Article:
Zhang, Xiaotian, Hu, Yihua, Zhang, Jingwei et al. (3 more authors) (2024) IGBT Module DPT Efficiency Enhancement via Multimodal Fusion Networks and Graph Convolution Networks. IEEE Transactions on Industrial Electronics. pp. 1-12. ISSN: 0278-0046

https://doi.org/10.1109/TIE.2024.3368165

IGBT Module DPT Efficiency Enhancement Via Multimodal Fusion Networks and Graph Convolution Networks

Xiaotian Zhang, Yihua Hu, *Senior Member*, *IEEE,* Jingwei Zhang, Mohammad Nasr Esfahani, *Member*, *IEEE*, Tim Tilford, Stoyan Stoyanov, *Senior Member*, *IEEE*

***Abstract*—The dynamic electrical characteristics of insulated-gate bipolar transistors (IGBT) are of great significance in practical high power electrical applications and are usually evaluated through double pulse test (DPT). However, DPTs of IGBTs under various working conditions is time-consuming and laborious. Traditional estimation methods are based on detailed physical parameters and complex formula calculations, making deployment process challenging. This paper proposes a novel DPT efficiency enhancement method based on graph convolution network (GCN) and feature fusion technology, which can estimate and supplement switching transient waveforms of all working conditions. Thereby, dynamic electrical characteristics of the IGBT are obtained by estimated waveforms of DPT. This method proposes a multimodal attention fusion network (MAFN) to capture and fuse the features of switching transient waveforms between different positions thereby improving the expressive power and performance of the model. Moreover, this method is novel in that it is the first to utilise GCN to embed DPT data under multiple working conditions into a graph structure, which can use the graph structure information to fuse the features of spatially correlated working conditions data to obtain reliable estimation results. The method has verified to be effective and accurate on real dataset collected on two batches of IGBTs.**

***Index Terms*—IGBT, Double pulse test, Graph convolutional network, Feature fusion.**

I. INTRODUCTION

Insulated-gate bipolar transistors (IGBTs) are now widely used in a variety of medium-power or high-power power electronics (PE) converters [1], such as full-bridge inverters [2], [3] and boost converters [4]. The IGBT is one of the key components in the PE system, and its performance directly affects the efficiency, response time, switching frequency, power density, reliability and stability of the system. Therefore, design engineers, researchers and manufacturers of PE systems need to fully verify the dynamic behavior of the IGBT to be able to estimate its power loss to judge whether the corresponding requirements are satisfied. However, the parameters given in the specification are measured under specific conditions. The external parameters in practical applications are application-specific and often vary, so some of these parameters cannot be used directly. Any suboptimal choice or improper use can directly or indirectly cause the failure of the IGBT, resulting in serious consequences. To evaluate the behavior of the switch of the IGBT and observe parameters, the most effective method is the double pulse test (DPT).

The DPT method is the standard method for characterizing and evaluating the dynamic electrical performance of IGBTs, which is achieved through use of an inductive load and a power supply [5]. The inductor is used to replicate circuit conditions in a converter design. The power supply is used to provide voltage to the inductor. An arbitrary function generator is used to output pulses that triggers the gate of the IGBT and thus turns it on to start conduction of current. Key parameters describing the behavior of the switch of the IGBT include the turn on delay time t_{don}, the current rise time t_{ri}, the turn off delay time t_{doff}, the current fall time t_{fi}, the turn on loss $\mathrm{E_{on}}$ and the turn off loss $\mathrm{E_{off}}$ [6]. Due to some limitations of DPT in practical applications, several estimation methods have been proposed to analyse and evaluate these electrical characteristics of IGBT.

Three methods are commonly used to obtain the static and dynamic characteristics of IGBT: Circuit simulation [7], [8], behavioral models [9] and electrothermal models [10], [11] and [12]. Device-level circuit simulation is based on IGBT physical model such as those of Kraus et al [13], Hefner et al [14] and Wang et al [7]. For example, the simulation program with integrated circuit emphasis (SPICE) is a popular IGBT circuit simulation tool that can improve the accuracy of the simulation by using more physics knowledge of devices [15], [16]. However, these methods always require detailed physical information about the IGBT model and the circuit topology. In addition, these methods bring computational burden while pursuing higher accuracy. In the case of behavioral models, the features are neglection of the device physics to a certain extent, but the advantage is that the models are simpler and easy to set, and involve fast simulation times [17]. Recent Bai et al [9] proposed an FPGA-based IGBT behavioral model to simulate IGBT fast transients in PE circuits without the need for an iterative solution algorithm. However, this type of method cannot represent the switching transients without information of parasitic parameters and diode reverse recovery. Although electrothermal models can use electrical and thermal couplings to solve the heat-flow calculation and consider the temperature effect, adding multi-dimensional thermal fields and considering packaging characteristics in the simulated estimation process will greatly increase the complexity of the model.

For accurate estimation of dynamic characteristics of IGBTs, researchers typically directly perform curve fitting on DPT results under specific working conditions, and then estimate switching transient waveforms based on measurement results and data sheets. For example, the electrical transient model (ETM) has proposed to simulate static and dynamic behaviors of IGBT-diode switching cell in order to determine semiconductor losses by using specially developed algebraic equations [18], [19]. However, this type of method ignores the effect of temperature on the device and relies heavily on mathematical formulas and parameter extraction. Until recently, [5] added temperature as a parameter to the original ETM to determine the losses of the IGBT diode during the simulation. However, this method can still not overcome the dependence on many physical parameters. The characteristics of related methods

are summarised in Table I.

TABLE I
SUPERIORITY AND WEAKNESS OF IGBT DPT ESTIMATION METHODS

Type of methods	Related papers	Superiorities	Weaknesses
Circuit simulation	[7], [8]	1. Relatively high accuracy 2. Strong interpretability	1. Relying on physics information of devices 2. High computational difficulty
Behavioural model	[9]	1. Ignoring device physics 2. More convenient 3. Fast simulation speed	1. Relying on parasitic parameters and diode reverse recovery
Electrothermal model	[10], [11], [12]	1. Considering temperature effect	1. High complexity 2. High computing requirement
ETM	[18], [19]	1. Accuracy limitation 2. Easy access to parameters	1. Involving complex mathematical formulas
Deep learning-based	The proposed	1. Ignoring device physics 2. High accuracy 3. Transferable model for other devices 4. Fast getting estimation result	1. Requiring training data

Under various working conditions, DPTs of IGBTs are time-consuming and laborious task. Traditional estimation methods are based on detailed physical parameters and complex calculation approaches, making deployment difficult. With the development of artificial intelligence techniques, deep learning-based algorithms are increasingly used in industrial applications. As DPT waveforms are various time-series signals that have correlation with input parameters such as gate voltage, temperature, and load current, data-driven deep learning-based methods have the potential to predict these signals and thus to obtain the dynamic parameters of the IGBT module.

In this article, graph convolution network (GCN) with multimodal attention fusion network (MAFN), referred to as MAFGCN, is proposed to estimate IGBT DPT waveforms. A novel MAFN is first proposed to capture the relationship between different positions in input signals, thereby improving the expression ability of the fusion feature. The GCN-based module is used to estimate switching transient waveforms of the IGBT diode and obtain the dynamic behavior of the switching unit under various working conditions. The highlight of this method is that the DPT results for all operating conditions can be quickly estimated, and the model can be generalised to other types of IGBTs using only DPT data at room temperature. The proposed method can greatly reduce the number of required DPTs, saving labour costs and time. The complete algorithm and model are implemented in Python and validated with experimental data obtained from a DPT rig.

There are three technical contributions in this paper:

1) This paper proposes a GCN-based approach to enhance IGBT DPT efficiency and fast evaluation of IGBT dynamic characteristic through estimated DPT results, which embeds DPT data under each working conditions into spatial graph structure.
2) MAFN is proposed to provide strong impressive features to GCN module so that graph convolutional layers can capture core feature of DPT waveforms successfully.
3) For generalizing the well-trained model to other IGBT modules, the graph transfer strategy is introduced.

The remaining sections of the paper are structured as follows. Section II presents the applicability analysis of DPT data embedded in a graph structure. Section III describes the proposed method. Section IV introduces the designed experimental platform, the test measurements and the data gathering. Experiment validations and results are reported in Section V. Finally, Section VI concludes this paper.

II. APPLICABILITY ANALYSIS OF GRAPH STRUCTURE EMBEDDING OF DPT DATA

A. Designed DPT Procedure

In the DPT experiment, the upper device is the freewheeling diode (FWD), the lower-side IGBT is the device under test (DUT). The diagram of the applied DPT circuit is shown in Fig. 1. A DPT is a tool that enables a power switch to be turned on and off at different current levels. By adjusting the switching times S_1, S_2 and S_3, the turn-on and turn-off waveforms of DUT can be controlled and measured over the full range of operating conditions. For the second pulse, it is important to build up current in the complementary device or diode so that when the switch turns on, the effects of any reverse recovery current can be evaluated. A load inductor limits the rate of di/dt.

