Letters

Computationally Efficient Dynamic Thermal Modeling Based on Dictionary Learning Reconstruction

Xinyue Zhang, *Student Member, IEEE*, Yi Zhang, *Member, IEEE*, Dao Zhou, *Senior Member, IEEE*, Xiaohua Wu, and Huai Wang, *Senior Member, IEEE*

***Abstract*—Accurate and rapid thermal estimation holds immense significance in the analysis of power semiconductors under long-term mission profile, reliability design, and real-time thermal assessment. This letter proposes a novel paradigm shift for thermal estimation of power semiconductors. First, long-term dissipation data are transformed into a limited set of base pulses through orthogonal decomposition. These base pulses are preconverted into corresponding base temperatures, enabling the simplification of long-term thermal estimation by efficient time-shifting and superposition of these base temperatures. Meanwhile, to achieve desired temperature estimation accuracy with a minimal set of base temperatures, we further employ dictionary learning for optimization. To validate the effectiveness of this approach, we compare it against a commercial simulation software and two existing methods. The proposed methodology demonstrates significant advantages in the analysis of long-term mission profile. In addition, we conduct experiments using three distinct standard driving cycles for electric vehicles, all demonstrating the accuracy under highly dynamic loading.**

***Index Terms*—Dictionary learning, long-term, shifting, superposition, thermal estimation.**

I. Introduction

RELIABILITY is a key performance for power electronic applications, such as photovoltaics, electric vehicles, and wind power [1]. An essential challenge to evaluate the reliability is the intensive computations of thermal estimation of long-term mission profiles.

Junction temperature fluctuations stand as a primary contributor to power device failures [2]. However, the effects of junction temperature fluctuations induced by the fundamental frequency are always omitted when assessing reliability in long-term mission profiles. In electric vehicles and traction systems, the magnitude and frequency of the current change dynamically as a result of sharp changes in speed and torque [3], [4], [5]. Moreover, in high-power applications, the junction temperature fluctuations caused by the fundamental frequency exhibit notable periodic thermal patterns [6], [7]. Consequently, it is imperative to consider these factors when conducting thermal estimations for power semiconductors.

In existing studies, the most typical method for long-term thermal estimation is based on a convolution of the power loss and thermal impedance, while this approach is computationally expensive. To address this challenge, the simplified methods were proposed in [8] and [9]. Their prerequisite assumptions are that the power loss is a periodic half-sine wave and can be divided into equivalent square pulses. However, these assumptions are not always valid in different applications [6]. In addition, the method in [8] does not essentially solve the computational burdens, which still relies on repeated convolution and exponential calculations, and the details will be discussed in Section II. A frequency-domain thermal modeling approach is proposed in [10], which can significantly reduce the computation time. Nevertheless, the performance is largely dependent on the fine-tuning of the low-pass filter parameters used in the modeling.

Essentially, the long-term mission profile analysis is processing a long time-series signal. Having this understanding enables us to broaden our perspective beyond the realm of power electronics and draw inspiration from other disciplines. For instance, a profile reconstruction technique is used to extract meaningful frequency features from the time series in mechanical engineering [11]. This method provides initial approximate profiles for shape reconstruction, and has potential to reduce the computing time. Moreover, dictionary learning is a powerful technique for analyzing highly overlapping time series data and uncovering the underlying structure [12]. By identifying the patterns and features associated with time series signals, it is possible to develop more accurate and efficient methods for reconstructing or predicting future values.

This letter proposes a novel paradigm shift for thermal estimation of power semiconductors. First, the limitations of existing thermal modeling methods are presented. Second, a shift- and

Manuscript received 11 July 2023; revised 22 August 2023; accepted 13 September 2023. Date of publication 19 September 2023; date of current version 23 October 2023. This work was supported by Innovation Fund Denmark (IFD) through the Project of Artificial Intelligence for Next-Generation Power Electronics (AI-Power). *(Corresponding author: Yi Zhang.)*

Xiaohua Wu is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: wxh@nwpu.edu.cn).

