IEEE Open Journal of
Power Electronics

Received 7 June 2020; revised 22 July 2020 and 2 August 2020; accepted 10 August 2020. Date of publication 13 August 2020; date of current version 23 August 2020. The review of this article was arranged by Associate Editor Zhixiang Zou.

Digital Object Identifier 10.1109/OJPEL.2020.3016296

Hierarchical Device-Level Modular Multilevel Converter Modeling for Parallel and Heterogeneous Transient Simulation of HVDC Systems

NING LIN (Member, IEEE), RUIMIN ZHU (Member, IEEE), AND VENKATA DINAVAHI (Fellow, IEEE)

Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta T6G 2V4, Canada

CORRESPONDING AUTHOR: NING LIN (e-mail: ning3@ualberta.ca)

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

ABSTRACT System-level electromagnetic transient (EMT) simulation of large-scale power converters with high-order nonlinear semiconductor switch models remains a challenge albeit it is essential for design preview. In this work, a multi-layer hierarchical modeling methodology is proposed for high-performance computing of the modular multilevel converter involving device-level IGBT/diode models. The computational burden induced by converter scale and model complexity is dramatically alleviated following the proposal of topological reconfiguration and network equivalence, which create a substantial number of identical circuit units that facilitate massively parallel processing on the graphics processing unit (GPU), using the kernel-based single-instruction multi-threading computing architecture. As the DC system brings significant inhomogeneity which dilutes parallelism, heterogeneous computing is investigated and the computational tasks are properly assigned to CPU and GPU to fully exploit their respective features. The separation of nonlinear device-level models from the rest of the system enables multi-rate implementation for further efficiency enhancement since the two parts allow distinct time-steps. A remarkable acceleration of over 50 times is achieved by the hybrid CPU/GPU platform over conventional CPU simulation, and the validity of the proposed modeling and computing method is confirmed by commercial EMT tools ANSYS/Simplorer and PSCAD/EMTDC.

INDEX TERMS Device-level modeling, electromagnetic transient, graphics processing unit, high-performance computing, high-voltage direct current, modular multilevel converter, parallel processing.

I. INTRODUCTION

The electromagnetic transients (EMTs) of semiconductor switches in a power converter are of increasing concern due to the need to develop efficacious control and protection strategies, and consequently, intensive tests are conducted during the design stage to ensure their safe operation once the converter is commissioned [1], [2]. Prior to prototype tests, commercial simulation tools with detailed nonlinear device-level models such as Pspice and Saber are frequently resorted to for preliminary studies [3]–[7], which offer a comprehensive insight into the converter operation status that aids a proper selection of devices.

When the power converter scales up to withstand a higher voltage and a larger current, the dozens or even hundreds of power semiconductor switches easily overwhelm the simulation capacity, especially in applications such as high-voltage direct-current (HVDC) transmission where the modular multilevel converter (MMC) contains numerous insulated-gate bipolar transistors (IGBTs). As a result, it forces the adoption of simpler models such as the Thévenin equivalent circuit based on two-state resistor model, or the average value model, which are largely adequate for system-level study following a tradeoff between accuracy and efficiency [8]–[10].

Nevertheless, detailed full-scale power converter modeling is still mandatory for design purposes and accurate system

interaction study. Some alternatives were proposed in an endeavor to reveal more device-level information, such as the curve-fitting model, which divides the overall IGBT switching waveform into multiple piecewise linear segments, a few discrete numbers exhibiting nonlinearity or some continuous curves whose duration could be automatically adjusted according to several parameters [11]–[14]. Based on these models, the power consumption can be estimated under normal operation. The fact that curve-fitting models are merely a rough approximation, rather than a substitution, of a power semiconductor switch and therefore could not reveal its exact spontaneous behaviors under various operating conditions in the EMT program manifests the significance of nonlinear high-order IGBT models [15], [16].

The MMC's sheer scale also exacerbates the processors' burden, even though simpler device-level models are adopted in the simulation, due to a huge submodule (SM) quantity that requires repeated calculations conducted in a sequential manner on the CPU, which is dominant in current commercial EMT-type solvers [17]–[19]. Therefore, focuses are diverted to multi-core CPU that supports parallel processing algorithms [20]. The graphics processing unit (GPU) is emerging in the massively parallel computation of electrical systems, and a remarkable speedup is desirable in some cases where an explicit homogeneity prevails [21], [22]. On the contrary, the CPU outperforms its many-core counterpart when the system is highly inhomogeneous, considering the latter processor has a lower clock frequency. As a result, for a complex system such as the MMC-based multi-terminal DC system which possesses both homogeneity and inhomogeneity, neither pure CPU or GPU could achieve the maximum efficiency.

In this work, a CPU/GPU co-simulation platform is proposed as a general solution for simulating complex device-level electrical systems, noticing that nowadays GPU is quite common in a workstation or a personal computer. The IGBT which has a huge quantity is proceeded by the GPU, whilst the less repetitive tasks can be tackled by the CPU. Topological reconfiguration (TR) using coupled voltage and current sources is adopted to reduce the corresponding admittance matrix dimension and consequently improve circuit solution efficiency in addition to avoiding numerical divergence. Meanwhile, as the SM still contains many nodes due to the high-order IGBT and its freewheeling diode (FWD) model, network equivalence (NE) is proposed for further computational burden alleviation.

The rest of the paper is organized as follows: Section II introduces the high-order nonlinear IGBT model, including its freewheeling diode. The detailed MMC multi-layer modeling is then specified in Section III, followed by Section IV where heterogeneous high-performance computing architecture design is illustrated. The CPU/GPU co-simulation results are presented, analyzed, and validated by commercial simulation tools in Section V, and Section VI provides the conclusions.

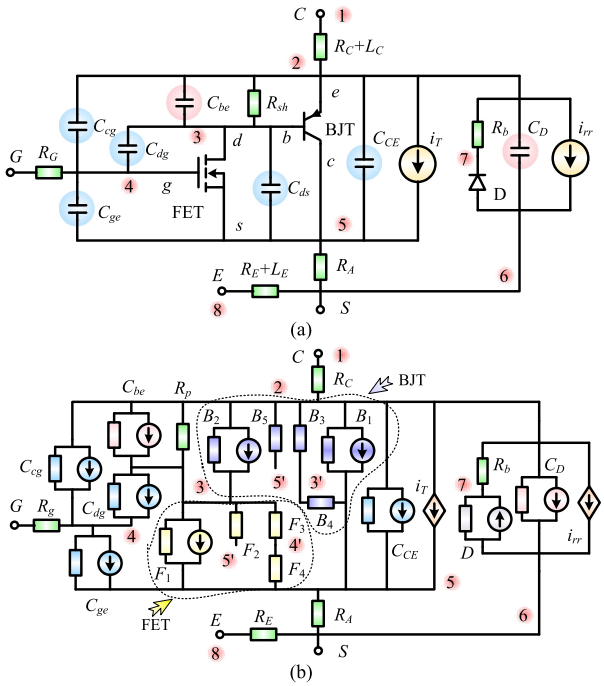

FIGURE 1. Physics-based IGBT modeling: (a) Nonlinear high-order model, (b) discrete time-domain EMT companion circuit.