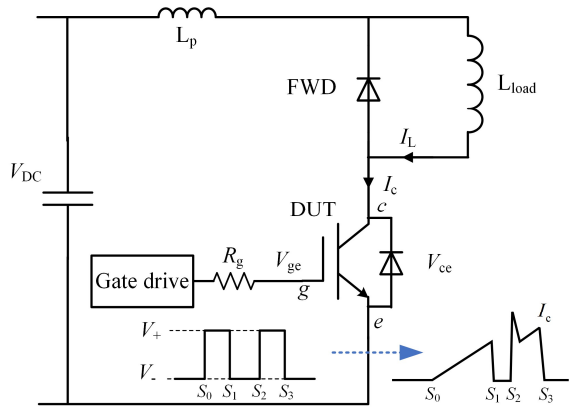


Fig. 1. The circuit diagram of DPT for deploying the experimental platform.

Two consecutive pulses generated by the pulse function generator are transmitted to the DUT. By adjusting the first pulse duration, the switching transient can be captured at the desired current level at the end of the first pulse and at the beginning of the second pulse. V_{DC} is the dc-link voltage, L_{load} is the load inductance, V_{ge} is the gate voltage, V_{ce} is the collector voltage, I_c is the collector current, and I_L is the load current. In the gate drive, the positive voltage V_+ = 15V, the negative voltage V_- = -10V, the gate resistance is R_g = 1.5Ω, 2.5Ω or 3.5Ω. Considering the influence of the parasitic parameters of the commutation circuit on the characteristics of the IGBT switch, the parasitic inductance was changed by changing the length of the copper bus. The two types of the parasitic inductance in our experiment $L_p \approx$ 120nH and 160nH.

During S_0 to S_1, the first pulse is applied to IGBT and the IGBT is turned on. A constant voltage V_{DC} is added to the load inductance L_{load} which makes the current through it to increase in linearly:

$$I_c = \frac{V_{DC} \cdot S}{L_{load}} \quad (1)$$

According to Eq. (1), I_L which is also I_c of the IGBT, depends on V_{DC}, L_{load} and pulse duration time S. It is feasible to set the current by the method of controlling S with fixed V_{DC} and L_{load}.

During S_1 to S_2, the IGBT is turned off, but the current flowing through L_{load} cannot be suddenly changed. So, the current will commutate to FWD of counterpart arm. The duration time from S_2 to S_3 of turning-off is not of importance during test condition. Therefore, in this experiment, we set it between 10μs and 20μs to meet the experimental requirements. After the reverse recovery process, the current flows into the IGBT through L_{load}, similarly to period S_0 to S_1. At S_3, the IGBT will be turned off again. Based on the loss P, the switching loss E_{on} in turn-on transient and E_{off} in turn-off transient are calculated by oscilloscope as:

$$E_{on/off} = P(t_2) - P(t_1) = \int_{t_1}^{t_2} V_{ce} I_c dt \tag{2}$$

where t_1 and t_2 represent the starting and ending moments of the turn-on or turn-off transient respectively. From switching transient waveforms generated by DPT, we can also get information relating to the IGBT's turn-on/turn-off delay ($t_{d(on)}/t_{d(off)}$), rise time (t_r), fall time (t_f), t_{on}/t_{off} (turn-on/turn-off time), dv/dt, and di/dt, etc.

B. External Influencing and Limitation Factors

When using the data-driven method, we hypothesise that the changes in the data are only related to the working conditions. In this case, the model obtained after training is more accurate and general. In this part, the main external influencing factors of IGBT DPT switching transient waveforms are discussed.

The first influencing factor is the measurement error. In fact, this is unavoidable in experiments due to the limitation of equipment precision, uncertainty and random error, etc.

Secondly, using different DC busbars (V_{DC}) also affects the switching transient voltage waveform. Different structures of the DC bus can lead to changes in the parasitic inductance L_S. At a high current change rate di/dt, the induced potential of L_S is superimposed on the IGBT, which will affect the waveform of V_{ce}:

$$V_{ce} = -L_S \frac{di}{dt} + V_{DC} \tag{3}$$

That is manifested that high L_S at the IGBT turn-on transient will result in low E_{on}, and the turn-on time will be extended; too high L_S at the turn-off transient will cause V_{ce} to overshoot the rated voltage of the IGBT, cause overvoltage breakdown fault, and reduce the tail current.

Furthermore, different FWD of the circuit will affect turn-on transient overcurrent peak I_{peak} and E_{on}. During the turn-on transient of the IGBT, the reverse recovery characteristic (RRC) of FWD can lead to I_{peak} in the collector current, which affects the turn-on rate and turn-on loss. In addition to being affected by di/dt, RRC is also related to the stored charge Q_{rr} of the diode. I_{peak} can be expressed as

$$I_{peak} = I_L + \sqrt{\frac{2\frac{di}{dt}Q_{rr}}{SF}} \tag{4}$$

where SF is the reverse recovery softness factor of FWD [20]. Q_{rr} is determined by the forward conduction current I_L before reverse recovery and the remaining carrier lifetime of the diode itself τ_H:

$$Q_{rr} \approx I_L \tau_H \tag{5}$$

Under a fixed I_L, different Q_{rr} of the diode will affect I_{peak}, resulting in a change in E_{on}.

However, we should notice that same users or same manufacturers usually tend to use the same DC busbars and FWD. In that cases, external influencing factors have little impact on the stability and accuracy of the data-driven method proposed in this paper. In other words, the proposed model is better personalised for a specific practical application from the training step. In order to comprehensively prove this method, this paper still considers different R_g in the gate driver and different DC link L_p.

C. Graph Structure Embedding for DPT Estimation Problem

This paper proposes to embed DPT data for all working conditions in a labeled graph of GCN. GCN belongs to the category of deep learning, which learns and trains based on graph structure data. GCNs enable prediction and inference tasks at both the node level and the graph level by effectively capturing the relationship between the topological structure of graph data and node features [21].

It can be seen from the experiment that the DPT transient waveforms under various working conditions are correlated, and the degree of correlation is different. We propose to treat each working condition of DPT as a node of the graph structure. The DPT switching transient waveforms of unmeasured nodes could be accurately estimated through the correlation between each node, so as to obtain the complete IGBT dynamic electrical characteristics. Fig. 2 demonstrates the principle and strategy of graph structure embedding for DPT estimation problem. It should be noted that if the difference in DC busbar is not considered, the graph structure will become simpler.

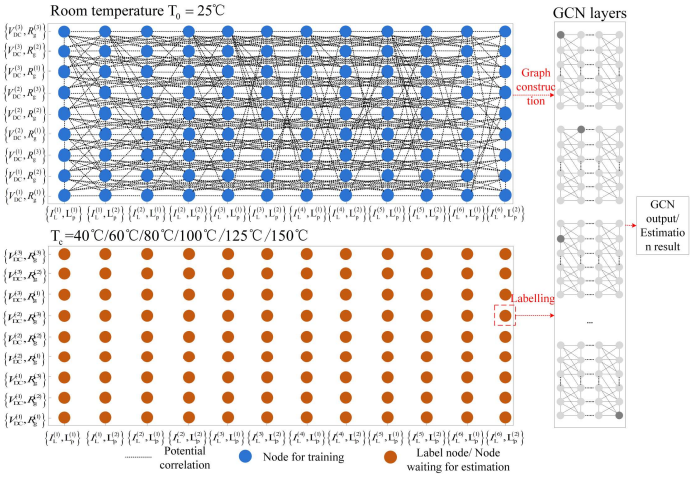


Fig. 2. The principle of graph structure embedding for DPT estimation problem.

In DPT data measurement, $V_{DC} = \left\{V_{DC}^{(1)}, V_{DC}^{(2)}, V_{DC}^{(3)}\right\}$, $I_L = \left\{I_L^{(1)}, I_L^{(2)}, I_L^{(3)}, I_L^{(4)}, I_L^{(5)}, I_L^{(6)}\right\}$, $R_g = \left\{R_g^{(1)}, R_g^{(2)}, R_g^{(3)}\right\}$, and $L_p = \left\{L_p^{(1)}, L_p^{(2)}\right\}$. Therefore, total measured working conditions $C_w = \left\{\left(V_{DC}^{(i)}, I_L^{(j)}, R_g^{(p)}, L_p^{(q)} \mid 1 \le i \le 3, 1 \le j \le 6, 1 \le p \le 3, 1 \le q \le 2\right)\right\}$, in which $i, j, p, q \in \mathbb{N}^+$. As shown in Fig. 2, the number C_w is 108. The case temperature $T_c = \left\{T_c^{(0)}, T_c^{(1)}, T_c^{(2)}, T_c^{(3)}, T_c^{(4)}, T_c^{(5)}, T_c^{(6)}\right\}$. The DPT dataset measured under $T_c^{(0)}$ is the training data, and the DPT dataset measured under $T_c = 40℃, 60℃, 80℃, 100℃, 125℃, 150℃$ are labels. That means the well-trained model can estimate DPT waveforms under any T_c.

