Component-Level Thermo-Electromagnetic Nonlinear Transient Finite Element Modeling of Solid-State Transformer for DC Grid Studies

Ning Lin, *Student Member, IEEE*, Peng Liu, *Student Member, IEEE*, and Venkata Dinavahi, *Fellow, IEEE*

***Abstract*—Highly-detailed equipment models for electromagnetic transient simulation provide an accurate insight into the system characteristics and behavior. In this article, a coupled field-circuit cosimulation employing detailed component-level models is proposed for the solid-state transformer. To reveal comprehensive thermo-electromagnetic information of the equipment, a high-order nonlinear insulated-gate bipolar transistor (IGBT) model is utilized for the modular multilevel converter, while the finite element method (FEM) is adopted in modeling the transformer. The heavy computational challenge posed by the complexity of these models is alleviated by exploiting model parallelism and the subsequent processing by massively parallel architecture of the graphics processing unit, e.g., a pair of coupled voltage-current sources is adopted for reducing the order of the matrix equation in the circuit part, while in the FEM-based models, a matrix-free nodal domain decomposition solution is utilized to parallelize the overall system to the maximum. A multirate scheme is applied for a further computational burden reduction of the cosimulation due to a large disparity in the appropriate time-steps between power semiconductor switches and the magnetic component. Simulation of a multiterminal dc system including the SST is carried out, and the accuracy of proposed models are validated by offline tools such as SaberRD, ANSYS, and PSCAD/EMTDC.**

***Index Terms*—Electromagnetic transients, electrothermal, finite element method, graphics processing unit, nodal domain decomposition, multiterminal direct-current, modular multilevel converter, parallel processing, solid-state transformer.**

I. Introduction

ELECTROMAGNETIC transient (EMT) simulation at system level is an effective approach to study transient behavior of conventional power systems, and the power electronic converters with a boom in construction of high-voltage direct current (HVdc) projects and the multiterminal dc (MTdc) grid. Currently, the commercial EMT-type power system tools can reveal the performance of the HVdc system with a sufficient accuracy, and thus, are widely used for validation of new topologies and models, control and protection algorithms, and system preview under various operating scenarios [1]–[3].

The involvement of high-frequency power converters composed of power semiconductor switches and magnetic components makes energy conversion efficiency a particular concern in these high-power applications [4]–[7], and the trend towards a more compact volume drives the adoption of detailed models for design evaluation—preferably in the overall system simulation that provides the exact electromagnetic environment where all linear and nonlinear elements become interactive. However, the system-level EMT simulation is incapable of revealing this sort of information because of the lumped and linearized models adopted for attaining fast computational speed. While complex component-level models exist in some simulation tools such as ANSYS Maxwell and ANSYS Simplorer for magnetic and circuit apparatus design [8]–[10], respectively, there are practical limitations on the system scale due to the computational burden, and a high chance of numerical divergence. Therefore, they are seldom used in a large-scale power system or power electronic system simulation.

The adoption of nonlinear high-order power switch models improves the simulation fidelity by providing all necessary information for converter design, including switching losses, instantaneous voltage and current stresses, and the junction temperature [11], [12]. However, the simulation quickly becomes untenable with an expansion in the scale that prolongs the execution time by a larger matrix equation and more iterations for convergence of nonlinear elements. It is also noticed that most component-level EMT simulation tools for power electronics focus mainly on detailed modeling of the semiconductor switches whilst excluding complicated phenomena in the magnetic components. An integrated circuit-field cosimulation is available in ANSYS, which shows finite element method (FEM) for magnetic components and device physics of power semiconductor switches [13]. An obvious drawback is the slow simulation speed due to the twofold complexity from components in the circuit—and most noticeably the field domain [14]–[16].

Manuscript received July 29, 2019; revised November 13, 2019; accepted January 7, 2020. Date of publication January 23, 2020; date of current version October 30, 2020. This work was supported by the Natural Science and Engineering Research Council of Canada. *(Corresponding author: Peng Liu.)*

The authors are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta T6G 2V4, Canada (e-mail: ning3@ualberta.ca; pliu3@ualberta.ca; dinavahi@ualberta.ca).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2020.2967687

0278-0046 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

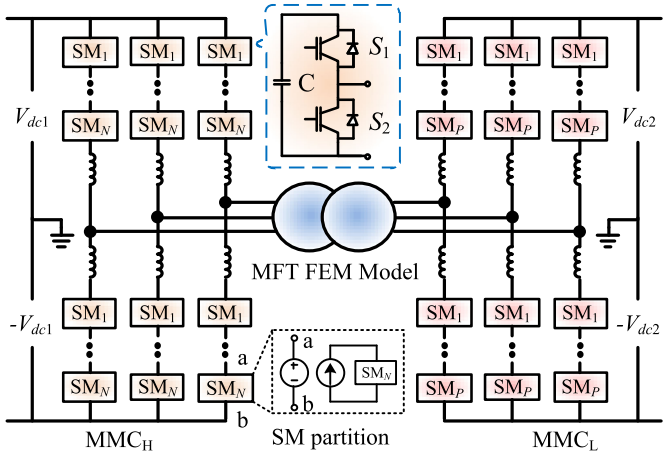

Fig. 1. Component-level solid-state transformer based on MMC.

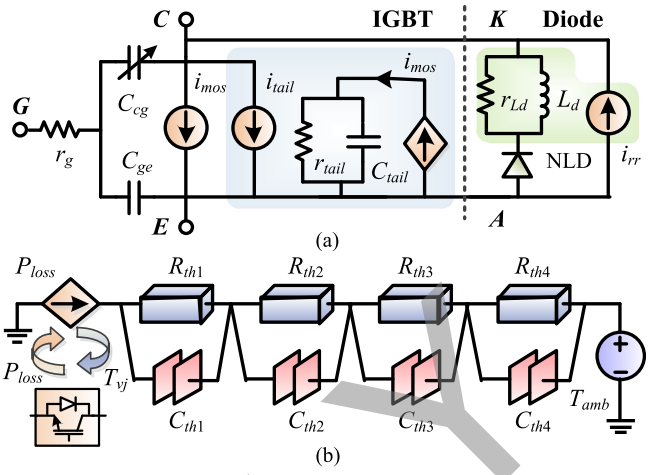

Fig. 2. Nonlinear IGBT electro-thermal model. (a) Fourth-order circuit model. (b) Transient electro-thermal impedance equivalent circuit.

The solid-state transformer (SST) based on the modular power converter topology and medium-frequency transformer (MFT) [17], [18] has a high homogeneity that provides parallelism, which is potentially a solution to efficient simulation. The modular multilevel converter (MMC) contains a large number of submodules sharing the same configuration; and it is intrinsic for FEM to have a substantial number of identical units. Motivated by the heavy computational burden of both the nonlinear component-level MMC model as well as the FEM computation of the MFT within the SST, in this article, a massively parallel cosimulation of these component-level models is established on the single instruction multiple data (SIMD) architecture of the graphics processing unit (GPU). Nonlinear device-level modeling and circuit partitioning are utilized to realize the massively parallel MMC model, which is interfaced with a massively parallel matrix-free solution of the FEM-based transformer model. The fidelity of this GPU-based cosimulation is verified against respective commercial simulation tools while attaining substantial gains in computational speed-up.

The rest of this article is organized as follows. Section II introduces the component-level MMC model, and Section III gives details of the transformer parallel FEM modeling. Section IV specifies the coupled field-circuit cosimulation of the SST. Section V presents the results and discussion of an MTdc grid case study. Finally Section VI concludes this article.

