pressure, flow, voltage, current, etc.) are transfered from Remote Terminal Units (RTUs) to the SCADA center through a communication system. These information are then visualized to the operators or passed to a control software which take decisions for corrective actions to be communicated back to the RTUs. In modern SCADA systems the refresh rate (ΔT) of these measurements is $2 - 5$ seconds [28].

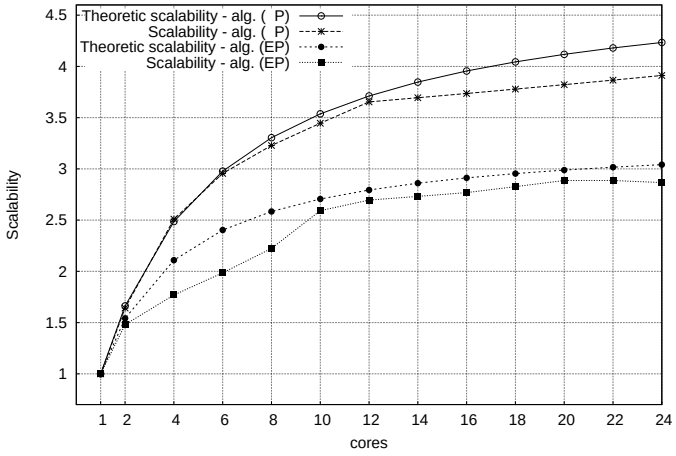


Figure 6. Case B: Scalability of algorithms (*P*) and (*EP*)

The simulator in these situations takes on the role of the real power system along with the RTU measurements and the communication system. It needs to provide the simulated "measurements" to the SCADA system with the same refresh rate as the real system. Thus, the concept of "real-time" for these applications translates to time deadlines.

In Fig. 7 we observe the performance of the algorithm for test-case A1. The straight diagonal line defines the limit of faster than real-time performance. When using six or more cores for the simulation, algorithm (*EP*) is able to simulate any of the tested disturbances on this system faster than real-time throughout the time horizon. That is, all simulation timings lay below the real-time line and all possible time deadlines can be met.

On the other hand, in Fig. 8 we observe that test-case B is faster than real-time after the initial 13 s. However, time deadlines with $\Delta T \geqslant 4$ s can still be met as the maximum lag between simulation and wall time is 4 s.

Overall, a wide range of available test-systems was evaluated in the same way for their real-time performance. A model of 8053 buses and 11292 injectors (totaling 74930 states) was found to be the limit of the proposed implementation, on platform (4), with faster than real-time execution.

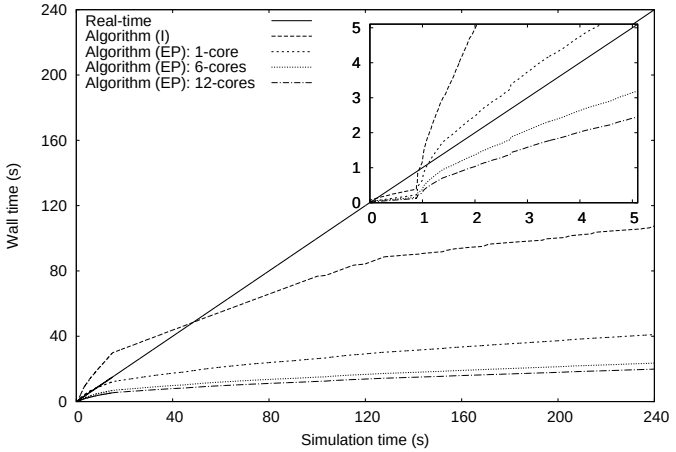


Figure 7. Case A1: Real-time performance (*EP*)

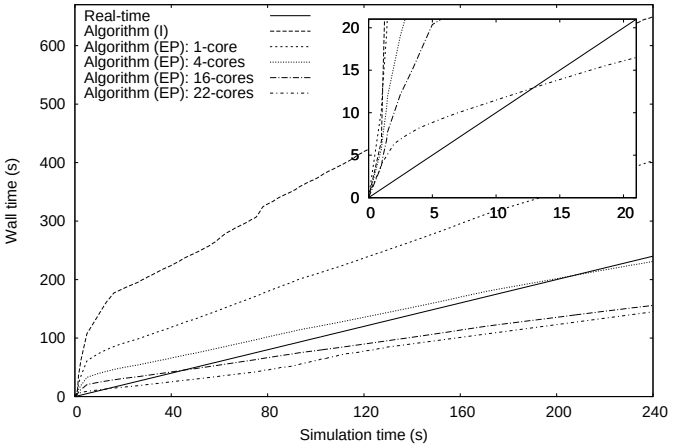


Figure 8. Case B: Real-time performance (*EP*)

7 Conclusions

In this paper a parallel Schur-complement-based algorithm for dynamic simulation of electric power systems is presented. It yields acceleration of the simulation procedure in two ways. On the one hand, the procedure is accelerated numerically, by exploiting the locality of the sub-domain systems and avoiding many unnecessary computations (factorizations, evaluations, solutions). On the other hand, the procedure is accelerated computationally, by exploiting the parallelization opportunities inherent to DDMs.

The proposed algorithm is accurate, as the original system of equations is solved exactly until global convergence. It is robust, as it can be applied on general power systems and has the ability to simulate a great variety of disturbances without dependency on the exact partitioning. It exhibits high numerical convergence rate, as each sub-problem is solved using a VDHN method with updated and accurate interface values during the whole procedure.

Along with the proposed algorithm, an implementation based on the shared-memory parallel programming model is presented. The implementation is portable, as it can be executed on any platforms supporting the OpenMP API. It can handle general power systems, as no hand-crafted, system specific, optimizations were applied and it exhibits good sequential and parallel performance on a wide range of inexpensive, shared-memory, multi-core computers.