II. HIGH-ORDER NONLINEAR IGBT MODEL

A. MATHEMATICAL DESCRIPTION

1) BASIC STATIC MODEL

The IGBT physical structure determines that it can be deemed as a metal-oxide-semiconductor field-effect transistor (MOSFET) sharing its drain with the base of a bipolar junction transistor (BJT), as shown in Fig. 1(a). The classical Schichman-Hodges model [23] is used to describe the N-channel MOSFET static characteristics in 3 different regions, namely the cutoff, linear, and saturation:

$$I_{ds}(V_{gs}, V_{ds}) = \begin{cases} 0, & V_{gs} \le V_{th} \\ \frac{V_{ds}}{V_{sat}} I_{sat}(V_{gs}, V_{ds}) \cdot \left(2 - \frac{V_{ds}}{V_{sat}}\right), & V_{ds} \le V_{sat} \\ I_{sat}(V_{gs}, V_{ds}), & V_{ds} > V_{sat} \end{cases} \tag{1}$$

where I_{ds} is the drain current, V_{ds} denotes the drain-source voltage, V_{gs}, V_{th} represent the voltages of gate-source and the channel threshold, respectively, and variables relating to saturation are

$$I_{sat} = K(V_{gs} - V_{th})^{N_{FET}} \cdot (1 + \lambda V_{ds}), \tag{2}$$

$$V_{sat} = A_{FET}(V_{gs} - V_{th})^{M_{FET}}, \tag{3}$$

$$K = \frac{\mu_n C_{ox}}{2} \frac{W_{\text{eff}}}{L_{\text{eff}}}, \tag{4}$$

where λ and μ_n are channel length modulation parameter and carrier mobility, C_{ox} is the gate capacitance per unit area, W_{eff} and L_{eff} are effective channel width and length, and the remaining symbols with subscript FET are constants.

In the PNP-BJT, with a current gain β, the collector current maintains a proportional relationship with the base current, which takes the form of

$$I_b = I_{S_{BJT}} \left(e^{\frac{V_{be}}{M_{BJT} V_T}} - 1 \right), \tag{5}$$

where V_{be} is the base-emitter voltage, V_T is the thermal voltage, and $I_{S_{BJT}}$ and M_{BJT} are coefficients.

2) PARASITIC DYNAMICS

The distribution of holes and electrons in the semiconductor yields parasitic capacitance. Following the changes of junction bias, a dynamic process when the positive and negative charges accumulate or neutralize with each other contributes to switching transients. Therefore, stray capacitors should be included in an integral nonlinear high-order model in Fig. 1(a).

The capacitors, including C_{gc}, C_{ce}, C_{ds} and C_{ge}, are generally under either the enhancement or depletion state, depending on how the junction is biased, i.e.,

$$V^*_{JNC} = M_s \cdot V_{\text{diff}} - V_{JNC}, \tag{6}$$

where V_{diff} and V_{JNC} are diffusion and junction voltages, respectively, M_s is a coefficient, and transition between the 2 states occurs when V^*_{JNC} crosses 0. Then, the capacitors, generally referred to as C_{xy}, can be written as [24]

$$C_{xy}(V^*_{JNC}) = \begin{cases} C_0 \left[(\kappa - 1) \cdot \left(1 - exp\left(\frac{\alpha \cdot V^*_{JNC}(1-\rho)}{(\kappa-1)V_{\text{diff}}} \right) \right) + 1 \right], & V^*_{JNC} \geq 0 \\ C_0 \left(\rho + \frac{1-\rho}{(1 - \frac{V^*_{JNC}}{V_{\text{diff}}})^{\alpha}} \right), & V^*_{JNC} < 0 \end{cases} \tag{7}$$

where C_0 is the reference capacitance, and α, κ, and ρ are factors. Taking the form of

$$R_s(V^*_{JNC}) = D_{dp} \sqrt{\frac{L_C + L_E}{C_{xy}(V^*_{JNC})}}, \tag{8}$$

in which D_{dp} is a coefficient, some additional damping resistances can be added in series to C-E and C-G capacitors so as to avoid possible oscillation between them and the terminal parasitic inductance L_C and L_E.

The diffusion capacitors C_{be} between the BJT base-emitter and C_D in the freewheeling diode, on the other hand, can uniformly be expressed as

$$C_{\text{diff}} = \tau \frac{i(t) + K(V_{gs} - V_{th})^{N_{FET}}}{M_{\text{diff}} \cdot V_T}, \tag{9}$$

where $i(t)$ is the BJT base current or diode anode current, τ denotes carrier lifetime, and M_{diff} is a coefficient.

3) FREEWHEELING DIODE

The freewheeling diode connected to the IGBT in an antiparallel position has the following typical static characteristics

$$I_D = I_{Dsat} \cdot \left(e^{\frac{V_D}{M_D V_T}} - 1 \right), \tag{10}$$

where M_D is an ideality factor, I_{Dsat} means saturation current, and V_D is the forward-biased voltage between anode and cathode. The current-dependent series resistor is calculated by

$$R_b = \frac{R_{Bk}}{\sqrt{1 + \frac{I_D}{I_{NOM}}}}. \tag{11}$$

In the above equation, R_{Bk} is the diode's bulk resistance, and I_{NOM} is the nominal collector current.

B. EMT MODEL DERIVATION

The IGBT/FWD model is described in continuous time-domain and therefore, discretization prior to digital simulation implementation on any processor is mandatory, as shown in Fig. 1(b). Among a variety of integration methods, the 1st-order implicit Backward Euler is preferred since it results in fewer computations without compromising accuracy due to a small time-step that is compulsory for device-level simulation.

Many internal components are determined by multiple factors, e.g., (1)-(3) indicates that the MOSFET current is not only dependent on its own terminal voltage V_{ds}, but also V_{gs} that exerts between gate and source. Therefore, partial derivatives which lead to both conductance and transconductance respectively are performed, e.g., for the former,

$$G_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}} = \begin{cases} 0, & V_{gs} \leq V_{th} \\ \frac{\partial I_{sat}}{\partial V_{ds}} \frac{2V_{sat}V_{ds} - V^2_{ds}}{V^2_{sat}} + 2I_{sat} \left(\frac{V_{sat} - V_{ds}}{V^2_{sat}} \right), & V_{ds} \leq V_{sat} \\ \frac{\partial I_{sat}}{\partial V_{ds}} = \lambda K (V_{gs} - V_{th})^{N_{FET}}. & V_{ds} > V_{sat} \end{cases} \tag{12}$$

The transconductance $G_{ds,gs}$ can also be derived by taking $\partial I_{ds}/\partial V_{gs}$ in the same manner. Then, the companion current in Branch F_1 can be found as

$$I_{dseq} = I_{ds} - G_{ds} V_{ds} - G_{ds,gs} V_{gs}. \tag{13}$$

Being nonreciprocal between 2 nodes is one distinct feature of transconductance, e.g., other than G_{ds} in F_1, Node 3 also links to Node 4 and 5 with -$G_{ds,gs}$ and $G_{ds,gs}$, respectively, but the latter 2 nodes do not have any transconductance pointing to Node 3. Therefore, virtual nodes 4′ and 5′ are introduced to reflect the unilateral relationship in branches F_3 and F_2. Similarly, Node 5 also has $G_{ds,gs}$ in F_4 whilst the reverse is not the case.

