A Thermo-electric Coupled Reduced Order Model for Real-time Prediction of Power Electronic Devices

Hao Tang, Run Hu
School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan, China

Abstract: Fast and accurate prediction method of the temperature field of power electronic devices are of increasing importance for monitoring and controlling the working state of electronics, especially when inserting temperature detectors is infeasible. Most existing prediction methods are only focused on heat transfer process and lack consideration of the coupling effect of the electric fields, making the prediction accuracy unsatisfactory. For this end, we develop a solution framework, named TE-ROM based on proper orthogonal decomposition and Galerkin projection method to solve the thermo-electric coupled field of power electronics. The model is validated on a self-developed half-bridge insulated gate bipolar transistor (IGBT) module. According to our tests with both static and dynamic conditions, our TE-ROM exhibits an outstanding speed-up ratio of 298 compared with full order model while the maximum absolute error being 1.64 K and mean absolute error being 0.18 K. Furthermore, TE-ROM facilitates a more accurate heat source distribution that results in accuracy improvement at bonding wire region since the mean absolute error reduces by 76.41% meanwhile the coefficient of determination improves from 0.6750 to 0.9842, demonstrating the importance of electrical self-heating for the temperature prediction of wires. Finally, our TE-ROM has achieved real-time temperature prediction for the full field at a time step of 10^{-1} s.

Keywords: Reduced Order Model; Power Electronic Device; Temperature Prediction

1. Introduction

Power electronic devices, who take charge of the circuit control and electric conversion in high-power electric equipment for their capability of carrying large voltage and current, have been employed extensively in the electronic manufacturing industry [1], playing a vital part in the fields like renewable energy generation [2], electric vehicle [3] and high-speed rail. Due to the evolution trend towards higher level of device integration and greater demand of electric energy utilization, the power density of the operating electric load carried by the device has increased incredibly as the industry updates, posing severe thermal power loss which causes material aging and device failure problems [4]. As those common thermal failures, such as bonding wire lift-off and solder layer defect [5], relate to hot spots within the devices which are inaccessible through contact measurement, it is necessary to monitor and predict their thermal information by simulation means in advance in order to adjust the electric and thermal inputs to protect them from overheating [6].

The ideal prediction method demands being swift, accurate and capable of full-field prediction

out of the requirements from actual application. To achieve these goals, there are basically several promising ways including numerical simulation [7], thermal network [3, 8] and neural network method [9, 10]. Numerical method is widely utilized and gains a commercial success for it can deal with a vast range of partial differential equations (PDEs) and uncover the entire physical field of the model precisely, yet the heavy cost of time and computation resources renders it inadequate for fast prediction. Thermal network method is commonly used by the device manufacturers for its simplicity but magnificent efficiency. The methodology behind it is thermo-electric analogy, analogizing the heat transfer routes within the device to the electric circuit. However, the method is only capable of obtaining information of specific scattered nodes which is not enough to support field analysis. Neural network method is a promising solution to address the issues mentioned in the above methods, who specializes in solving PDEs through deep learning method. A well-trained neural network model could accelerate the thermal design and optimization of the device, yet the only drawback is the immense consumption of time and computation resources for dataset accumulation and model training.

In order to overcome these limitations, great efforts have been devoted to reduce the scale of the problem while keeping the solution accurate and that's when the idea of reduced order model (ROM) came out as a suitable path [11, 12]. The fundamental principal of ROM, a relative concept of full order model (FOM), is to find an approximate solution of the studied problem in a reduced solution space whose dimension is much smaller than the former FOM. Among the reduced order modeling, proper orthogonal decomposition (POD) has been applied widely in engineering analysis due to its low computation cost and good compatibility with numerical methods [13, 14]. The POD method was firstly introduced in statistics which was also called principal component analysis (PCA) to extract the dominant features from the metadata [15], applied in the fields concerning flow pattern analysis like fluid mechanics [16] and meteorology [17, 18] such as capturing the large eddy structures in the turbulence. During early times, the POD alone was only capable of feature analysis based on the dataset whereas couldn't compute solution without combining with specific physical mechanisms [19]. After incorporating with numerical methods like finite volume method (FVM) or finite element volume (FEM), POD was renovated and extended with the ability of extrapolation. As a result, POD was used to predict different physical fields in the context of electro-magnetic problems [20] and fluid dynamics-heat transfer problems [21].

In recent years, with the growing need of temperature monitoring and prediction, POD method has also been utilized for the thermal field reconstruction in electronic devices. State-of-the-art research has applied POD-FEM method to predict the temperature field of 3-D three-phase insulated gate bipolar transistor (IGBT) module and has achieved a 100-fold acceleration compared to traditional FEM while maintaining a maximum absolute error of 1 K [22]. Nowadays, it is concluded that the main bottleneck of POD-based prediction methods is the heavy computation cost for reduced matrix construction that lags down the efficiency. Therefore, diverse studies have retrofitted POD with other methods to boost efficiency. [23] introduces a unified index matrix to reduce the amount of processed data when reducing the matrix order, resulting in a maximum speed-up of in 101 times. [24] converts the ROM into equivalent thermal network model to calculate temperature data of designated points under 10-3 s per step. [25] reports a conservation-based method that introduces POD basis into Galerkin projection method, which calculates 1600 moments of the full temperature field with only 61.3 seconds (0.038 s per moment), showing 440 times faster than traditional FVM.

