



Aalborg Universitet

AALBORG UNIVERSITY  
DENMARK

## A Sparsity-Promoting Time Domain Evaluation Method for Thermal Transient Measurement of Power Semiconductors

Zhang, Yi; Evgrafov, Anton; Zhao, Shuai; Kalker, Sven; Doncker, Rik W. de

*Published in:*  
IEEE Transactions on Power Electronics

*DOI (link to publication from Publisher):*  
[10.1109/TPEL.2024.3367854](https://doi.org/10.1109/TPEL.2024.3367854)

*Creative Commons License*  
CC BY 4.0

*Publication date:*  
2024

*Document Version*  
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*  
Zhang, Y., Evgrafov, A., Zhao, S., Kalker, S., & Doncker, R. W. D. (2024). A Sparsity-Promoting Time Domain Evaluation Method for Thermal Transient Measurement of Power Semiconductors. *IEEE Transactions on Power Electronics*, 39(6), 7525-7535. Article 10440461. <https://doi.org/10.1109/TPEL.2024.3367854>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.




- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.



# A Sparsity-Promoting Time Domain Evaluation Method for Thermal Transient Measurement of Power Semiconductors

Yi Zhang , *Member, IEEE*, Anton Evgrafov , Shuai Zhao , *Member, IEEE*, Sven Kalker, *Student Member, IEEE*, Rik W. de Doncker, *Fellow, IEEE*,

**Abstract**—This paper investigates evaluation methods of thermal transient measurements to obtain the internal thermal structure of semiconductor devices. First, the study uncovers the limitations of a widely accepted standard method that uses frequency-domain deconvolution. An important finding is that the sideband of the time constant spectrum by the standard method has no physical meaning despite it has been interpreted as a continuous spectrum for a long time. Second, by understanding the limitations of the existing method, the paper proposes an alternative method that remodels the frequency-domain deconvolution as a regularized least squares problem in the time domain. With the benchmark of the true values of several thermal networks based on simulation, the proposed sparsity-promoting method demonstrates several advantages, including a better ability to identify adjacent parameters in the time-constant spectrum and the obtained structure function reducing relative error by an order of magnitude. The influence of varying noise levels has also been evaluated. Finally, a proof-of-concept experiment using a commercial power semiconductor device validates its effectiveness.

**Index Terms**—Reliability, system identification, semiconductor device packaging, transient analysis, thermal analysis.

## I. INTRODUCTION

**T**HERMAL modeling and characterization are crucial in various fields of electronics [1]–[3]. Around two decades ago, a non-destructive method known as network identification by deconvolution (NID) was introduced for evaluating the inner structure of electronic devices (see Fig. 1). By analyzing measurable thermal transient responses, the NID method produces a detailed heat-flow map of the device’s internal structure. This transformative method has revolutionized several areas, including identification of inner structure [4], [5], non-destructive failure analysis [6], thermal model estimation [7], [8], thermal characterization of wide band-gap devices [9], [10],

junction temperature estimation [11], [12], and characterization of materials such as thermal grease [13]. Thanks to its effectiveness, the NID method has been standardized in [14] and widely adopted in industry today.

Despite the widespread acceptance of the NID method, research on this approach over the past two decades has predominantly focused on various applications [6]–[8], [11], new devices [9], [10], [16], and thermal coupling effects [17]. However, little attention has been paid to the limitations or challenges of the NID method itself [18], and the development of this method appears to have stagnated for years. Therefore, it is crucial to revisit the origin of this method. By investigating the first publication in [15] and a series of subsequent efforts [6], [7], [14], [19], the existing NID method exhibits limitations in solving the crucial thermal time constant spectrum as shown in Fig. 1. Specifically, the existing method uses an ill-posed frequency-domain deconvolution. This leads to an indistinct spectrum and the blurred structure function which are introduced as follows.

First of all, the NID method based on a frequency-domain inversion is an ill-posed problem. The unknown structure function is obtained as an inverse Fourier transform of the quotient, where the denominator is the Fourier transform of the convolution kernel. Since the convolution kernel for this problem is a very smooth function, the resulting denominator rapidly tends to zero for higher frequencies, making the deconvolution into an ill-posed problem. As a result, the measurement noise will be strongly enhanced in the identified structure information. This apparent challenge was pointed out when the NID method was originally proposed in [15]. The standard approach to alleviate this issue is to utilize a low-pass filter [14]. This does indeed suppress the effect of the noise but inevitably lowers the resolution of identification [20].

As a result, the essential intermediate of the conventional NID method, i.e., the time constant spectrum, has limited resolution. For instance, the effectiveness of the time constant spectrum is often validated by  $RC$  networks. When the distance between two time constants is more than a decade, see [15], the conventional NID method performs well. In contrast, time constants which are close to each other are more difficult to identify [19].

Meanwhile, reference [15] also constructs a graphic representation of the heat-flow path with the thermal resistance and capacitance map of the structure, which is known as

Manuscript received Month xx, 2xxx; revised Month xx, xxxx; accepted Month x, xxxx. This work was supported in part by Innovation Fund Denmark under the Project of Artificial Intelligence for Next-Generation Power Electronics (AI-Power) and in part by Independent Research Fund Denmark with the grant number 1031-00024B. An earlier preprint version was available in TechRxiv [DOI: 10.36227/techrxiv.21739460.v1].

Y. Zhang and S. Zhao are with AAU Energy, Aalborg University, 9220 Aalborg, Denmark. A. Evgrafov is with the Department of Mathematical Sciences, Aalborg University, 9220 Aalborg, Denmark (e-mail: yiz@energy.aau.dk).

S. Kalker and R. W. de Doncker are with the Institute for Power Electronics and Electrical Drives (ISEA), RWTH Aachen University, 52066 Aachen, Germany (e-mail: post@isea.rwth-aachen.de).

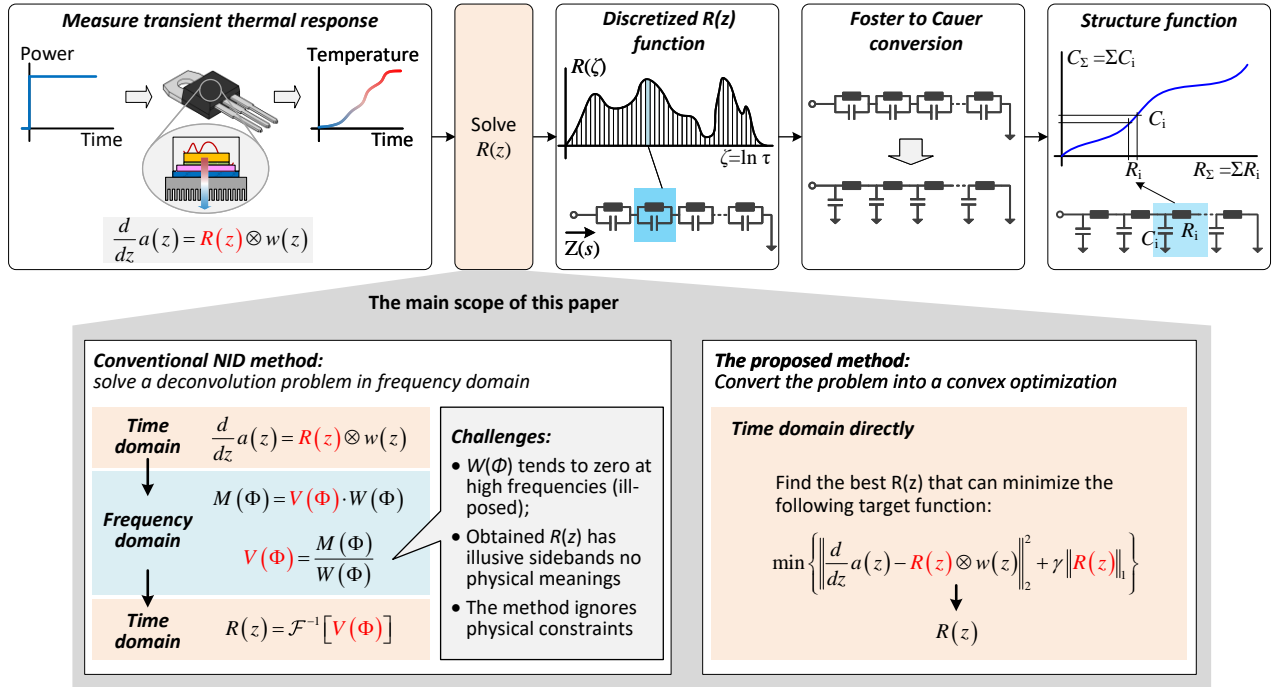


Fig. 1. The framework for internal structure identification in power semiconductors through thermal transient response analysis and the main scope of this paper. (NID represents network identification by deconvolution [15].  $R(z)$  is thermal time constant spectrum and  $V(\Phi)$  is its frequency-domain expression.  $a(z)$  is a normalized step temperature response.  $w(z)$  is a conventional kernel as referred to (6).)