III. The Proposed Methodology

The proposed method MAFGCN consists of three main parts: graph construction module, MAFN block, GCN-based module and optional graph transfer module: 1) In order to capture more comprehensive DPT waveform information and enhance the

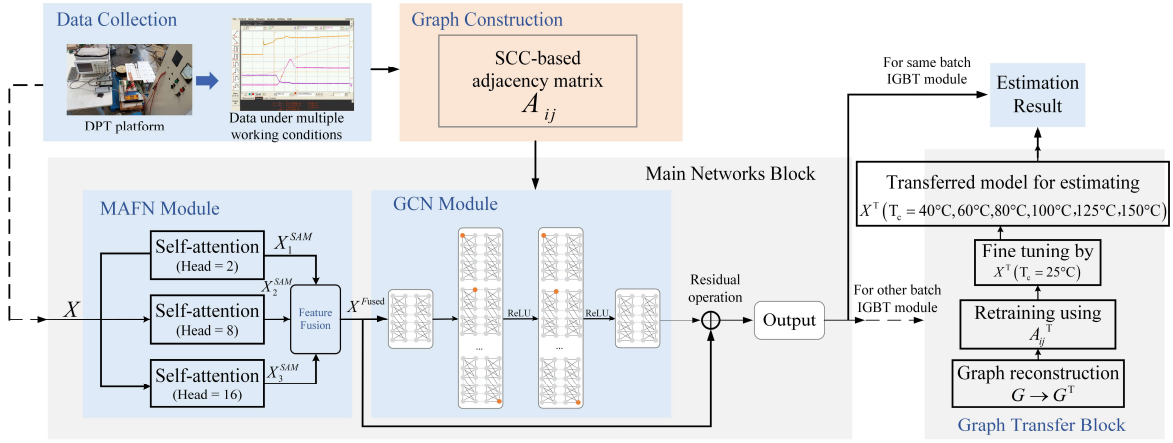


Fig. 3. Structure of the proposed (MAFGCN) method (the GCMCN method includes a MAFN module, a graph construction block, a GCN-based spatial module and an optional graph transfer block, in which a residual operation between the input and output of the of GCN module for improving the effect of feature fusion and preventing the gradient disappearance problem during training).

overall performance of the model, the input data needs to be input to the proposed MAFN for feature extraction and fusion. The overall input data is denoted by X; 2) SCC-based graph construction block conveys an adjacency matrix A_{ij} to GCN module; 3) Spatial GCN-based module is trained by output of the MAFN and A_{ij} to integrate the information of graph nodes to provide spatial dependence for the DPT waveform estimation; 4) The graph transfer block can generalise the model so that it can be used to predict DPT results of other batch IGBT modules and even other brand IGBT modules. Fig. 3 illustrates the complete architecture of the proposed method.

A. Data Input and MAFN Block

Appropriate input variables and their input forms should be selected from the data acquisition module to ensure the accuracy and reliability of DPT waveform estimation method. The total input data collected by this method from the DPT system is:

$$X = \left\{X_{V_{ge}}, X_{V_{ce}}, X_{I_c}, X_P\right\} \tag{6}$$

These input variables are the core waveforms in DPT and the final estimated target waveforms. For each training of the model, one type of data in X should be as input data to the next step. Dataset X contains a total of 4 variables, which means that the model needs to be completely trained 4 times to obtain a complete DPT estimation result.

Attention mechanism can solve the deficiencies that RNN cannot perform parallel computing and CNN cannot capture the long-distance relationships in sequence [22]. The designed MAFN as a multimodal feature fuser consists of three self-attention modules with different number of attention head as shown in Fig. 4. Fewer attention heads provide more stable and consistent feature representations, while more attention heads can capture more fine-grained relations and patterns of the input sequence. Selected dataset will firstly be input into MAFN module to get most expressive feature representations as the input of GCN module. The details are presented in the following.

Self-attention (SA) as a special attention mechanism. This module uses multi-head self-attention (MSA) mechanism network. Input X needs to go through positional encoding process. For the hth head, the weights of SA can be calculated by

$$\mathrm{SA}(Q_h, K_h, V_h) = \mathrm{softmax}\left(\frac{Q_h {K_h}^{\mathrm{T}}}{\sqrt{d_k}}\right) V_h \tag{7}$$

where $Q_h = XW_q^h$, $K_h = XW_k^h$, $V_h = XW_v^h$. W_q^h, W_k^h and W_v^h are linear transformation matrix of Q_h, K_h and V_h, respectively. Weights of SA of each head are weighted and spliced to obtain the representation of MSA:

$$\mathrm{MSA}(Q, K, V) = \mathrm{Concat}(head_1, \cdots, head_h) W^{\mathrm{O}} \tag{8}$$

$$head_h = \mathrm{SA}(Q_h, K_h, V_h) \tag{9}$$

where W^{O} donates the output weight matrix of MSA. The output of MSA (X_i^{MSA}) needs to go through a feed-forward network (FFN) as the final representation X^{SAM}. FFN includes two fully-connected (FC) layers:

$$\mathrm{FFN} = \mathrm{ReLU}(W^{\mathrm{FFN}} X_i^{\mathrm{MSA}} + b) \tag{10}$$

where $\mathrm{ReLU}(\cdot)$ is the activation function, W^{FFN} is weight matrix, b is the bias. Outputs of three self-attention modules X_i^{SAM} should be fused following strategy:

$$X^{\mathrm{Fused}} = \sum_{i=1}^{3} w_i^{\mathrm{Fusion}} X_i^{\mathrm{SAM}} \tag{11}$$

where w_i^{Fusion} donates the weight of the ith self-attention module. X^{Fused}, as the most expressive feature representation of DPT sequence, will be the input of the GCN-based module.

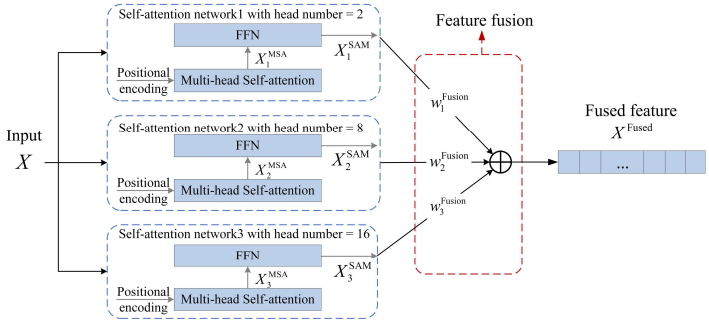


Fig. 4. Detailed structure of the MAFN block.

B. Graph Construction Module

As described in Section II.*C*, the whole DPT dataset can be embedded into a graph structure with multiple nodes. In this paper, this method applies the Spearman correlation coefficient technique to calculate and visualise the spatial correlation between subdatasets under each working condition, which is constrained by linear relationships so as to does not lose nonlinearity information.

According to the training data, we should firstly calculate the SCC between each two nodes. Given two input sequences $D_1 = (x_1, x_2, \cdots, X_{l_s})$ and $D_2 = (y_1, y_2, \cdots, y_{l_s})$, the Spearman correlation coefficient ρ can be calculated by

$$\rho = 1 - \frac{6\sum_{i=1}^{l_s} d_i^2}{l_s(l_s^2 - 1)} \tag{12}$$

where l_s is the length of the input sequence, $d_i = R_x - R_y$ is the difference between ranks of variables D_1 and D_2. The calculation principle of R_x and R_y can refer to [23], which will not be explained in detail here.

This method uses ρ as the similarity index between nodes to build an adjacency matrix. In GCN, the graph can be referred as $G = (P, E)$, in which $P = \{p_1, p_2, p_3, \cdots, p_N\}$ is the set of N nodes and E is the edge set. In the set E, element $e_{ij} \in E$ is the edge between node p_i and node p_j. To presents the correlations between nodes, the adjacency matrix can be identified as $A \subset \mathbb{R}^{N\times N}$, and elements A_{ij} in A can be computed by

$$A_{ij} = \begin{cases} w_{ij}, & if(p_i, p_j) \in E \\ 0, & otherwise \end{cases} \tag{13}$$

where $w_{ij} \in [0,1]$ is the degree of relevance between node p_i and p_j. In this method, $w_{ij} = \rho_{ij}$. Due to $\rho_{ij} \in [-1,1]$, further normalization process is not required here.

To reduce the number of edges in the labeled graph and reduce computational cost, we preserve strong correlations between nodes and eliminate weak correlations. Therefore, A_{ij} can be updated as

$$A_{ij} = \begin{cases} A_{ij}, & if\, A_{ij} < \eta \\ 0, & otherwise \end{cases} \tag{14}$$

where η is a preset threshold so that any A_{ij} less than η will be reset to 0. If the device has sufficient computing power, this step can be ignored.