References

[1] D. Koester, S. Ranka, and G. Fox, "Power systems transient stability-a grand computing challenge," *Northeast Parallel Architectures Center, Syracuse, NY, Tech. Rep. SCCS*, vol. 549, 1992.

[2] P. Kundur, *Power system stability and control*. McGraw-hill New York, 1994.

[3] B. Wohlmuth, *Discretization methods and iterative solvers based on domain decomposition*. Springer Verlag, 2001.

[4] A. Toselli and O. Widlund, *Domain decomposition methods–algorithms and theory*. Springer Verlag, 2005.

[5] G. Kron, *Diakoptics: the piecewise solution of large-scale systems*. MacDonald, 1963.

[6] M. Ilic'-Spong, M. L. Crow, and M. A. Pai, "Transient Stability Simulation by Waveform Relaxation Methods," *Power Systems, IEEE Transactions on*, vol. 2, no. 4, pp. 943 –949, nov. 1987.

[7] M. La Scala, A. Bose, D. Tylavsky, and J. Chai, "A highly parallel method for transient stability analysis," *Power Systems, IEEE Transactions on*, vol. 5, no. 4, pp. 1439 –1446, nov 1990.

[8] D. Guibert and D. Tromeur-Dervout, "A Schur Complement Method for DAE/ODE Systems in Multi-Domain Mechanical Design," *Domain Decomposition Methods in Science and Engineering XVII*, pp. 535–541, 2008.

[9] E. Lelarasmee, A. Ruehli, and A. Sangiovanni-Vincentelli, "The waveform relaxation method for time-domain analysis of large scale integrated circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 1, no. 3, pp. 131–145, 1982.

[10] Y. Saad, *Iterative methods for sparse linear systems*, 2nd ed. Society for Industrial and Applied Mathematics, 2003.

[11] D. Fabozzi, A. Chieh, B. Haut, and T. Van Cutsem, "Accelerated and localized newton schemes for faster dynamic simulation of large power systems," *Power Systems, IEEE Transactions on*, 2013.

[12] B. Stott, "Power system dynamic response calculations," *Proceedings of the IEEE*, vol. 67, no. 2, pp. 219–241, 1979.

[13] D. Tylavsky, A. Bose, F. Alvarado, R. Betancourt, K. Clements, G. Heydt, G. Huang, M. Ilic, M. La Scala, and M. Pai, "Parallel processing in power systems computation," *Power Systems, IEEE Transactions on*, vol. 7, no. 2, pp. 629 –638, may 1992.

[14] J. Chai and A. Bose, "Bottlenecks in parallel algorithms for power system stability analysis," *Power Systems, IEEE Transactions on*, vol. 8, no. 1, pp. 9–15, 1993.

[15] L. Yalou, Z. Xiaoxin, W. Zhongxi, and G. Jian, "Parallel algorithms for transient stability simulation on PC cluster," in *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, vol. 3, 2002, pp. 1592 – 1596 vol.3.

[16] K. Chan, R. C. Dai, and C. H. Cheung, "A coarse grain parallel solution method for solving large set of power systems network equations," in *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, vol. 4, 2002, pp. 2640–2644.

[17] J. S. Chai, N. Zhu, A. Bose, and D. Tylavsky, "Parallel newton type methods for power system stability analysis using local and shared memory multiprocessors," *Power Systems, IEEE Transactions on*, vol. 6, no. 4, pp. 1539–1545, 1991.

[18] V. Jalili-Marandi, Z. Zhou, and V. Dinavahi, "Large-Scale Transient Stability Simulation of Electrical Power Systems on Parallel GPUs," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, no. 7, pp. 1255 –1266, july 2012.

[19] M. Ten Bruggencate and S. Chalasani, "Parallel Implementations of the Power System Transient Stability Problem on Clusters of Workstations," in *Supercomputing, 1995. Proceedings of the IEEE/ACM SC95 Conference*, 1995, p. 34.

[20] D. Fang and Y. Xiaodong, "A new method for fast dynamic simulation of power systems," *Power Systems, IEEE Transactions on*, vol. 21, no. 2, pp. 619–628, 2006.

[21] J. Shu, W. Xue, and W. Zheng, "A parallel transient stability simulation for power systems," *Power Systems, IEEE Transactions on*, vol. 20, no. 4, pp. 1709 – 1717, nov. 2005.

[22] V. Jalili-Marandi and V. Dinavahi, "SIMD-Based Large-Scale Transient Stability Simulation on the Graphics Processing Unit," *Power Systems, IEEE Transactions on*, vol. 25, no. 3, pp. 1589 –1599, aug. 2010.

[23] X. Zhang, R. H. Byrd, and R. B. Schnabel, "Parallel Methods for Solving Nonlinear Block Bordered Systems of Equations," *SIAM Journal on Scientific and Statistical Computing*, vol. 13, no. 4, p. 841, 1992.

[24] HSL(2011). A collection of Fortran codes for large scale scientific computation. [Online]. Available: http://www.hsl.rl.ac.uk

[25] D. P. Koester, S. Ranka, and G. Fox, "A parallel Gauss-Seidel algorithm for sparse power system matrices," in *Supercomputing '94., Proceedings*, Nov, pp. 184–193.

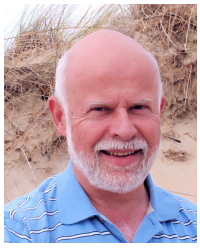
[26] B. Chapman, G. Jost, and R. Van Der Pas, *Using OpenMP: Portable Shared Memory Parallel Programming*. MIT Press, 2007.

[27] D. Gove, *Multicore Application Programming: For Windows, Linux, and Oracle Solaris*. Addison-Wesley Professional, 2010.