structure function. Theoretically, the jumps in the slope of the structure function correspond to material or geometry changes. Thermal resistance and capacitance, geometrical dimensions, and material parameters may therefore in theory be directly read from the structure function. In practice, however, computed structure functions by the NID method often do not contain obvious slope changes, or instead have many illusive inflection points, which are difficult to relate to structure properties [6]. Further improvements over the state-of-the-art identification methods are thus highly desirable.

Some prior studies have partially investigated the aforementioned challenges. For instance, the utilization of a Bayesian iterative approach [20] for deconvolution computation has been explored, effectively eliminating the necessity for low-pass filtering and achieving better resolution. However, this method necessitates a substantial number of iterations and lacks a clear understanding of convergence properties [21], particularly in scenarios involving complicated spectra.

In contrast with the conventional NID method or the Bayesian method for carrying out the deconvolution in the frequency domain, this paper proposes a sparsity-promoting time domain evaluation method to solve the problem. This work has following contributions:

- 1) This paper comprehensively investigates the limitations of the standard NID method. Apart from the aforementioned ill-posed nature and under-performances, an important finding is that the sideband/lobe of the time constant spectrum by the standard method holds no physical meanings. However, this sideband has been interpreted as a continuous spectrum for a long time.
- 2) Based on a better understanding of the limitations of the

existing approach, this paper proposes an alternative method to compute the measured thermal transient response in the time domain directly. The proposed sparsity-promoting method respects the underlying physical constraints of the problem and does not prevent sharp (high frequency) variations in the recovered quantities.

- 3) With benchmarked against true values of different thermal networks, the proposed method has exhibited a better resolution in the time constant spectrum and a better accuracy in the structure function. For instance, two adjacent time constants with a distance less than a decade can be clearly identified in the time constant spectrum while the conventional method fails to do so. The obtained structure function by the proposed method reduces the relative error by an order of magnitude. A proof-of-concept experiment based on a commercial power semiconductor device is also presented to verify its viability.

## II. LIMITATIONS OF THE STANDARD METHOD AND THE PROPOSED APPROACH

This section briefly revisits the theory of the thermal transient measurement and the conventional NID method. An important contribution is to unveil the three limitations of the conventional approach. In particular, this paper mathematically verifies that the sideband of the thermal time constant spectrum has no physical meanings, although it is often interpreted as a continuous spectrum. Based a better understanding of the limitations, a time-domain evaluation method is proposed to address aforementioned challenges. The proposed method

does not have artificial sidebands and respects the underlying physical constraints of the problem well in the solution.

### A. Mechanism of Thermal Transient Measurement

Thermal transient measurement is a widely accepted method for evaluating the inner structure of electronic devices. By an abrupt power dissipation step onto the device under test, the measured thermal transient response is utilized to generate a heat-flow map of the device structure nondestructively. The following part briefly introduces the mechanism of this method.

For a semiconductor package with multiple materials and layers, the unit step response  $a(t)$  of the thermal structure can be represented as a sum of  $N$  individual exponential terms with different time constants  $\tau_i$  and magnitudes  $r_i$ , that is,

$$a(t) = \sum_{i=1}^N r_i \cdot \left(1 - e^{-t/\tau_i}\right). \quad (1)$$

It should be noted that the parameters  $\tau_i$  and  $r_i$  here represent the inner structure physically, which is fundamentally different from a Foster network obtained by curving fitting [22].

Considering the heat propagation along the semiconductor packaging across a wide range of time constants, a logarithmic time is introduced as

$$z = \ln(t) \quad \text{and} \quad \zeta = \ln(\tau). \quad (2)$$

A continuous expression of (1) is now introduced

$$a(z) = \int_{-\infty}^{\infty} R(\zeta) \left(1 - e^{-\exp(z-\zeta)}\right) d\zeta, \quad (3)$$

where  $R(\zeta)$  is a density function with respect to the logarithmic time, also known as the thermal *time constant spectrum* [15]. Note that (1) is formally recovered by putting

$$R(z) = \sum_{i=1}^N r_i \delta(z - \ln \tau_i), \quad (4)$$

that is a linear combination of Dirac  $\delta$ -functions.

By differentiating both sides of the convolution integral (3), the obtained integral equation is given by

$$\frac{da}{dz}(z) = \int_{-\infty}^{\infty} w(z - \zeta) R(\zeta) d\zeta, \quad (5)$$

where

$$w(z) = e^{z - \exp(z)}. \quad (6)$$

For the system identification of the semiconductor package, the problem essentially becomes to compute  $R(z)$  from (5) given the measured transient response  $a(z)$ . Mathematically, equation (5) is the most iconic example of an ill-posed problem, the Fredholm integral equation of the first kind. To make things even worse, the measurement noise in  $a(z)$  is further amplified by the presence of the differentiation operator.

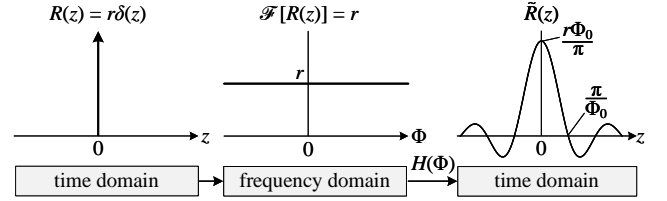


Fig. 2. The process of converting a single time constant into frequency domain and re-inverting into time domain by conventional NID method. The regenerated time-constant spectrum is modulated by the applied low-pass filter  $H(\Phi)$ .

### B. The Conventional NID Method in Frequency Domain

Owing to the convolution structure of (5), the conventional NID method utilizes a frequency-domain method based on Fourier transform to solve  $R(z)$ . Namely, the computation of the Fourier transform of both sides of (5) is given by

$$M(\Phi) = V(\Phi) \cdot W(\Phi), \quad (7)$$

where  $\Phi$  is the generalized frequency corresponding to the logarithmic time  $z$ .  $M(\Phi)$ ,  $V(\Phi)$ , and  $W(\Phi)$  are the Fourier transforms of  $da/dz$ ,  $R(z)$ , and  $w(z)$ , respectively. Equipped with this knowledge the computation of the  $R(z)$  is expressed as the inverse Fourier transform  $\mathcal{F}^{-1}$  of a quotient in the frequency domain, that is

$$\begin{aligned} V(\Phi) &= M(\Phi)/W(\Phi), \\ R(z) &= \mathcal{F}^{-1}[V(\Phi)]. \end{aligned} \quad (8)$$

### C. Investigation of the Limitations of the NID Method

The standard NID method of (7) and (8) based on the frequency domain has three limitations as follows.