Xinyue Zhang is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: zhang_xin_yue@mail.nwpu.edu.cn).

Yi Zhang, Dao Zhou, and Huai Wang are with AAU Energy, Aalborg University, 9220 Aalborg, Denmark (e-mail: yiz@ieee.org; zda@energy.aau.dk; hwa@energy.aau.dk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2023.3317249.

Digital Object Identifier 10.1109/TPEL.2023.3317249

superposition-based profile reconstruction method is proposed, and dictionary learning is used to achieve optimal profile reconstruction. Even though the part of results have previously outlined in a conference paper [13], this letter provides a comprehensive discussion of proposed thermal profile reconstruction method. Simultaneously, the following new results have been incorporated:

1) two-step based profiles reconstruction method;
2) performance of dictionary learning results in accuracy and efficiency;
3) efficiency comparison with three existing methods under high precision;
4) dynamic profiles applied to case study.

II. Limitations of Conventional Thermal Modeling Methods

Existing thermal models have been facing several problems in accurately predicting the thermal behavior of electronic devices, including the complexity of the calculations and the inaccuracy of the models. In this section, two typical thermal modeling methods and their limitations are presented.

A. Thermal Estimation by the Convolution Calculation

The typical thermal estimation can be represented as a convolution of the power loss and thermal impedance, which is given by

$$\Delta T_{ja}(t) = \int_0^t \left[\frac{\mathrm{d}P(z)}{\mathrm{d}z}\right] \cdot Z\mathrm{th}(t-z)dz \tag{1}$$

where $Z\mathrm{th}$ is the junction-to-ambient thermal impedance, P is the instantaneous loss of IGBT or diode, and $\Delta T_{ja}(t)$ is the temperature fluctuation from junction to ambient.

However, the convolution calculation is computationally heavy, especially for long mission profiles. This is because with the increase of the set length, a longer convolution kernel hinders the computational speeds, which is challenging for long-term reliability analysis.

B. Simplified Thermal Modeling Method

In [6], a simplified transient thermal modeling method was proposed. In this method, the half-sine loss curve is equivalently divided into square pulses using the principle of equal area. This division enables a simplified convolution operation since the power loss is constant in each time interval. For a single layer Foster thermal network, the temperature at the nth time period is given by

$$\Delta T_{ja}(t) = \Delta T_{ja(n-1)} e^{-\frac{t}{\tau\mathrm{th}}} + P_n R\mathrm{th}(1 - e^{-\frac{t}{\tau\mathrm{th}}}) \tag{2}$$

where $R\mathrm{th}$ and $\tau\mathrm{th}$ are the thermal resistance and thermal time constant, respectively. P_n is the dissipated power in the nth period, and $\Delta T_{ja(n-1)}$ is the temperature fluctuation at the end of $(n-1)$th period.

It can be noted that in (2), the simplified method does not essentially overcome the challenge of convolution, which still needs to repeatedly calculate the previous states and exponential operations for each time point. As power consumption occurs