The BJT companion model is conceptually the same, with its conductance expressed by

$$G_{BJT} = \frac{\partial I_b}{\partial V_{be}} = \frac{I_{S_{BJT}}}{M_{BJT} V_T} e^{\frac{V_{be}}{M_{BJT} V_T}}, \tag{14}$$

and the accompanying current contribution in Branch B_1 as

$$I_{BJTeq} = I_b - G_{BJT} V_{be}. \tag{15}$$

As another 3-terminal device, the BJT discretization also yields transconductance that demonstrates unidirectionality. Among the 5 branches, B_1 and B_2 are reciprocal between nodes 2-5 and 2-3, respectively; however, to Node 3 and 5, B_3-B_5 are unilateral and therefore, 2 virtual nodes 3′ and

TABLE I. Admittance Contributed by Mosfet and bjt

BJT	B_1	B_2	B_3	B_4	B_5
G	βB_2	$G_{BJT}//I_{BJT_{eq}}$	βG_{BJT}	$-\beta G_{BJT}$	$-\beta G_{BJT}$
FET	F_1	F_2	F_3	F_4	
G	$G_{ds}//I_{dseq}$	$G_{ds,gs}$	$-G_{ds,gs}$	$G_{ds,gs}$	–

Node-to-Node Relation

Node	**N-2**	**N-3**	**N-4**	**N-5**
N-2	$\sum_1^3 B_i + B_5$	$B_2 + B_3$	0	$B_1 + B_5$
N-3	B_2	$B_2 + F_1$	F_3	$F_1 + F_2$
N-4	0	0	0	0
N-5	B_1	$F_1 + B_4$	F_4	$F_1 + F_4 + B_1 + B_4$

5′ are introduced. Table I lists all admittance G contributed by the 2 3-terminal nonlinear devices and the node-to-node relations where N-m denotes node m. The nonreciprocal relationship between different nodes is fully shown, e.g., when establishing the admittance matrix, Node 3 contributes branch B_3 to Node 2, while the reverse is not the case. Therefore, the admittance matrix turns out to be asymmetric.

The nonlinear capacitors also distinguish themselves from linear elements. Since C_{xy} is solely dependent on its terminal voltage, the companion current is calculated in a manner given in (15), whilst the actual current is calculated by

$$I_{C_{xy}}(t) = \frac{Q_{C_{xy}}(t) - Q_{C_{xy}}(t - \Delta t)}{\Delta t}, \tag{16}$$

$$Q_{C_{xy}}(t) = \int_{t_0}^{t} C_{xy}(V^*_{JNC}) dV^*_{JNC}, \tag{17}$$

and the equivalent conductance is expressed as

$$G_{C_{xy}}(t) = \frac{C_{xy}(V^*_{JNC})}{\Delta t}, \tag{18}$$

just as other linear and nonlinear capacitors. For diffusion capacitors, the charge is first derived using integration

$$Q_{C_{\text{diff}}}(t) = \tau K \left(V_{gs} - V_{th}\right)^{N_{FET}} \left(e^{V \cdot V_T M^{-1}} - 1\right), \tag{19}$$

and then the current can be calculated according to variation of the charges in two neighboring time-steps, i.e., (16).

As annotated, the entire IGBT model has 8 nodes, meaning a single such device corresponds to an admittance matrix of 8×8 and a current vector with 8 elements in the process of nodal voltages solution using the following equation:

$$\mathbf{G}^k(t) \cdot \mathbf{U}^{k+1}(t) = \mathbf{J}^k(t), \tag{20}$$

where k denotes the Newton-Raphson iteration count, since the nonlinearity requires the admittance matrix and the current vector be updated in every calculation based on the nodal voltage vector at the same iteration $\mathbf{U}^k(t)$.

III. HIERARCHICAL MMC DEVICE-LEVEL MODELING

A general 3-phase MMC is depicted in Fig. 2(a), where the submodule, depending on its type, contains a few high-order nonlinear IGBT/FWDs. All SMs should be presented once device-level power semiconductor switch models are involved to reveal their actual operation status in addition to avoiding numerical divergence. The instant consequence, as the circuit implies, is that there could be tens of thousands of nodes in the power converter. Therefore, in EMT simulation, an admittance matrix of the same order generated accordingly will undergo repetitive Newton-Raphson iterations in each time-step, and numerical convergence may not be attained due to extensive nonlinearity, in addition to an extremely low computation efficiency, which accounts for the fact that device-level large scale converters are not seen in commercial EMT-type solvers.

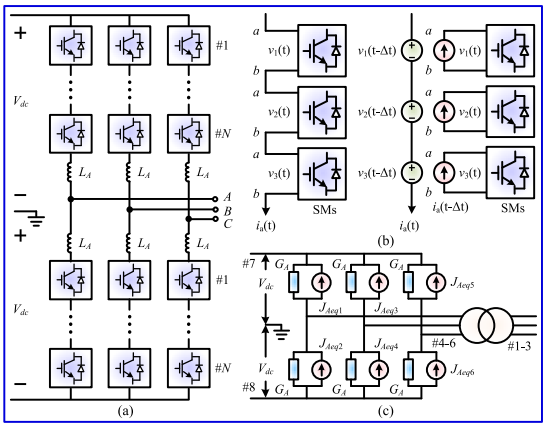

FIGURE 2. Modular multilevel converter modeling: (a) Three-phase configuration, (b) nonlinear submodule partitioning, and (c) linear circuit.

A. TOPOLOGICAL RECONFIGURATION (TR)

A topologically reconfigured EMT model is proposed to mitigate the computational burden by reducing the order of the corresponding admittance matrix, as shown in Fig. 2(b). The submodules can always be separated from the arm to constitute individual subsystems regardless of their type, for the fact that the arm currents have a much lower frequency than that of EMT computation, and consequently can be deemed as constant in two neighboring time-steps.

Coupled by a pair of voltage-current sources, the reconfigured MMC largely has 2 parts, i.e., the nonlinear submodules, and the remaining linear main circuit which interfaces with the external AC and DC grids. Their respective solution turns out to be independent, as the only bond is the interaction which occurs when the 2 sides exchange arm currents and submodules' voltages. Noticing that the main circuit part is purely system-level, a large time-step in the range of a few microseconds is applicable, so that the simulation can proceed faster. In the meantime, each SM undergoes individual Newton-Raphson iterations, and the lower nonlinearity implies better and faster numerical convergence compared with solving the entire MMC as one circuit.