1) *The ill-posed limitation*: It is evident that formally expressing the solution of (5) into (8) does not alleviate the ill-posed nature of the problem. Indeed, in the context of (8) the ill-posedness manifests itself through the fact, that owing to the smoothness of the convolution kernel (6) its Fourier transform appearing in the denominator of (8) will be diminishingly small for large frequencies  $\Phi$ . In this way, any high-frequency measurement and discretization noise present in  $M(\Phi) = \mathcal{F}[da/dz] = 2\pi j\Phi \mathcal{F}[a]$  will be significantly amplified in the inversion process [23]. Note that the deteriorating effect of the differentiation operator present in (5) is particularly apparent in this representation.

2) *Physically meaningless sidebands in the time-constant spectrum*: To alleviate the ill-posed limitation, the standard method [14] applies a low-pass filter in the formula (8) prior to evaluating the inverse Fourier transform. While this deals with the high-frequency noise amplification, it causes artificial side bands in the obtained time-constant spectrum  $R(z)$ . More importantly, these artificial side bands have been regarded as a continuous form of the time-constant spectrum in existing studies [15], [23], which is a misunderstanding. As shown in Fig. 2, to demonstrate this effect, a simple system with a single time constant is assumed, that is

$$R(z) = r\delta(z) \quad (9)$$

and an ideal low-pass filter  $H(\Phi)$ , which is expressed as

$$H(\Phi) = \begin{cases} 1, & |\Phi| \leq \Phi_0 \\ 0, & \text{else} \end{cases}, \quad (10)$$

where  $\Phi_0$  is the cut-off frequency of the filter. By converting (9) into the frequency domain and applying the low-pass filter, the re-generated time-constant spectrum  $\tilde{R}(z)$  via the inverse Fourier transformation is expressed as

$$\tilde{R}(z) = \mathcal{F}^{-1} \{ \mathcal{F} [R(z)] \cdot H(\Phi) \} = \frac{r \sin(\Phi_0 z)}{\pi z}, \quad (11)$$

which is clearly not identical to (9). The original amplitude  $r$  is modulated by the filter parameter  $\Phi_0/\pi$ . Meanwhile, the generated side band with a width of  $2\pi/\Phi_0$  is purely related to the filter without any physical meaning. However, in the existing studies, this side band is often explained as a continuous time-constant spectrum, which is a misunderstanding.

3) *Ignoring physical constraints:* In passing, we also note that the deconvolution formula (8) does not respect the physical constraints of the problem. As a result, the obtained spectrum  $R(z)$  may have negative amplitudes which has no physical meaning. It can be seen from  $\tilde{R}(z)$  in Fig. 2 and following simulations and experiments.

#### D. Proposed Sparsity-Promoted Time Domain Evaluation Method

Armed with a better understanding of the limitations of the conventional NID method, this paper proposes to apply inversion of (5) directly in time domain. Compared to the inversion in the frequency domain using Fourier transformation, the proposed method is able to address the aforementioned challenges. Thus, the problem (5) is converted into

$$\inf_{R(\cdot) \geq 0} [F(R) + G(R)], \quad (12)$$

where

$$F(R) = \int_{-\infty}^{\infty} \left[ \frac{da}{dz}(z) - \int_{-\infty}^{\infty} w(z - \zeta) R(\zeta) d\zeta \right]^2 dz, \quad (13)$$

$$G(R) = \gamma \int_{-\infty}^{\infty} |R(\zeta)| d\zeta,$$

that is, the problem (5) is converted into a standard  $\ell_1$ -regularized constrained least squares formulation. It is noted that one needs to seek  $R(\cdot)$  in the space of measures to guarantee that the infimum in (12) is attained, which is of course consistent with (4). Aside from this technical difficulty, classical splitting-based convex optimization algorithms [24], [25] are available for solving discretized versions of (12). The first term,  $F$ , in (12) represents the data fidelity, which keeps the consistency between the data predicted by the model and the measured data. The second term,  $\ell_1$ -regularization term  $G$ , is sparsity-promoting, so that we expect that the obtained constant spectrum  $R(\cdot)$  will be zero apart from narrow peaks around  $\ln \tau_i$ , see (4). In this way, the proposed method is able to improve the resolution over the standard NID method. Parameter  $\gamma$  is adjusted to balance the two terms. In this paper, we use L-curve to select a proper  $\gamma$ . L-curve is a log-log plot of the norm of  $G(R)$  versus  $F(R)$ , which displays

the trade-off between the regularized solution and the data fidelity. For a better balancing of these two terms, the parameter  $\gamma$  near the corner of the L-curve is selected. More theories about L-curve should be referred to [26].  $R(\cdot) \geq 0$  guarantees the underlying physical constraints of the problem that the spectrum density function is non-negative. The details about the numerical solutions utilized in this work are provided in Appendix. A similar work [27] also applied optimization efforts to thermal problems, but they aimed to a good curve fitting of the Foster network and had a different target from the proposed method.

### III. COMPARISON OF THE CONVENTIONAL NID METHOD AND THE PROPOSED APPROACH BASED ON FOUR DIFFERENT CASES

To compare the effectiveness of the conventional NID method and the proposed method, a lumped circuit is modeled in simulation with precisely known values of the thermal resistance and capacitance. This verification method is also used in [15]. However, in contrast to [15] using a Foster network, this paper uses a Cauer network to associate the thermal resistance and capacitance to the different physical regions of semiconductors. Moreover, the different noise levels are also evaluated. The corresponding thermal time constant spectrum and the structure function of the NID method can be easily obtained by the software attached in the standard JESD 51-14 [28] or the commercial tool [29], thus the process is not repeated here.

#### A. Test of the Conventional NID Method on Three-Layer Thermal Networks

As shown in Fig. 3(a), a three-layer Cauer network with two sets of parameters is used, which is referred to as case 1 and case 2, respectively. The corresponding thermal transient responses are shown in Fig. 3(b), where the curves are across seven decades from  $10^{-6}$  s to 10 s. The parameter variation of the second-order RC pair of case 2 causes a larger steady-state thermal impedance and a different transient process.

By applying the conventional NID method, the transient curve of case 1 is converted into a time constant spectrum  $R(z)$  as shown in Fig. 3(c). The time constant spectrum has three peaks, which essentially depict the three dominated poles of the evaluated network. However, each peak in the spectrum  $R(z)$  has a side band with a width of around a decade. For instance, the first peak in Fig. 3(c) has a side band of around  $1.1 \times 10^{-4}$  s, with the other two peaks demonstrating similar behavior. It should be noted that the original network has three RC pairs only. The sidebands of  $R(z)$  with intensive minor magnitudes around the three peaks reveals that the conventional NID method does not reconstruct the original system well. Moreover, the negative amplitudes of the time constant spectrum are also in conflict with the underlying physics. The causes of these under-performances are investigated in Section II-C.

The structure function of the NID method is further obtained as shown in Fig. 3(d). The true values of the RC network are marked with red dots. The inflection points of the structure function should in theory allow us to identify the parameter