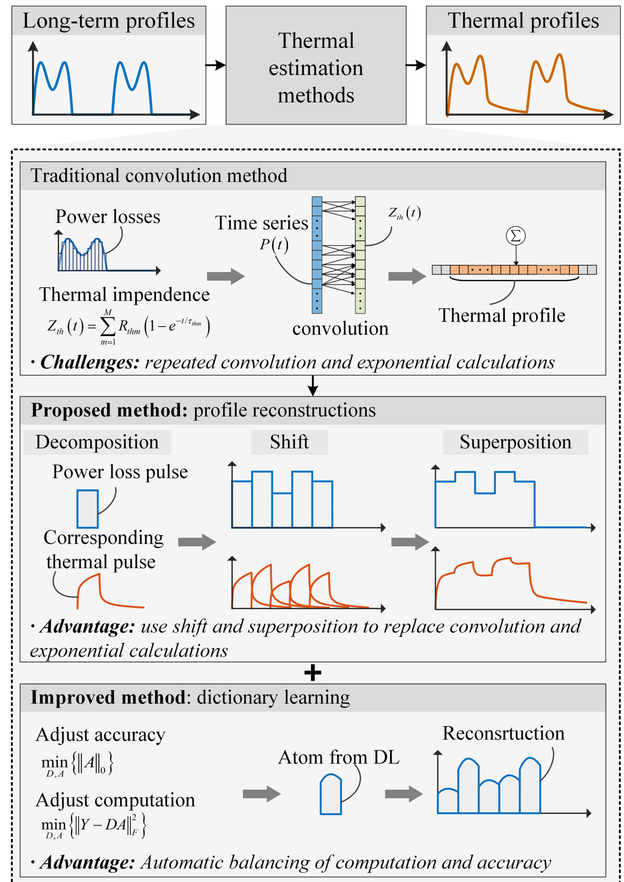

Fig. 1. Schematic diagram of the proposed thermal modeling method and its improved method, compared with the conventional convolution method.

periodically, the temperature of each heating period is determined through iteration and has a time ordered and continuous pattern. Calculating temperature rise for longer time scales can become increasingly complex and require a greater amount of computation time. As shown in Fig. 1, to streamline the thermal analysis process, this letter introduces a dictionary learning-based thermal analysis method that is specifically designed for long-term profiles.

III. Profile Reconstruction-Based Thermal Modeling Method

A. Proposed Thermal Modeling Method

The power loss and thermal models, being linear systems, have time-invariant properties and satisfy the superposition principle, which making it well suited for reconstruction using sparse linear combinations of short pulses [14]. Thus, this letter proposes that thermal profile estimation can be achieved by simply shifting and superimposing some base pulses, which can essentially overcome the repeated convolution and exponential calculations in the existing methods. To model this method, the power loss $P(t)$ can be represented as a linear combination of N base pulses, which is given by

$$P(t) = \sum_{i=0}^{N-1} A_i \cdot p(t - i * \tau) \tag{3}$$

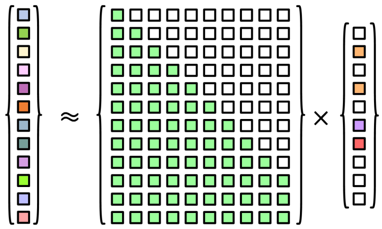

Fig. 2. Schematic of signal sparse reconstruction and dictionary learning.

where $p(t)$ is the base pulse of power loss $P(t)$, τ is the corresponding duration, and A_i is the amplitude of the ith base pulse. It is worth noting that (3) is essentially different from the simple division of (2). The proposed method is set to find the orthogonal base pulses of a period of loss profile. The developed pulse-based method is not limited to equal division or square pulses as (2).

By identifying the base pulses, the corresponding base temperature $b(t)$ can be calculated through the thermal impedance, which is expressed as

$$p(t) \xrightarrow{\text{thermal impedance}} b(t). \tag{4}$$

The convolution and exponential calculations are only applied for the conversion of the base pulses. The final thermal response $\Delta T_{ja}(t)$ can be reconstructed as a linear combination of a set of $b(t)$ with the different amplitudes, which is given by

$$\Delta T_{ja}(t) = \sum_{i=0}^{N-1} A_i \cdot b(t - i * \tau). \tag{5}$$

According to (1) and (5), it is clear that using the proposed method, the convolution and exponential calculation between power loss and thermal impedance can be largely eliminated, reducing the computational burden of the thermal modeling process.

Usually, the square pulse is a commonly used base pulse and is obtained according to the principle of area equivalence, as shown in Fig. 1. However, there is no evidence that this method of base pulse selection is optimal. To further improve the computational efficiency and accuracy of the proposed method, an improved method based on dictionary learning is proposed.