However, the above introduced existing studies still have room for further improvement. In

general, the literature focuses on the temperature evolution where the heat transfer process is considered and accelerated only, while the coupling effect of electric field is excluded. In the article, we refer to this solution framework as thermal-based reduced order model (T-ROM). In T-ROM, the heat dissipation rate is determined by the power loss functions which only cover the main heat sources such as chips and capacitors, as shown in Figure 1, but omit the Joule heat generated by other parts like wires. It's worth pointing out that such a simplification makes the calculation easier for avoiding the current but brings error which might lead to inaccurate prediction. However, deriving a detailed heat source requires solving the electric field, whereas a reduced order model that integrates the temperature and electric fields, along with a comprehensive assessment of its efficiency and accuracy, remains unexplored.

Based on this reason, given the purpose to predict a more precise temperature field premised on the swiftness, we propose a thermo-electric coupled reduced order model (TE-ROM) to predict the three-dimensional electric and temperature field quickly and accurately. In the TE-ROM framework shown in Figure 1, both electric and thermal field are reduced and accelerated by the POD-Galerkin method and the heat source during the thermal solution is acquired from the electric result through Joule's law to ensure its accuracy. To validate the performance of our TE-ROM, we built a schematic IGBT module from SOLIDWORKS and tested its calculation speed and accuracy by simulation means through different methods. The results have proven that the proposed TE-ROM could solve the temperature field more precisely than T-ROM while still possess an outstanding calculation speed, realizing real-time prediction for the devices.

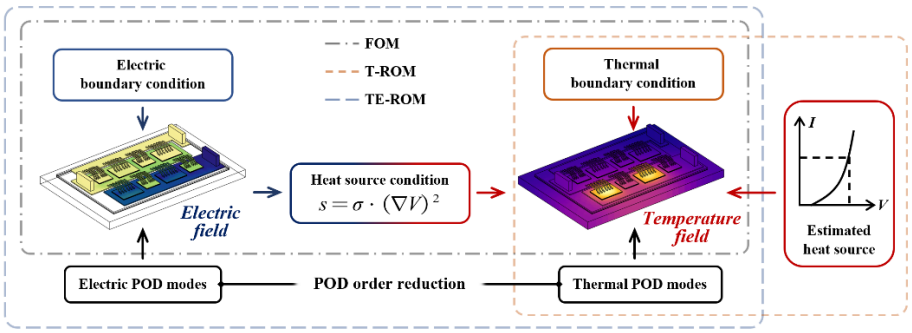

Figure 1: Strategic differences between the workflows of various methods. s means the heat source term calculated by Joule's law, where σ is the conductivity and V is the electric potential.

2. Methodology

2.1 POD Order Reduction

Given a physical field to be solved, denoted as $\boldsymbol{F}$, discretized into grids of N, the time-relevant solution of $\boldsymbol{F}$ can be regarded as the trajectory of a point moving through time in an N-dimensional solution space, which we also call the full order model (FOM):

$$F_{N\times 1}(t) = \sum_{i=1}^{N} \sigma_i f_{x_i}(t) \tag{1}$$

where σ_i is the standard basis vector whose i-th component is 1 while all the other components are 0. f_{x_i} represents the value of F at point $\boldsymbol{x}_i$, also referred to as the projection coefficient of $\boldsymbol{F}$

under the basis σ_i.

Nowadays, to solve a complex FOM through numerical method such as finite volume method with sufficient precision could easily make the scale of grid number N escalate to hundreds of thousands or even millions, resulting in excessively redundant solution space and lagging down the computational performance.

Rather than choosing the standard bases $\{\sigma_i\}$, the main concept of POD method is to select a set of characteristic orthogonal bases properly based on the amount of physical information it embodies about the system $\boldsymbol{F}$ to scale down the solution space. These bases, also referred to as POD modes, is utilized to reconstruct and approximate the original physical field. The more information obtained in this modal combination, the more accurate the reconstructed physical field will be.

Using POD bases, the original physical field F can be expressed as:

$$\boldsymbol{F}(t) = \sum_{i=1}^{N} \boldsymbol{\varphi}_i \bullet a_i(t) \tag{2}$$

where $\boldsymbol{\varphi}$ denote these POD modes, every one of which is geometrically perpendicular with each other and $\boldsymbol{a_i}$ denotes the projection coefficient of F under POD mode:

$$a_i = prj_{\boldsymbol{\varphi}_i}\boldsymbol{F} = \frac{\boldsymbol{F} \bullet \boldsymbol{\varphi}_i}{|\boldsymbol{\varphi}_i|} = \boldsymbol{F}^T\boldsymbol{\varphi}_i \tag{3}$$

Remaining the sum of the first n terms of Eq. (2), the approximation of $\boldsymbol{F}$, denoted as $\hat{\boldsymbol{F}}$ and the truncation error E can be expressed as:

$$\hat{\boldsymbol{F}} = \sum_{i=1}^{n} \boldsymbol{\varphi}_i \bullet a_i \tag{4}$$

$$E = \left(F - \hat{F}\right)^2 = \left(\sum_{n+1}^{N} \varphi_i \bullet a_i\right)^2 = \sum_{n+1}^{N} a_i^2 \tag{5}$$

Substitute Eq. (3) into Eq. (5) for further deduction:

$$E = \sum_{n+1}^{N} a_i^T \bullet a_i = \sum_{n+1}^{N} \varphi_i^T F F^T \varphi_i = \sum_{n+1}^{N} \varphi_i^T R \varphi_i \tag{6}$$