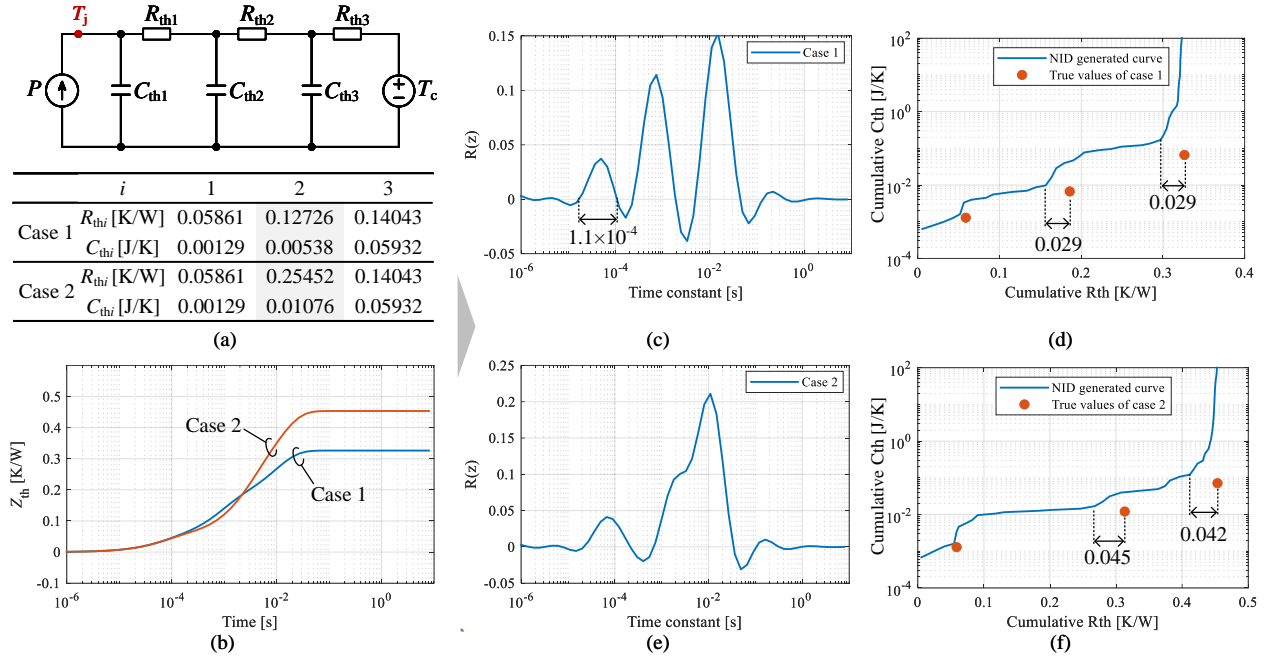


Fig. 3. Two cases of the three-layer Cauer networks based on the conventional NID method: (a) the network and the  $RC$  values, (b) thermal transient responses, (c) time constant spectrum of case 1, (d) comparison of the structure function and the true values of case 1, (e) and (f) are the corresponding time constant spectrum and structure function of case 2.

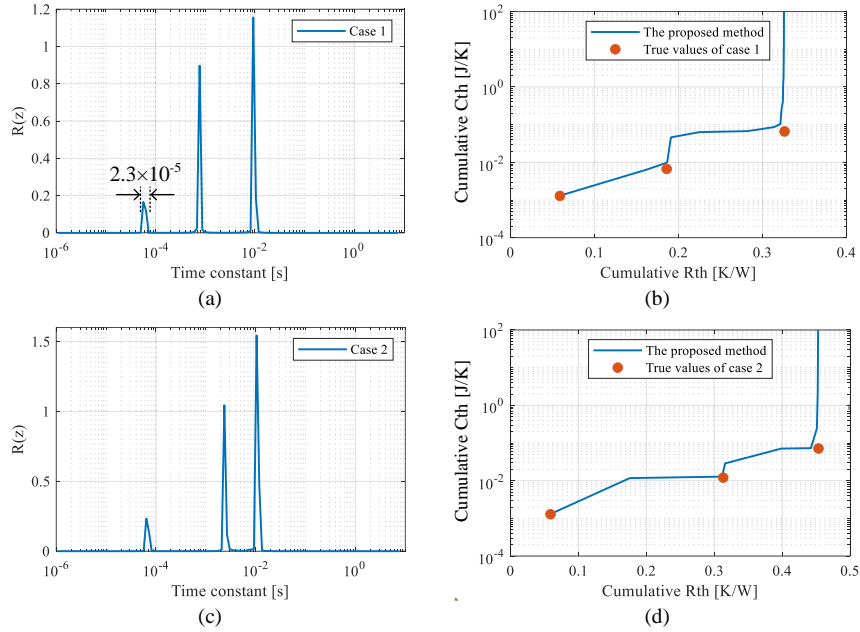


Fig. 4. Proposed method for the three-layer Cauer networks in Fig. 3: (a) time constant spectrum of case 1, (b) obtained structure function vs. the true values of case 1, (c) and (d) are corresponding results for case 2.

variations, that is, thermal properties of different layers of semiconductor packages. When comparing the generated structure function with the true values for this benchmark, it can be seen that the first true value coincides with the inflection point of the curve. However, the structure function curve starts to separate for the next two true values, which have around 0.029 K/W (9% relative to the total thermal resistance) distance from their nearest major inflection points. Meanwhile, multiple small illusive inflection points are visible along the curve but

without any physical meaning. It is therefore challenging to accurately infer the parameter values and their variations simply by the inflection points of the structure function curve.

The time constant spectrum and the corresponding structure function of case 2 are shown in Figs. 3(e) and (f), respectively. With the increase of the second-layer  $RC$  values in case 2, the last two peaks of the spectrum  $R(z)$  are merged, which means that the limited resolution of the conventional NID method makes it challenging to identify two adjacent time constants.

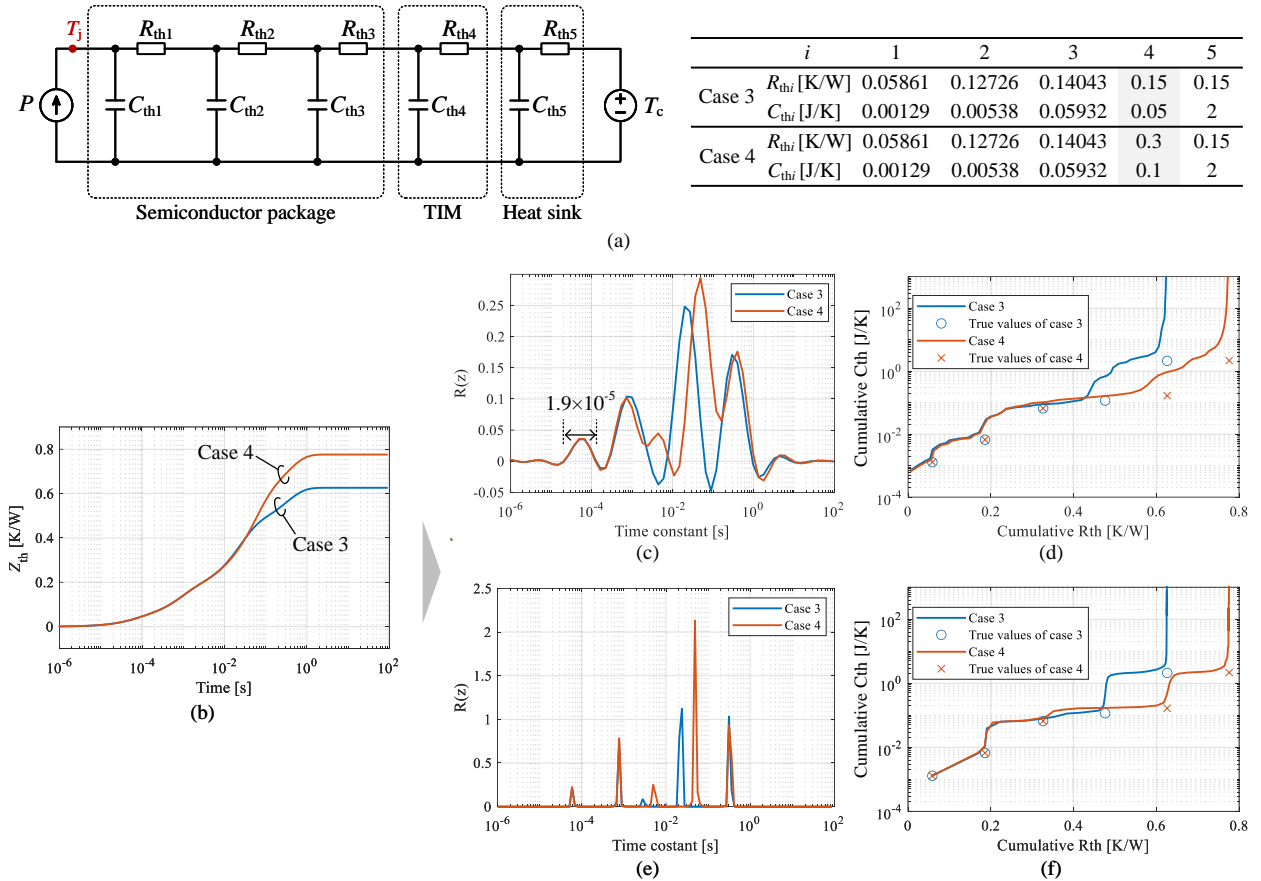


Fig. 5. Comparison of the conventional NID method and the proposed method in two more cases for consistency verification: (a) the five-layer thermal network and two sets of parameters, (b) transient thermal responses, (c) time constant spectrums with the NID method, (d) the structure functions vs. the true values with the NID method, (e) and (f) are corresponding results with the proposed method.