II. MMC Nonlinear Component-Level Modeling

The SST comprising of two front-to-front connected MMCs is given in Fig. 1, where MMC_H and MMC_L are subjected to different dc voltages. Power loss is concerned in the medium-frequency transformer, which is able to operate above the frequency of ac grid, and in the two converters as it affects the operation status of the power semiconductor switch insulated gate bipolar transistor (IGBT). Therefore, the two devices are modeled at component level for gaining an accurate insight.

A. IGBT Nonlinear Thermo-Electromagnetic Transient Model

1) High-Order Circuit Model: As Fig. 2 shows, a datasheet-driven IGBT component-level model contains the equivalent circuit of the IGBT and its antiparallel diode, and an electro-thermal network that is solved independently.

The component realizing basic turn-ON and -OFF behaviors is the essential part that can be expressed as following [19]:

$$i_{\text{mos}} = \begin{cases} 0, & (v_{Cge} < V_t) \| (v_d \le 0) \\ a_2 \cdot v_d^z - b_2 \cdot v_d^{(z+1)}, & v_d < (y \cdot (v_{Cge} - V_t))^{\frac{1}{x}} \\ \frac{(v_{Cge}-V_t)^2}{a_1 + b_1 \cdot (v_{Cge}-V_t)}, & (\text{others}) \end{cases} \tag{1}$$

where a_1, a_2, b_1, b_2 are constants, V_t is the threshold voltage, and v_d is the voltage of the current source. Calculated by

$$i_{\text{tail}} = \left(\left| \frac{v_{\text{tail}}}{r_{\text{tail}}} - i_{\text{mos}} \right| + \left(\frac{v_{\text{tail}}}{r_{\text{tail}}} - i_{\text{mos}} \right) \right) \cdot \frac{i_{\text{rat}}}{2} \tag{2}$$

the tail current appears only during the turn-OFF process and is, then, superimposed onto the result of the MOSFET part.

The freewheeling diode takes an exponential form in the continuous time-domain, as

$$i_D = I_s \cdot \left(e^{\frac{v_D}{nV_T}} - 1 \right) \tag{3}$$

where I_s is the saturation current, V_T is the thermal voltage, and n is the emission coefficient. In addition, the reverse recovery current realized by the voltage-controlled current source is proportional to the voltage over the R-L pair v_{rr} with a coefficient of K_{rr}, as

$$i_{rr} = K_{rr} \cdot v_{rr}. \tag{4}$$

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

Discretization of each component is carried out for EMT simulation. For example, it yields conductance and transconductance when the current source i_{mos} is discretized

$$G_{vd}^{i\text{mos}} = \frac{\partial i_{\text{mos}}}{\partial v_d} = a_2 \cdot z \cdot v_d^{z-1} - b_2(z+1) \cdot v_d^z \tag{5}$$

$$G_{vcge}^{i\text{mos}} = \frac{\partial i_{\text{mos}}}{\partial v_{cge}} = \frac{1}{b_1} - \frac{a_1^2}{b_1(a_1 + b_1(v_{cge} - V_t))^2} \tag{6}$$

and the current contribution in the companion model is

$$I_{\text{moseq}} = i_{\text{mos}} - G_{vd}^{i\text{mos}} v_d - G_{vcge}^{i\text{mos}} v_{cge}. \tag{7}$$

A single switch corresponds to a 4×4 admittance matrix comprising elements from both the IGBT and the freewheeling diode, written as (8) shown at the bottom of this page, and the history currents forms a vector as given in (9) shown at the bottom of this page.

2) Electro-Thermal Model: The IGBT power loss produces heat that is diffused through its case which, as an integral part of the switching device, can be modeled as the transient electro-thermal impedance taking the form of the following analytical function:

$$Z_{\text{th}} = \sum_{i=1}^{n} R_{\text{th}(i)} \left(1 - e^{-\frac{t}{\tau_i}}\right) \tag{10}$$

where the impedance $R_{\text{th}(i)}$ along with the time constant τ_i can be realized by a paralleled R-C pair for EMT simulation, as given in Fig. 2(b) where

$$C_{\text{th}(i)} = \frac{\tau_i}{R_{\text{th}(i)}}. \tag{11}$$

In the equivalent circuit of the transient thermal impedance, the input current source is numerically equal to the IGBT power loss P_{loss}, and its terminal voltage is deemed as the semiconductor's junction temperature T_{vj}. With the inherent cooling mechanism, the IGBT is exposed to the air and, therefore, the other terminal of the R-C chain is a constant voltage source denoting the ambient temperature T_{amb} which is 25 °C. The electro-thermal chain has a total of five nodes, but the acquisition of the junction temperature does not rely on solving matrix equation; instead, knowing that the network injection current provides an instantaneous solution

$$T_{vj}(t) = \left(\sum_{i=1}^{4} \frac{P_{\text{loss}}(t) + I_{\text{his}(i)}(t - \Delta t)}{G_{Ci} + R_{\text{th}i}^{-1}}\right) + T_e \tag{12}$$

where $G_{Ci} = \Delta t / 2C_{\text{th}(i)}$, and the capacitor history current is

$$I_{\text{his}(i)}(t) = P_{\text{loss}}(t) + G_{Ci} \frac{P_{\text{loss}}(t) + I_{\text{his}(i)}(t - \Delta t)}{G_{Ci} + R_{\text{th}i}^{-1}} \tag{13}$$

according to the Trapezoidal rule with a time step of Δt. The outcome T_{vj} is, then, used to update temperature-dependent parameters of the data-sheet-driven IGBT nonlinear model, e.g., V_t, a_1, b_1, which are linearly interpolated since in the data sheet all tests are conducted at two different temperatures, i.e., 25 °C and 125 °C.

B. Fine-Grained MMC Partitioning

The direct solution of an entire MMC is extremely inefficient and—as a matter of fact—impractical since it tries to acquire the circuit information via the Newton–Raphson (NR) iteration of a high-dimension matrix equation, which would be subjected to numerical divergence. Circuit partitioning enabling all MMC submodules to be computed separately is adopted since the arm current has a much lower frequency compared with that of EMT simulation, as shown in Fig. 1. Then, the overall converter is split into two types of subcircuits, the submodule where the high-order nonlinear elements requiring an iterative computation process are confined to, and the purely linear main circuit. The quantity of submodules in a practical application such as the SST indicates a high parallelism, which is a key factor that the GPU is able to gain a remarkable speedup over the CPU in circuit simulation.

1) MMC Nonlinear Submodule: One prominent feature of the MMC-based dc–dc converter is the fault blocking capability even when the submodules are half bridge. After partitioning, the submodules become physically independent and from a mathematical point of view, they can be solved simultaneously. An arbitrary half-bridge submodule corresponds to a six-dimensional (6-D) matrix equation, where the admittance matrix and the current vector are constructed as

$$\mathbf{G_{SM}} = \begin{bmatrix} \mathbf{G^u_{S4\times4}} & \mathbf{0}_{4\times2} \\ \mathbf{0}_{2\times4} & \mathbf{0}_{2\times2} \end{bmatrix} + \begin{bmatrix} \mathbf{0_{3\times3}} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{G^d_{S3\times3}} \end{bmatrix} \tag{14}$$

$$\mathbf{J_{SM}} = \begin{bmatrix} \mathbf{J^u_{S1\times4}} & \mathbf{0}_{1\times2} \end{bmatrix}^T + \begin{bmatrix} \mathbf{0_{1\times3}} & \mathbf{J^d_{S1\times3}} \end{bmatrix} + \begin{bmatrix} \mathbf{0_{1\times3}} & J_A & \mathbf{0}_{1\times2} \end{bmatrix}. \tag{15}$$