Where $\boldsymbol{R} = \boldsymbol{F}\boldsymbol{F}^{\boldsymbol{T}}$. Therefore, the set of POD modes can be derived by solving the optimization problem of minimizing E. The solution of the optimization problem is expressed as follows which reveals the condition that the optimal POD modes conform to:

$$\boldsymbol{R}\boldsymbol{\varphi}_i = \lambda_i \boldsymbol{\varphi}_i \tag{7}$$

Finally, after substituting Eq. (7) into Eq. (6), the truncation error can be quantified as:

$$E = \sum_{n+1}^{N} \boldsymbol{\varphi}_i^T \lambda_i \boldsymbol{\varphi}_i = \sum_{n+1}^{N} \lambda_i \tag{8}$$

Eq. (7) and Eq. (8) reveal that every eigenvector of $\boldsymbol{R}$ turns out to be the candidate of POD modes and its corresponding eigenvalue λ_i, which is also named the "energy" of the $i-th$ mode, quantifies the physical information captured in $\boldsymbol{\varphi}_i$ and is regarded as the criterion for mode selection. The higher the energy sum from chosen modes takes up, the more accurate the reconstruction will be achieved.

When it comes to practical means, however, as the temporal evolution of $\boldsymbol{F}(t)$ remains to be solved, the theoretical deduction of time-continuous functions listed above should be adjusted to a series of operations exerted on the time-discrete snapshot matrix $\boldsymbol{U}$:

$$\boldsymbol{U}_{N\times M} = [\boldsymbol{U}_{t_1} \quad \boldsymbol{U}_{t_2} \quad \cdots \quad \boldsymbol{U}_{T_M}] = \begin{bmatrix} u_{x_1,t_1} & u_{x_1,t_2} & \cdots & u_{x_1,t_M} \\ u_{x_1,t_1} & u_{x_2,t_2} & \cdots & u_{x_2,t_M} \\ \vdots & \vdots & \ddots & \vdots \\ u_{x_N,t_1} & u_{x_N,t_2} & \cdots & u_{x_N,t_M} \end{bmatrix} \tag{9}$$

where $\boldsymbol{U}_{t_i}$, a column of matrix $\boldsymbol{U}$, is the vector-shaped snapshot of the entire field $\boldsymbol{F}$ at $\boldsymbol{t}_i$. In brief, the snapshot matrix $\boldsymbol{U}$ is the collection composed of the full-field data records of the physical field $\boldsymbol{F}$ under different cases and moments in time, which contains the spatial and temporal evolution characteristics of $\boldsymbol{F}$. It is utterly the practical operation to obtain the POD modes to substitute $\boldsymbol{U}$ as $\boldsymbol{F}$ into Eq. (6) ~ (8).

2.2 Galerkin Projection Method

POD bases acquired by snapshot dataset offer a new norm to reconstruct the solution space in a more concise and efficient way, yet only with that it's still unable to figure out the movement of the system unless combined with its governing mechanisms such as heat transfer equation. Therefore, the Galerkin projection method is introduced to obtain the projection coefficients under the POD coordinate system.

In the context of IGBT module electric-thermal field prediction, the main governing equations are the heat transfer equation (Eq. (10)) and the electric current conservation equation (Eq. (11)):

$$\rho c_p \frac{\partial T}{\partial t} = \nabla \bullet (\kappa \nabla T) + s \tag{10}$$

$$\nabla \bullet J = 0 \tag{11}$$

where T and $\boldsymbol{J}$ denote the temperature and current density vector, ρ, c_p and κ are respectively the density, specific thermal capacity and coefficient of thermal conductivity of the material, s represents the heat source term. It's worth mentioned that Eq. (11) demonstrates the steady-state form of current conservation, which is a simplification assuming that the electric field would swiftly converge to steady state according to the shift of boundary condition at a negligible time cost.

Taking the deduction of heat transfer equation using POD-Galerkin projection for instance, every one of the grid elements in the system satisfies the discrete form of Eq. (12):

$$\rho_i c_{pi} \frac{T_i^* - T_i}{\Delta t} \bullet dV_i = \sum_{j \in nei(i)} C_{i,j}\left(T_j - T_i\right) \bullet dA_{i,j} + s_j \bullet dV_i \tag{12}$$

Here, i and j represent the index of the studied element and its adjacent ones respectively, $nei(i)$ indicates the set of neighbor elements of element i. T_i^* denotes the temperature of the i-th element T_i in the next time step. dV_i denotes the volume of elment i and $dA_{i,j}$ is the interface area between element i and j. $C_{i,j}$ is defined as thermal conductivity per unit area between element i and j, which equals to the reciprocal of thermal resistance per unit area:

$$C_{i,j} = \frac{\kappa_{i,j}}{\delta_{i,j}} \tag{13}$$

Where $\delta_{i,j}$ is the distance between the center points of element i and j.

The POD-Galerkin projection mainly operates two steps: projecting the equation on the POD basis and integration, which is expressed as follows:

$$\sum_{i=1}^{N} \rho_i c_{pi} \frac{T_i^* - T_i}{\Delta t} \varphi_i^{(k)} \bullet dV_i = \sum_{i=1}^{N} s_i \varphi_i^{(k)} \bullet dV_i + \sum_{i=1}^{N} \sum_{j \in nei(i)} C_{i,j}\left(T_j - T_i\right) \varphi_i^{(k)} \bullet dA_{i,j} \tag{14}$$

where N is the total number of elements and $\varphi_i^{(k)}$ denotes the value of the k-th POD mode on element i. The three terms in Eq. (14) are interpreted respectively as heat capacitance term, heat source term and heat conduction term from left to right.