[28] J. Giri, D. Sun, and R. Avila-Rosales, "Wanted: A more intelligent grid," *Power and Energy Magazine, IEEE*, vol. 7, no. 2, pp. 34–40, March-April.

Petros Aristidou (S'10) received his Diploma in 2010 from the Dept. of Electrical and Computer Engineering of the National Technical University of Athens, Greece. He is currently pursuing his Ph.D. in Analysis and Simulation of Power System Dynamics at the Dept. of Electrical and Computer Engineering of the Univ. of Liège, Belgium. His research interests are in power system dynamics, control and simulation. In particular investigating the use of domain decomposition algorithms and parallel computing techniques to provide fast and accurate time-domain simulations.

Davide Fabozzi (S'09) received the B.Eng. and M.Eng. degrees in Electrical Engineering from the Univ. of Pavia, Italy, in 2005 and 2007, respectively. In 2007 he joined the Univ. of Liège, where he received the Ph.D. degree in 2012. Presently, he is a Marie Curie Experienced Researcher Fellow at Imperial College London, UK. His research interest include simulation of differential-algebraic and hybrid systems, power system dynamics and frequency control.



Thierry Van Cutsem (F'05) graduated in Electrical-Mechanical Engineering from the Univ. of Liège, Belgium, where he obtained the Ph.D. degree and he is now adjunct professor. Since 1980, he has been with the Fund for Scientific Research (FNRS), of which he is now a Research Director. His research interests are in power system dynamics, security, monitoring, control and simulation, in particular voltage stability and security. He is currently Chair of the IEEE PES Power System Dynamic Performance Committee.

Supplemental Material for "Dynamic Simulation of Large-scale Power Systems Using a Parallel Schur-complement-based Decomposition Method"

Petros Aristidou, *Student Member, IEEE,* Davide Fabozzi, and Thierry Van Cutsem, *Fellow Member, IEEE*

APPENDIX A
DOMAIN DECOMPOSITION METHODS

DDMs were originally used due to the lack of memory in computing systems: data needed for smaller portions of a problem could fit entirely to the memory while for the whole problem they could not. They lost their appeal as larger and cheaper memory became available, only to resurface in the era of parallel computing. These methods are inherently suited for execution on parallel architectures and many parallel implementations have been presented on multi-core computers, clusters and lately Graphics Processing Units (GPUs) [1], [2].

They are mainly distinguished by three features: sub-domain partitioning, problem solution over sub-domains and sub-domain interface variable processing [3].

A.1 Sub-domain Partitioning

Sub-domain partitioning has to be chosen based on the desired sub-domain characteristics for the given problem. This includes choosing the number of sub-domains, the type of partitioning, and the level of overlap between the sub-domains. Each of these choices depend on a variety of factors such as size, type, and geometry of the problem domain, the number of parallel processors, communication cost, and the actual system being solved.

When considering spatial domain problems, such as PDEs, the decomposition is usually given by the geometrical data and the order of the discretization scheme used. Conversely, in state domain problems, such as DAE, no a priori knowledge of the coupled variables is available since there are no regular data dependencies. Furthermore, each system model can be composed by several sub-models which are sometimes hidden, too complex, or used as black boxes. Hence, an automatic decomposition of the system is not trivial [4]. In fact, they usually have to rely on problem specific techniques which require good knowledge of the underlying system, the models composing it and the interaction between them.

Petros Aristidou *is with the Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium, e-mail: p.aristidou@ieee.org.*
Davide Fabozzi *is with the Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium.*
Thierry Van Cutsem *is with the Fund for Scientific Research (FNRS) at the Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium, e-mail: t.vancutsem@ulg.ac.be.*

A.2 Solution and Interface Variable Processing

Each sub-domain problem is then solved exactly or approximately before exchanging information with other sub-domains. The frequency at which information is exchanged with other sub-domains leads to a compromise between numerical convergence and data exchange rate.

Exchanging information frequently leads to faster convergence, as sub-domain solution methods always use recent values of interface variables, but higher data exchange rate. Exchanging information infrequently or keeping them constant during the whole solution leads to smaller data exchange rates but might degrade the global convergence as sub-domain solution methods use older interface values. It is obvious that when the sub-domains are weakly connected or disjoint, thus interface variables do not affect strongly the sub-domain solution, infrequent updating is better. This kind of partitioning, though, might be very difficult or even impossible.

The choice on the processing of the interface variables dictates the method used for solving the decomposed problem. The two principle methods are: Schwartz alternating and Schur complement.

A.2.1 Schwarz Alternating Method

Among the simplest and oldest techniques are the Schwarz alternating procedures. These methods work by freezing the interface variables during the solution of each sub-domain, hence the sub-domain problems are totally decoupled and no exchange of information is needed. This formulation is very attractive for parallel

implementations since the data exchange rate is minimum. On the contrary, if the sub-domains are not weakly coupled the algorithm can suffer from degraded convergence or even divergence [5], [6], [7], [8]. Other variants of this method can be found in literature depending on how often and in which order the interface variables are updated, for instance the additive or multiplicative Schwartz procedures [3].

A.2.2 Schur Complement Method

When applying the Schur complement DDM, also called iterative sub-structuring, non-overlapping sub-domain partitioning is employed. The sub-domain problems usually involve interior (coupled only through local equations), local interface (coupled through both local and non-local equations) and external interface (belong to other sub-domains) variables. Next, a numerical method (e.g. Newton's) is used to solve the sub-problems.

The Schur complement technique is a procedure to eliminate the interior variables in each sub-domain and derive a global, reduced in size, linear system involving only the interface variables. This reduced system is then solved to obtain the interface variables before each sub-domain iterative solution.