$$\mathbf{G_S} = \begin{bmatrix} G_{vd}^{i\text{mos}} + G_{ccg} + K_{rr} + G_{Ld} + G_{RLd}, & G_{vcge}^{i\text{mos}} - G_{ccg}, & -K_{rr} - G_{Ld} - G_{RLd}, & -G_{vd}^{i\text{mos}} - G_{vcge}^{i\text{mos}} \\ -G_{ccg}, & G_{cge} + G_{ccg} + G_g, & 0, & -G_{cge} - G_g \\ -G_{Ld} - G_{RLd}, & 0, & G_{vj} + G_{Ld} + G_{RLd}, & -G_{vj} \\ -G_{vd}^{i\text{mos}} - K_{rr}, & -G_{cge} - G_{vcge}^{i\text{mos}} - G_g, & K_{rr} - G_{vj}, & G_{cge} + G_{vd}^{i\text{mos}} + G_{vcge}^{i\text{mos}} + G_g + G_{vj} \end{bmatrix} \tag{8}$$

$$\mathbf{J_S} = \left[I_{Ldeq} - I_{\text{moseq}} - I_{ccgeq}, I_{ccgeq} + V_g G_g - I_{cgeeq}, I_{jeq} - I_{Ldeq}, I_{\text{moseq}} + I_{ccgeq} - V_g G_g - I_{jeq} \right] \tag{9}$$

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

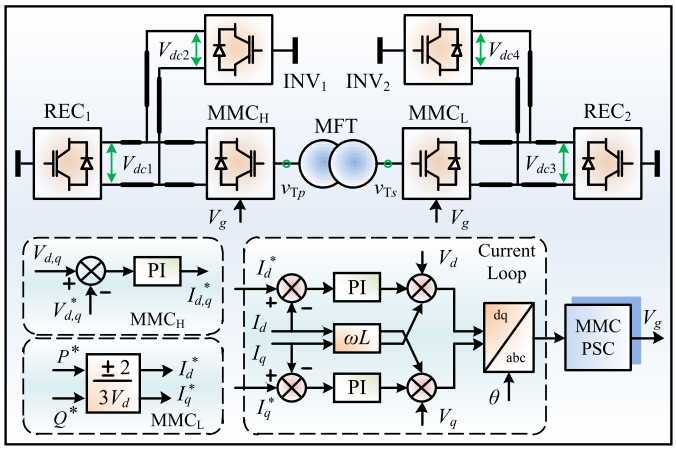

Fig. 3. SST for formation of a four-terminal dc grid and the controller.

The superscripts u and d denote elements from the upper and lower switches, respectively. The last node of the lower IGBT is taken as the virtual ground, and therefore, elements related to that node can all be ignored. The current vector J_{SM} is composed of the arm current J_A in addition to two vectors contributed by the IGBTs. An iterative NR process is involved in solving the six nodal voltages in the highly nonlinear submodule, and the IGBT elements are updated repeatedly; nevertheless, the arm current maintains till the next time-step when a new value is received. The involvement of a lot of nonlinearities requires a small time-step of dozens of nanoseconds in order to prevent an early termination of the simulation due to numerical divergence.

2) Linear Main Circuit: The MMC main circuit turns out to be linear after separating all the submodules. The arm, composed of voltage sources coupled to the submodules and the inductor, can be expressed as

$$v_{\mathrm{arm}}(t) = (i_{\mathrm{arm}} - I_{eq}) \cdot Z_{L\mathrm{arm}} + \sum_{i=1}^{N} v_{pi}. \tag{16}$$

Unlike the submodule, the main circuit having a direct connection to the MFT can be solved by a large time-step in the range of dozens of microseconds. Therefore, the multirate scheme is adopted to reduce the computational burden: a time step of 100 ns is used on the submodule side for capturing switching transients, while it is 20 μs on the main circuit side.

C. MMC-Based DC Grid

An MMC-based 4-terminal dc grid composing of two HVdc links interconnected by the SST is taken as the testbench, as shown in Fig. 3. The HVdc system having a direct connection to MMC_H has a rated dc voltage of $\pm$400 kV, while its counterpart on the other side has a rating of $\pm$200 kV. The two rectifier stations REC_1 and REC_2 controls the power, and their inverter counterparts INV_1 and INV_2 are in charge of dc voltage.

The medium frequency transformer in the center of the SST has a turn ratio of 2 so as to interconnect the two HVdc systems. At the high-voltage side, MMC_H aims to produce a stable, sinusoidal three-phase voltage on the primary windings, and the other MMC regulates the power flow. All the six MMCs share a similar control algorithm, which mainly contains three parts, i.e., the outer-loop controller where the main difference among various types of MMCs exhibits, the current-loop and the inner-loop adopting phase-shift control (PSC) strategy specifically designed for balancing the capacitor voltages of all active MMC submodules [20]. For example, in the outer loop, MMC_H compares the d-q frame voltages with their references, and for the inverter, the dc voltage becomes the control object; on the other hand, the two rectifiers are virtually identical to MMC_L other than the latter is able to operate under a much higher frequency in its ac supply.

III. Finite Element Transformer Model

A. Finite Element Model and Field-Circuit Coupling

The Ampere's law, which states that the magnetic field in space can be generated by both external impressed current and eddy current, is the working principle of a transformer. A finite element transformer model, which provides more accurate and comprehensive information for simulation by considering the factors from a design perspective such as geometries, winding parameters, and material nonlinearity, is governed by the following partial differential equation [21] :

$$\nabla \times (\upsilon \nabla \times A) = J - \sigma \frac{\partial A}{\partial t} \tag{17}$$

where A is the magnetic potential and satisfies $B = \nabla \times A$, υ is the nonlinear magnetic reluctivity, and the right-hand terms are the impressed current and the eddy current, respectively.

The air has poor magnetic conductivity, and the boundary conditions for the magnetic field usually uses $A = 0$ on the artificial boundary of the problem domain. Note that the winding zone with external impressed current density is assumed nonconducting and the eddy current term is 0 in the winding zone.

Since the impressed current density J is associated with the winding currents $J = NI/S$ with N the number of winding turns, I the winding current, and S the winding area, the unknown magnetic potentials $A(i_p, i_s)$, which is a function of primary and secondary winding currents, can be solved by the finite element method with given winding currents. However, in real application, transformers are usually not connected directly to current sources; to interface the finite element model with the external networks, an indirect field-circuit coupling scheme was also proposed, which is very suitable for parallel processing and multirate simulation. According to the Faraday's law, the following equation can be used to calculate the induced winding voltage with the magnetic potentials:

$$U = rI + \frac{Nl}{\Delta_S} \int_S \frac{\partial A}{\partial t} dS \tag{18}$$

where I is the winding current, r the winding resistance, N the number of turns, l the axial length of each filament, S the winding zone, and Δ_S the area of the winding zone.

Since $A(i_p, i_s, t)$ is a function of winding currents and time, according to the partial differential chain rule

$$\frac{\partial A}{\partial t} = \frac{\partial A}{\partial i_p}\frac{\partial i_p}{\partial t} + \frac{\partial A}{\partial i_s}\frac{\partial i_s}{\partial t} \tag{19}$$

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

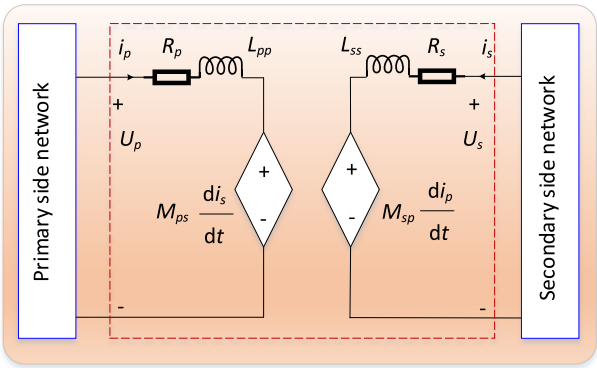

Fig. 4. Self and mutual inductances extracted from the finite element model for electrical network coupling.