Similarly, the structure function of case 2 is shown in Fig. 3(e). The inflection points of the curve are not well aligned with the true values with a even larger distance (around 0.045 K/W, or 10% of the total thermal resistance) than in case 1.

### B. Test of the Proposed Method on Three-Layer Networks

With applying the proposed method, the computed time constant spectrum  $R(z)$  and the structure functions are shown in Fig. 4. For case 1, the obtained spectrum  $R(z)$  has three peaks as shown in Fig. 4(a). Compared to the conventional NID method, each peak of the proposed method is much more pronounced and without any negative amplitudes. It reveals that  $R(z)$  is computed with a much better resolution using the present approach, which has thus better potential for resolving thermal time constants located closer to each other. Moreover, the generated structure function is more distinctive as shown in Fig. 4(b). The major inflection points coincide well with the true values. The distance between the inflection points and the true values is less than 1% of the total thermal resistance.

Similarly, the computations for case 2 based on the proposed method are shown in Figs. 4(c) and (d), respectively. In contrast to the conventional NID method, which cannot distinguish the second and the third peaks in case 2, the proposed method clearly depicts the two peaks without any overlaps. It confirms our hypothesis that the proposed method has better resolution

for identifying two thermal time constants that are close to each other. Meanwhile, the corresponding structure function also reflects the true values as shown in Fig. 4(d).

### C. Comparison based on Five-Layer Thermal Networks

To further examine the consistency of the proposed method, we investigate two more cases with different structures and parameters as shown in Fig. 5(a). The five-layer Cauer network is typically used to represent a simple power semiconductor assembly, including the semiconductor package, thermal interface material (TIM), and the heat sink. In the standard JEDEC JESD 51-14 [14], a transient dual interface test method is proposed to identify the junction-to-case thermal resistance of semiconductor devices. Two thermal impedance measurements with different TIM conditions are applied. The structure function at the separation point of these two measurements is defined as junction-to-case thermal resistance. Thus, the varied parameters of  $R_{th4}$  and  $C_{th4}$  are used to mimic the different TIM conditions. The simulated transient thermal impedances are therefore shown in Fig. 5(b).

Two time constant spectra  $R(z)$  based on the conventional NID method are shown in Fig. 5(c). First, the obtained  $R(z)$  has four major peaks for case 3 while five peaks for case 4 although both cases are five-layer Cauer networks. Meanwhile, each peak has a wide spread, making it possible for two neighboring peaks



TABLE I  
COMPARISON OF THE RESULTS OF THE STUDIED FOUR CASES.

	Layers of the network	Number of peaks in $R(z)$		Maximum distance <sup>1</sup> [K/W]		Relative error <sup>2</sup>	
		Conventional	Proposed	Conventional	Proposed	Conventional	Proposed
Case 1	3	3	3	$2.9 \times 10^{-2}$	$5.0 \times 10^{-4}$	8.9%	0.2%
Case 2	3	2	3	$4.5 \times 10^{-2}$	$1.7 \times 10^{-3}$	9.9%	0.4%
Case 3	5	4	5	$5.7 \times 10^{-2}$	$3.7 \times 10^{-3}$	9.1%	0.6%
Case 4	5	5	5	$8.7 \times 10^{-2}$	$8.0 \times 10^{-3}$	11.2%	1.0%

<sup>1</sup>Distance between the true value and the inflection point of the structure function curve

<sup>2</sup>Relative error = maximum distance  $\div$  total thermal resistance

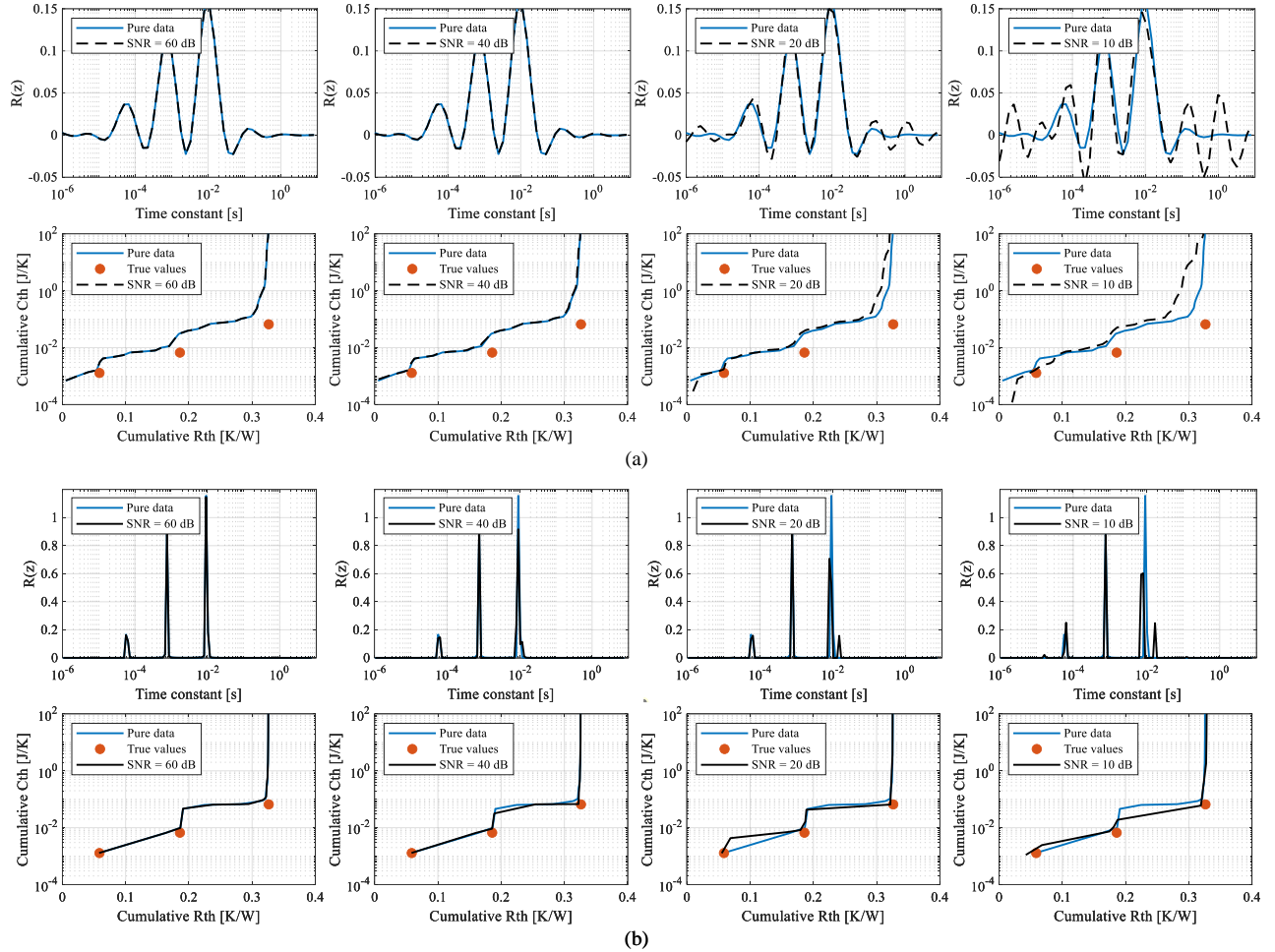


Fig. 6. Influences of varying noise levels in the thermal transient response on the resulting thermal time constant spectrum and structure function based on the parameters of case 1: (a) using the conventional NID method and (b) employing the proposed approach. (SNR: signal-to-noise ratio)

along the spectrum to merge and become indistinguishable. Meanwhile, apart from the time constants within the range from  $10^{-5}$  s to 1 s, some small peaks exist outside this range. For example, a minor peak exists around 5 s, yet neither of the considered systems has a time constant around 5 s since the transient response has reached the steady-state after 2 s. Thus, the minor peaks above 1 s are caused by the methodological artifacts of the NID method instead of any physical meaning.