Once the interface variables are computed, the sub-problems are decoupled and the remaining, interior to each sub-domain, variables can be computed independently. In many cases, building and solving the reduced system involves high computational cost. Many methods are used to speed up the procedure, such as approximately solving the system [9], assembling the matrix in parallel using the "local" Schur complements [3], using Krylov solvers [4] or, exploiting the structure of the decomposition to simplify the problem [4], [3].

The formulation and update of the sub-domain solution systems, the elimination of the interior variables, the formulation of the reduced system and the solution of the sub-domain systems can be done in parallel. Unfortunately, this method introduces a bottleneck to the solution algorithm: the sequential computation of the global reduced system to update the interface values. The ratio between the sequential and the parallel part of the algorithm dictates the scalability of the algorithm. However, due to the continuous update of interface variables, the numerical convergence of the algorithm is significantly better than that of Schwarz methods.

APPENDIX B
DYNAMIC SIMULATION ALGORITHMS

B.1 VDHN Algorithm

One of the most common sequential algorithms used in simulation software [10] is the Very DisHonest Newton (VDHN) which belongs to the quasi-Newton family [11]. The algorithm solves directly the integrated DAE system with the use of a Newton method over discretized time. At each discrete time instant the non-linear DAE equations are discretized and algebraized to acquire a system of linear equations $\mathbf{J}\triangle\mathbf{y} = \mathbf{b}$, where $\mathbf{J}$ is the Jacobian matrix, $\mathbf{y}$ is the vector of unknowns ($\mathbf{x}$ and $\mathbf{V}$) and $\mathbf{b}$ is the vector of mismatch values of the non-linear, algebraized equations. The linear system is then solved using a sparse linear solver and the values of $\mathbf{y}$ and $\mathbf{b}$ are updated. Using the updated values, a new linear system is formulated and solved until the procedure converges.

Usually, due to the high computational cost of updating the Jacobian matrix $\mathbf{J}$ after each solution, the latter is kept constant for many consecutive iterations or even time-steps. If correctly implemented, these methods do not affect the accuracy but only the trajectory of the iterative solution. The convergence of the method can be checked on the computed correction $\Delta\mathbf{y}$, on the mismatch values $\mathbf{b}$ or a combination of both. When the method has converged, the solution algorithm proceeds to the next time instant, formulates and solves the new DAE system.

This algorithm is employed by many industrial and academic software and its capabilities and performance are well known. For this reason, it is usually used as the benchmark for proposed algorithms [10].

B.2 Fine-grained Parallel Methods

In order to accelerate the simulation, researchers tried to employ fine-grained parallelization with the use of customized parallel linear solvers. Some methods, like parallel VDHN [12], Newton W-matrix [13] and parallel LU [14], divide the independent vector and matrix operations involved in the linear system solution over the available computing units. Other methods, like parallel successive over relaxed Newton [15] and Maclaurin-Newton [12], use an approximate (relaxed) Jacobian matrix with more convenient structure for parallelization.

While the fine-grained parallelization methods provide some speedup, they are don't exploit the full potential of parallel architectures. A more coarse-grained way of exploiting parallelization was sought, and for that, researchers redirected their attention to DDMs.

B.3 Coarse-grained Parallel Methods

As described in App. A, the main idea of DDMs is to partition the original system, into smaller interconnected sub-systems and employ some form of parallel algorithm to solve them. The first to envisage this application on power system was probably Kron [16] with the diakoptics method, where the domain is "teared" into sub-problems, solved independently and joined back together. At the time, parallel computing was not an option and the target was to address memory issues, but, this method provided the ignition for many of the parallel methods to follow.

Later methods, like waveform relaxation [17] and parallel-in-time [18], introduced the idea of exploiting parallelization in time to increase the granularity of the

parallel tasks. Following, several methods were proposed inspired by different hardware platforms, memory models and partitioning schemes [2], [19], [20], [21]. Some recently proposed methods make use of both coarse-grained and fine-grained parallelization in a nested way [1] to increase performance. Several characteristics differentiate these methods. The most important being the partitioning scheme, the interface variables processing method and the relaxation of interface variables.

As discussed in App. A, automatic partitioning of DAE systems, such as power systems, is not trivial. Some methods, like coherency analysis [22], epsilon decomposition [23] and graph partitioning [7] have been proposed in literature, each with its own benefits and problems. The choice of the decomposition plays big role in the speed of convergence, the load balancing among parallel tasks and the overall performance of the method.

A common characteristic of the already proposed decomposition schemes is the partitioning of the network to interconnected sub-networks and the application of Schwartz-based methods for the full parallelization of the solution avoiding the sequentiality of Schur-complement-based methods. This comes at the cost of computing the partition of a network, which can change according to the topology of the system or even the disturbance to be simulated. Moreover, the Schwartz-based treatment of interface variables can initiate several new iterations, especially if partitioned sub-networks are closely coupled [7].

APPENDIX C
PARALLEL COMPUTING

C.1 Selecting a Parallel Programming Model

Several options are available when developing a parallel implementation. The main candidates considered for our application were:

- distributed memory model, mainly using Message Passing Interface (MPI)
- General Purpose computing on GPUs (GPGPU)
- partitioned global address space, mainly using Fortran Co-array
- shared-memory model, mainly using OpenMP.

The main factors considered to select the appropriate model were: synchronization cost, data exchange rate, hardware cost and easiness to program.

MPI was rejected because its high cost of communication makes it more suitable for coarse-grained parallel algorithms. Algorithms with high rate of data exchange among parallel tasks, as the one proposed, are not likely to be efficient on distributed memory architectures.

GPUs are really good at crunching numbers and can deliver huge peak performance, but they are not as good in handling the irregular computation patterns (unpredictable branches, looping conditions, irregular memory access patterns, etc.) that most engineering software deal with. Moreover, the CPU to GPU data transfer link has relatively high latency introducing a significant bottleneck in the execution of the program. Additionally, there is a high effort needed to develop and maintain GPGPU code and low portability as no default standard exists among GPU vendors. Thus, it was rejected.