(18) can be rewritten as

$$U = rI + \frac{Nl}{\Delta_S}\int_S \frac{\partial A}{\partial i_p}dS\frac{\partial i_p}{\partial t} + \frac{Nl}{\Delta_S}\int_S \frac{\partial A}{\partial i_s}dS\frac{\partial i_s}{\partial t}. \quad (20)$$

Applying (20) to the primary and secondary windings, respectively

$$U_p = r_p i_p + L_p\frac{\partial i_p}{\partial t} + M_{ps}\frac{\partial i_s}{\partial t} \quad (21)$$

$$U_s = r_s i_s + L_s\frac{\partial i_s}{\partial t} + M_{sp}\frac{\partial i_p}{\partial t} \quad (22)$$

where the self and mutual inductances of the transformer can be extracted as

$$L_p = \frac{N_p l_p}{\Delta_{S_p}}\int_{S_p}\frac{\partial A}{\partial i_p}dS,\ M_{ps} = \frac{N_p l_p}{\Delta_{S_p}}\int_{S_p}\frac{\partial A}{\partial i_s}dS$$
$$M_{sp} = \frac{N_s l_s}{\Delta_{S_s}}\int_{S_s}\frac{\partial A}{\partial i_p}dS,\ L_s = \frac{N_s l_s}{\Delta_{S_s}}\int_{S_s}\frac{\partial A}{\partial i_s}dS. \quad (23)$$

Thus, the finite element transformer model can be represented by the self and mutual inductances in Fig. 4 to interface with external networks. Since the winding currents are time varying, these nonlinear inductance values should be updated by the finite element computation at each time-step for different winding currents.

Once the winding currents i_p and i_s are updated from the networks, the inductance calculation in (23) can be performed after solving $A(i_p, i_s)$, $A(i_p + \Delta i_p, i_s)$, and $A(i_p, i_s + \Delta i_s)$ with the finite element method. The choice of Δi_p and Δi_s can be a small increment value such as 0.001 A, and the partial differentials can be calculated based on the Euler's rule

$$\frac{\partial A}{\partial i_p} = \frac{A(i_p + \Delta i_p, i_s) - A(i_p, i_s)}{\Delta i_p} \quad (24)$$

$$\frac{\partial A}{\partial i_s} = \frac{A(i_p, i_s + \Delta i_s) - A(i_p, i_s)}{\Delta i_s} \quad (25)$$

and the integral over the winding zone is intrinsically a weighted summation of the values in the triangular elements. Note that the winding currents are also time varying, which means the inductances from the previous time-step are utilized to calculate the current time-step. The error introduced by the indirect coupling depends on the time step: the smaller the time step, the more accurate the results will be.

Compared with the A-V formulation that solves the magnetic and electrical potentials simultaneously in a global system [23], the indirect coupling scheme extracting the coupling coefficients from the magnetic field computation is more flexible and scalable. Also, it works for multirate simulation using different time-steps.

B. Matrix-Free Finite Element Solution

Fig. 5 shows a 2-D transformer model with triangular mesh, and the problem domain includes the air, transformer core, and winding area. Applying the Galerkin's FEM to (17) in each triangular, the elemental equation can be obtained as

$$KA + T\frac{dA}{dt} = J \quad (26)$$

where A and J are 3×1 vectors representing the unknown magnetic potential and the impressed current source at the vertexes of the triangle, respectively; K and T are 3×3 matrices whose entries are determined by the material properties and shape functions N_i, and within a specific element e they can be written as follows:

$$K_{ij}^e = \upsilon^e \int_{\Omega^e} \nabla N_i \cdot \nabla N_j d\Omega \quad (27)$$

$$T_{ij}^e = \sigma^e \int_{\Omega^e} N_i \cdot N_j d\Omega. \quad (28)$$

The reluctivity υ is a function of the unknown magnetic potential, thus, the problem is nonlinear. Traditional FE solution of such a nonlinear problem usually involves assembly of the elemental equations followed by the NR iteration to solve the global system.

For finite element transient simulation with nonlinearities, the repeat factorization of the Jacobian matrix is very intensive computation, which can be potentially solved by the prevalent high-performance computing (HPC) resources. However, the main challenge lies in the parallelism of a program, which is a prerequisite to utilized these HPC resources according to Amdahl's law [24].

The matrix factorization is sequential and can be hardly vectorized for massively parallel processing. In order to accommodate the SIMD paradigm of modern high-performance computing resources, the nodal domain decomposition scheme is utilized in our article.

Imagine that in the domain decomposition scheme, what will happen if the domain is partitioned in such a way that each subdomain contains only one unknown? Fig. 1 in [25] shows such a case with the extreme domain partition. Each subdomain is solved independently in an iterative manner. The subdomains overlap with each other and information exchanges between them during the iteration. Fig. 6 shows the detailed subdomain solver (SDS) illustrating how the elemental equations are converted to a linear or nonlinear equation with only one unknown.

As is shown in Fig. 5, each finite element node and its direct neighbors make up a subdomain and calculation of each node is mapped to each computing core in a massively parallel manner.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

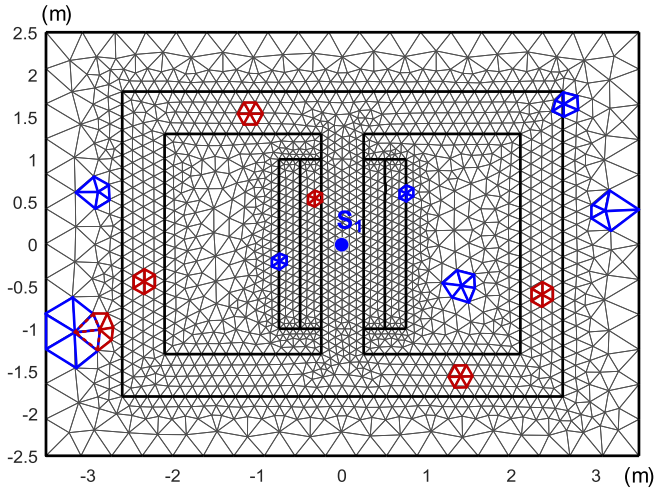

Fig. 5. Finite element transformer model and the node-based subdomains.

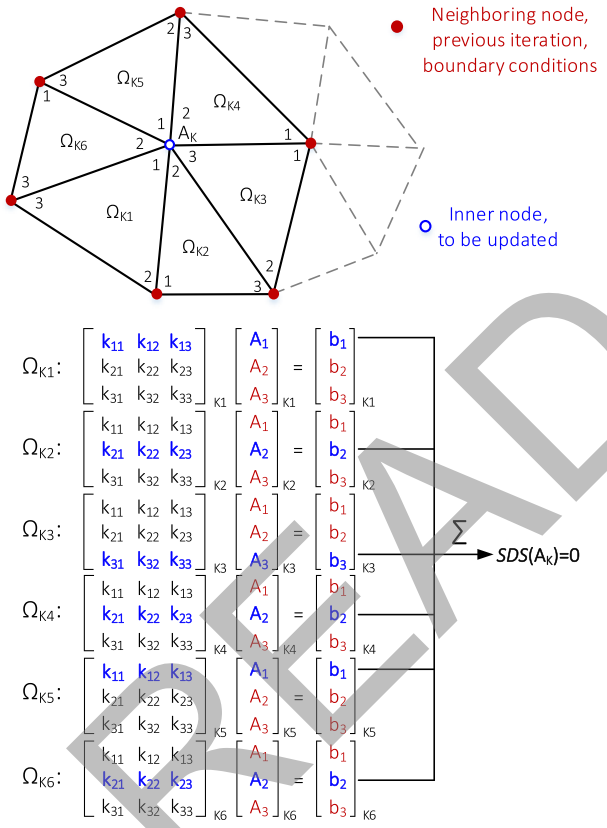

Fig. 6. Inside a SDS.