C. Spatial GCN-based Estimation Module

Instead of applying regular convolutional and recurrent networks, this method formulates the problem on graphs and build the model with complete convolutional structures, which enable much faster training speed with fewer parameters [24].

After embedding nodes according to the strategy in Section II.*C* and building an adjacency matrix [25], GCN is used to integrate the information of neighbor nodes into the target node to provide spatial dependence for the DPT waveform estimation task. The calculation process of graph convolution is:

$$Z^{(k+1)} = \mathrm{ReLU}\left(L^{\mathrm{norm}} Z^{(k)} W^{(k)}\right) \tag{15}$$

$$L^{\mathrm{norm}} = D^{-\frac{1}{2}} \tilde{A} D^{-\frac{1}{2}} \tag{16}$$

$$\tilde{A} = A + I_N \tag{17}$$

where $Z^{(k)}$ is the input of the $(k+1)$th graph convolution layer, $W^{(k)}$ is the layer-specific trainable weight matrix of the $(k+1)$th layer, $\mathrm{ReLU}(\cdot)$ is the activation function, L^{norm} is the normalised Laplace matrix, D is the degree matrix of node, I_N is the identity matrix. In this paper, the number of GCN layer is set as 2. The output layer is a FNN layer. The mean squared error (MSE) function is used as loss function $\mathcal{L}$ for network training. The structure of the spatial GCN module is shown in Fig. 3.

D. Graph Transfer Block

For generalizing the model to estimate other IGBTs and less data (only DPT data at room temperature $T_c^{(0)}$), the transfer learning strategy are proposed. After our experiments, to achieve the best transfer effect, the proposed graph transfer method is advocated to be used on different IGBTs of the same batch, or on IGBTs of different batches with similar rated parameters (such as IGBTs in Table II). There are three main steps of this block: graph reconstruction, pretraining and fine tuning. Fig. 5 shows the detailed flowchart of this graph transfer block.

In the graph reconstruction step, the graph structure of the target IGBT $G^{\mathrm{T}} = (P^{\mathrm{T}}, E^{\mathrm{T}})$ should be redefined, in which $P^{\mathrm{T}} = \{p_1^{\mathrm{T}}, p_2^{\mathrm{T}}, p_3^{\mathrm{T}}, \cdots, p_N^{\mathrm{T}}\}$, E^{T} is the edge set. The adjacency matrix A_{ij}^{T} should be recalculated according to G^{T}:

$$A_{ij}^{\mathrm{T}} = \begin{cases} w_{ij}^{\mathrm{T}}, & if\left(p_i^{\mathrm{T}}, p_j^{\mathrm{T}}\right) \in E^{\mathrm{T}} \\ 0, & otherwise \end{cases} \tag{18}$$

where $w_{ij}^{\mathrm{T}} = \rho_{ij}^{\mathrm{T}}$ is degree parameters. To obtain a pre-trained model with G^{T}, the whole should be retrained using X. Then the retrained model is fine-tuned by the DPT dataset of the target IGBT $X^{\mathrm{T}} = \left\{X_{V_{\mathrm{ge}}}^{\mathrm{T}}, X_{V_{\mathrm{ce}}}^{\mathrm{T}}, X_{I_c}^{\mathrm{T}}, X_P^{\mathrm{T}}\right\}$. We should note it is allowed if the amount of data in $X^T \ll X$ because X^{T}does not need DPT data under $T_c^{(1)}, T_c^{(2)}, T_c^{(3)}, T_c^{(4)}, T_c^{(5)}, T_c^{(6)}$, only need data under $T_c^{(0)}$.

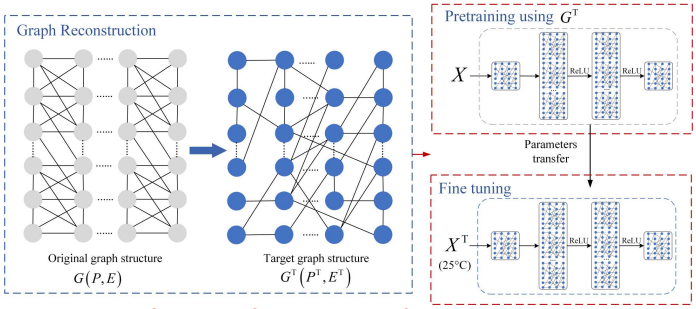


Fig. 5. Detailed flowchart of the graph transfer learning block.

E. Implement Procedure of MAFGCN

The implementation process of the proposed MAFGCN is divided into training stage and verification stage. The entire algorithm updates parameters with each iteration. The pseudocode of the whole procedure of MAFGCN is shown in Algorithm 1.

Algorithm 1: The proposed MAFGCN with multimodal feature fusion for DPT estimation.

Input: The DPT dataset $X = \left\{X_{V_{\mathrm{ge}}}, X_{V_{\mathrm{ce}}}, X_{I_c}, X_P\right\}$ obtained from real DPT experiment that is divided into training set X_{train}, validating set X_{val} and testing set X_{test}. The DPT dataset from similar IGBT module $X^{\mathrm{T}} = \left\{X_{V_{\mathrm{ge}}}^{\mathrm{T}}, X_{V_{\mathrm{ce}}}^{\mathrm{T}}, X_{I_c}^{\mathrm{T}}, X_P^{\mathrm{T}}\right\}$.

Output: The estimated DPT results at X_{est} and $X_{\mathrm{est}}^{\mathrm{T}}$ and estimation accuracy of testing data.

In the training and validating stages of MAFGCN:

Collect dataset from real IGBT modules and obtain X and X^{T}. Initialise parameters of the whole MAFGCN;
Calculate A_{ij} (Eq. (13)) through Eq. (12);
Compute and obtain updated A_{ij} through Eq. (14);

Training the MAFN module:

Input dataset X_{train} into three MSA networks to extract features X_i^{MSA};
Forward propagation through three MSA networks with parameters Q_h, K_h, V_h and W^{O};
Calculate X_i^{SAM} by using Eq. (10);
Forward propagation through FFNs with parameters W^{FFN} and b;
Calculate X^{Fused} by using Eq. (11);

Training the GCN module with residual operation:

Input MAFN result X^{Fused} into GCN;
Calculate GCN output by using Eq. (15)-(17);
Forward propagation through $Z^{(k)}$ with parameters $W^{(k)}$;

Compute the loss function $\mathcal{L}$ and update parameters of the entire model;
Backpropagation with RMSprop optimiser to minimise $\mathcal{L}$;
Update the parameters of the entire MAFGCN model;
Compute the performance indexes of the model on X_{val};

In the graph transfer block:

Calculate A_{ij}^{T} (Eq. (18)) through Eq. (12);
Retraining MAFGCN with the input $X^{\mathrm{T}}(T_c^{(0)})$;

14 Fine-tune by the dataset $X^{T}(T_c^{(0)})$ and update the parameters of the entire MAFGCN model;

In the testing stage:

15 Evaluate the model with testing data X_{test} and X^{T} under $T_c^{(1)}, T_c^{(2)}, T_c^{(3)}, T_c^{(4)}$, and compute accuracy of estimated DPT results.

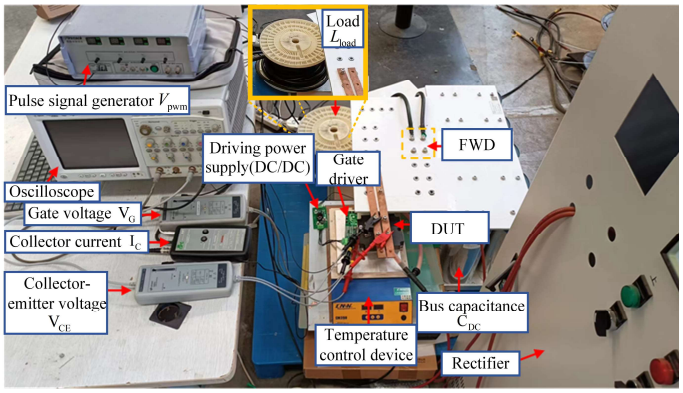


Fig. 6. Setup of the DPT measurement platform.

IV. Setup of DPT Platform and Data Acquisition

This section will provide information about the setup of the DPT platform. The experimental and validation datasets are collected from this described DPT bench.

In order to collect accurate and reliable data, we have carried a series of real DPT experiments. Based on the circuit in Fig. 1 in Section II.*A*, a DPT platform was designed and implemented for characterizing IGBTs and diodes. As shown in Fig. 6, the DPT measurement device is composed of a high voltage power supply, capacitor bank, laminated DC busbar, load inductances, DPT and drive circuit, measuring devices (voltage probe and current probe), oscilloscope and pulse generator.