Co-array was recently introduced in the Fortran standard as an integrated parallel programming model. It was rejected as the existing support by compilers is minimal and the available user experience and supporting material almost non-existing.

OpenMP, the selected model, is an shared-memory API aiming to facilitate shared-memory parallel programming. OpenMP is not an official standard but it is supported by most hardware and software vendors and it provides a portable, user-friendly, and efficient approach to shared-memory parallel programming. It is intended to be suitable for a broad range of symmetric multiprocessing architectures.

It consists of a set of compiler directives, library routines, and environment variables that influence run-time behavior. A set of predefined directives are inserted in Fortran, C or C++ programs to describe how the work is to be shared among threads that will execute on different processors or cores and to order accesses to shared data.

C.2 OpenMP Parallel Work Scheduling

OpenMP offers some mechanisms for the assignment of loop iterations to threads through the *schedule* clause. Very often, the best load balancing strategy depends on the target architecture, the actual data input, and other factors not known at programming time. In the worst case, the best strategy may change during the execution time due to dynamic changes in the behavior of the loop or changes in the resources available in the system. Even for advanced programmers, selecting the best load balancing strategy is not an easy task and can potentially take a large amount of time.

Three default strategies to assign loop iterations (where each iteration treats a sub-domain) to threads. With the *static* strategy, the scheduling is predefined and one or more successive iterations are assigned to each thread rotationally *prior* to the parallel execution. This decreases the overhead needed for scheduling but can introduce load imbalance if the work inside each iteration is not the same. With the *dynamic* strategy, the scheduling is dynamic during the execution. This introduces a high overhead cost for managing the threads but provides the best possible load balancing. Finally, with the *guided* strategy, the scheduling is again dynamic but the number of successive iterations assigned to each thread are progressively reduced in size. This way, scheduling overheads are reduced at the beginning of the loop and good load balancing is achieved at the end. Of course, many other, non-standard, scheduling strategies have been proposed in literature [24].

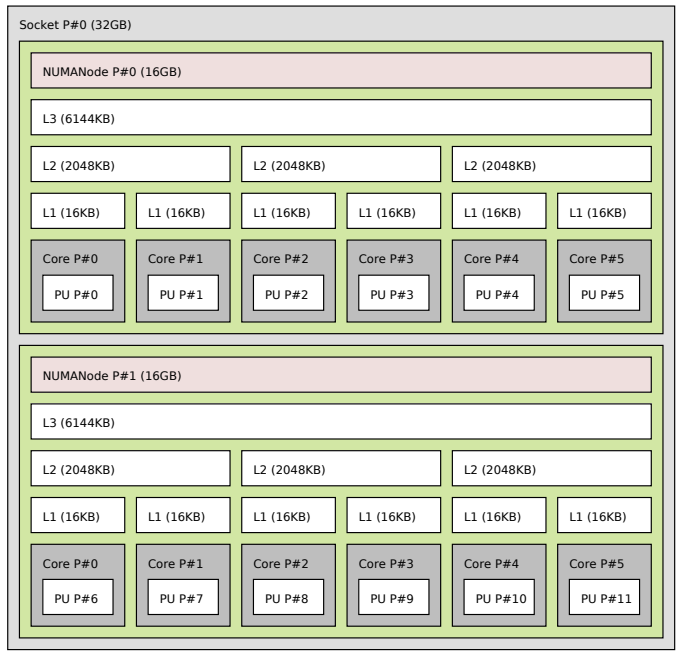


Figure 1. cc-Numa architecture used in our tests

C.3 NUMA Architecture Computers

The proposed implementation targets small and medium, shared-memory parallel computers. Small shared-memory machines (e.g. multi-core laptops and office desktops) have UMA architecture, thus each individual processor can access any memory location with the same speed. On the other hand, larger shared-memory machines usually have NUMA architecture, hence some memory may be "closer to" one or more of the processors and accessed faster by them [24].

The main benefit of NUMA computers over UMA is scalability, as it is extremely difficult to scale UMA computers beyond 8-12 cores. At that number of cores, the memory bus is under heavy contention. NUMA is one way of reducing the number of CPUs competing for access to a shared memory bus by having several memory buses and only having a small number of cores on each of those buses.

The cache coherent NUMA (cc-NUMA) nodes presented in Fig. 1 are part of a 24-core NUMA parallel computer, based on 6238 AMD Opteron Interlagos, used in our tests (see Section 5.2, machine (4)). The computer has two identical sockets, each hosting two NUMA nodes with six cores. So, even though the system physically has two CPU sockets with 12 cores each, they are in fact four NUMA nodes with six cores each.

Resources within each node are tightly coupled with a high speed crossbar switch and access to them inside a NUMA node is fast. Moreover, each core has dedicated L1 cache, every two cores have shared L2 cache and the L3 cache is shared between all six cores. These nodes are connected to each other with HyperTransport 3.0 links. The bandwidth is limited to 12GB/s between the two nodes in the same socket and 6GB/s to other nodes.

Parallel applications executing on NUMA computers need special consideration to avoid high overhead costs. First, given the large remote memory access latency, obtaining a program with a high level of data locality is of the utmost importance. Hence, in addition to choosing the appropriate scheduling strategy, some features of the architecture and the OS affect the application's performance (binding threads to particular CPUs, arranging the placement and dynamic migration of memory pages, etc.) [24].