B. Improved Thermal Model Based on Dictionary Learning

To find the most appropriate base pulse to represent the power loss and thermal profile, a shift-invariant dictionary (SID) learning approach is applied [12]. These base pulses are designed to capture the important features of the time-series signals and can be used to reconstruct the original signal in a simplified and efficient way, as shown in Fig. 2. Based on the different objective functions, dictionary learning can be divided into two ways: sparsity-based and error-based.

$$\min_{D,A} \left\{ \|Y - DA\|_F^2 \right\}, \quad \text{s.t. } \forall i, \|A\|_0 \le k \tag{6}$$

$$\min_{D,A} \{\|A\|_0\}, \quad \text{s.t. } \|Y - DA\|_F^2 \le \varepsilon \tag{7}$$

where Y is a set of training samples and $D = [d_1, d_2, \ldots, d_N]$ is a dictionary learned from this training sample, $A = [a_1, a_2, \ldots, a_N]^T$ is the coefficient matrix. ε is the maximum allowable errors, and $\|A\|_0$ represents the zero-order norm, which indicates the number of nonzero coefficients in A, $\|\cdot\|_F$ is the Frobenius norm, and k_0 is the allowable number of nonzero elements in A.

The objective functions used for dictionary learning can either be sparsity-based or error based. The sparsity-based optimization reduces the computational burden, while the error-based optimization increases accuracy. By using a given objective, dictionary learning can automatically strike a balance between these two factors, achieving both computational efficiency and accurate predictions.

The error ε between the training sample Y and the linear combination of atoms $X_P = \sum_{i=1}^{N} a_i \cdot d_i$ in a fundamental period can be defined as

$$\varepsilon = \frac{\max(Y) - \max(X_P)}{\max(Y)}. \tag{8}$$

The time series data can be decomposed into a set of base functions that represent precharacterized behavior, which only needs to be done once during the whole modeling process. It cannot be artificially designed, but rather is obtained through the learning process.

To make the loss model closer to reality, a half saddle wave can be used instead of a half sine wave. When comparing the performance of traditional square pulses and dictionary-learned atomic pulses in reconstructing thermal curves, it was observed that both methods require dividing the loss into five pulses in one cycle in Fig. 3(a). However, the base pulses learned from the dictionary showed a more accurate reconstruction with a smaller error compared with the traditional square pulses. In Fig. 3(b), the improved proposed method using dictionary learning requires only five base pulses to achieve the same level of reconstruction error compared with the conventional method, which requires eight base pulses. This reduction in the number of base pulses results in fewer shifting and stacking operations, leading to the improvement of computational efficiency.

In summary, dictionary learning is used to capture the properties of the profiles, so that no repeated calculations are required when performing junction temperature estimation. The results show that, with proper implementation, dictionary learning can be a useful tool for reducing the computational burden and improving the accuracy of thermal modeling.

IV. Application in Long-Term Thermal Profile

To appraise the accuracy and computation cost of the proposed method, dynamic mission profiles covering current amplitudes from 0 to 30 A and frequency from 0 to 50 Hz were employed. For a fixed time interval, the calculation error decreases as the frequency decreases. Consequently, utilizing variable time interval pulses may offer a feasible method to optimize computation