Furthermore, the NID method generates the structure functions as shown in Fig. 5(d). Obviously, the last two true values are not well aligned with the inflection points of the curves. The distance between the nearest inflection point and the true

value is more than 0.02 K/W for both cases. Moreover, the structure function has many minor inflection points as well. In the standard [14], the separation point of the two structure functions is used to identify the location of the parameter starting to change. Comparing the structure functions of case 3 and case 4, the separation point occurs slightly ahead of the third true value.

The proposed method also generates the time constant spectrum and the structure functions as shown in Figs. 5(e) and (f), respectively. In Fig. 5(e), both cases have five peaks within the range from  $5 \times 10^{-5}$  to 1 s. Beyond this range, there is no noticeable peaks. The parameter change of the 4<sup>th</sup>-layer

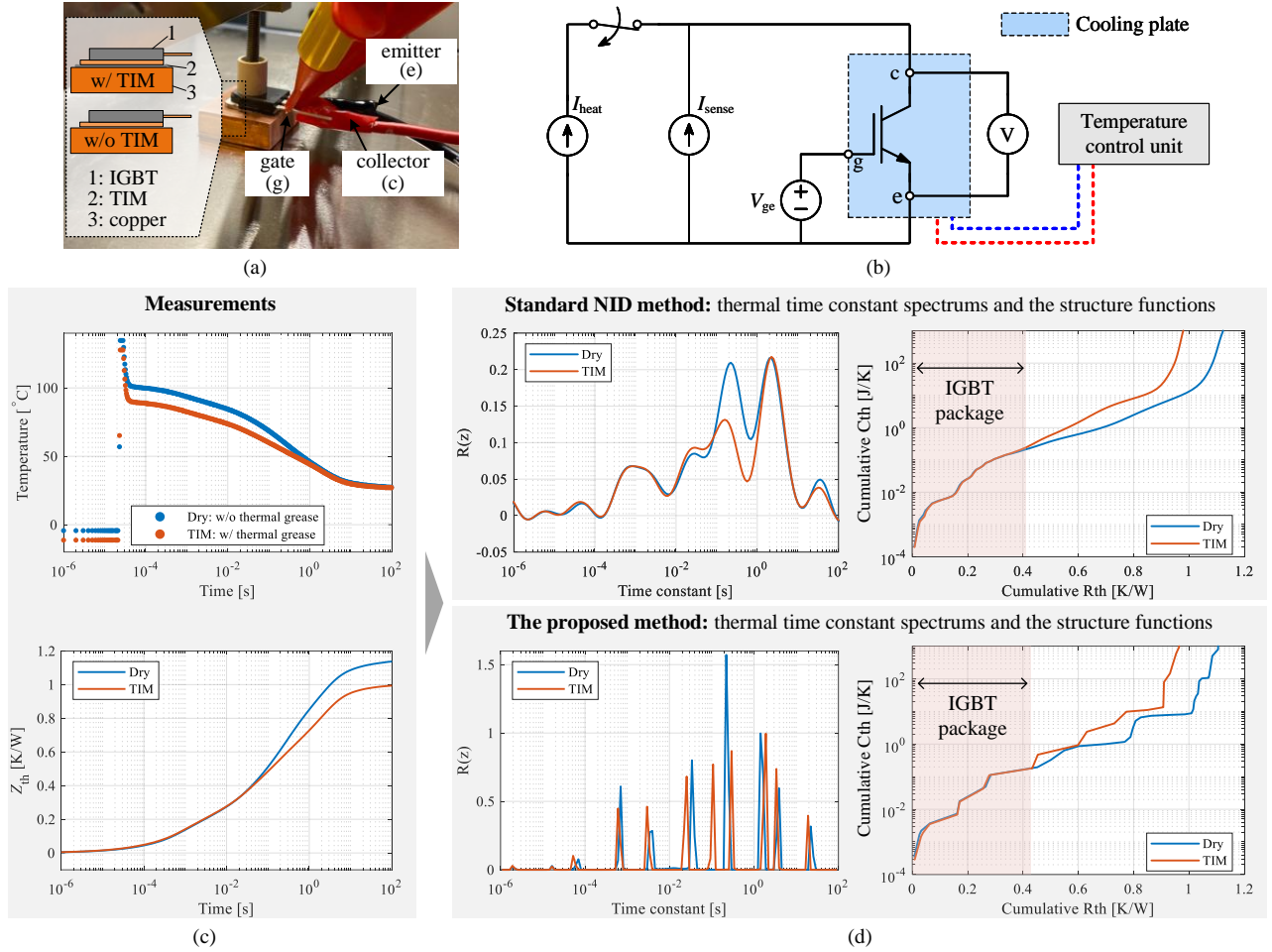


Fig. 7. Experimental setup and results: (a) the photo of the experimental setup, (b) the circuit configuration of the setup, (c) measured transient junction temperatures and their thermal impedance curves, and (d) the thermal time constant spectrums and the structure functions based on the conventional NID method and the proposed method, respectively. (IGBT: insulated-gate bipolar transistor, TIM: thermal interface material)

of the Cauer network causes different time constants in  $R(z)$ . Meanwhile, the structure functions are shown in Fig. 5(f). All true values agree well with the inflection points and the generated curves are clean and with few illusive inflections. When we take the separation point into account, the two curves are exactly divergent at the third true value.

#### D. Summary of the Four Different Benchmark Cases

The key results of the four aforementioned cases are summarized in Table I in terms of both the  $R(z)$  and the structure functions. Case 1 and case 2 are three-layer networks while case 3 and case 4 are five-layer networks. Comparing the time constant spectrum of the two methods, the conventional NID method generated  $R(z)$  has three and five peaks for case 1 and case 4, respectively, but only two and four peaks for case 2 and case 3. In contrast, the obtained  $R(z)$  by the proposed method has the exact number of peaks as the provided network for all cases. Moreover, the relative errors between the generated structure functions and the true values are also compared. The proposed method reduces the relative error to around 1% from 10% of the conventional method. It suggests that the proposed method is better in reflecting the parameter

changes of the thermal network. The four cases with different networks and parameters verify the consistency of the proposed method.

#### E. Impacts of Noise on the Obtained Results

Apart from the above validations based on pure data, this part evaluates the impact of noise on the obtained results. Four different artificial noise levels are added into the thermal transient response of case 1. Fig. 6 shows the obtained thermal time constant spectrum and structure functions based on the conventional NID and the proposed methods. The NID method is more robust when the noise level is low (e.g., SNR = 60 or 40 dB) because of applying a low-pass filter. However, as the noise level increases (e.g., SNR = 20 or 10 dB), the inherent ill-posed nature may amplify leakage noise from the filter, resulting in numerous artificial peaks in the time constant spectrum and distortion of the structure function [see Fig. 6(a)]. Conversely, the proposed method is indeed more sensitive to noise. Although a slight increase of the noise from SNR = 60 to 40 dB, a small deviation in the obtained structure function is visible as shown in Fig. 6(b). With higher noise levels (e.g., SNR = 20 or 10 dB), an artificial peak emerges around  $10^{-2}$  s.

TABLE II  
EXPERIMENTAL PARAMETERS.