Data accessed more frequently by a specific thread should be allocated "close" to that thread. *First Touch* memory allocation policy, which is used by many OS, dictates that the thread initializing an object gets the page associated with that item in the memory local to the processor it is currently executing on. This policy works surprisingly well for programs where the updates to a given data element are typically performed by the same thread throughout the computation. Thus, if the data access pattern is the same throughout the application, the initialization of the data should be done inside a parallel segment using the same pattern so as to have a good data placement in memory. This data initializing procedure is followed in our parallel implementation, with each thread initializing the data of the sub-domains statically assigned to it.

Some further consideration is needed when large amount of data are read from files to avoid page migration during the initialization. This problem usually affects NUMA machines with low link speed and applications with intensive i/o procedures. In power system dynamic simulations the data reading is usually done once and then used numerous times to asses several different contingencies on the same system, thus this feature is not critical to their overall performance.

The second challenge on a cc-NUMA platform is the placement of threads onto the computing nodes. If during the execution of the program a thread is migrated from one node to another, all data locality achieved by proper data placement is destroyed. To avoid this we need some method of binding a thread to the processor it was executing during the initialization. In the proposed implementation, the OpenMP environment variable *OMP_PROC_BIND* is used to prevent the execution environment from migrating threads. Several other vendor specific solutions are also available, like *kmp_affinity* in Intel OpenMP implementation, *taskset* and *numactl* under Linux, *pbind* under Solaris, *bindprocessor* under IBM AIX, etc.

APPENDIX D
DYNAMIC RESPONSE AND ACCURACY

D.1 Test-case A

Figure 2 shows the voltage evolution on a transmission bus close to the disturbance for test-case A1. It can be seen that the system is stable in the short-term and long-term. This is a marginally stable simulation. That is, after

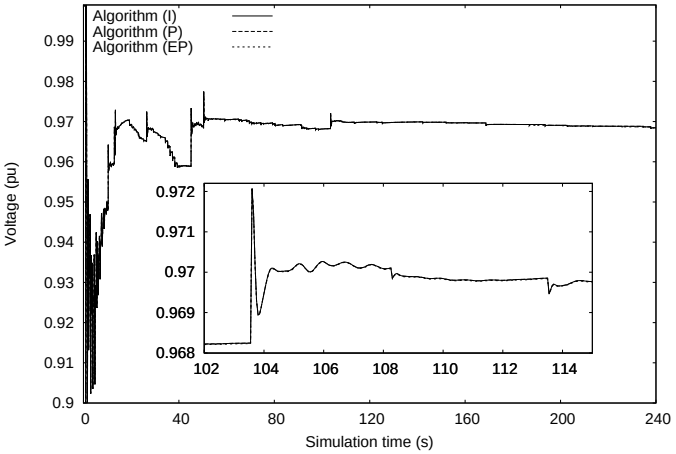


Figure 2. Case A1: Transmission bus voltage

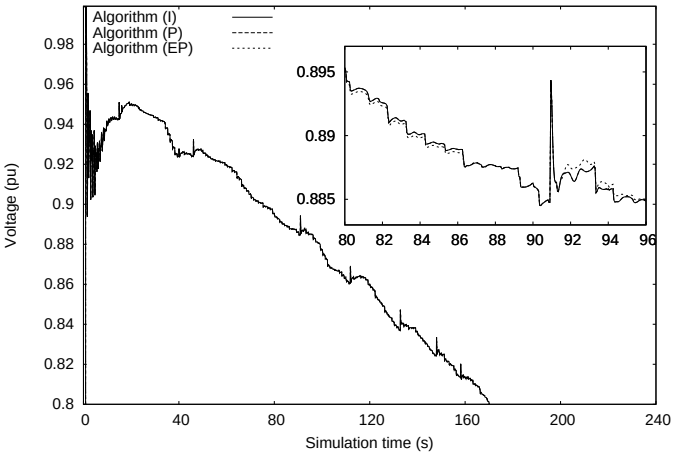


Figure 3. Case A2: Transmission bus voltage

the short-term dynamics the system starts collapsing but is stabilized in the long-term by the actions of the ASCS devices. Such test-cases are the most computationally demanding as they need to be simulated for the whole time horizon to decide whether the disturbance is critical. Moreover, the actual trajectory of the system states is very important, hence static simulations cannot conclude for their stability.

Figure 3 shows the voltage evolution on the same transmission bus for test-case A2. This time, the system is stable in the short-term but long-term voltage unstable. In test-case A1 the voltage collapse is averted by the actions of the ASCS devices deactivated in A2. All three algorithms provide the same results concerning the stability and response of the test-cases.

Figures 2 and 3 show the same responses simulated with all three algorithms. (*P*) offers exactly the same response as (*I*) as they are numerically equivalent. On the other hand, algorithm (*EP*) shows some small deviations from the other two (see Figs. 2 and 3, zoom). As explained in Section 3.5, (*EP*) allows converged subdomains to stop being computed but keeps checking that they satisfy the convergence criteria throughout the remaining solution. Therefore, its response is almost