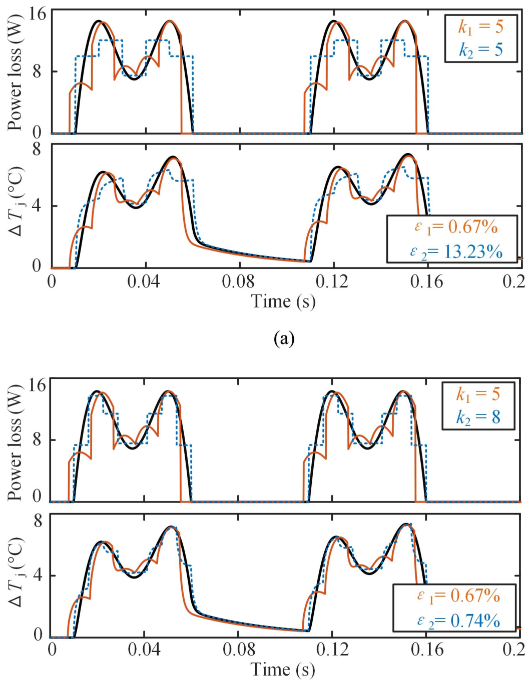

Fig. 3. Comparison of the reconstruction results of dictionary learning pulses (orange) and square pulses (blue). (a) Under the same time. interval. (b) Under a similar error. ($k_{1,2}$: number of the base pulses, $\varepsilon_{1,2}$: error).

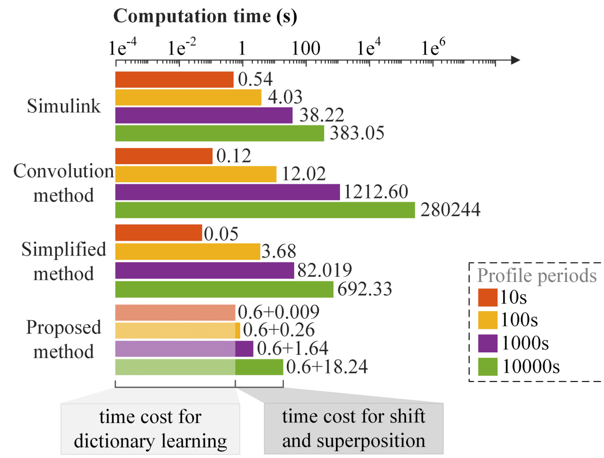

Fig. 4. Computation time of four thermal estimation methods under different load profile periods.

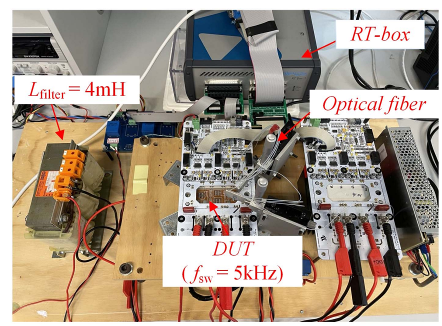

Fig. 5. Photo of the experimental platform.

time during dynamic profile reconstruction. In this letter, the fixed time interval pulse is used to simplify the modeling process. To ensure a maximum allowable error of 1%, this letter utilize the loss profile at 50 Hz as input for the process of dictionary learning. By employing this input, the base pulses from the SID are acquired, and the reconstructed profile is applied by shifting and superposing the obtained pulses.

A. Benchmark of the Proposed Method With Three Conventional Methods

The proposed method is compared with three existing methods, including commercial software tool (MATLAB/Simulink), the traditional convolution method in (1), and the simplified thermal model in [8]. Except for Simulink simulation, the other three methods are implemented using MATLAB scripts, and all evaluations are conducted on an i7-11800H laptop with 24 GB RAM.

The comparison results are shown in Fig. 4. The time cost in the proposed method involves two distinct time-consuming steps: dictionary learning and superposition. Hence, the correlation between computation time and profile periods is not a straightforward ratio. For short-time load profiles (e.g., 10 s and 100 s), the proposed method does not show advantages due to the initial dictionary learning time of 0.6 s. However, as the increase in the length of the load profiles, the computation time has an almost negligible increase compared to the other three methods. It strongly demonstrates the superiority of the proposed method for analyzing long-term mission profiles.

B. Verification Under Long-Term Profiles

To further verify the computational efficiency of the proposed method, the dynamic profiles [15] are also applied to the experimental platform, as shown in Fig. 5. A thermo-optical fiber is used to measure the junction temperatures of the semiconductor chips. Fig. 6 shows the comparisons of the junction temperatures obtained from the proposed method and the measured under the three different dynamic driving cycles HWFET, NYCC, and UDDS.