Compared with the traditional nonlinear FE solver based on the NR algorithm, the nodal domain decomposition (NDD) scheme has the following features [25].

1) There exists only one unknown in each subdomain and, thus, no matrices are necessary, i.e., the NDD is matrix free.
2) As an evolution-type algorithm, the problem is solved by iteratively updating each node based on the values of its neighbors in a massively parallel manner.
3) Each subdomain is solved independently following the same pattern and the computation is concise since there is only one unknown. Therefore, the NDD shows perfect modularity for Kernel programming on GPU architectures.
4) No matrices are involved and the required memory has a linear increase with number of unknowns; and the only data communication between subdomains is an implicit synchronization at the end of each iteration.

C. Multidomain Interfacing

To consider the thermal effect of the transformer windings, a linear heat conduction equation with the same problem domain is also solved

$$\nabla \cdot (\lambda \nabla T) = \rho C \frac{\partial T}{\partial t} - q \tag{29}$$

where λ is the thermal conductivity; ρ is the volumetric mass; C is the heat capacity; q is the heat source.

The elemental equation has similar form with (26), and the natural convective boundary conditions are employed

$$\lambda \frac{\partial T}{\partial n} = h(T - T_0) \tag{30}$$

where n denotes the outward normal direction on the boundary, h is the convection coefficient, and T_0 is the external environmental temperature. In this article, we utilized a convection coefficient of 50W/m^2 K to the surface of the transformer, which is equivalent to forced air cooling with moderate speed or natural oil cooling conditions.

The winding material is copper and the transformer core is made of electrical steel, and their conductivities $\sigma(T)$ altered by the temperature can be represented with

$$\frac{1}{\sigma(T)} = \frac{1}{\sigma_0} + \alpha(T - T_0) \tag{31}$$

where T_0 is the ambient temperature 20°C, for copper, $\alpha = 3.8 \times 10^{-11} \Omega$ m/K and $\sigma_0 = 5.8 \times 10^7$S/ m; and for steel, $\alpha = 5 \times 10^{-11}$ Ω m/K and $\sigma_0 = 9.6 \times 10^6$ S/m.

Thus, the integrated model consists of three physics-bases subsystems: external electrical networks, the magnetic field, and the thermal field, thus, there exist three kinds of interfaces between each other, as shown in Fig. 7.

The interfaces of the thermal field with the other two domains are quite direct: the external electrical networks can provide the time-varying winding currents, thus, the Joule-type loss in the copper windings is available; the time-varying magnetic field will induce electrical field in the steel transformer core, and the eddy current losses can also be obtained after postprocessing. Naturally, both the winding loss and eddy current loss are fed to the thermal field as heat sources. In turn, the thermal field updates the temperature distribution at each time-step, and the conductivities of the copper winding and steel transformer core are altered due to change of temperature.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

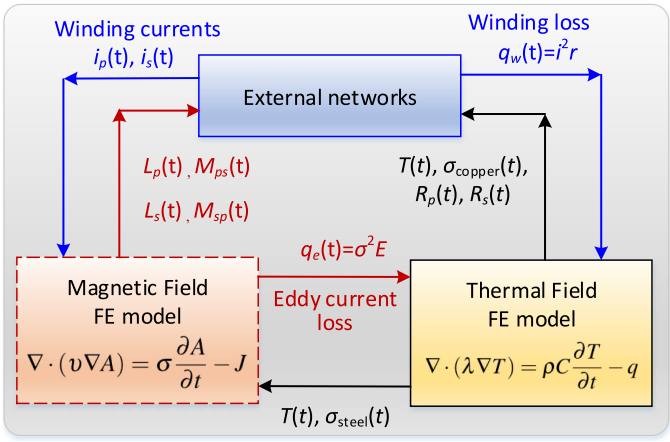

Fig. 7. Coefficients to couple the magnetic field, thermal field, and the external electrical networks.

IV. Coupled Field-Circuit SST Kernel

The parallel cosimulation executed on the GPU conforms to the single-instruction, multiple-thread (SIMT) mode where system components with the same attribute are written as a kernel—the terminology of global function in CUDA C programming—which is able to invoke a designated number of threads simultaneously [26]. Therefore, all electrical component models and mathematical functions need to be categorized according to their properties prior to being written into various kernels, as given in Fig. 8(a), where a general classification is available in the SST computational architecture.

It has been noticed that the SIMT implementation mode neither forces the adoption of exactly identical functions and circuit component models, or their quantities, meaning that a maximum possible parallelism can be achieved to enhance the computational efficiency. It is notable in the SST that its two MMCs have distinct control targets, and their submodule numbers are most likely unequal, which is the major factor that makes the GPU program design intricate. As can be seen from Fig. 8(a), all the submodules in both MMCs are written into one kernel with a total thread number of 6 (N+P) in order to maintain an efficient computational architecture. The probable difference in the submodule numbers results in distinct structure in the inner-loop controllers employing the phase-shift control strategy, which requires a proper thread-based identification of various submodules, MMC arms and legs, and even the two converters. The PSC is intentionally split into two kernels, i.e., the averaging (Ave) control and the balancing control for a thorough utilization of the massively parallelism noticing that the latter part is able to invoke the same large number of threads to the SM kernel, both of which are, as can be observed, the dominant factor in the circuit domain in determining the parallel simulation efficiency; on the contrary, the Ave kernel accounts for merely six threads in the SST and, therefore, is not a determinative factor in the simulation efficiency.

While the GPU is in charge of the major computational tasks of conducting the SST field-circuit cosimulation, the involvement of CPU, termed as host, is still mandatory, as shown in

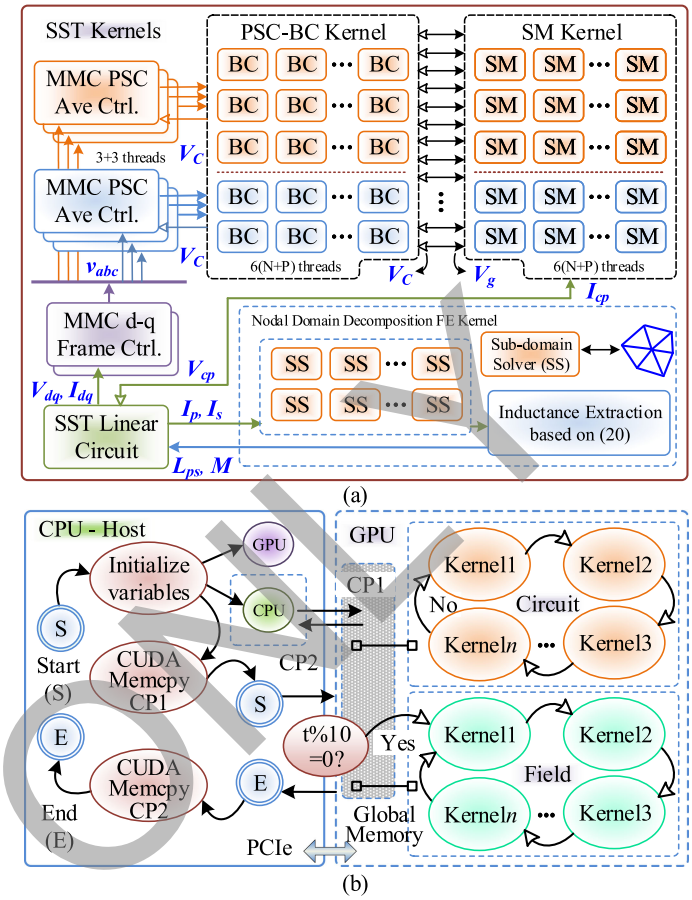

Fig. 8. SST computation architecture on the GPU. (a) Massively parallelism realization. (b) GPU cosimulation implementation process.

the general computation process depicted in Fig. 8(b). All the variables transmitting among kernels need to be initialized in the CPU and, then, copied via the PCIe bus to the global memory of the GPU, which takes over once the cosimulation starts. As can be seen, the independence between the field domain and its circuit counterpart enables their concurrent computations, with the only interaction taking place when one time-step is completed in either domain, and the distinct transient times between them promotes the multirate implementation, i.e., the MMC submodules with nonlinear IGBT/diode model require a much smaller time-step in the range of hundreds of nanoseconds to ensure proper IGBT/diode device-level results and numerical convergence while the FEM transformer tolerates a much larger time-step.