TABLE II
PARAMETERS OF IGBT MODULES

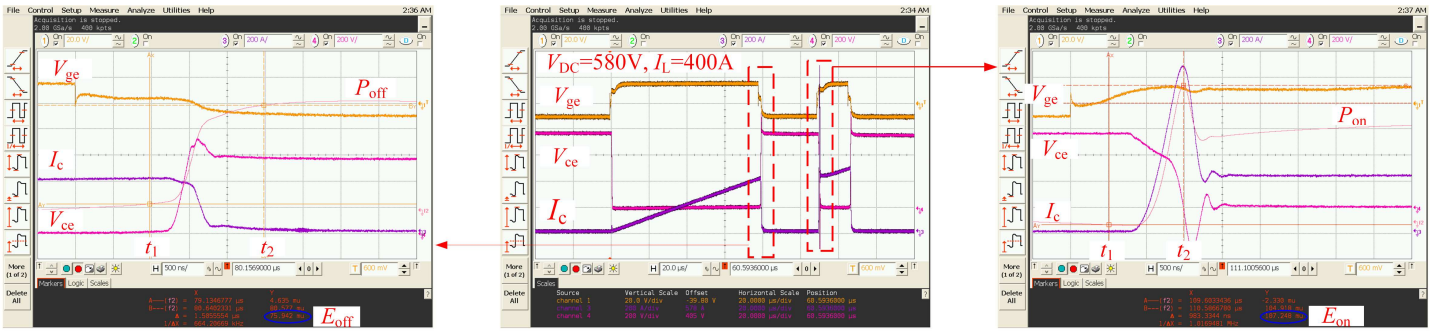
Batch	DIM1200FSS12-A000	DIM1200FSS12-A076
V_{CES}	1.2kV	1.2kV
$V_{CE(sat)}$ (typ)	2.2V	2.2V
I_C (max)	1200A	1200A
$I_{C(PK)}$ (max)	2400A	2400A

The main DUTs are DIM1200FSS12-A000 and DIM1200FSS12-A076 from Dynex semiconductor. Both IGBTs are single switch 1.2kV, *n*-channel enhancement mode, IGBT module. This IGBT has a wide reverse bias safe operating area (RBSOA) plus 10μs short circuit withstand. The main applications of these IGBTs are high power inverter and motor controller. The parameters of the tested IGBTs are summarised in Table II.

The variables of test conditions are T_c, V_{DC}, I_L, R_g, L_p and turn-on/turn-off transient:

1) T_c: 25°C, 40°C, 60°C, 80°C, 100°C, 125 °C and 150 °C;
2) V_{DC}: 200V, 400V and 600V;
3) I_L: 50A, 200A, 400A, 600A, 800A and 1000A;
4) R_g: 1.5Ω, 2.5Ω and 3.5Ω.
5) L_p: 120nH and 160nH.
6) Transient condition: turn-on, turn-off.

where the number of total test conditions is 180. Each IGBT was subjected to 10 double-pulse tests for each test condition. Measurement data of each batch includes 1800 groups, and a total of 4800 groups of data are obtained. Test time *t*, V_{ge}, V_{ce}, I_c, P_{on} and P_{off} are included in each group data. We apply four channels oscilloscope to observe and output waveforms of V_{ge}, V_{ce}, I_c, P_{on} and P_{off}. Waveforms of channels 1,3 and 4 are respectively represented as V_{ge}, I_c and V_{ce} as shown in Fig. 7. There are examples of collected waveforms on oscilloscope with the working condition of T_c=25°C, V_{DC}=580V, I_c =400A, R_g=2.5Ω, L_p=160nH.



(a) T_c=25 °C, V_{DC}=580V, I_c =400A and in turn-off transient (b) The complete DPT waveforms at T_c=25 °C, V_{DC}=580V, I_c =400A (c) T_c=25 °C, V_{DC}=580V, I_c =400A and in turn-on transient

Fig. 7. Typical example waveforms of the measured and collected DPT result on oscilloscope at working condition T_c=25 °C, V_{DC}=580V, I_c =400A, R_g=2.5Ω, L_p=160nH ((a) in turn-off transient. (b) the complete DPT waveforms (c) in turn-on transient).

V. Experimental Result

This section describes the experimental validations based on the proposed DPT estimation approach in Section III and the collected dataset in Section IV.

A. Performance Criteria

The proposed method is evaluated using three quantitative indexes: mean absolute error (MAE), root mean square error (RMSE) and Logcosh error (LE), which are represented by Eq. (19), Eq. (20), and Eq. (21), respectively. These evaluation metrics provide a clear indication of the accuracy of the DPT waveform estimation task. The MSE and MAE measures assess the level of deviation between the estimated waveform and the ground truth, while the CS captures the similarity between the two waveforms in terms of their shape and direction. Together, these indexes provide a comprehensive assessment of the accuracy and fidelity of the proposed method in estimating the DPT waveform.

$$\mathrm{MAE} = \frac{1}{L_s}\sum_{n=1}^{L_s}|y^n - \hat{y}^n| \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{L_s}\sum_{n=1}^{L_s}(y^n - \hat{y}^n)^2} \tag{20}$$

$$\mathrm{LE} = \frac{1}{L_s}\sum_{n=1}^{L_s}\left(\frac{e^{\hat{y}^n - y^n} + e^{-(\hat{y}^n - y^n)}}{2}\right) \tag{21}$$

where y^n and $\hat{y}^n$ are the true value and estimated value of the DPT waveform amplitude, n is the n^{th} sample point, and L_s is the length

of DPT signals.

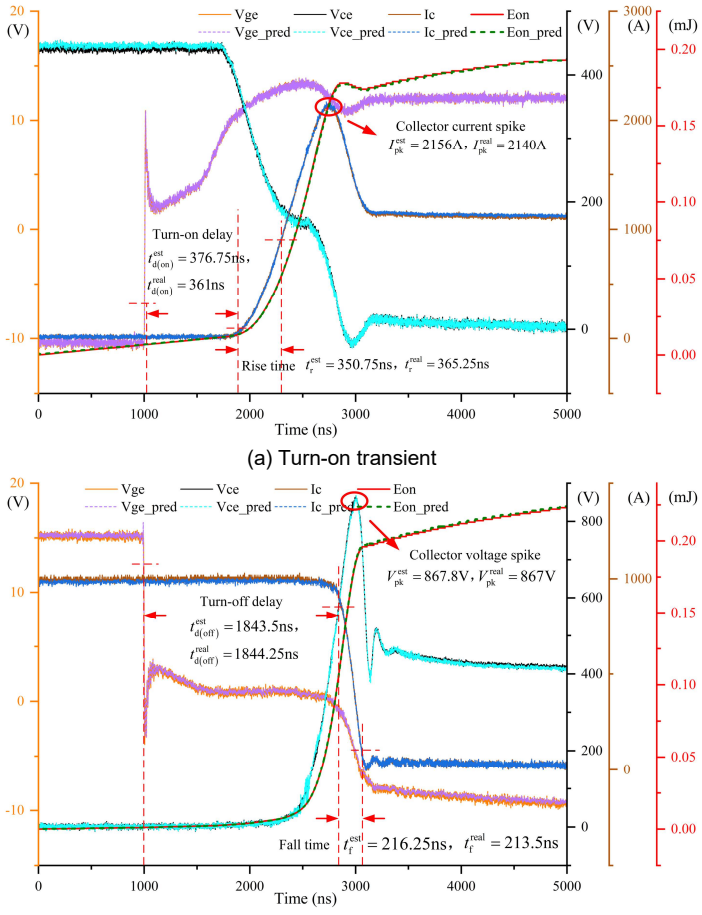
B. Experiments Settings and Result

Firstly, we should number the data collected from the two types of IGBTs in Section IV. The 1800 groups collected from DIM1200FSS12-A000 are referred as Dataset A; the 1800 groups collected from DIM1200FSS12-A076 are referred as Dataset B. This method can estimate DPT results under 1296 different working conditions (except 216 kinds under 25℃). The use condition of these datasets is listed in Table III.

TABLE III
THE USE OF TWO MEASURED DPT DATASETS

Dataset	Data division	Detailed use
A	90% in T_c = 25℃, 40℃, 60℃, 80℃, 100℃, 125℃ and 150℃	Training and validating data
	10% in T_c = 40℃, 60℃, 80℃, 100℃, 125℃ and 150℃	Testing data
B	100% in T_c = 25 ℃	Graph transfer data
	10% in T_c = 40℃, 60℃, 80℃, 100℃, 125℃ and 150℃	Testing data

Due to limited space, turn-on and turn-off transient detailed experimental result under one working condition (T_c=100℃, V_{DC}=400V, I_c=1000A, R_g=2.5Ω, L_p=160nH) in Dataset A and one (T_c=80℃, V_{DC}=400V, I_c=600A, R_g=2.5Ω, L_p=160nH) in Dataset B are illustrated in Fig, 8 and Fig, 9 respectively. Batch size is set as 64 and training epoch is 100. The RMSprop optimiser is chosen with initial learning rate 1×10^{-5}. All data are normalised before model training.