Parameter	Value
Heating current $I_{\text{heat}}$	30 A
Sensing current $I_{\text{sense}}$	100 mA
Gate voltage $V_{\text{ge}}$	15 V
Heating time	100 s
Cooling time	100 s
Coolant temperature	25°C
Temperature sampling frequency	1 MHz

However, the inflection points of the structure function remain more consistent compared to the NID method when SNR = 10 dB. Instead of adding a purely numerical study with artificially added noise we have opted out to presenting an experimental study in the following, which thereby includes the realistic noise.

#### IV. EXPERIMENTAL VERIFICATION

To verify the effectiveness of the proposed method, a commercial 1200-V/20-A IGBT device with TO-3P package is selected (FGA20N120FTD [30]). The experimental setup is shown in Fig. 7(a). The copper block with a large thermal capacitance provides a constant temperature reference for the device under test. The temperature of the copper block is managed by a liquid-cooled system. Two different interface conditions have been applied between the device and the copper block, namely, with and without thermal grease, which are denoted as TIM and dry, respectively.

The measuring circuit is shown in Fig. 7(b). The heating current  $I_{\text{heat}}$  is used to heat up the device and provides sufficient temperature gradients within the device. The sensing current  $I_{\text{sense}}$  is a small current that is used to measure the junction temperature by utilizing the forward voltage  $V_{\text{ce}}$  as a temperature-sensitive electrical parameter. The gate voltage  $V_{\text{ge}}$  keeps the channel of the IGBT activated. An external temperature control unit keeps the cooling plate at a constant temperature. The exact parameters are listed in Table II. Subsequently, the measured transient thermal responses under two different interface conditions are shown in Fig. 7(c). The varied interface conditions cause a separation between the two curves, which is similar to the aforementioned simulation cases.

Based on the conventional NID method, the obtained time constant function  $R(z)$  and the structure function are shown in the top of Fig. 7(d). The function  $R(z)$  has positive amplitudes almost along the entire spectrum from  $10^{-6}$  to  $10^2$  s, which is in contrast to the fact that the tested power semiconductor device is composed of limited materials and three major layers. As the obtained structure functions, the two curves separate around 0.4 K/W, but the exact separating point is difficult to determine. Within the IGBT package, the two curves have multiple inflection points which are difficult to identify the parameter changes in the system.

On the other hand, the proposed method generates the time constant spectrum  $R(z)$  and the structure function are shown in the bottom of Fig. 7(d). The function  $R(z)$  only has around eight major peaks across the time constants from  $10^{-6}$  to  $10^2$  s, which emphasizes the tested device with the cooling

system has a few numbers of dominant thermal constants. Furthermore, the generated structure functions exhibit three clear inflection points within the package, which agree with that the tested device consists of three major layers. Moreover, the obtained structure functions have a more distinct separation point between the two interface measurement, which is easier for identifying separating boundary.

#### V. CONCLUSIONS

The present study sets out to unveil the limitations of the conventional NID method and proposes an alternative method for the thermal transient measurements. A series of simulation and experimental case studies have shown that the conventional NID method is limited in its ability to estimate the parameters of the original thermal network based on the thermal transient measurements. The evidence includes that the obtained thermal time constant spectrum has smeared distributions, which makes it difficult to identify the adjacent time constants within a decade, and the corresponding structure function cannot reflect the thermal network parameters well.

This paper remodeled the frequency domain deconvolution of the conventional NID method into a regularized least squares problem in the time domain directly, which brings several advantages. In particular, the proposed method has a better resolution for time constant spectrum, which in turn implies better capability for identifying the neighboring thermal time constants. Furthermore, the relative error in terms of the structure function based on benchmark studies reduces to 1% from the 10% provided by the conventional method. These results suggest that the proposed methods is a superior way of identifying the thermal transient measurements.

Finally, due to the limited scope of this article, some open questions still remain to be answered within the framework of the thermal transient measurement as shown in Fig. 1. It is imperative to revisit this framework thoroughly, identify potential challenges, and address them with new perspectives, which would be a fruitful direction.

#### APPENDIX

This appendix provides details about the numerical solutions utilized in this work. The proposed  $\ell_1$ -regularized least squares method is proceed as follows. First, the logarithmic time interval  $[z, \bar{z}]$  is subdivided into  $K$  subintervals uniformly. The endpoints of the time interval correspond to the first and the last measurement points. The unknown time constant spectrum is thus represented as the form

$$R(z) \approx \sum_{k=1}^K R_k N_k(z),$$

where  $\{N_k\}_{k=1}^K$  form the standard basis for the piece-wise constant polynomials. Note that the non-negativity requirement on  $R$  simply amounts to the lower bounds  $R_k \geq 0$  for the

underlying physics. With these definitions, the functions  $F$  and  $G$  in (13) are approximated as

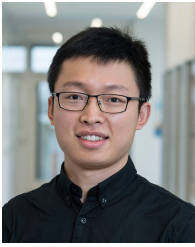
$$F(R) \approx \sum_{k_1=1}^K \sum_{k_2=1}^K R_{k_1} R_{k_2} \times \int_{-\infty}^{\infty} \int_{\bar{z}_{k_1}}^{\bar{z}_{k_1}} \int_{\bar{z}_{k_2}}^{\bar{z}_{k_2}} w(z - \zeta_1) w(z - \zeta_2) N_{k_1}(\zeta_1) N_{k_2}(\zeta_2) d\zeta_2 d\zeta_1 dz - 2 \sum_{k=1}^K R_k \int_{\bar{z}_k}^{\bar{z}_k} \int_{-\infty}^{\infty} \frac{da}{dz}(z) w(z - \zeta) N_k(\zeta) dz d\zeta + \text{const},$$

$$G(R) \approx \sum_{k=1}^K R_k \int_{\bar{z}_k}^{\bar{z}_k} N_k(z) dz,$$

where const is the term independent from the vector of unknowns  $\{R_k\}_{k=1}^K$ , and therefore does not have to be computed. In the formula above, the notation  $(\bar{z}_k, \bar{z}_k)$  is introduced for the support of the basis function  $N_k$ . The derivative  $da/dz$  is approximated using central finite differences based on the measurement points. All integrals in the formulas above are computed using numerical quadratures. In this way, the problem is reduced into a convex quadratic optimization problem with bound constraints, which can be efficiently solved using standard optimization solvers. We utilize Matlab's `quadprog` interior point solver for this purpose.