The temperature field under POD bases can be expressed as:

$$T_i = \sum_{m=1}^{n} \varphi_i^{(m)} \cdot a_m + T_e \tag{15}$$

where n is the number of chosen POD modes, also named as the order of the POD space and equal to the degrees of freedom of the subspace, T_e denotes the environmental temperature. Substituting Eq. (15) into Eq. (14) reveals the algebraic equations of projection coefficients governed by the heat transfer law in matrix form:

$$\frac{A_{n\times n}}{\Delta t} \bullet a^*_{n\times 1} = \left(\frac{A_{n\times n}}{\Delta t} - K_{n\times n}\right) \bullet a_{n\times 1} + S_{n\times 1}$$

$$A_{k,m} = \sum_{i=1}^{N} \rho_i c_{pi} \varphi_i^{(k)} \varphi_i^{(m)} \bullet dV_i$$

$$K_{k,m} = \sum_{i=1}^{N} \sum_{j\in nei(i)} C_{i,j}\left(\varphi_i^{(m)} - \varphi_j^{(m)}\right) \bullet dA_{i,j}$$

$$S_k = \sum_{i=1}^{N} s_i \varphi_i^{(k)} \bullet dV_i \tag{16}$$

where the vector a is the coordinate of temperature field under POD bases and a^* is a at the next time step. Obviously, compared with traditional FVM methods, the dimension of the solution space, equal to the number of unknowns, has been scaled down from N to n (magnitude reduction usually from 10^5 to 10^0), leading to a prodigious computational acceleration and time saving.

The deduction of the steady-state current conservation equation follows the same path as heat transfer equation, so the final matrix-formed equations to obtain the projection coefficients of electric field, leaving out all the intermediate detailed deductions, is expressed as follows:

$$0_{n\times 1} = D_{n\times n} \bullet b_{n\times 1}$$

$$D_{k,m} = \sum_{i=1}^{N} \sum_{j\in nei(i)} \left(\phi_j^{(m)} - \phi_i^{(m)}\right) \phi_i^{(k)} c_{i,j} \bullet dA_{i,j} \tag{17}$$

$$V_i \sum_{m=1}^{n} \phi_i^{(m)} \bullet b_m \tag{18}$$

$V, \phi_i^{(m)}$ and b_m are respectively the voltage value, the m-th electric POD mode and its projection coefficient. $c_{i,j}$ is the electric conductivity per unit area, equal to the electric conductivity rate $\sigma_{i,j}$ divided by $\delta_{i,j}$.

2.3 Workflow of the TE-ROM

The schematic order of TE-ROM is divided into offline stage and online stage. Offline stage is responsible for preparing POD modes in advance and needs to be executed once. On the other hand, the online stage needs to quickly calculate the electric field and temperature field based on the input boundary conditions. During online stage, the TE-ROM is executed by the following order: electric field solution, heat source calculation and temperature field solution. The electric field is solved ahead of temperature field because the heat generation rate from the electronic device is calculated by Joule's law, which is the bridge connecting electric and temperature field and the very difference between TE-ROM and T-ROM shown in Figure 1. To obtain a more precise heat source distribution, the TE-ROM considers the heating effect of the whole electric domain and studies its influence to the temperature field.

3. Performance Validation of TE-ROM for Power Device Prediction

3.1 Structure of the Power Device

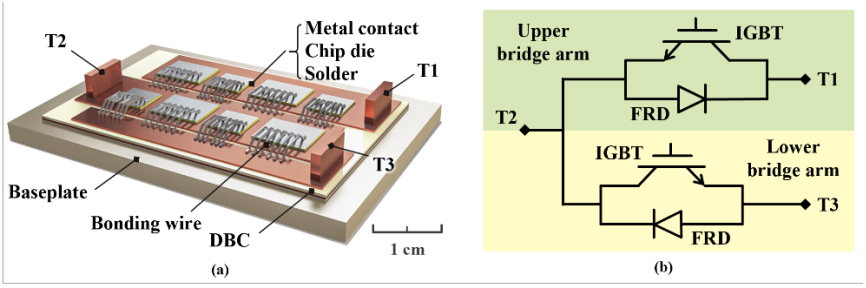

Figure 2: The 3-D model (a) and Circuit Topologic Diagram (b) of IGBT Module.

To validate the performance of the proposed TE-ROM, a self-developed half-bridge IGBT module is built by SOLIDWORKS for example. Figure 2 shows the three-dimensional structure and the circuit diagram of the studied IGBT module. The module has two bridge arms and each arm incorporates two IGBT chips and two fast recovering diode (FRD) chips paralleled with each other through binding wires and direct bonding copper (DBC). To achieve current modulation, the two arms turn on alternatively by switching IGBTs on and off. The size of the bottom surface of the baseplate is 60 mm × 40 mm. The properties of the IGBT module are listed in Table 1.

Table 1: The Physical Property Parameters of IGBT Module.

Material	Density (kg/m^3)	Specific heat capacity (J/kg/K)	Thermal conductivity (W/m/K)	Electric conductivity (m/S)
Cu	8940	390	398	5.8×10^7
Al	2700	900	238	3.7×10^7
Si	2330	715	135	100 (on) 0 (off)
AlN	3900	900	27	0
SAC305	7300	230	54	9.1×10^6

The schematic diagram in Figure 3 illustrates the practical application scenario of an IGBT module. In this setup, the current input flows through one of the IGBT bridge arms, generating power loss that is finally dissipated at the bottom of the baseplate to the heat sink. The heat transfer process follows the third kind boundary condition, with a uniformed ambient temperature (set to 293.15 K) and a heat transfer coefficient. Therefore, each operation case is uniquely defined by a combination of three parameters: the active bridge arm, input current load and heat transfer coefficient.