In each domain, the kernels are computed one after another, with their inputs read from the global memory where the outputs are also stored. When an intact loop representing one time-step finishes, the time moves forward by one step. When the simulation ends on the GPU, the CPU takes over again to deal with the signals, e.g., transferring the data back to the CPU.

At the end of each kernel, the synchronization order is carried out to ensure that all threads launched by the same kernel with exactly the same or slightly different contents are completed before computing the next kernel either in the following NR

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

iteration or a new time-step; otherwise, a potential data asynchronization will lead to incorrect simulation results or even numerical divergence.

V. Massively Parallel Cosimulation Results

The inclusion of IGBT fourth-order nonlinear transient model and the transformer FEM model provides a comprehensive and accurate view of the SST. The component-level results are validated by SaberRD and ANSYS while the system-level performance is verified by PSCAD/EMTDC. The power electronic switch models in SaberRD have been subjected to rigorous experimental tests [27], and thus, the tool provides indirect but reliable validation. In the meantime, for power equipment design evaluation, or system-scale performance study, the abovementioned industrial-standard offline simulation tools are widely accepted and consequently extensively referred to. The fact that those commercial tools become extremely slow in computing when too many details, especially in component level, are involved prompts exploitation of the GPU parallel cosimulation in this article, which achieves functions beyond a single commercial offline EMT-type solver, let alone with a much faster simulation speed.

A. Comprehensive SST Component-Level Tests

As the scale of power converters with high-order nonlinear power semiconductor switch models that current component-level simulation tools able to solve is very limited due to numerical instability, a single-phase nine-level MMC with a reasonable dc-bus voltage of 8 kV, which is the maximum scale that SaberRD can solve is employed for the proposed IGBT model validation.

The switching transients peculiar to a device-level model, e.g., the IGBT turn-ON currents and voltages, are provided. In Fig. 9(a), a 5 μs dead time is set for the two complimentary IGBTs, and their current waveforms demonstrate tolerable overshoots; on the other hand, when the dead time is canceled, a tremendous surge which exceeds the capacity of the BSM300GA160D IGBT module is observed in both switches, as given in Fig. 9(b). More apparent results are provided by showing the junction temperatures in Fig. 9(c) and (d). As can be seen, in the former case, both of the IGBTs maintain a normal temperature of around 55 °C, while without an appropriate dead time the junction temperatures could reach over 100 °C, meaning the IGBTs cannot operate in such scenarios. The good match between the proposed fourth-order model and the SaberRD fifth-order model indicates that the proposed model can be used in real applications such as the SST with dozens or even hundreds of voltage levels to provide a thorough insight into component-level details that otherwise current commercial EMT-type solvers unable to achieve.

B. SST in DC Grid

In the four-terminal dc grid shown in Fig. 3, the rectifier REC_1 initially delivers a total amount of 800 MW, around 600 MW of which is fed to the inverter INV_1, and the remaining is

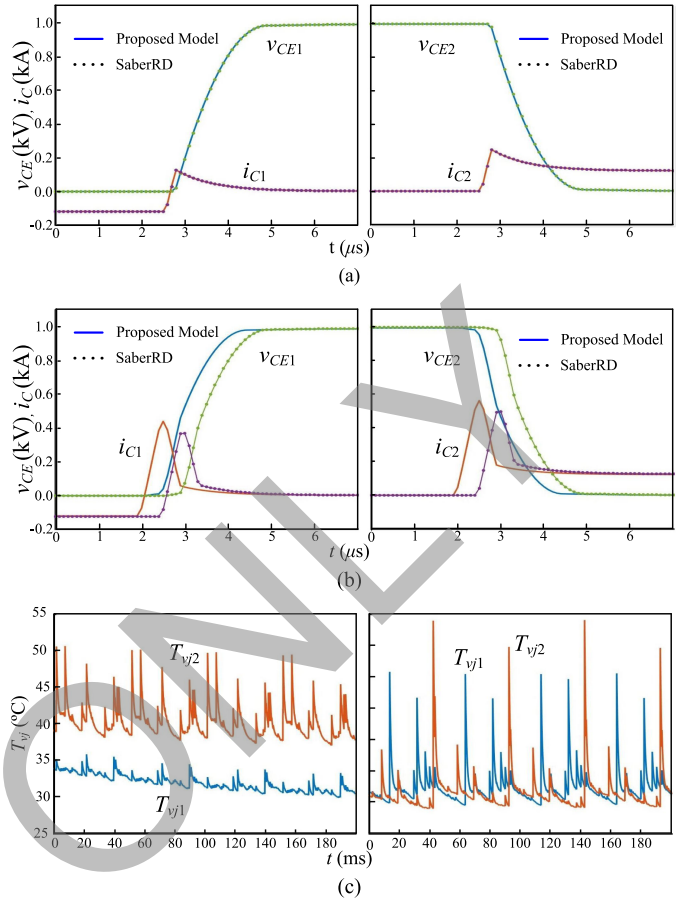

Fig. 9. IGBT component-level model validation. (a) Switching dead time 5 μs. (b) No switching dead time. (c) Junction temperatures when the dead time is 5 μs (left) and 0 (right).

transferred via the SST to the other inverter, which also receives a total amount of approximately 600 MW. Therefore, the dc currents at the two inverter stations are about 1.44 kA and 2.92 kA, respectively, as given in Fig. 10. Then, at $t = 6$ s, the load demand of the second HVdc link surges to 800 MW, and consequently the power command in MMC_L increases by another 200 to 400 MW. After a transition lasting a few hundred milliseconds, the dc current at the two inverters change to nearly 0.9 and 3.9 kA, or the powers are roughly 370 and 780 MW, respectively. As can be seen, a surge in the power command in the SST controller causes insignificant perturbations to the dc grid. For example, the dc voltages in HVdc link 1 witness a 5 kV—or 1.25%—dip momentarily; while the voltages on both windings of the MFT are kept sinusoidal with an amplitude ratio of 2. The abovementioned results are validated by PSCAD/EMTDC with a lower MMC level because otherwise it is unable to simulate.