The results in Fig. 6 indicate that the error in junction temperature estimation varies across different driving cycles due to the dynamic nature of speed and torque. However, all the errors remain within the range of 1%. For example, for the profile over 2000 s under the HWFET, the proposed method takes 3.4 s only to complete the temperature estimation considering thermal coupling. Meanwhile, the estimated result only has 0.66% mean absolute percentage error (MAPE) with the measured temperature.

V. Discussion

A. Errors of Thermal Estimation

The essence of the proposed method centers on providing a novel paradigm shift to achieve a computationally efficient thermal estimation method. However, there are some limitations that warrant further investigation. This letter primarily focuses on the

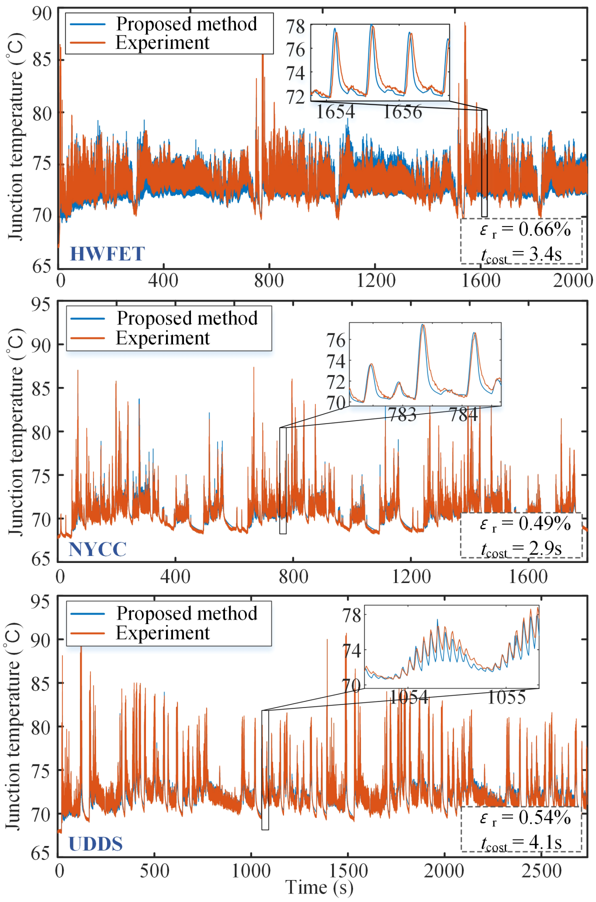

Fig. 6. Comparison of measured and reconstructed junction temperature profiles under dynamic driving cycle profiles. (HWFET: the Highway Fuel Economy Driving Schedule, NYCC: the New York City Cycle, UDDS: the EPA Urban Dynamometer Driving Schedule, MAPE: $\varepsilon_r = \frac{100\%}{N_{\text{end}}} \sum_{t=0}^{N_{\text{end}}} \frac{|T_{j\_\text{exp}}(t) - T_{j\_\text{cal}}(t)|}{T_{j\_\text{exp}}(t)}$.)

thermal transient estimation based on the provided loss profiles and thermal impedance parameters. Therefore, does not delve into discussions regarding aging or the temperature-dependent material effects on the modeling.

B. Feasibility Analysis of Thermal Estimation on Microcontroller Unit (MCU)

This letter assesses the performance of the proposed method in an offline mode, showcasing its notable computational efficiency. Simultaneously, this fast calculation of junction temperature paves the way for real-time online temperature prediction. Consequently, we briefly investigate the feasibility of implementing the proposed method in hardware.

First, given that dictionary learning primarily involves obtaining features from the loss profile, its practical application may use dictionary learning as a preprocessing step on the computer side. Subsequently, the learned base pulses are encoded into an array format in the MCU for subsequent access during temperature calculations. Consequently, real-time temperature predictions can be generated by retrieving temperature values from the corresponding positions of the base pulses and the provided loss data.