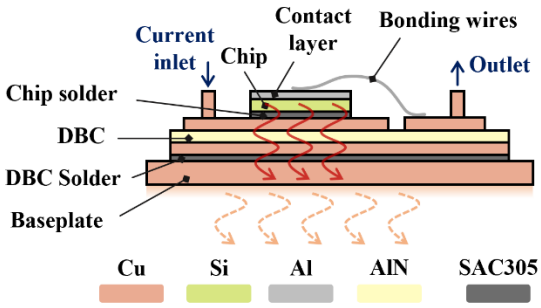

Figure 3: The Physical Condition of IGBT Operation Case.

To mitigate the computational burden of model construction and solution, some assumptions are made for simplification, which are listed below:

1) Chip structure simplification: The IGBT chip is considered homogenous and left out its micro-structure.

2) Gate structure simplification: The gate area is omitted in the module for its negligible influence to the temperature and electric solution.

3) Steady-state simplification for electric field: As mentioned before, we choose stationary form of electric field function because the time scale of electric field variance is negligible compared with the shifting period of electric boundary condition.

3.2 Snapshot Dataset Preparation

Snapshot dataset prepared for the mode extraction process is of great importance to the prediction ability of the ROM. The cases in snapshots must be diverse enough to encompass the whole prediction range in order to obtain the physical information as much as possible. Therefore, a parameter scanning simulation containing 18 time-dependent studies with different parameter combinations are computed in advance by finite-volume simulation.

Table 2: Case Parameters of Snapshot Dataset.

Parameters	Values
Operating bridge arm	Upper bridge arm / Lower bridge arm (See Figure 2(b))
Input load (A)	50 / 100 / 150
Heat transfer coefficient ($W/m^2/K$)	1250 / 2000 / 2500
Switching frequency (Hz)	0 / 1 /50

The parameters of each study are chosen from Table 2 according to the real loading stress and liquid cooling scenario for IGBT, where the current inlet is chosen from 50 A to 150 A, the heat transfer coefficient ranges from 1500 $W/m^2/K$ to 2500 $W/m^2/K$ and the active region shifts between the upper bridge and the lower bridge with different frequency. In the row of switching frequency parameters of Table 2, 0 Hz indicates that the current load remains unshifted. After processing all the simulations, the results of the temperature field and the electric field are vectorized and concatenated to form the snapshot matrix as the dataset for mode extraction.

3.3 Validation Method

To validate the prediction performance of the TE-ROM, two types of tests are implemented including constant case test (Test A) and time-variant case test (Test B). Both tests use the Robin boundary condition to solve heat transfer equation with fixed heat transfer coefficient h. The initial temperature field is applied uniform and the same as the ambient temperature of 293.15 K.

Table 3: Case Configurations for Test A

Case	Operating bridge arm	Input load (A)	Heat transfer coefficient ($W/m^2/K$)
1	Upper bridge	60	1000
2	Lower bridge	60	1000
3	Upper bridge	75	1200
4	Lower bridge	75	1200
5	Upper bridge	90	1500
6	Lower bridge	90	1500
7	Upper bridge	120	2000
8	Lower bridge	120	2000

In Test A, where the case parameters are set fixed, the current load ranges from 60 A to 120 A to simulate all scenarios of IGBT from mild load to heavy load, meanwhile the heat transfer coefficient variates from 1000 $W/m^2/K$ to 2000 $W/m^2/K$, representing the typical liquid cooling performance of IGBT from weak to strong. The total 8 case configurations are detailed in Table 3. For each case the simulation lasts 100 s with a time step of 0.1 s, yielding 1000 iterations per case.

Table 4: Case configurations for Test B.

Case	Shifting Frequency (Hz)	Input load (A)	Heat transfer coefficient ($W/m^2/K$)
1	1	75	1200
2	1	90	1500
3	1	120	2000
4	50	75	1200
5	50	90	1500
6	50	120	2000

Meanwhile, Test B is designed to see the prediction ability of TE-ROM under dynamic occasions as it is more conformed to practical application of IGBT. During Test B, the current is exerted on the two bridge arms alternatively with a constant frequency, leading to heat source varying through time. A total of 6 cases are arranged and categorized into two groups based on switching frequency: second-scaled switching tests (Case 1~3) and microsecond-scaled switching tests (Case 4~6). The specific configurations are summarized in Table 4.

All the tests were conducted on a platform equipped with an Intel Xeon E5-2690 v4 CPU and a single NVIDIA Quadro P2000 GPU. All the models are coded and tested on MATLAB R2022b.

4. Results and Discussions

4.1 Accuracy Analysis

Figure 4 displays the visualization result of the temperature and electric field distribution at 100 s under Case 7 of Test A when reaching the stationary state. The errors are mainly distributed in high-temperature areas as well as at the edges of different layers, which may be caused by the grid size. Nevertheless, by comparing these figures we can find a high consistency between the results of FOM and TE-ROM since the max absolute deviation between them is 1.476 K in the temperature result and 0.0013 V in the electric result. Divided by the range, the max relative error under Case 7 is 2.67% in the temperature field. Among the results in Test A, he highest max absolute error and mean absolute error of the temperature field under all cases is 1.64 K and 0.18 K from Case 8, accounting for 2.86% and 0.34% compared with its range, showing good agreement between the two methods.