C. Finite Element Results

Compared with lumped transformer model, the finite element model trades computational efficiency for accuracy and detailed physical information such as the distributions of magnetic potential, flux density, eddy current density, temperature, etc.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

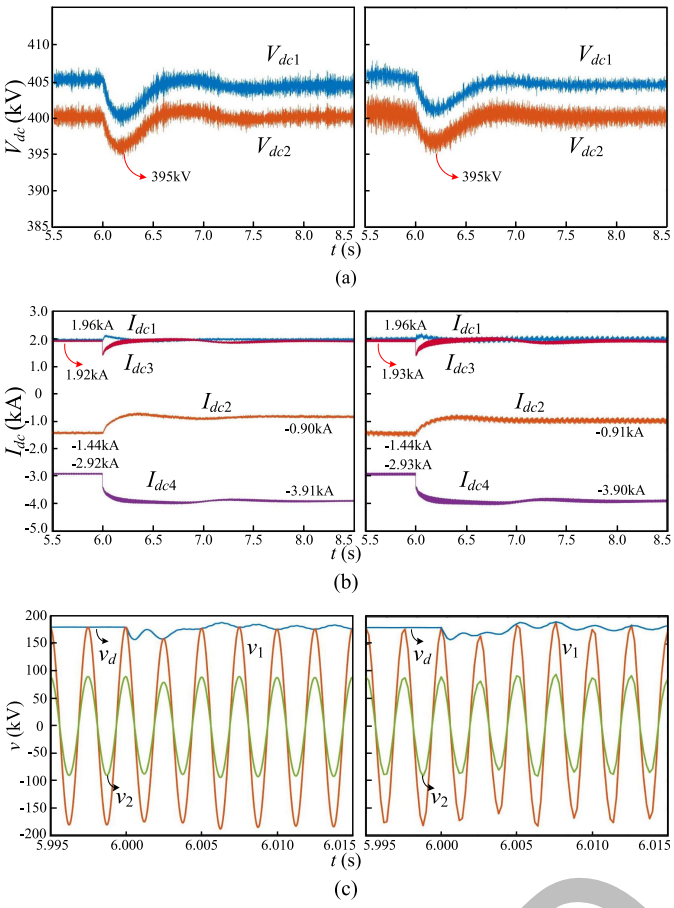

Fig. 10. DC grid reaction to SST power step (Left: Proposed model; Right: PSCAD/EMTDC). (a) DC voltages of HVdc link 1. (b) DC currents. (c) MFT 400 Hz voltages.

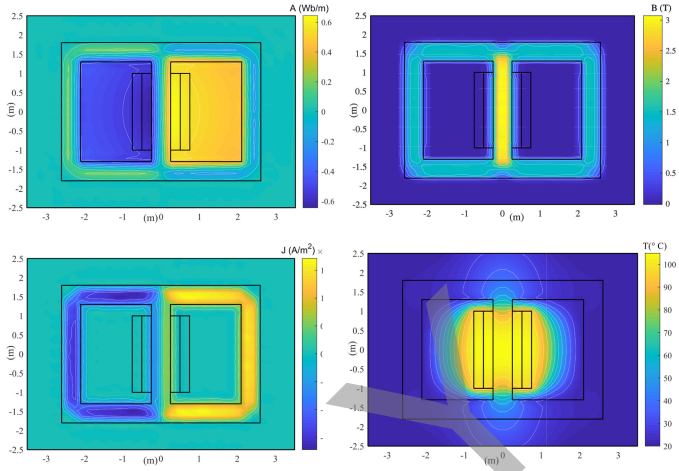

Fig. 11. Field distributions at time point t=2000 s including: the magnetic potential A, the flux density B, the eddy current J, and the temperature T.

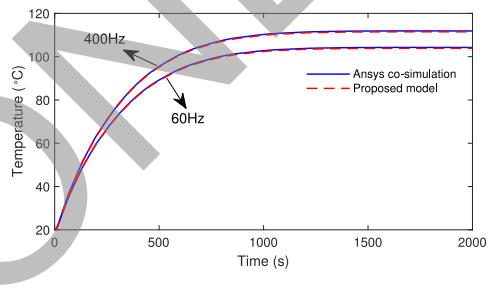

Fig. 12. Temperature rise on the sample point for different case studies.

Fig. 11 shows these field distributions obtained from the proposed model at certain time point, and these information including magnetic saturation, temperature rise, and eddy current loss will help engineers on designing and testing equipments. The flux density distribution shows the transformer core is saturated at this time point. The magnetizing current $I_m = I_p - I_s N_s/N_p$ from the circuit determines the time-varying field distributions; and the field distributions in turn influence the inductances fed to the circuit simulation to calculate the winding current for the next time-step. Normally, the direction of the flux lines change with the magnetizing current I_m, with a phase delay caused by the eddy current term.

Fig. 12 provides the temperature-time curve on the sample point noted in Fig. 5. It takes approximately 20 min for the thermal field to achieve steady state, and the maximum temperature within the winding areas can reach 110° with the convection coefficient 50 W/m^2 K.

As mentioned previously, the NDD scheme is perfectly suited for massively parallel architectures since each subdomain can be solved independently within each iteration. The NDD scheme is implemented on a parallel workstation using the many-core GPU while the ANSYS cosimulation is executed using the multicore CPUs. Specifically, the workstation has dual Intel Xeon E5-2698 v4 CPUs, 20 cores each, 2.2 GHz clock frequency, and 128 GB RAM. The GPU is the NVIDIA Tesla V100-PCIE-16 GB with 5120 Cuda cores.

Since the nodal domain decomposition is an iterative scheme, the number of iterations determined the accuracy. In our case study, the relative tolerance between two successive iterations is set to 10^{-5} as the convergence criterion, and the relative error compared with ANSYS Maxwell is less than 0.5%. It takes 1.8 s per time-step for ANSYS Maxwell–Simplorer cosimulation to compute the nonlinear FE transformer with 1273 nodes, while the execution time per time-step of the proposed NDD solution on GPU is only 65 ms, implying a speedup of 27.7. In other words, to run 10^5 time steps, it takes two days for ANSYS and 1.8 h on GPU. Note that the ANSYS Maxwell also has built-in HPC license and the number of available cores are set to 40 for the workstation. Meanwhile, the cosimulation achieves computing two high-level MMCs with component-level IGBT/diode details, which is infeasible for ANSYS, or even PSCAD/EMTDC without proper optimization such as circuit partitioning proposed in this article. Therefore, the component-level parallel cosimulation of a complex electric-magnetic system features the function that ANSYS incapable of achieving.

In the future work, we will implement the 3-D finite element model with edge element using this massively parallel scheme and deploy the computation on multiple GPUs and cluster of workstations.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

VI. Conclusion

In this article, the component-level nonlinear modeling of a solid-state transformer, including the power converter and the FEM model of transformer, for field-circuit thermal-electromagnetic transient cosimulation on GPU using kernel-based massively parallel computation approach was presented. The nonlinear high-order IGBT model provided accurate switching transients and, thus, revealed electro-thermal information for converter design evaluation purpose, and so does the FEM-based transformer model, which provided an insight into detailed phenomena such as magnetic saturation, eddy current, and temperature distribution that were otherwise not available from system-level models. An interface between the FEM model and the MMC required for the field-circuit cosimulation of the SST was introduced. Modeling optimization algorithms such as circuit partitioning and the nodal domain decomposition were proposed to exploit parallelism and reduced the computational burden brought about by a lot of nonlinearities and a huge number of nodes at both circuit and the field level. The multirate implementation scheme was adopted to further expedite the simulation by avoiding calculating power system components at an unnecessarily high frequency while the accuracy of power semiconductor switches was retained. Using the massively parallel feature, the proposed cosimulation on GPU was demonstrated a computational capability far beyond that of commercial offline solvers based on CPU, and the results were sufficiently accurate for both equipment design evaluation and system preview. With an increasingly higher demand for simulation accuracy and a more comprehensive insight into circuit information, the conventional CPU-based simulation exhibits a dramatic inefficiency, whilst the proposed modeling and parallel processing concept can be further extended for detailed analysis of other types of complex electrical and magnetic systems where the computational burden posed by the existence of a substantial number of elements can be alleviated by the kernel-based implementation on the GPU.

Appendix

The MMCs forming SST have following parameters.