VI. Conclusion

In this letter, the profile reconstruction method as a novel paradigm shift is used to simplify the thermal modeling process and provide a more efficient solution for predicting temperature rise in power modules over extended periods. The proposed method significantly reduces the computational burden while maintaining the accuracy of the junction temperature predictions, by avoiding the repetitive convolution and exponential calculations. The sparse matrix is obtained by iterative of the allowable error or sparsity, and linearly superimposed on the dictionary to obtain the temperature response signals. To validate the effectiveness, simulation and experiments are carried out, and the results verified the feasibility of the method. In summary, the proposed method performs well in improving the efficiency of transient thermal analysis of power modules.

References

[1] K. Fischer et al., "Reliability of power converters in wind turbines: Exploratory analysis of failure and operating data from a worldwide turbine fleet," *IEEE Trans. Power Electron.*, vol. 34, no. 7, pp. 6332–6344, Jul. 2019.

[2] S. Yang, A. Bryant, P. Mawby, D. Xiang, L. Ran, and P. Tavner, "An industry-based survey of reliability in power electronic converters," *IEEE Trans. Ind. Appl.*, vol. 47, no. 3, pp. 1441–1451, May/Jun. 2011.

[3] S. Lin, X. Fang, F. Lin, Z. Yang, X. Wang, and T. Taku, "Lifetime prediction of IGBT modules based on mission profiles in traction inverter application," in *Proc. IEEE Veh. Power Propulsion Conf.*, 2019, pp. 1–6.

[4] F. Blaabjerg, H. Wang, I. Vernica, B. Liu, and P. Davari, "Reliability of power electronic systems for EV/HEV applications," *Proc. IEEE*, vol. 109, no. 6, pp. 1060–1076, Jun. 2021.

[5] H. Xia et al., "Impact of loss model selection on power semiconductor lifetime prediction in electric vehicles," in *Proc. IEEE 48th Annu. Conf. Ind. Electron. Soc.*, 2022, pp. 1–7.

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

[7] K. Ma, F. Blaabjerg, and M. Liserre, "Thermal analysis of multilevel grid-side converters for 10-MW wind turbines under low-voltage ride through," *IEEE Trans. Ind. Appl.*, vol. 49, no. 2, pp. 909–921, Mar./Apr. 2013.

[8] Y. Zhang, H. Wang, Z. Wang, Y. Yang, and F. Blaabjerg, "A simplification method for power device thermal modeling with quantitative error analysis," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 7, no. 3, pp. 1649–1658, Sep. 2019.

[9] K. Ma, A. S. Bahman, S. Beczkowski, and F. Blaabjerg, "Complete loss and thermal model of power semiconductors including device rating information," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2556–2569, May 2015.

[10] K. Ma, N. He, M. Liserre, and F. Blaabjerg, "Frequency-domain thermal modeling and characterization of power semiconductor devices," *IEEE Trans. Power Electron.*, vol. 31, no. 10, pp. 7183–7193, Oct. 2016.

[11] S. Zhang et al., "Model-based analysis and quantification of bearing faults in induction machines," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2158–2170, May/Jun. 2020.

[12] K. Skretting and K. Engan, "Sparse approximation by matching pursuit using shift-invariant dictionary," in *Proc. Image Anal. Scand. Conf.*, 2017, pp. 362–373.

[13] X. Zhang, Y. Zhang, D. Zhou, H. Wang, and X. Wu, "Transient thermal modeling of power semiconductors for long-term load profiles," in *Proc. IEEE Appl. Power Electron. Conf. Expo.*, 2023, pp. 2569–2574.

[14] S. Y. Heng, Y. Asako, T. Suwa, and K. Nagasaka, "Transient thermal prediction methodology for parabolic trough solar collector tube using artificial neural network," *Renew. Energy*, vol. 131, pp. 168–179, Feb. 2019.

[15] "Dynamometer Drive Schedules," 2022. [Online]. Available: https://www.epa.gov/vehicle-and-fuel-emissions-testing/dynamometer-drive-schedules