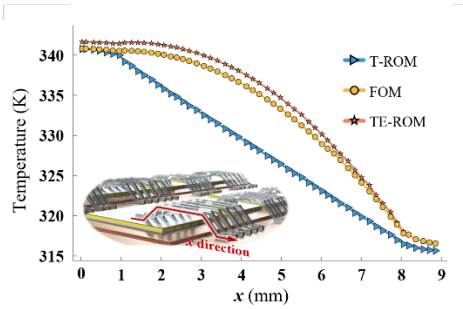

Figure 4: The electric and temperature field distribution results (100 s) under of Test A acquired by FOM and TE-ROM and the deviations between the two methods.

Figure 5 shows the time-dependent temperature variation of four critical points using TE-ROM and FOM under Test A. The curves of the two methods overlap well in each case since the biggest absolute error is 1.58 K on p2 at 2.8 s of Case 7, proving the prediction accuracy of TE-ROM. Figure 6 depicts the simulated time-dependent temperature of the same reference points under Test B, where the biggest error is 1.65 K at point p2 (0.55 s) in Case 3 of the second-scaled switching tests and 0.48 K at point p2 (0.091 s) in Case 6 of the microsecond-scaled switching tests. These results indicate that the thermal reduced solution of TE-ROM consistently aligns with that of FOM, regardless of whether the switching occurs on a microsecond or second scale.

It is noteworthy that although the operating conditions for Tests A (Table 3) and B (Table 4) fall outside the distribution of the POD dataset (Table 2), the model maintains a remarkably high level of predictive accuracy as discussed in figure 4~6, demonstrating the robust generalization capabilities of the TE-ROM.

4.2 Comparison Analysis Between TE-ROM and T-ROM

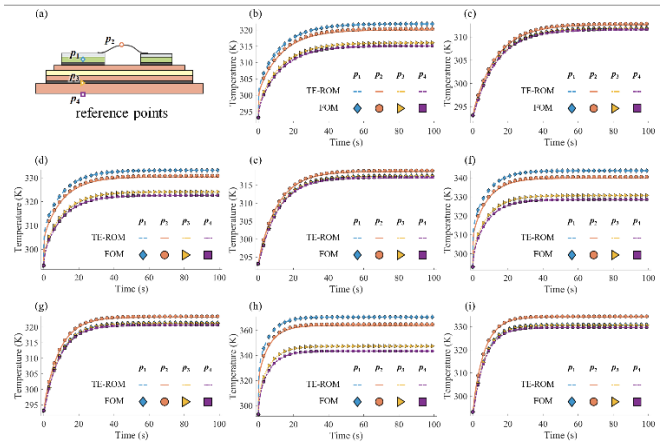

Figure 5: The simulated time-dependent temperature of four reference points using TE-ROM and FOM under 8 cases of Test A. (a) The four critical points respectively refer to chip (p1), bonding wire (p2), solder layer (p3) and base plate (p4). (b) ~ (i) refer to results of Case 1~8 respectively.

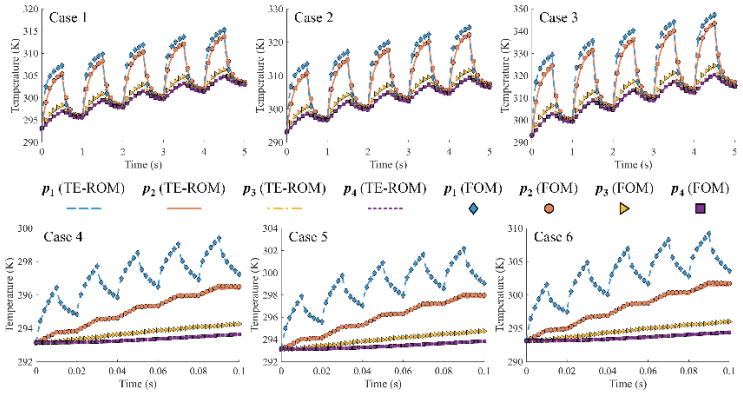

Figure 6: The simulated time-dependent temperature of reference points using TE-ROM and FOM under 6 cases of Test B. The four detector points are set the same as Figure 5, referring to chip (p1), bonding wire (p2), solder layer (p3) and base plate (p4).

As depicted in Figure 1, T-ROM uses estimated heat source based on IGBT output curve for fast prediction, leaving out the electric field solution. To demonstrate the accuracy difference between

TE-ROM and T-ROM, Figure 7 shows the detailed temperature distribution along the length direction of the bonding wires under Test A. In Figure 7, the prediction result of TE-ROM is more consistent with the reference group of FOM compared with T-ROM. In T-ROM, the heat generation of bonding wires is omitted so that the temperature along x is linear distributed, resulting in a great deviation from real temperature distribution since the maximum absolute error reaching 7 K compared to the total wire temperature drop of 25.14 K. The better agreement of TE-ROM result with FOM in both temperature and temperature gradient distribution can be justified by two indexes: mean absolute error (MAE) and coefficient of determination (R^2). The MAE between TE-ROM and FOM is reduced by averagely 76.41% compared with that between T-ROM and FOM, meanwhile the R^2 between TE-ROM and FOM has improved from 0.6750 to 0.9842. This accuracy advantage is due to a more comprehensive and precise heat source for TE-ROM considers the heat generation of the whole electric field including bonding wires rather than T-ROM who only considers the chip heating. The study also proves that although the heat generation of the bonding wires is only 1.8% of that by

the chips, it still plays a decisive role in the temperature distribution of the wires, which should not be omitted. Such an accurate temperature reconstruction of bonding wires could furthermore enhance the prediction accuracy of the wires' thermal stress, which results in the peeling and breakage of the wires directly and affect their aging and life-span.