(a) Turn-on transient

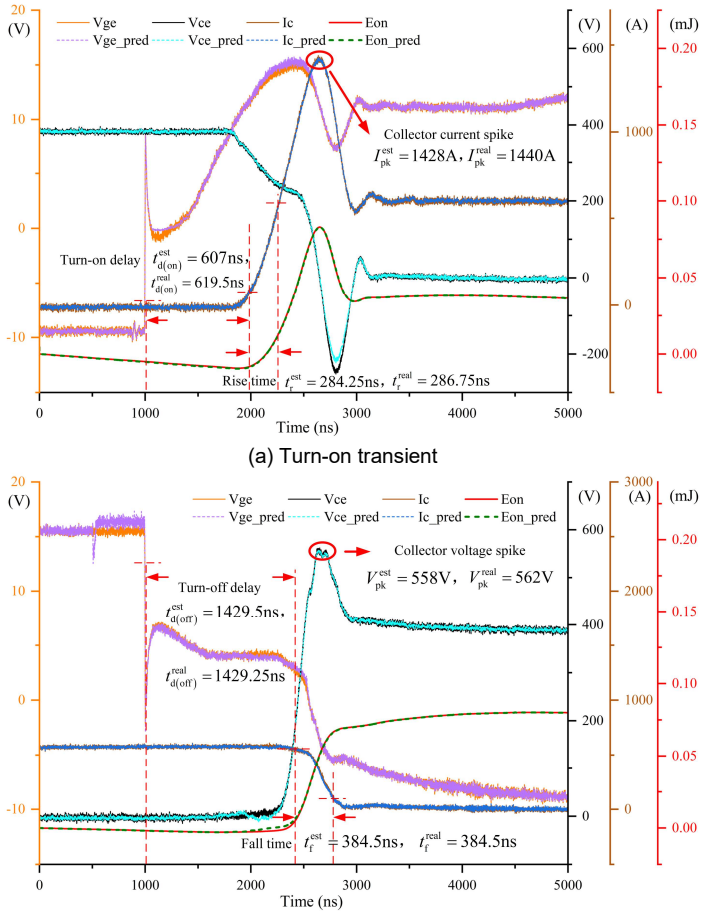
(b) Turn-off transient

Fig. 8. Turn-on and turn-off transient waveforms of IGBT DIM1200FSS12-A000 DPT estimating results at the working condition T_c=100℃, V_{DC}=400V, I_c=1000A, R_g=2.5Ω, L_p=160nH ((a) turn-on transient, (b) turn-off transient).

Fig. 8(a) and (b) demonstrates the performance of the proposed method on turn-on transient and turn-off transient, respectively. It can be found that when using the data test of DIM1200FSS12-A000, the DPT waveform estimation is highly accurate. As demonstrated in Fig. 8(a), the real turn-on delay $t_{d(on)}^{real} = 376.75$ns and the estimated value $t_{d(on)}^{est} = 361$ns, the real rise time $t_r^{real} = 365.25$ns and the estimated value $t_r^{est} = 350.75$ns, the real value of collector current spike $I_{pk}^{real} = 2140$A and the estimated value $I_{pk}^{est} = 2156$ A. In Fig. 8(b), the real turn-off delay $t_{d(off)}^{real} = 1844.25$ns and the estimated value $t_{d(off)}^{est} = 1843.5$ns, the real fall time $t_f^{real} = 213.5$ns and the estimated value $t_f^{est} = 216.25$ns, the real value of collector voltage spike $V_{pk}^{real} = 867$ V and the estimated value $V_{pk}^{est} = 867.8$V.

C. Graph Transfer Estimation Result

After the graph transfer operation in Section III.*D*, we performed a complete DPT result estimation verification for the IGBT batch DIM1200FSS12-A076. At the same time, the dynamic parameters of this IGBT module are also calculated using real measurement data and estimated data.



(a) Turn-on transient

(b) Turn-off transient

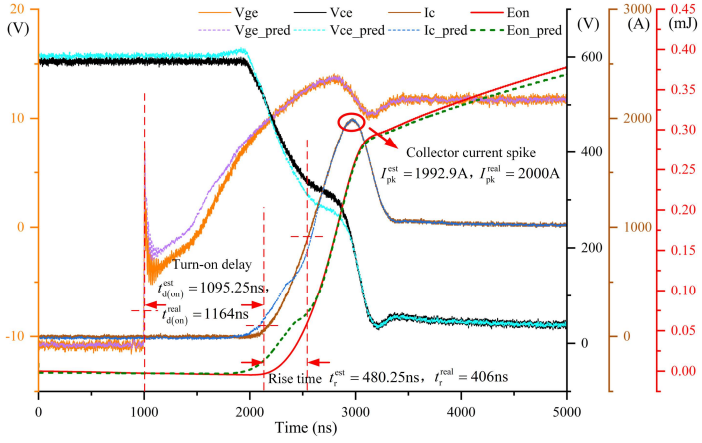
Fig. 9. Turn-on and turn-off transient waveforms of IGBT DIM1200FSS12-A076 DPT estimating results at the working condition T_c=80℃, V_{DC}=400V, I_c=600A, R_g=2.5Ω, L_p=160nH ((a) turn-on transient, (b) turn-off transient).

As shown in Fig. 9, the transferred model can correctly estimate the DPT results of DIM1200FSS12-A076. In Fig. 9(a), $t_{d(on)}^{real} = 619.5$ ns and $t_{d(on)}^{est} = 607$ ns, $t_r^{real} = 286.75$ ns and $t_r^{est} = 284.25$ ns, $I_{pk}^{real} = 1440$ A and $I_{pk}^{est} = 1428$ A. In Fig. 9(b), $t_{d(off)}^{real} = 1429.25$ns and $t_{d(off)}^{est} = 1429.5$ns, $t_f^{real} = 384.5$ns and $t_f^{est} = 384.5$ns, $V_{pk}^{real} = 562$ V and $V_{pk}^{est} = 558$ V. Additionally, we also present the estimation result and predicted dynamic

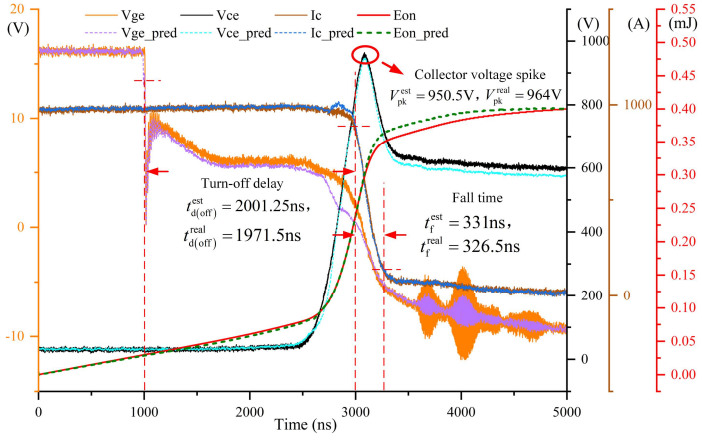
TABLE IV
PERFORMANCE OF THE PROPOSED METHOD TESTED ON TWO BATCH OF IGBTS (AVERAGE VALUE OF 216 WORKING CONDITIONS AT EACH T_c)

T_c	40°C			60°C			80°C		
Batch	DIM1200FSS12-A000								
Indexes	MAE	RMSE	LE	MAE	RMSE	LE	MAE	RMSE	LE
V_{ge}	0.0264	0.0853	1.0015	0.0378	0.1632	1.0071	0.0419	0.0974	1.0546
V_{ce}	1.1511	1.9497	1.734e+12	0.9513	1.0927	3.134e+16	0.7641	1.3183	4.243e+8
I_c	3.1633	5.1987	4.984e+19	2.9119	5.8146	4.745e+26	2.9373	6.9872	5.947e+28
E_{on}/E_{off}	0.0001	0.0003	1.0000	0.0002	0.0004	1.0000	0.0002	0.0004	1.0000
Batch	DIM1200FSS12-A076								
V_{ge}	0.2742	0.8712	3.4521	0.0923	0.2743	1.8326	0.2043	0.2896	1.8362
V_{ce}	0.3342	0.9838	4.163e+4	1.9878	3.9473	8.723e+16	0.2746	1.9473	1.4375e+5
I_c	2.2335	10.1221	3.778e+34	4.8122	10.7352	1.946e+32	2.7462	12.8364	1.743e+37
E_{on}/E_{off}	0.0003	0.0005	1.0000	0.0004	0.0005	1.0000	0.0004	0.0008	1.0000
T_c	100°C			125°C			150°C		
Batch	DIM1200FSS12-A000								
V_{ge}	0.0419	0.0812	1.0089	0.0412	0.1544	1.0103	0.0634	0.2177	1.0542
V_{ce}	1.2689	2.1723	2.342e+15	0.2234	1.4359	2.393e+8	0.2978	1.3927	9.235e+12
I_c	7.1649	9.7883	8.813e+36	3.7581	7.2134	4.894e+28	0.4677	1.8354	6.783 e+14
E_{on}/E_{off}	0.0004	0.0005	1.0000	0.0004	0.0005	1.0000	0.0004	0.0006	1.0001
Batch	DIM1200FSS12-A076								
V_{ge}	0.2001	0.3306	1.0510	0.0425	0.1693	1.0226	0.0854	0.2956	1.0879
V_{ce}	1.8469	4.0545	4.549e+19	1.5542	3.0467	1.521e+21	0.7127	4.0035	3.206e+22
I_c	6.0917	11.9801	3.534e+37	5.0381	7.9205	3.431e+58	0.5312	2.8188	1.575e+25
E_{on}/E_{off}	0.0006	0.0008	1.0002	0.0170	0.0413	1.0008	0.0115	0.0349	1.0006

electrical parameters at the working condition T_c=150°C, V_{DC}=600V, I_c=1000A, R_g=3.5Ω, L_p=120nH.