In the main circuit, each arm is comprised of cascaded voltage sources and an inductor, and it can be properly converted into the Norton equivalent circuit in Fig. 2(c), with an equivalent conductance and its companion current of

$$G_{L_A} = \Delta t / (2L_A), \tag{21}$$

$$J_{Aeq}(t) = G_{L_A} \cdot \left(\sum_{k=1}^{N} v_k(t) + 2v^i_{L_A}(t) \right), \tag{22}$$

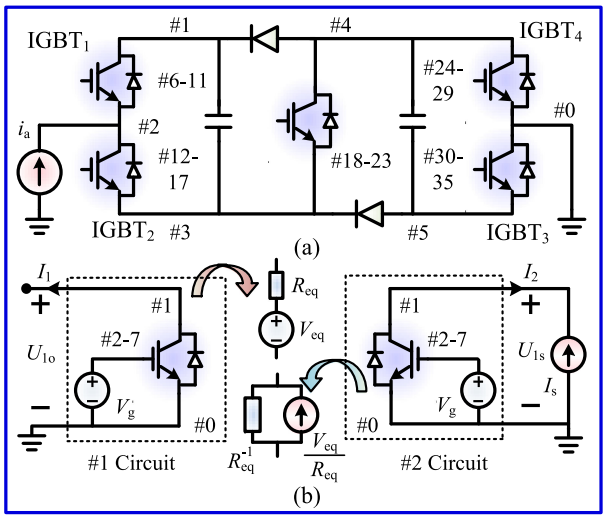

FIGURE 3. Network equivalence: (a) Partitioned CDSM, (b) two-port Thévenin/Norton equivalent circuit derivation.

where v_{LA}^{i} means the incident pulse of a lossless transmission line that could represent reactive components [25]. Then, with a 3-phase transformer and transmission line on its AC and DC sides, the converter linear part has 8 nodes, whose voltage can be solved without iteration following the establishment of the admittance matrix and companion current vector.

B. NETWORK EQUIVALENCE (NE)

The separation of nonlinear submodules from each other achieves a remarkable computational burden alleviation and therefore, for the predominant half-bridge SM (HBSM) MMC, HVDC system-level simulation involving high-order IGBT/FWD model becomes feasible. Nevertheless, the partitioned subsystems may still be beyond the computational capability, especially when many high-order power switches exist, e.g., the clamped double submodule (CDSM) contains 5 IGBTs and 7 diodes, as shown in Fig. 3(a). Therefore, the partitioned circuit with the proposed IGBT/FWD model has up to 35 nodes which results in a large admittance matrix of 35×35, accompanied by excessive iterations and a high chance of numerical divergence.

It can be found that of all the nodes in a single partitioned CDSM, 30 of them are IGBT/FWD internal nodes, while the remaining 5, i.e., Node 1-5, are external nodes. The evident fact that solving a 5-node circuit is remarkably faster than its 35-node counterpart prompts a further simplification. During the iterative solution of a submodule, the admittance matrix and current vector are updated in each calculation following a temporary solution of nodal voltages, meaning that the IGBT/FWD EMT model, including the conductance, transconductance, and companion current sources, is known. Then, the 3-terminal device can be treated as a two-port network since the gate voltage keeps constant in a time-step and consequently does not change any circuit parameters.

Two sets of inputs are sufficient for deriving the IGBT/FWD network's equivalent circuit, as given in Fig. 3(b). The first one is open-circuit which seeks the Thévenin internal voltage V_{eq}. Noticing that one of IGBT/FWD's 8 nodes should be grounded when solving the 2-port circuit, the actual dimension of the admittance matrix reduces to 7. Using the following expanded form of (20), the terminal voltage U_{1o}, which is also V_{eq}, can be obtained:

$$\begin{bmatrix} \boxed{U_{1o}} \\ U_2 \\ \vdots \\ U_7 \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{17} & G_{18} \\ G_{21} & G_{22} & \cdots & G_{27} & G_{28} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ G_{71} & G_{72} & \cdots & G_{77} & G_{78} \\ G_{81} & G_{82} & \cdots & G_{87} & G_{88} \end{bmatrix}^{-1} \begin{bmatrix} J_1 \\ J_2 \\ \vdots \\ J_7 \\ J_8 \end{bmatrix}. \quad (23)$$

In the above equation, elements G_{ij} and J_i belong solely to IGBT/FWD and no external circuit is involved since the two-port network is under the open-circuit state.

In the second case, a DC current source I_s is imposed on the semiconductor switch. Thus, the external current should be added to the 1st element of $\mathbf{J}$, and $U_{1\,s}$, the first element of $\mathbf{U}$, can be obtained from the same equation (23). Taking the two-port IGBT/FWD network as a Norton circuit, its resistance can be derived as

$$R_{eq} = \frac{U_{1s} - U_{1o}}{I_s}. \quad (24)$$

Solving a 7th-order matrix equation, rather than direct components merging, is adopted since a number of suspended nodes in the EMT model make it incomplete.

Following the establishment of a hierarchical MMC structure of system-level main circuit, device-level submodule, and a single IGBT/FWD, the iterative Newton-Raphson process evolves to calculating each IGBT/FWD with a dimension of 7, rather than the overall power converter, or a subsystem involving too many nonlinear high-order semiconductor switches. For example, the solution of CDSM-MMC after the multilayer modeling becomes solving 5 individual IGBTs first, and then the 5-node SM in conjunction with the main circuit.

IV. HETEROGENEOUS COMPUTING ARCHITECTURE

A. BOUNDARY DEFINITION

The MMC-based HVDC system exhibits both homogeneity and inhomogeneity that requires coordinated processing of GPU and CPU respectively so that the computing resources can be thoroughly utilized. Prior to that, a proper boundary that defines the subsystems that the two processors will handle needs to be created. Based on the GPU's hardware features, e.g., the core number and its memory, two main principles are proposed to appropriately divide the DC system, i.e., the component quantity, and the number of signals that will be exchanged between the 2 processors as a result. Fig. 4 presents a complete reconfigured (N+1)-level MMC, whose 6N SMs and corresponding PWM units fully address massive parallelism, whilst the main circuit and the controller are an epitome of

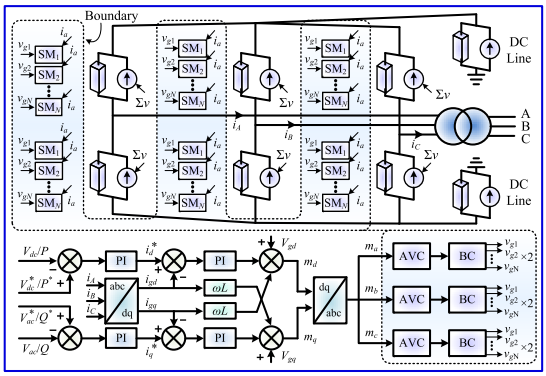

FIGURE 4. Boundary definition for MMC-HVDC high-performance computing.