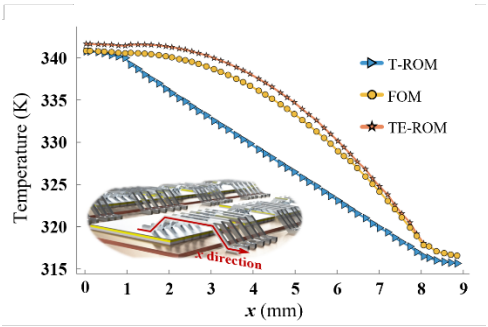

Figure 7: The temperature distribution in the bonding wires, where the x axis is established along the length of the wire.

4.3 Performance Efficiency Analysis

The order reduction of POD-Galerkin method brings dramatic acceleration to the calculation of physical field compared with FOM. Table 5 shows the total time cost along with detailed comparison of each operation within TE-ROM and FOM. Here, the speed up ratio is defined as the time cost of FOM divided by that of TE-ROM. Comparing each item, the performance of TE-ROM is 432 times faster in electric matrix construction and 162 times faster in the field solution because of the matrix scale reduction from 160742 (grid number of the electric field) to 5. As for temperature field, the speed up ratio of matrix construction and solution is respectively 254 and 169 due to the order reduction from 396567 (grid number of the temperature field) to 10. As a reflection of comprehensive performance, TE-ROM takes 15.18 s to calculate the temperature field of the IGBT module over 1000 iterations, representing 1/298 of the FOM's cost. As shown in Table 5, it takes TE-ROM only 0.0138 s per iteration to calculate the full temperature field, which means a real-time prediction under a timescale of 10-1 s can be achieved when the simulation time step is set above 0.0138 s.

Table 5: The Time Cost During Each Step Between TE-ROM and FOM.

Operation		FOM result		TE-ROM result		Speed up ratio
		Full order	Time cost (s)	Order applied	Time cost (s)	
Stationary electric field	Matrix construction	160742	228.5	5	0.5271	433
	Solution	160742	0.5290	5	0.0032	163
Temperature field	Matrix construction	396567	1963	10	7.726	254
	Solution (per iteration)	396567	2.342	10	0.0138	169
Total time cost of 1000 iterations (s)		4534.5		15.18		298

5. Conclusions

To achieve real-time temperature prediction in power electronic devices, we have presented a reduced order model through POD-Galerkin projection method to solve coupling field of both the electric and thermal field. The key of our methodology is the integration of the heating effect of the electric field into thermal ROM, resulting in a more comprehensive and precise heat source distribution calculated by the Joule effect of the electric field.

First, according to our performance tests, our model has demonstrated the novel accuracy and efficiency for approximating the FOM. Compared to recent advancements [22], the accuracy of TE-ROM is proved comparable to the latest research, with a maximum error of 1.65 K. However, our model demonstrates a superior computational efficiency, achieving a speed-up ratio of 298, nearly three-fold compared to the existing literature.

Then, an accuracy comparison between TE-ROM and existing T-ROM is discussed. Notably, the TE-ROM has been proved more precise in prediction details compared with T-ROM, as the MAE of wire-region reduced by 76.41% and R2 improvement from 0.6750 to 0.9842. The advancement of our model is due to considering a more comprehensive and accurate heat source term from electric heating effect, which also proves that the self-heating of bonding wires plays a vital role in their own temperature field and should not be left out of prediction. The accuracy novelty could furthermore facilitate thermal stress prediction of wires.

Finally, the absolute computation time of solution per iteration is a key index to test the ability of real-time prediction. Our model has achieved an average time cost of 0.0138s for full-field output, enabling real-time thermal prediction with execution times on the magnitude of 10-1 s for the global field that outperforms the state-of-the-art results from the existing literatures.

Funding

The work is supported financially by the National Key R & D Project from Ministry of Science and Technology of China under grant 2022YFA1203100.

Data Availability

The processed data supporting the results of this study are available in the Zenodo repository, https://doi.org/10.5281/zenodo.19217240. The original code is not publicly available, but available from the corresponding author on reasonable request.

Conflicts of Interest

The authors have no conflict of interest to declare.

References

[1] Zhang, J., Shen, H., Sun, H., & Wang, Z. (2025). Review on the thermal models applications in the reliability of power semiconductor device. IET Power Electronics, 18(1), e70065. https://doi.org/10.1049/pel2.70065

[2] Kalay, M. S., & Akpolat, A. N. (2025). A systematic review on reliability and lifetime evaluation of power converters in power generation systems. IET Renewable Power Generation, 19(1), e70111. https://doi.org/10.1049/rpg2.70111

[3] Ma, M., Guo, W., Yan, X., Yang, S., Zhang, X., Chen, W., & Cai, G. (2020). A three-dimensional boundary-dependent compact thermal network model for IGBT modules in new energy vehicles. IEEE Transactions on Industrial Electronics, 68(6), 5248–5258. https://doi.org/10.1109/TIE.2020.2991926

[4] Morel, C., & Morel, J.-Y. (2024). Power semiconductor junction temperature and lifetime estimations: A review. Energies, 17(18), 4589. https://doi.org/10.3390/en17184589