(a) Turn-on transient



(b) Turn-off transient

Fig. 10. Turn-on and turn-off transient waveforms of IGBT DIM1200FSS12-A076 DPT estimating results at the working condition T_c=150°C, V_{DC}=600V, I_c=1000A, R_g=3.5Ω, L_p=120nH ((a) turn-on transient, (b) turn-off transient).

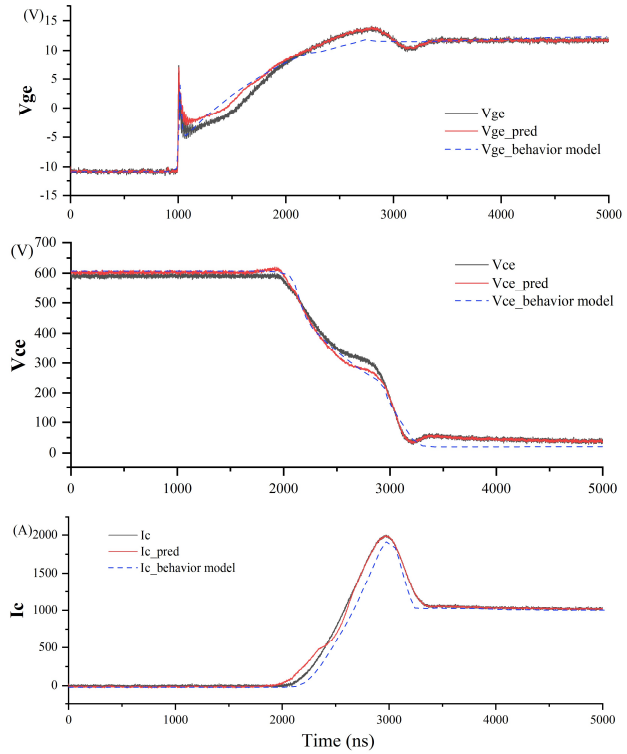
With the working condition T_c=150°C, in Fig. 10(a), $t_{d(on)}^{real}$ = 1164 ns and $t_{d(on)}^{est}$ = 1095.25 ns, t_r^{real} = 406 ns and t_r^{est} = 480.25 ns, I_{pk}^{real} = 1992.9 A and I_{pk}^{est} = 2000 A. In Fig. 10(b), $t_{d(off)}^{real}$ = 1971.5ns and $t_{d(off)}^{est}$ = 2001.25ns, t_f^{real} = 326.5ns and t_f^{est} = 331ns, V_{pk}^{real} = 964V and V_{pk}^{est} = 950.5V.

In order to make the experimental results more reliable, we conducted DPT result estimation experiments under all working conditions. The experimental performance of the proposed MAFGCN method tested on 40°C, 60°C, 80°C, 100°C, 125°C and 150°C (average value of 216 working conditions at each T_c) is summarised in Table IV. However, MAFGCN is the first deep learning-based method for IGBT DPT waveform estimation so that we cannot compare with similar methods in other literature.

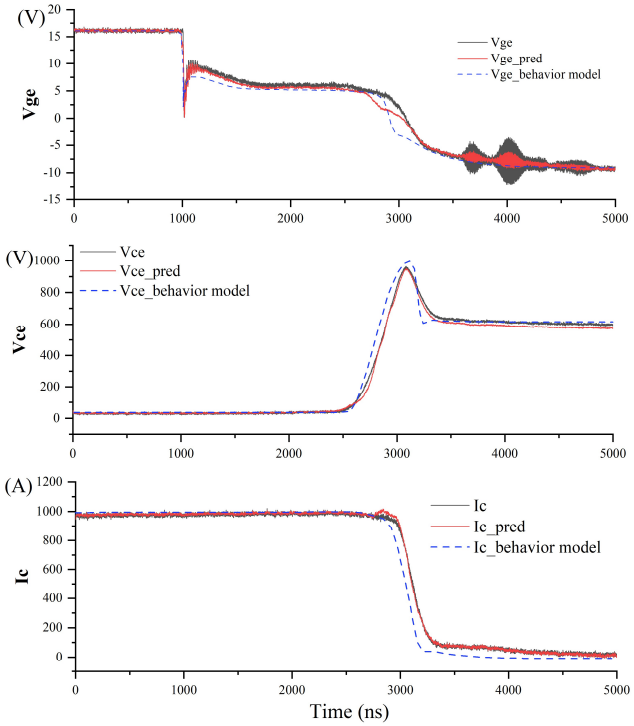
D. Comparison Experiments

Traditional behavioral models can also obtain IGBT DPT results through mathematical calculations to obtain dynamic electrical characteristic parameters. This section compares the estimation performance between the proposed method and a conventional IGBT behavior model. The dynamic behavior model is built by ANSYS Simplorer software according to the parameters given on the DIM1200FSS12-A076 datasheet provided by the manufacturer.

We compare the complete DPT waveform estimation result of the IGBT batch DIM1200FSS12-A076 on the working condition T_c=150°C, V_{DC}=600V, I_c=1000A, R_g=3.5Ω, L_p=120nH.



(a) Turn-on transient



(b) Turn-off transient

Fig. 11. Estimation performance between the proposed method and the conventional behavior model at the working condition T_c=150°C, V_{DC}=600V, I_c=1000A, R_g=3.5Ω, L_p=120nH ((a) turn-on transient, (b) turn-off transient).

As shown in Fig. 10, the deviation between DPT waveforms obtained by the well-configured behavior model and the actual measured value is significantly larger than that of the proposed method. Many simulation parameters need to be set, including device parameters and circuit parameters, to ensure that the simulation environment can reflect the conditions in actual applications. If more physical effects need to be considered, the modeling needs to be more refined to improve the accuracy of the simulation. The configuration of complex device parameters and simulation parameters during the simulation process may not be easy for some users. The data used in the proposed method are collected from real experiments, so the results obtained are more reliable in actual use than using device model simulations.

VI. Conclusion

This paper proposes a novel DPT waveform result estimating method, referred to as MAFGCN. The MAFGAN comprehensively considers the correlation between each working condition to build a spatial graph structure and embeds DPT waveform data. This method also designs MAFN block connected to GCN to ensure that the model can capture excellent temporal feature representations and preform high accuracy. Through the above experimental analysis, the following conclusions can be drawn:

1) This MAFGAN based on deep learning and graph neural network is accurate for IGBT DPT estimation. These reliable estimates can be used to calculate the dynamic electrical characteristics of the IGBT module.
2) The estimation of the test datasets of the same and similar batches of IGBTs has high accuracy through graph transfer process. For IGBT batches with limited difference, the proposed method can still accurately estimate the correct DPT results (as shown in Table IV).
3) The dynamic parameters $t_{d(on)}$, $t_{d(off)}$, t_r, t_f, I_{pk} and V_{pk} calculated using the DPT estimation results show high accuracy compared with real values.

In the feature work, we will try to use the proposed MAFGCN with graph migration method to estimate other brands of IGBT module (e.g., Infineon) DPT results at all working condition with only DPT data at room temperature. In the case of the same graph structure, the proposed method has great potential.

Reference

[1] S. Mohsenzade, "A Gate Driving Strategy for the Series-Connected IGBTs to Improve the Resilience Against IGBTs Short-Circuit Failures," *IEEE Transactions on Industrial Electronics,* vol. 69, no. 10, pp. 9961-9971, Oct. 2022.

[2] X. Chen, L. Wang, Y. Liu and Q. Wu, "Bang-Bang Funnel Control of Three-Phase Full-Bridge Inverter Under Dual-Buck Scheme," *IEEE Transactions on Industrial Electronics,* vol. 70, no. 6, pp. 5399-5409, June 2023.

[3] R. Li, C. Ho and C. Xu, "Active virtual ground - single phase transformerless grid-connected voltage source inverter topology," *IEEE Trans. Power Electron.,* vol. 33, no. 2, pp. 1335-1346, Feb. 2018.