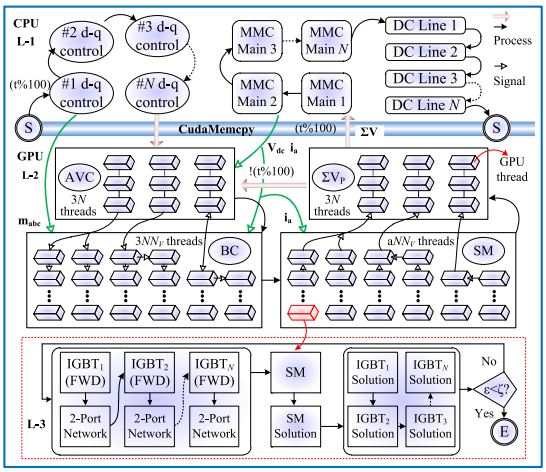

FIGURE 5. Hierarchical multi-layer HVDC system EMT simulation heterogeneous CPU/GPU computing architecture.

inhomogeneity. Thus, conforming to the first principle, an initial boundary can be ascertained.

To be more specific, the MMC controller is twofold: the outer-loop in d-q frame regulates DC voltage or active power on the d-axis, and the AC voltage or reactive power on the other axis; the inner-loop using phase-shift control (PSC) [26] is specifically designed for MMC SM capacitor voltage management. The balancing control (BC) units which generate IGBT gate signals belong to the same category as submodules noticing its huge quantity, whereas the averaging control (AVC) redefines the boundary by moving to the GPU domain considering it minimizes the number of signals exchanged between the two processors.

As a result, the existing boundary leads to an exchange of following signals, the arm currents i_a, the summation of submodule voltages Σv, the DC voltage, and three-phase modulation signals $m_{a,b,c}$ that PSC needs.

B. CPU/GPU COMPUTING ARCHITECTURE

The units classified above are implemented on NVIDIA Tesla V100 GPU [27] installed in a server with 192 GB RAM and 20-core Intel Xeon E5-2698 v4 CPU for the co-simulation, based on the heterogeneous framework where GPU exhibits superiority in computing numerous identical elements which otherwise would be a huge burden to CPU, whose higher processing frequency, in turn, caters to circuits with low regularity. Thus, as shown in Fig. 5, the MMC model yields a hierarchical multi-layer computing structure involving both sequential and parallel processors.

The CPU is in charge of the basic system-level layer L-1 where the MMC linear part, the DC transmission lines, and the d-q frame controller are allocated. In a point-to-point HVDC transmission system or a multi-terminal DC grid, their numbers are limited to a few, whereas the mathematical operations are potentially significant, especially the controller which could include a variety of control and protection strategies. The other 2 layers are dealt with by the GPU using single-instruction multi-threading (SIMT) massive parallelism [28]. The kernel-based layer L-2 is interactive with CPU and constitutes an individual simulation loop with a time-step of 100 ns.

Once the submodule kernel *SM* is invoked, each thread conducts concurrent computation, and a number of GPU device functions are called. The implementation of the last layer is sequential, as shown in the figure: each IGBT/FWD function is first called and then discounted into a 2-port equivalent circuit, followed by solving the submodule and an individual power switch. The iterative Newton-Raphson process ends when the maximum error between two neighboring calculations is below the preset threshold ζ, meaning that 1 thread of SM kernel completes. However, this by no means indicates that all corresponding threads complete as SIMT allows slight diversity which maximizes parallelism. Therefore, a synchronization order is included in each kernel so that the disparity among threads will not induce any signal mismatch between two kernels.

Noticing that system-level simulation tolerates a large time-step, multi-rate is employed in the computing scheme, and a 10 μs step size is applied in L-1. Therefore, the simulation circulates among the 4 major GPU kernels in L-2 100 times before CPU functions intervene and constitute an intact CPU/GPU co-simulation loop.

V. HETEROGENEOUS CPU/GPU HPC RESULTS AND VALIDATION

As one CDSM is equivalent to 2 HBSMs regarding generating the number of voltage levels, the 2 types can share the same control strategy and consequently the performance of HBSM-MMC under normal operation status is representative. In Fig. 6(a)-(b), the upper IGBT switching transients under 2 different freewheeling times, as well as the power loss in one cycle, are compared. Commanding ON/OFF states of the IGBT without any freewheeling time as normally seen in the ideal switch model leads to a dramatic collector current overshoot that is detrimental to the converter. On the other hand, the inclusion of a 10% freewheeling time mitigates the turn-on surge, and the IGBT switching power loss reduces almost by

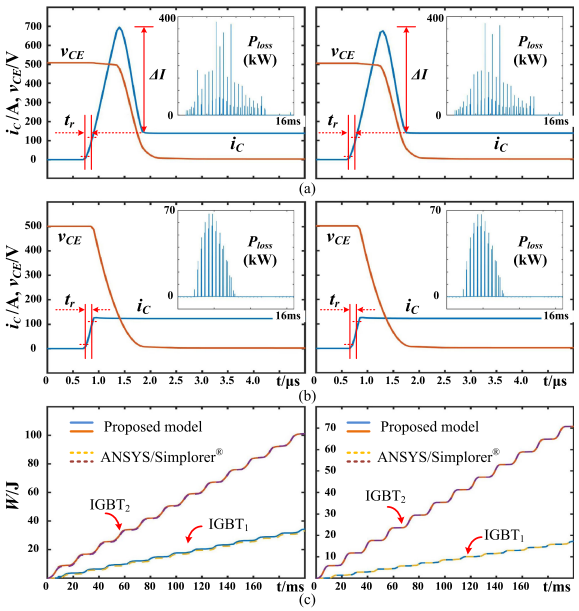

FIGURE 6. IGBT switching transients ((a)-(b) left: physics model; right: ANSYS/Simplorer model): (a) without freewheeling time, (b) 10% freewheeling time, and (c) accumulative energy loss with (L) and without (R) freewheeling time.

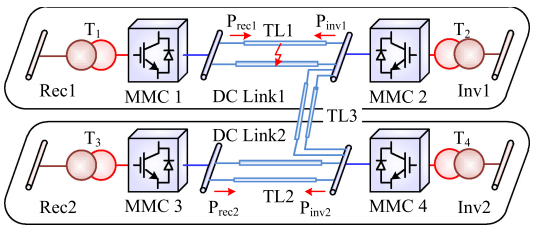

FIGURE 7. A 4-terminal DC grid composed of 2 links.

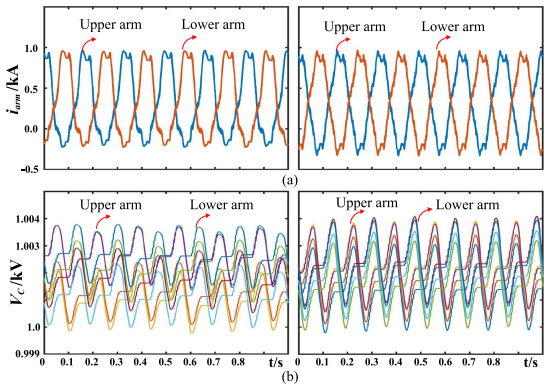

FIGURE 8. MMC steady-state operation (L: physics model; R: PSCAD/EMTDC switch model): (a) rectifier arm currents, (b) submodule DC voltages.