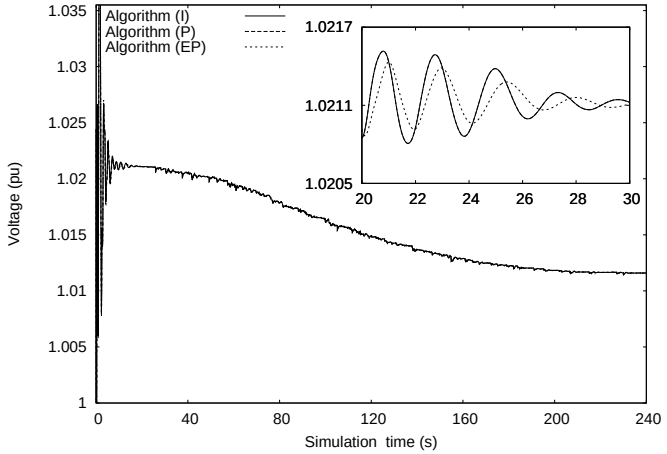


Figure 4. Case B: Transmission bus voltage

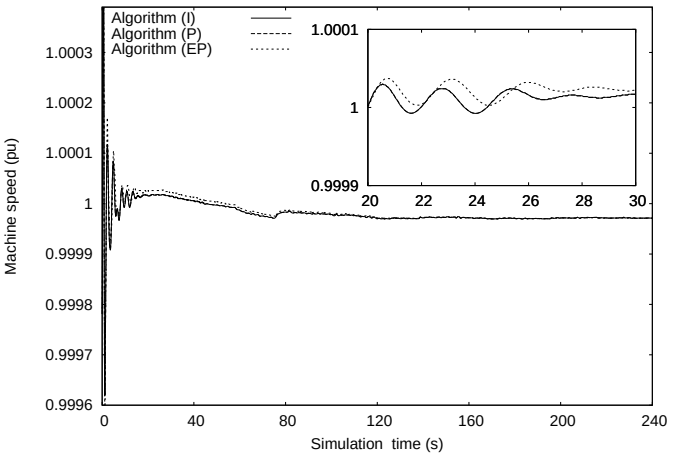


Figure 5. Case B: Generator speed

indistinguishable from the other two and any small deviations at each discrete time computation are bounded by the convergence tolerance.

D.2 Test-case B

Figures 4 and 5 show the voltage evolution of a transmission bus and the machine speed of a synchronous generator, respectively. This test-case exhibits short-term as well as long-term stability. Similarly to test-case A1, this is a marginally stable simulation. That is, after the electromechanical oscillations have died out, the system evolves in the long-term under the effect of LTC devices acting to restore distribution voltages. The decision about the stability of the system can only be made after the simulation of the whole time horizon.

The figures display the responses simulated with all three algorithms. The same observations hold, as with test-cases A1 and A2, concerning the accuracy of the proposed algorithms.

APPENDIX E
ASSESSING THE SCALABILITY OF PARALLEL IMPLEMENTATIONS

E.1 Performance Evaluation

To asses the performance of an existing parallel implementation or the potential of a proposed algorithm, Amdahl's law is often used [25]. It is based on the observation that any parallel implementation consists of a sequentially computed portion S and a parallel portion P that can be split and assigned to M computational units. Furthermore, it assumes perfect load balancing and a perfect parallel machine without any parallelization overhead. The most well known variant is:

$$Runtime(M) = S + \frac{P}{M} \tag{1}$$

Of course, the sum $P+S$ has to account for the sequential execution time of the implementation ($M = 1$).

Based on (1), the scalability of parallel algorithm can be defined as:

$$Theoretic\ scalability(M) = \frac{S+P}{Runtime(M)} \tag{2}$$

It is called *theoretic scalability* as it can never be reached in real applications because of parallelization overhead costs, imbalances in load scheduling, etc. Figure 6 displays the theoretic scalability for several percentage values of parallel work P. It is noticeable that even small differences in the percentage of parallel work lead to big differences in the scalability of the algorithm.

To accommodate for overhead cost of making the code run in parallel (managing threads, communication, memory latency, etc.) Amdahl's law can be modified to:

$$Runtime(M) = S + \frac{P}{M} + OHC(M) \tag{3}$$

where OHC is the overhead cost as a function of the number of computational units used. The modified formula can be used to provide a more realistic prediction of scalability and can be directly linked to the formulas presented in Section 5.1.

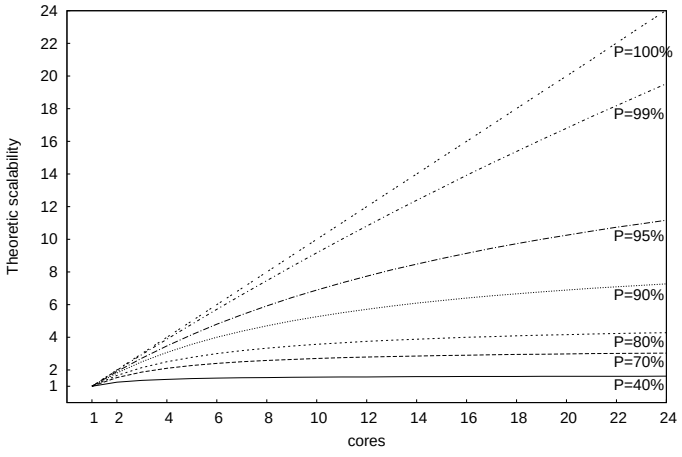


Figure 6. Theoretic scalability based on Amdahl's law

Table 1
Profiling results: Test-case B / algorithm (*P*)

	%	Parallel
Time step initialization	10.02	NO
Injector sub-domain discretization Jacobian calculation and factorization	12.51	YES
Schur complement contributions to simplified reduced system	2.84	YES
Factorization and solution of simplified reduced system (Section 3.3, Eq. 8)	7.12	NO
Injector sub-domain solution for interface and interior variables (Section 3.2, Eq. 4)	61.75	YES
Sub-domain convergence check	3.15	YES
Various (bookkeeping, etc.)	2.61	NO
Total	**100.00%**	**80.25%**

E.2 Profiling Example

Equation 3 shows that scalability can be increased either by increasing the parallel work percentage P or by reducing the OHCs. Finally, it can explain situations where increasing the number of available computational units degrades the performance due to increased OHCs. That is, the value of $\left[\frac{P}{M} - \frac{P}{M+1}\right] - [OHC(M+1) - OHC(M)]$ becomes negative.