[5] Kang, Y., Dang, L., Yang, L., Wang, Z., et al. (2023). Research progress in failure mechanism and health state evaluation index system of welded IGBT power modules. Electronics, 12(15), 3248. https://doi.org/10.3390/electronics12153248

[6] Hu, Z., Du, M., & Wei, K. (2017). Online calculation of the increase in thermal resistance caused by solder fatigue for IGBT modules. IEEE Transactions on Device and Materials Reliability, 17(4), 785–794. https://doi.org/10.1109/TDMR.2017.2746571

[7] Hefner, R., & Blackburn, D. L. (1993). Simulating the dynamic electrothermal behavior of power electronic circuits and systems. IEEE Transactions on Power Electronics, 8(4), 376–385. https://doi.org/10.1109/63.261007

[8] Bahman, S., Ma, K., & Blaabjerg, F. (2018). A lumped thermal model including thermal coupling and thermal boundary conditions for high-power IGBT modules. IEEE Transactions on Power Electronics, 33(3), 2518–2530. https://doi.org/10.1109/TPEL.2017.2694548

[9] Yang, Y., et al. (2025). A parameterized thermal simulation method based on physics-informed neural networks for fast power module thermal design. IEEE Transactions on Power Electronics, 40(7), 9200–9210. https://doi.org/10.1109/TPEL.2025.3547390

[10] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2018). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

[11] Bader, E., Kärcher, M., Grepl, M. A., & Veroy, K. (2016). Certified reduced basis methods for parametrized distributed elliptic optimal control problems with control constraints. SIAM Journal on Scientific Computing, 38(6), A3921–A3946. https://doi.org/10.1137/16M1059898

[12] German, P., Tano, M. E., Fiorina, C., & Ragusa, J. C. (2021). Data-driven reduced-order modeling of convective heat transfer in porous media. Fluids, 6(8), 266. https://doi.org/10.3390/fluids6080266

[13] Xiang, L., Lee, C. W., Zikanov, O., & Hsu, C.-C. (2021). Efficient reduced order model for heat transfer in a battery pack of an electric vehicle. Applied Thermal Engineering, 201, 117641. https://doi.org/10.1016/j.applthermaleng.2021.117641

[14] Akhtar, I., Nayfeh, A. H., & Ribbens, C. J. (2009). On the stability and extension of reduced-order Galerkin models in incompressible flows. Theoretical and Computational Fluid Dynamics, 23, 213–237. https://doi.org/10.1007/s00162-009-0112-y

[15] Lever, J., Krzywinski, M., & Altman, N. (2017). Principal component analysis. Nature Methods, 14, 641–642. https://doi.org/10.1038/nmeth.4346

[16] Karim, A., Kim, Y. J., & Kim, J.-H. (2021). Two-dimensional flow boiling characteristics with wettability surface in microgap heat sink and heat transfer prediction using artificial neural network. Journal of Heat Transfer, 143(9), 091601. https://doi.org/10.1115/1.4051602

[17] Selten, F. M. (1997). Baroclinic empirical orthogonal functions as basis functions in an atmospheric model. Journal of the Atmospheric Sciences, 54(16), 2099–2114. https://doi.org/10.1175/1520-0469(1997)054<2099:BEOFAB>2.0.CO;2

[18] Berkooz, G., Holmes, P., & Lumley, J. L. (1993). The proper orthogonal decomposition in the analysis of turbulent flows. Annual Review of Fluid Mechanics, 25, 539–575. https://doi.org/10.1146/annurev.fl.25.010193.002543

[19] Towne, Schmidt, O. T., & Colonius, T. (2018). Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. Journal of Fluid Mechanics, 847, 821–867. https://doi.org/10.1017/jfm.2018.283

[20] Gu, Y., Bottrell, N., & Green, T. C. (2017). Reduced-order models for representing converters in power system studies. IEEE Transactions on Power Electronics, 33(4), 3644–3654. https://doi.org/10.1109/TPEL.2017.2711267

[21] Wang, L., Pan, C., Wang, J., & Gao, Q. (2022). Statistical signatures of u component wall-attached eddies in proper orthogonal decomposition modes of a turbulent boundary layer. Journal of Fluid Mechanics, 944, A26. https://doi.org/10.1017/jfm.2022.495

[22] Bosnjic, Z., Koenigseder, F., & Hartmann, M. (2025). Efficient thermal simulation of power modules using proper orthogonal decomposition. 2025 IEEE Energy Conversion Congress & Exposition Asia (ECCE-Asia), Bengaluru, India. https://doi.org/10.1109/ECCE-Asia63110.2025.11112460

[23] Xiang, L., Zhang, B., Zha, Y., Xing, G., Yang, X., Wang, Z., Cheng, Y., Yu, X., Hu, R., & Luo, X. (2025). Physics-informed proper orthogonal decomposition for accurate and superfast prediction of thermal field. ASME Journal of Heat Transfer, 147(7), 073301. https://doi.org/10.1115/1.4068266

[24] Zhao, Y., Wang, Z., Luo, D., Chen, C., Ji, B., & Li, G. (2024). Multitimescale thermal network model of power devices based on POD algorithm. IEEE Transactions on Power Electronics, 39(4), 3906–3924. https://doi.org/10.1109/TPEL.2023.3340675

[25] Feng, X., Bai, F., Ding, H., & Tao, W. (2024). An improved POD-Galerkin method for rapid prediction of three-dimensional temperature field for an IGBT module. International Communications in Heat and Mass Transfer, 152, 107241. https://doi.org/10.1016/j.icheatmasstransfer.2024.107241