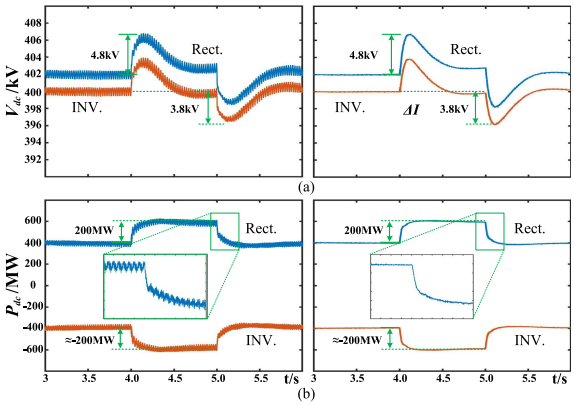

FIGURE 9. DC link 1 power step test (L: physics model; R: PSCAD/EMTDC switch model): (a) Line-line DC voltages, (b) rectifier and inverter power.

one order of magnitude. These switching transients demonstrate a good agreement with results from ANSYS/Simplorer, whose IGBT/diode model has been experimentally validated against real devices and shows a decent accuracy from both static and dynamic perspectives [29], [30]. The accumulated energy losses under these 2 scenarios also show a good agreement with the validation tool in Fig. 6(c). In contrast, the ideal switch model is unable to reveal power loss, neither could it reflect the relationship between current waveform shape and freewheeling time. Therefore, a high-order device-level model is unique for being able to assess a converter design.

A. HBSM-MMC-BASED DC SYSTEMS

System-level simulation with device-level MMC model is carried out, based on the DC system in Fig. 7, which shows a 4-terminal DC grid composing of 2 individual links.

Fig. 8 gives the 5-level rectifier arm currents and submodule DC voltages in a 4 kV DC system. It shows that the MMC is internally well balanced and the SM voltages are regulated with small ripples under steady-state. These waveforms are validated by PSCAD/EMTDC which is a highly-regarded EMT-type solver specialized for electric utility systems simulation, and it has a rich model library that precisely describes the characteristics of various electrical components. Numerous works have been carried out based on this de facto longest-standing, industry-leading transient study tool [31].

As point-to-point HVDC is prevalent, in Fig. 9, DC link 1 is chosen and some system-level simulation results are given, based on the proposed device-level modeling and heterogeneous computing method. After entry into the steady-state, a step order of 200 MW to the power is issued at t=4 s and retrieved 1 sec later. As can be seen, the power strictly follows the order because of a proper controller, and some perturbations to the DC voltages are witnessed in the proposed physic model-based MMC, as well as the two-state resistor case for comparison. The variations in both cases are virtually the same, e.g., the DC voltage first has a sudden rise of approximately 4.8 kV and then gradually reaches a new steady-state at around 4.8 s before sagging. However, the ripples in both voltage and power waveforms reveal that the proposed physics-based IGBT model has a wider frequency band, whereas the conventional ideal switch model adopted in PSCAD/EMTDC leads to much smoother curves since the

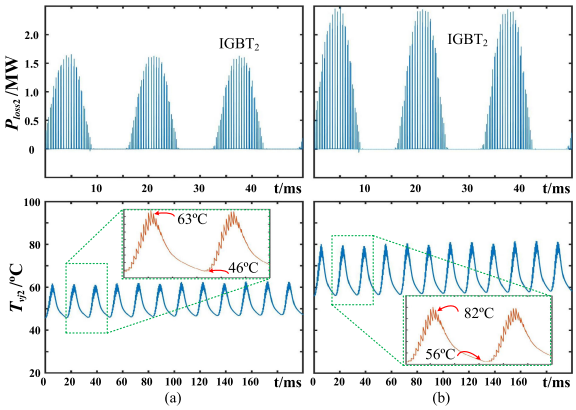

FIGURE 10. Inverter-side HBSM lower IGBT power loss and junction temperature when HVDC DC power is: (a) $P_{dc} = 400$ MW, (b) $P_{dc} = 600$ MW.

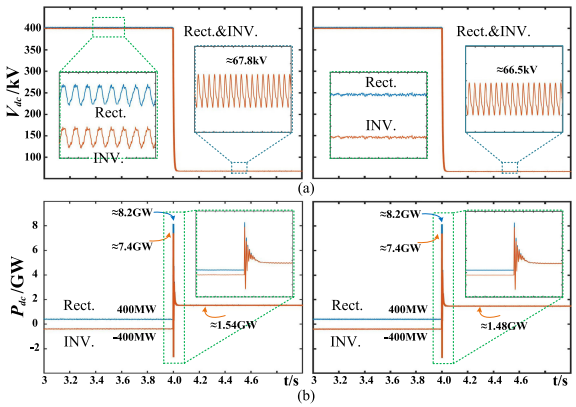

FIGURE 11. HBSM-HVDC line fault (L: physics model; R: PSCAD/EMTDC switch model): (a) Line-line DC voltages, (b) rectifier and inverter power.

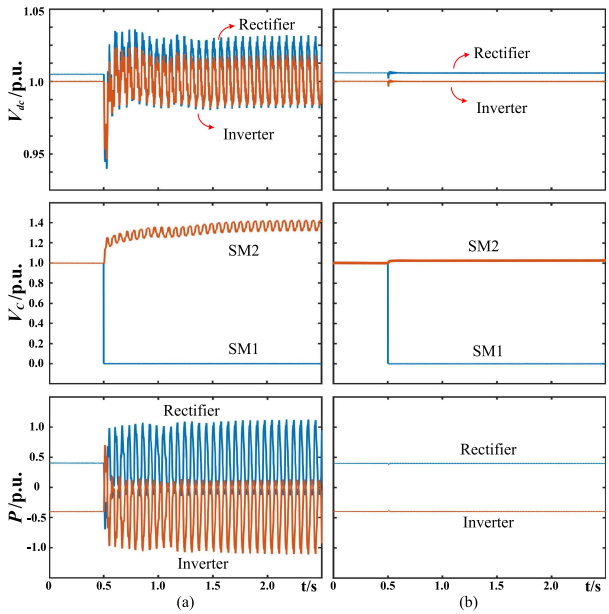

FIGURE 12. HVDC rectifier HBSM 1 inner short-circuit fault: (a) 5-level MMC, (b) 51-level MMC.

TABLE II. Processor Execution Times With 2-t hbsm-hvdc

	TR			(TR)+(NE)		
Level	CPU(s)	Hybrid(s)	Sp	CPU(s)	Hybrid(s)	Sp
5	142	412	0.3	167	127	1.3
51	2187	533	4.1	2725	353	7.7
101	4432	565	7.8	5148	406	12.7
201	8536	540	15.8	10146	427	23.8
401	18794	535	35.1	22928	434	52.8

transition between ON and OFF states completes instantly. As the physics model has been validated in Fig. 6, its involvement in system-level simulation yields more accurate results.