[4] J. Kim, S. -W. Ryu, M. Kim and J. -W. Jung, "Triple-Mode Isolated Resonant Buck–Boost Converter Over Wide Input Voltage Range for Residential Applications," *IEEE Transactions on Industrial Electronics,* vol. 68, no. 11, pp. 11087-11099, Nov. 2021.

[5] Y. Xu, C. N. M. Ho, A. Ghosh and D. Muthumuni, "An Electrical Transient Model of IGBT-Diode Switching Cell for Power Semiconductor Loss Estimation in Electromagnetic Transient Simulation," *IEEE Transactions on Power Electronics,* vol. 35, no. 3, pp. 2979-2989, March 2020.

[6] J. Qi et al., "Temperature Dependence of Dynamic Performance Characterization of 1.2-kV SiC Power mosfets Compared With Si IGBTs for Wide Temperature Applications," *IEEE Transactions on Power Electronics,* vol. 34, no. 9, pp. 9105-9117, Sept. 2019.

[7] W. Wang, Z. Shen and V. Dinavahi, "Physics-Based Device-Level Power Electronic Circuit Hardware Emulation on FPGA," *IEEE Transactions on Industrial Informatics,* vol. 10, no. 4, pp. 2166-2179, Nov. 2014.

[8] C. Liu, H. Bai, S. Zhuo, X. Zhang, R. Ma and F. Gao, "Real-Time Simulation of Power Electronic Systems Based on Predictive Behavior," *IEEE Transactions on Industrial Electronics,* vol. 67, no. 9, pp. 8044-8053, Sept. 2020.

[9] 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 Transactions on Industrial Electronics,* vol. 66, no. 8, pp. 6581-6591, Aug. 2019.

[10] P. Górecki and D. Wojciechowski, "Accurate Electrothermal Modeling of High Frequency DC–DC Converters With Discrete IGBTs in PLECS Software," *IEEE Transactions on Industrial Electronics,* vol. 70, no. 6, pp. 5739-5746, June 2023.

[11] Z. Wang and W. Qiao, "A Physics-Based Improved Cauer-Type Thermal Equivalent Circuit for IGBT Modules," *IEEE Transactions on Power Electronics,* vol. 31, no. 10, pp. 6781-6786, Oct. 2016.

[12] R. Wu et al., "A Temperature-Dependent Thermal Model of IGBT Modules Suitable for Circuit-Level Simulations," *IEEE Transactions on Industry Applications,* vol. 52, no. 4, pp. 3306-3314, July-Aug. 2016.

[13] R. Kraus and H. J. Mattausch, "Status and trends of power semiconductor device models for circuit simulation," *IEEE Transctions on Power Electronics,* vol. 13, no. 3, pp. 452-465, May 1998.

[14] A. R. Hefner and D. M. Diebolt, "An experimentally verified IGBT model implemented in the saber circuit simulator," *IEEE Transctions on Power Electronics,* vol. 9, no. 5, pp. 532-542, Sep. 1994.

[15] 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 Journal of Emerging and Selected Topics in Power Electronics,* vol. 7, no. 1, pp. 62-70, March 2019.

[16] K. Górecki and P. Górecki, "Nonlinear Compact Thermal Model of the IGBT Dedicated to SPICE," *IEEE Transactions on Power Electronics,* vol. 35, no. 12, pp. 13420-13428, Dec. 2020.

[17] J. L. Tichenor, S. D. Sudhoff and J. L. Drewniak, "Behavioral IGBT modeling for predicting high frequency effects in motor drives," *IEEE Transctions on Power Electronics,* vol. 15, no. 2, pp. 354-360, Mar. 2000.

[18] A. D. Rajapakse, A. M. Gole and R. P. Jayasinghe, "An improved representation of FACTS controller semiconductor losses in EMTP-type

programs using accurate loss-power injection into network solution," *IEEE Trans. Power Del.,* vol. 24, no. 1, pp. 381-389, Jan. 2009.

[19] A. D. Rajapakse, A. M. Gole and P. L. Wilson, "Electromagnetic transients simulation models for accurate representation of switching losses and thermal performance in power electronic systems," *IEEE Trans. Power Del.,* vol. 20, no. 1, pp. 319-327, Jan. 2005.

[20] H. Luo, W. Li and X. He, "Online High-Power P-i-N Diode Chip Temperature Extraction and Prediction Method With Maximum Recovery Current di/dt," *IEEE Transactions on Power Electronics,* vol. 30, no. 5, pp. 2395-2404, May 2015.

[21] J. Xu, H. Ke, Z. Chen, X. Fan, T. Peng and C. Yang, "Oversmoothing Relief Graph Convolutional Network-Based Fault Diagnosis Method With Application to the Rectifier of High-Speed Trains," *IEEE Transactions on Industrial Informatics,* vol. 19, no. 1, pp. 771-779, Jan. 2023.

[22] A. Vaswani et al., "Attention is all you need," arXiv.org, 06 Dec 2017. [Online]. Available: https://arxiv.org/abs/1706.03762.

[23] K. Jia, Z. Yang, L. Zheng, Z. Zhu and T. Bi, "Spearman Correlation-Based Pilot Protection for Transmission Line Connected to PMSGs and DFIGs," *IEEE Transactions on Industrial Informatics,* vol. 17, no. 7, pp. 4532-4544, July 2021.

[24] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A Deep Learning Framework for traffic forecasting," arXiv.org, 12 Jul. 2018. [Online].

[25] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," arXiv.org, 22 February 2017. [Online]. Available: https://arxiv.org/abs/1609.02907.

[26] S. Hochreiter and J. Schmidhuber, Long Short Term Memory, München: Inst. für Informatik, TUM,, 1994.

Xiaotian Zhang was born in Shandong, China, in 1996. He received the B.Sc. degree in applied physics from Hohai University, Nanjing, China, in 2018, and the M.Sc. degree in electronic engineering from the Kings College London, London, U.K., in 2020. He is currently pursuing the Ph.D. degree in electrical engineering with the Department of Electronic Engineering, University of York, Heslington, York, UK. His research interests include health monitoring (fault diagnosis and condition monitoring) of electric powertrain systems, forecasting and estimation of time-series parameters in engineering applications.

Yihua Hu (Senior Member) received the B.S. degree in electrical engineering and the Ph.D. degree in power electronics and drives from China University of Mining and Technology, Xuzhou, China, in 2003 and 2011, respectively. Between 2011 and 2013, he was with the College of Electrical Engineering, Zhejiang University as a Postdoctoral Fellow. He is a reader at the King's College London, and he was the head of the Electrical Engineering Group at the University of York. His research interests include renewable generation, power electronics converters & control, electric vehicle, more electric ship/aircraft, smart energy system. He is the associate editor of IEEE Transactions on Industrial Electronics. He is a fellow of Institution of Engineering and Technology (FIET). He was awarded Royal Society Industry Fellowship.

Jingwei Zhang was born in Xuzhou, Jiangsu,China, in 1991. He received the B.Eng. degree from the School of Electrical and Power Engineering, China University of Mining and Technology (CUMT), China, in 2014, where he is currently pursuing the Ph.D. degree in electrical engineering. His current research interests include gate drive technologies and application for IGBT and silicon carbide MOSFET modules.

Mohammad Nasr Esfahani (Member, IEEE) received the M.S. degree in material engineering from Sharif University of Technology and the Ph.D. degree in mechanical engineering from Koc University in 2010 and 2017, respectively. From 2017 to 2018, he was with Mechanical Engineering, Koc University as a Postdoctoral Fellow. He is currently a Senior Lecturer at the School of Physics, Engineering and Technology, The University of York, UK. His research interests include advanced materials, sensors, safety and failure, and smart energy systems.

Tim Tilford received the B.Eng. degree in aeronautical engineering from Queen Mary College, University of London, London, U.K., in 1998, the M.Sc. degree in computational fluid dynamics and the Ph.D. degree in computational mechanics from the University of Greenwich, London, U.K., in 1999 and 2013, respectively. He has been a Research Fellow and subsequently a Senior Lecturer with the University of Greenwich, since 2002. His current research interests include numerical analysis of multiphysics/multiscale problems and high-performance parallel computing. Dr. Tilford is a Fellow of the Institute of Mathematics and its Applications, a Chartered Mathematician, and a Chartered Scientist.

Stoyan Stoyanov (Senior Member, IEEE) received the BSc degree in mathematics from Sofia University, Bulgaria in 1994 and the MSc degree in applied mathematics from the same institution in 1996. He has obtained his PhD degree in computational engineering at University of Greenwich, London, UK, in 2004. In 2004, he joined the Computational Mechanics and Reliability Group (CMRG) at the School of Computing and Mathematical Sciences at University of Greenwich where he advanced his professional career. Dr Stoyanov is currently Head of CMRG at the University of Greenwich. His research interests include the development and application of modeling and simulation tools for numerical analysis of the performance and reliability of electronics products and microsystems, physics-of-failure modeling, computational intelligence for data-driven and prognostics modeling, additive manufacturing, and design optimization.