In Table 1 a sample profiling performed on the sequential execution of algorithm (*P*) for test-case B is presented. Consequently, the theoretic scalability on 24 cores can be computed as $\frac{S+P}{S+\frac{P}{24}} = 4.3$, with $P = 0.8025$ the parallel and $S = 0.1975$ the sequential portion of the implementation as defined in App. E.1. This means that, if the profiled simulation is executed on a 24 core parallel computer, without any OHC and with perfect load balancing, a scalability of 4.3 could be expected. This value can be compared to the actual scalability of 3.9 achieved (Section 5.4, Table 3). As expected, the actual scalability is smaller than the theoretical due to the overhead costs discussed in Apps. C.3 and E.1.

REFERENCES

[1] V. Jalili-Marandi, Z. Zhou, and V. Dinavahi, "Large-Scale Transient Stability Simulation of Electrical Power Systems on Parallel GPUs," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, no. 7, pp. 1255 –1266, july 2012.

[2] M. Ten Bruggencate and S. Chalasani, "Parallel Implementations of the Power System Transient Stability Problem on Clusters of Workstations," in *Supercomputing, 1995. Proceedings of the IEEE/ACM SC95 Conference*, 1995, p. 34.

[3] Y. Saad, *Iterative methods for sparse linear systems*, 2nd ed. Society for Industrial and Applied Mathematics, 2003.

[4] D. Guibert and D. Tromeur-Dervout, "A Schur Complement Method for DAE/ODE Systems in Multi-Domain Mechanical Design," *Domain Decomposition Methods in Science and Engineering XVII*, pp. 535–541, 2008.

[5] B. Wohlmuth, *Discretization methods and iterative solvers based on domain decomposition*. Springer Verlag, 2001.

[6] A. Toselli and O. Widlund, *Domain decomposition methods–algorithms and theory*. Springer Verlag, 2005.

[7] CRSA, RTE, TE, and TU/e, "D4.1: Algorithmic requirements for simulation of large network extreme scenarios," Tech. Rep. [Online]. Available: http://www.fp7-pegase.eu/download.html

[8] Z. Jackiewicz and M. Kwapisz, "Convergence of waveform relaxation methods for differential-algebraic systems," *SIAM Journal on Numerical Analysis*, vol. 33, no. 6, pp. 2303–2317, 1996.

[9] Y. Saad, "Schur complement preconditioners for distributed general sparse linear systems," *Domain Decomposition Methods in Science and Engineering XVI*, pp. 127–138, 2007.

[10] D. Tylavsky, A. Bose, F. Alvarado, R. Betancourt, K. Clements, G. Heydt, G. Huang, M. Ilic, M. La Scala, and M. Pai, "Parallel processing in power systems computation," *Power Systems, IEEE Transactions on*, vol. 7, no. 2, pp. 629 –638, may 1992.

[11] J. E. Dennis Jr and J. J. Moré, "Quasi-Newton methods, motivation and theory," *SIAM review*, vol. 19, no. 1, pp. 46–89, 1977.

[12] J. Chai and A. Bose, "Bottlenecks in parallel algorithms for power system stability analysis," *Power Systems, IEEE Transactions on*, vol. 8, no. 1, pp. 9–15, 1993.

[13] L. Yalou, Z. Xiaoxin, W. Zhongxi, and G. Jian, "Parallel algorithms for transient stability simulation on PC cluster," in *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, vol. 3, 2002, pp. 1592 – 1596 vol.3.

[14] K. Chan, R. C. Dai, and C. H. Cheung, "A coarse grain parallel solution method for solving large set of power systems network equations," in *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, vol. 4, 2002, pp. 2640–2644.

[15] J. S. Chai, N. Zhu, A. Bose, and D. Tylavsky, "Parallel newton type methods for power system stability analysis using local and shared memory multiprocessors," *Power Systems, IEEE Transactions on*, vol. 6, no. 4, pp. 1539–1545, 1991.

[16] G. Kron, *Diakoptics: the piecewise solution of large-scale systems*. MacDonald, 1963.

[17] M. Ilic'-Spong, M. L. Crow, and M. A. Pai, "Transient Stability Simulation by Waveform Relaxation Methods," *Power Systems, IEEE Transactions on*, vol. 2, no. 4, pp. 943 –949, nov. 1987.

[18] M. La Scala, A. Bose, D. Tylavsky, and J. Chai, "A highly parallel method for transient stability analysis," *Power Systems, IEEE Transactions on*, vol. 5, no. 4, pp. 1439 –1446, nov 1990.

[19] D. Fang and Y. Xiaodong, "A new method for fast dynamic simulation of power systems," *Power Systems, IEEE Transactions on*, vol. 21, no. 2, pp. 619–628, 2006.

[20] J. Shu, W. Xue, and W. Zheng, "A parallel transient stability simulation for power systems," *Power Systems, IEEE Transactions on*, vol. 20, no. 4, pp. 1709 – 1717, nov. 2005.

[21] V. Jalili-Marandi and V. Dinavahi, "SIMD-Based Large-Scale Transient Stability Simulation on the Graphics Processing Unit," *Power Systems, IEEE Transactions on*, vol. 25, no. 3, pp. 1589 –1599, aug. 2010.

[22] D. Koester, S. Ranka, and G. Fox, "Power systems transient stability-a grand computing challenge," *Northeast Parallel Architectures Center, Syracuse, NY, Tech. Rep. SCCS*, vol. 549, 1992.

[23] A. Zecevic and N. Gacic, "A partitioning algorithm for the parallel solution of differential-algebraic equations by waveform relaxation," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 46, no. 4, pp. 421 –434, apr 1999.

[24] B. Chapman, G. Jost, and R. Van Der Pas, *Using OpenMP: Portable Shared Memory Parallel Programming*. MIT Press, 2007.

[25] D. Gove, *Multicore Application Programming: For Windows, Linux, and Oracle Solaris*. Addison-Wesley Professional, 2010.