The power losses of a lower IGBT in HBSM can be estimated by the device-level simulation when the 401-level HVDC link delivers 400 MW and 600MW, respectively, as shown in Fig. 10. The peak power loss, caused by IGBT switching transients, could momentarily reach 1.5 MW and 2.5 MW in every cycle for these 2 transmission conditions. The junction temperatures – derived based on the power loss and thermal impedance given in Appendix – are also provided. In addition to a larger fluctuation, a 200 MW increase in the power could also result in an approximate 20 °C temperature rise. Nevertheless, it demonstrates that the converter will be able to maintain a normal function since the junction temperatures in both cases are below the maximum operating temperature 125 °C, as specified in the datasheet [32].

Fig. 11 gives the DC system's response when its ±200 kV lines touch each other in the center at $t = 4$ s and the MMCs block 200 μs later. Initially, the rectifier station delivers 400 MW to the inverter. Following the contingency, a momentary power injection of 8.2 GW and 7.4 GW from the 2 stations into the fault is witnessed, and eventually, they – both working as rectifiers – have a stable provision of 1.5 GW. As mentioned, the pre-fault DC voltage waveform shapes of the two models are different, but they become the same after fault since only freewheeling diodes conduct. The ON-state resistance of the ideal switch is calibrated as 1 mΩ so that the results could match those of the proposed model.

Fig. 12 shows the responses of different MMCs when the inner short-circuit fault occurs in an HBSM in the rectifier. When the MMC level is 5, the DC system's voltage and power are severely affected, and other SMs witness an obvious voltage rise; on the contrary, a sufficient number of remaining SMs in the 51-level MMC can remedy the loss of a single submodule. It implies that conducting experiments with a low-level MMC prototype may lead to incorrect results and therefore, full-scale modeling and simulation are crucial for studying a practical system.

Table II gives a summary of times that pure CPU and hybrid CPU/GPU require to execute 1 million steps of a 2-terminal DC system. The involvement of GPU achieves a few to dozens of times speedup over pure CPU simulation due to its massively parallel processing capability, and the speedup is roughly proportional to the voltage level. Though

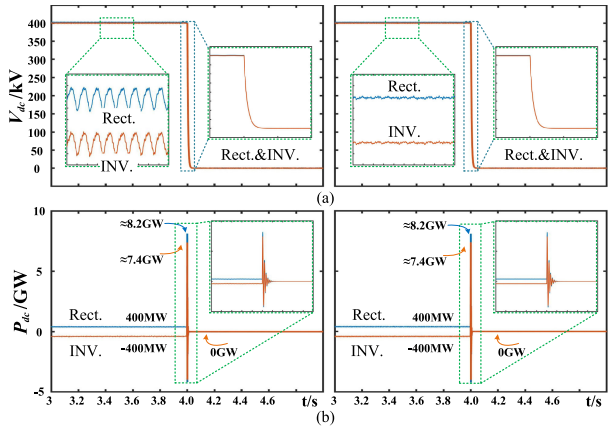

FIGURE 13. **CDSM-HVDC line fault (L: physics model; R: PSCAD/EMTDC switch model): (a) Line-line DC voltages, (b) rectifier and inverter power.**

TABLE III. **Execution Times of cdsm-hvdc With TR+NE**

	2-Terminal			4-Terminal		
Level	CPU(s)	Hybrid(s)	Sp	CPU(s)	Hybrid(s)	Sp
5	189	252	0.8	387	511	0.8
51	2946	573	5.1	6053	690	8.8
101	6480	703	9.2	13862	754	18.4
201	13564	827	16.4	27864	904	30.8
401	31117	891	34.9	58400	1505	38.8

pure TR method leads to a higher simulation efficiency than the other scheme when CPU is the only processor, adopting GPU enables the introduction of NE to yield an additional 20% efficiency improvement, and thereby becomes the most efficient simulation method among the 4 schemes.

B. CDSM MMC-HVDC

Fig. 13 shows the CDSM-HVDC performance when the same DC line fault occurs as Fig. 11. Following the issue of converter block order 200 μs after the fault, the DC voltage instantly drops to 0, and there is no power exchange between AC and DC grids, indicating that CDSM-MMC has the capability to prevent the hazardous line fault from affecting the remaining system. A similar magnitude of power surge emerges as the transmission path discharges the energy into the fault location. Since the physics-based IGBT model is adopted in the proposed CPU/GPU co-simulation, the voltage waveform, as well as the power, has a larger ripple than that of the ideal switch case; nevertheless, their averages are identical.

Table III summarizes the execution times of the CPU/GPU platform over 1 million steps with proposed TR and NE schemes. The CPU computation time is approximately proportional to the number of DC terminals and the converter output voltage level, whilst the duration increases steadily for hybrid simulation until the voltage level reaches 401 in the 4-terminal case, or from a computational point of view when the total number of threads have a step from the 2 neighboring cases of 4800 to 9600 for the 5120-core GPU. Nevertheless, the speedup increases to around 40 albeit the comparisons are made within the same TR+NE scheme, not to mention the mere TR or original configuration, both of which fail to proceed with the simulation.

VI. CONCLUSION

This paper proposes a hierarchical multi-layer modeling architecture for heterogeneous CPU/GPU computing of electromagnetic transients of MMC-based HVDC systems. The device-level EMT model of a nonlinear high-order IGBT/FWD which yields an asymmetric admittance matrix is described in detail and is included in system-level studies. As the direct solution of a dramatic rise in number of IGBT/FWDs in the MMC posing a severe computational challenge, submodule separation which consequently reduced the scale of nonlinearity and distributes it to individual subsystems was adopted. The two-port network equivalence further improves computation performance noticing that a mere topological reconfiguration could not guarantee numerical calculation efficiency and convergence. The existence of numerous elements with identical attributes such as submodules enabled efficient kernel-based massively parallel processing on the GPU; meanwhile, CPU is superior in handling the remaining inhomogeneity. A proper boundary definition of the converter, including its controller, is critical to achieving high-performance computing with thorough utilization of the 2 processors' capacities and a minimum data exchange latency between them. Test results show that the proposed modeling and computing approach overcomes numerical divergence which accounts for an early simulation abortion in conventional circuit solvers, and device-level system study becomes feasible since the integration of CPU and GPU attains a remarkable speedup over pure CPU computation.

APPENDIX

TABLE IV. **DC System Parameters**

MMC 5-401, V_{dc}=±(4-200)kV, C_{SM}=3-15mF, L_A=50mH, L_{dc}=0-100mH; l_{len}=100km, l_l=0.05mH/km, r_l=0.01Ω/km, c_l=0.15μF/km
5SNA 2000K450300 StakPak IGBT
R_{th1-4}=1.601, 1.765, 0.358, 0.328K/kW, τ_{1-4}=0.581, 0.059, 0.006, 0.001sec;I_{NOM}=3000A, V_{th}=6V, A_{FET}=0.4, M_{FET}=1.2, N_{FET}=1.7, λ=10^{-3}, K=5, β=15, $I_{S_{BJT}}$=10^{-9}A, M_{BJT}=2, τ=50ns, M_D=1.21855, I_{Dsat}=10^{-6}A, $C_{0_{CG}}$=50nF, $C_{0_{CE}}$=20nF, R_C=R_E=R_A=1.16×10^{-5}, α=0.5, κ=2, ρ=10^{-4}, R_{Bk}=0.1mΩ

REFERENCES

[1] 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.