Arm inductor $L_{u,d} = 30$ mH/1Ω; MMC$_H$: 257-level, SM capacitor 25 mF; MMC$_L$: 129-level, SM capacitor 10 mF.

The medium-frequency transformer parameters are as follows.

1000 MVA/180 kV/90 kV; 5.2 m × 3.6 m for the outer rectangle and 1.85 m × 2.6 m for the two inner rectangles; the coil size is 0.25 m × 2 m, and the coil ratio is 400:200.

References

[1] M. K. Bucher and C. M. Franck, "Fault current interruption in multiterminal HVDC networks," *IEEE Trans. Power Del.*, vol. 31, no. 1, pp. 87–95, Feb. 2016.

[2] J. Xu, C. Zhao, Y. Xiong, C. Li, Y. Ji, and T. An, "Optimal design of MMC levels for electromagnetic transient studies of MMC-HVDC," *IEEE Trans. Power Del.*, vol. 31, no. 4, pp. 1663–1672, Aug. 2016.

[3] S. Debnath and M. Chinthavali, "Numerical-stiffness-based simulation of mixed transmission systems," *IEEE Trans. Ind. Electron.*, vol. 65, no. 12, pp. 9215–9224, Dec. 2018.

[4] E. L. Barrios, A. Ursa, L. Marroyo, and P. Sanchis, "Analytical design methodology for litz-wired high-frequency power transformers," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2103–2113, Apr. 2015.

[5] Y. Tang, L. Ran, O. Alatise, and P. Mawby, "A model assisted testing scheme for modular multilevel converter," *IEEE Trans. Power Electron.*, vol. 31, no. 1, pp. 165–176, Jan. 2016.

[6] F. Hahn, M. Andresen, G. Buticchi, and M. Liserre, "Thermal analysis and balancing for modular multilevel converters in HVDC applications," *IEEE Trans. Power Electron.*, vol. 33, no. 3, pp. 1985–1996, Mar. 2018.

[7] Y. Dong, H. Yang, W. Li, and X. He, "Neutral-point-shift-based active thermal control for a modular multilevel converter under a single-phase-to-ground fault," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 2474–2484, Mar. 2019.

[8] R. Wu. *et al.*, "A temperature-dependent thermal model of IGBT modules suitable for circuit-level simulations," *IEEE Trans. Ind. Appl.*, vol. 52, no. 4, pp. 3306–3314, Jul. 2016.

[9] Y. Duan, F. Xiao, Y. Luo, and F. Iannuzzo, "A lumped-charge approach based physical SPICE-model for high power soft-punch through IGBT," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 7, no. 1, pp. 62–70, Mar. 2019.

[10] Y. Zhang, H. Wang, Z. Wang, Y. Yang, and F. Blaabjerg, "Simplified thermal modeling for IGBT modules with periodic power loss profiles in modular multilevel converters," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 2323–2332, Mar. 2019.

[11] A. R. Hefner, "A dynamic electro-thermal model for the IGBT," *IEEE Trans. Ind. Appl.*, vol. 30, no. 2, pp. 394–405, Mar. 1994.

[12] M. Cotorogea, "Physics-based SPICE-model for IGBTs with transparent emitter," *IEEE Trans. Power Electron.*, vol. 24, no. 12, pp. 2821–2832, Dec. 2009.

[13] ANSYS simplorer technical data sheet. 2016. [Online]. Available: http://www.ansys.com/resource-library

[14] Y. Huangfu, S. Wang, J. Qiu, H. Zhang, G. Wang, and J. Zhu, "Transient performance analysis of induction motor using field-circuit coupled finite-element method," *IEEE Trans. Magn.*, vol. 50, no. 2, pp. 873–876, Feb. 2014.

[15] G. J. Li *et al.*, "Excitation winding short-circuits in hybrid excitation permanent magnet motor," *IEEE Trans. Energy Convers.*, vol. 29, no. 3, pp. 567–575, Sep. 2014.

[16] W. Liang, J. Wang, T. Lu, and W. Fang, "A new method for multiple finite-element models in cosimulation with electrical circuit using machine multiloop modeling scheme," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 6583–6590, Dec. 2014.

[17] H. Qin and J. W. Kimball, "Solid-state transformer architecture using AC–AC dual-active-bridge converter," *IEEE Trans. Ind. Electron.*, vol. 60, no. 9, pp. 3720–3730, Sep. 2013.

[18] B. Zhao, Q. Song, J. Li, and W. Liu, "A modular multilevel DC-link front-to-front DC solid-state transformer based on high-frequency dual active phase shift for HVDC grid integration," *IEEE Trans. Ind. Electron.*, vol. 64, no. 11, pp. 8919–8927, Nov. 2017.

[19] M. Zhang, A. Courtay, and Z. Yang, "An improved behavioral IGBT model and its characterization tool," in *Proc. IEEE Electron Devices Meeting*, Hong Kong, Jun. 2000, pp. 142–145.

[20] M. Hagiwara and H. Akagi, "Control and experiment of pulsewidth-modulated modular multilevel converters," *IEEE Trans. Power Electron.*, vol. 24, no. 7, pp. 1737–1746, Jul. 2009.

[21] Jian-Ming Jin, *The Finite Element Method in Electromagnetics*. Hoboken, NJ, USA: Wiley, 2015.

[22] P. Liu and V. Dinavahi, "Real-time finite-element simulation of electromagnetic transients of transformer on FPGA," *IEEE Trans. Power Del.*, vol. 33, no. 4, pp. 1991–2001, Aug. 2018.

[23] H. De Gersem and K. Hameyer, "A finite element model for foil winding simulation," *IEEE Trans. Magn.*, vol. 37, no. 5, pp. 3427–3432, Sep. 2001.

[24] M. D. Hill and M. R. Marty, "Amdahl's law in the multicore era," *Computer*, vol. 41, no. 7, pp. 33–38, Sep. 2008.

[25] P. Liu and V. Dinavahi, "Matrix-free nodal domain decomposition with relaxation for massively parallel finite-element computation of EM apparatus," *IEEE Trans. Magn.*, vol. 54, no. 9, pp. 1–7, Sep. 2018.

[26] CUDA C programming guide, Apr. 2018. [Online]. Available: http://docs.nvidia.com/cuda/

[27] A. R. Hefner and D. M. Diebolt, "An experimentally verified IGBT model implemented in the Saber circuit simulator," *IEEE Trans. Power Electron.*, vol. 9, no. 5, pp. 532–542, Sep. 1994.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.

Ning Lin (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 2008 and 2011, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Alberta, Edmonton, AB, Canada, in 2018.

From 2011 to 2014, he was an Engineer with FACTS and HVdc. He is currently an Assistant Researcher with the University of Alberta. His research interests include real-time simulation of power systems and high-performance computing.

Peng Liu (Student Member, IEEE) received the B.Sc. and M.Eng. degrees in electrical engineering from the Harbin Institute of Technology, Harbin, China, in 2013 and 2015, respectively. He is currently working toward the Ph.D. degree in energy system with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada.

His research interests include computational electromagnetics, high-performance finite element analysis, and parallel and distributed processing.

Venkata Dinavahi (Fellow, IEEE) received the B.Eng. degree in electrical engineering from the Visveswaraya National Institute of Technology, Nagpur, India, in 1993, the M.Tech. degree in electrical engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 1996, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2000.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. His research interests include real-time simulation of power systems and power electronic systems, electromagnetic transients, device-level modeling, large-scale systems, and parallel and distributed computing.

READ ONLY

Authorized licensed use limited to: UNIVERSITY OF ALBERTA. Downloaded on April 18,2022 at 16:31:43 UTC from IEEE Xplore. Restrictions apply.