[2] M. Morati, D. Girod, F. Terrien, V. Peron, P. Poure, and S. Saadate, "Industrial 100-MVA EAF voltage flicker mitigation using VSC-based STATCOM with improved performance," *IEEE Trans. Power Del.*, vol. 31, no. 6, pp. 2494–2501, Dec. 2016.

[3] A. Raciti, D. Cristaldi, G. Greco, G. Vinci, and G. Bazzano, "Electrothermal PSpice modeling and simulation of power modules," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6260–6271, Oct. 2015.

[4] M. Turzynski and W. J. Kulesza, "A simplified behavioral MOSFET model based on parameters extraction for circuit simulations," *IEEE Trans. Power Electron.*, vol. 31, no. 4, pp. 3096–3105, Apr. 2016.

[5] D. Kumar and F. Zare, "Harmonic analysis of grid connected power electronic systems in low voltage distribution networks," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 4, no. 1, pp. 70–79, Mar. 2016.
[6] I. Zeltser and S. Ben-Yaakov, "On SPICE simulation of voltage-dependent capacitors," *IEEE Trans. Power Electron.*, vol. 33, no. 5, pp. 3703–3710, May 2018.
[7] Y. Jia, F. Xiao, Y. Duan, Y. Luo, B. Liu, and Y. Huang, "PSpice-COMSOL based 3D electro-thermal-mechanical modeling of IGBT power module," *IEEE J. Emerg. Sel. Topics Power Electron.*, doi: 10.1109/JESTPE.2019.2935037, 2019.
[8] J. Xu, C. Zhao, W. Liu, and C. Guo, "Accelerated model of modular multilevel converters in PSCAD/EMTDC," *IEEE Trans. Power Del.*, vol. 28, no. 1, pp. 129–136, Jan. 2013.
[9] A. Beddard, M. Barnes, and R. Preece, "Comparison of detailed modeling techniques for MMC employed on VSC-HVDC schemes," *IEEE Trans. Power Del.*, vol. 30, no. 2, pp. 579–589, Apr. 2015.
[10] L. Yang, Y. Li, Z. Li, P. Wang, S. Xu, and R. Gou, "A simplified analytical calculation model of average power loss for modular multilevel converter," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 2313–2322, Mar. 2019.
[11] C. Wong, "EMTP modeling of IGBT dynamic performance for power dissipation estimation," *IEEE Trans. Ind. Appl.*, vol. 33, no. 1, pp. 64–71, Jan. 1997.
[12] N. Lin and V. Dinavahi, "Dynamic electro-magnetic-thermal modeling of MMC-based DC–DC converter for real-time simulation of MTDC grid," *IEEE Trans. Power Del.*, vol. 33, no. 3, pp. 1337–1347, Jun. 2018.
[13] H. Bai, C. Liu, A. K. Rathore, D. Paire, and F. Gao, "An FPGA-based IGBT behavioral model with high transient resolution for real-time simulation of power electronic circuits," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6581–6591, Aug. 2019.
[14] C. Liu, R. Ma, H. Bai, Z. Li, F. Gechter, and F. Gao, "FPGA-based real-time simulation of high-power electronic system with nonlinear IGBT characteristics," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 7, no. 1, pp. 41–51, Mar. 2019.
[15] K. Sheng, B. W. Williams, and S. J. Finney, "A review of IGBT models," *IEEE Trans. Power Electron.*, vol. 15, no. 6, pp. 1250–1266, Nov. 2000.
[16] B. J. Baliga, "Analytical modeling of IGBTs: Challenges and solutions," *IEEE Trans. Electron Devices*, vol. 60, no. 2, pp. 535–543, Feb. 2013.
[17] 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.
[18] G. P. Adam and B. W. Williams, "Half- and full-bridge modular multilevel converter models for simulations of full-scale HVDC links and multiterminal DC grids," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 2, no. 4, pp. 1089–1108, Dec. 2014.
[19] E. Sánchez-Sánchez, E. Prieto-Araujo, A. Junyent-Ferré, and O. Gomis-Bellmunt, "Analysis of MMC energy-based control structures for VSC-HVDC links," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 6, no. 3, pp. 1065–1076, Sep. 2018.
[20] S. Debnath, "Parallel-in-time simulation algorithm for power electronics: MMC-HVdc system," *IEEE J. Emerg. Sel. Topics Power Electron.*, doi: 10.1109/JESTPE.2019.2947411, 2019.
[21] S. Yan, Z. Zhou, and V. Dinavahi, "Large-scale nonlinear device-level power electronic circuit simulation on massively parallel graphics processing architectures," *IEEE Trans. Power Electron.*, vol. 33, no. 6, pp. 4660–4678, Jun. 2018.
[22] N. Lin and V. Dinavahi, "Exact nonlinear mcromodeling for fine-grained parallel EMT simulation of MTDC grid interaction with wind farm," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6427–6436, Aug. 2019.
[23] H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," *IEEE Trans. Circuits Syst.*, vol. 3, no. 3, pp. 285–289, Sep. 1968.
[24] ANSYS, Inc., *Simplorer Components: Basic Elements*, Jan. 2020.
[25] H. Selhi, C. Christopoulos, A. F. Howe, and S. Y. R. Hui, "The application of transmission-line modelling to the simulation of an induction motor drive," *IEEE Trans. Energy Convers.*, vol. 11, no. 2, pp. 287–297, Jun. 1996.
[26] 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.
[27] NVIDIA Corp., *NVIDIA Tesla V100 GPU Architecture*, Aug. 2017.
[28] NVIDIA Corp., *CUDA C++ Programming Guide*, Aug. 2020.
[29] T. Musikka *et al.*, "Improvement of IGBT model characterization with experimental tests," in *Proc. 15th Conf. EPE*, Sep. 2013, pp. 1808–1813.
[30] ON Semiconductor, *Physically Based, Scalable SPICE Modeling Methodologies for Modern Power Electronic Devices*, Oct. 2019.
[31] [Online]. Available: https://www.pscad.com/about-us
[32] ABB Switzerland Ltd., *5SNA 2000K450300 StakPak IGBT Module*, Jan. 2018.

NING LIN (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Zhejiang University, 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 worked as an Engineer on FACTS and HVDC. His research interests include electromagnetic transient simulation, transient stability analysis, real-time hardware-in-the-loop emulation, integrated AC/DC grids, massively parallel processing, heterogeneous high-performance computing of power systems and power electronics.

RUIMIN ZHU (Member, IEEE) received the B.Eng. and Ph.D. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2012 and 2020 respectively. He is currently a Lecturer with the School of Electrical and Electronic Engineering, North China Electric Power University, Baoding, China. In 2018 and 2019, he was a Joint Ph.D. Student with RTX-LAB, University of Alberta, Edmonton, AB, Canada. His research interests include EMI analysis of power systems, power electronics, device-level modeling, and parallel simulation.

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