## REFERENCES

- [1] X. Dong, A. Griffio, D. Hewitt, and J. Wang, "Reduced-order thermal observer for power modules temperature estimation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 12, pp. 10085–10094, Dec. 2019.
- [2] Y. Peng, Q. Wang, H. Wang, and H. Wang, "An on-line calibration method for TSEP-based junction temperature estimation," *IEEE Trans. Ind. Electron.*, vol. 69, no. 12, pp. 13616–13624, Dec. 2022.
- [3] Y. Zhang, H. Wang, Z. Wang, F. Blaabjerg, and M. Saeedifard, "Mission profile-based system-level reliability prediction method for modular multilevel converters," *IEEE Trans. Power Electron.*, vol. 35, no. 7, pp. 6916–6930, Jul. 2020.
- [4] Y. Yang, R. Master, G. Refai-Ahmed, and M. Touzelbaev, "Transient frequency-domain thermal measurements with applications to electronic packaging," *IEEE Trans. Components, Packag. Manuf. Technol.*, vol. 2, no. 3, pp. 448–456, Mar. 2012.
- [5] D. Schweitzer, H. Pape, and L. Chen, "Transient measurement of the junction-to-case thermal resistance using structure functions: Chances and limits," in *Proc. 24th IEEE Semicond. Therm. Meas. Manag. Symp.*, pp. 191–197, 2008.
- [6] P. Szabó, O. Steffens, M. Lenz, and G. Farkas, "Transient junction-to-case thermal resistance measurement methodology of high accuracy and high repeatability," *IEEE Trans. Components Packag. Technol.*, vol. 28, no. 4, pp. 630–636, Dec. 2005.
- [7] G. Farkas, D. Schweitzer, Z. Sarkany, and M. Rencz, "On the reproducibility of thermal measurements and of related thermal metrics in static and transient tests of power devices," *Energies*, vol. 13, no. 3, 2020.
- [8] J. Zhang, X. Du, Y. Yu, S. Zheng, P. Sun, and H.-M. Tai, "Thermal parameter monitoring of IGBT module using junction temperature cooling curves," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 8148–8160, Oct. 2019.
- [9] Y. Zhang, Y. Zhang, Z. Xu, Z. Wang, H. Wong, Z. Lu, and A. Caruso, "A guideline for silicon carbide MOSFET thermal characterization based on source-drain voltage," in *Proc. IEEE Applied Power Electron. Conf. Expo. (APEC)*, pp. 378–385, 2023.
- [10] Y. Zhang, Y. Zhang, Z. Xu, Z. Wang, V. H. Wong, Z. Lu, and A. Caruso, "Guideline for reproducible SiC MOSFET thermal characterization based on source-drain voltage," *IEEE Trans. Ind. Appl.*, Early Access 2024.
- [11] W. Guo, M. Ma, H. Wang, N. Xiang, H. Wang, Z. Chen, and W. Chen, "Real-time average junction temperature estimation for multichip IGBT modules with low computational cost," *IEEE Trans. Ind. Electron.*, vol. 70, no. 4, Apr. 2023.
- [12] M. Ma, W. Guo, X. Yan, S. Yang, X. Zhang, W. Chen, and G. Cai, "A three-dimensional boundary-dependent compact thermal network model for IGBT modules in new energy vehicles," *IEEE Trans. Ind. Electron.*, vol. 68, no. 6, pp. 5248–5258, Jun. 2020.
- [13] S. Gao, K. D. Ngo, and G. Q. Lu, "Two-dimensional mapping of interface thermal resistance by transient thermal measurement," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4448–4456, May. 2021.
- [14] JEDEC, "JESD51-14: Transient dual interface test method for the measurement of the thermal resistance junction-to-case of semiconductor devices with heat flow through a single path," 2010.
- [15] V. Szekely, "A new evaluation method of thermal transient measurement results," *Microelectronics J.*, vol. 28, pp. 277–292, 1997.
- [16] Y. Zhang, R. Wu, F. Iannuzzo, and H. Wang, "Aging investigation of the latest standard dual power modules using improved interconnect technologies by power cycling test," *Microelectron. Reliab.*, vol. 138, no. 114740, Nov. 2022.
- [17] H. Wang, Z. Zhou, Z. Xu, X. Ge, Y. Yang, Y. Zhang, B. Yao, and D. Xie, "A thermal network model for multichip power modules enabling to characterize the thermal coupling effects," *IEEE Trans. Power Electron.*, Early Access 2024.
- [18] M. Salleras, M. Carmona, and S. Marco, "Issues in the use of thermal transients to achieve accurate time-constant spectrums and differential structure functions," *IEEE Trans. Adv. Packag.*, vol. 33, no. 4, pp. 918–923, Nov. 2010.
- [19] V. Székely, "Identification of RC networks by deconvolution: chances and limits," *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, vol. 45, no. 3, pp. 244–258, 1998.
- [20] K. A. Pareek, C. Grosse, M. Sternberg, D. May, M. A. Ras, and B. Wunderle, "Effect of different deconvolution methods on structure function calculation," in *Proc. 26th Int. Workshop Therm. Invest. ICs Syst. (THERMINIC)*, pp. 97–104, 2020.
- [21] T. Kennett, W. Prestwich, and A. Robertson, "Bayesian deconvolution i: Convergent properties," *Nucl. Instrum. Methods*, vol. 151, no. 1-2, pp. 285–292, 1978.
- [22] Y. Zhang, H. Wang, Z. Wang, and F. Blaabjerg, "Simplified multi-time scale thermal model considering thermal coupling in IGBT modules," in *Proc. IEEE Appl. Power Electron. Conf. Expo. (APEC)*, pp. 319–324, 2019.
- [23] S. Fukunaga and T. Funaki, "Transient thermal network model identification for power module packages," *Nonlinear Theory and Its Applications, IEICE*, vol. 11, no. 2, pp. 157–169, 2020.
- [24] A. Beck, *First-order methods in optimization*. SIAM, 2017.
- [25] M. Grant, S. Boyd, and Y. Ye, "Disciplined convex programming," *Global optimization: From theory to implementation*, pp. 155–210, 2006.
- [26] P. C. Hansen, *The L-curve and its use in the numerical treatment of inverse problems*. Citeseer, 1999.
- [27] L. De Tommasi, A. Magnani, V. d'Alessandro, and M. de Magistris, "Time domain identification of passive multiport rc networks with convex optimization: An application to thermal impedance macromodeling," in *Proc. IEEE 18th Workshop Signal Power Integr. (SPI)*, pp. 1–4, 2014.
- [28] TDIM Master Software. [Online]. Available: <https://www.jedec.org/standards-documents/docs/jesd51-14-0>
- [29] Simcenter T3STER. [Online]. Available: <https://www.plm.automation.siemens.com/global/en/products/simcenter/t3ster.html>
- [30] Datasheet FGA20N120FTD. [Online]. Available: <https://www.onsemi.com/products/discrete-power-modules/igbts/fga20n120ftd>



**Yi Zhang** (Member, IEEE) received the B.S. and M.S. degrees from Harbin Institute of Technology, China, in 2014 and 2016, respectively, and the Ph.D. degree from Aalborg University, Denmark, in 2020. All degrees are in electrical engineering. He is currently an Assistant Professor with the AAU Energy, Aalborg University, Denmark.

During 2020-2023, he was affiliated with multiple institutions as a postdoctoral researcher with the support of the Danish Research Council for Independent Research, including RWTH-Aachen University, Germany, Swiss Federal Institute of Technology Lausanne, Switzerland, and Massachusetts Institute of Technology, USA. He was also a visiting scholar with Georgia Institute of Technology, USA, in 2018. His researches include reliability of power electronics, thermal modeling, and multilevel converters.

Dr. Zhang is the Guest Associate Editor of IEEE Transactions on Power Electronics. He is recipient of the First Place Prize Paper Award of the IEEE Transactions on Power Electronics in 2021, and the IEEE Power Electronics Society Ph.D. Thesis Award in 2020.



**Rik W. De Doncker** (Fellow, IEEE) received his Ph.D. degree in electro-mechanical engineering from the KULeuven, Belgium. In 1987, he was appointed Visiting Associate Professor at the University of Wisconsin, Madison, where he developed the DAB converter. In 1988, he joined the GE Corporate Research and Development Center, Schenectady, NY. In November 1994, he joined Silicon Power Corporation (formerly GE-SPCO) as Vice President Technology, developing world's first medium-voltage static transfer switch. Since Oct. 1996, he is professor

at RWTH Aachen University, Germany, where he leads the Institute for Power Electronics and Electrical Drives (ISEA). In Oct. 2006 he was appointed director of the E.ON Energy Research Center at RWTH Aachen University. Since 2014, he is director of the German Federal Government BMBF Flexible Electrical Networks (FEN) Research CAMPUS. He has a doctor honoris causa degree of TU Riga, Latvia. He has published over 800 technical papers and is holder of more than 60 patents.

Dr. De Doncker is recipient of the IAS Outstanding Achievements Award, the 2013 Newell Power Electronics IEEE Technical Field Award, and the 2014 IEEE PELS Harry A. Owen Outstanding Service Award. In 2015 he was awarded Fellow status at RWTH University. In 2016 he became member of the German Academy of Science and Technology (ACATECH). 2020 he received the IEEE Medal in Power Engineering.



**Anton Evgrafov** is an applied mathematician with an MSc (2000) from Saint Petersburg Electrotechnical University (LETI), St. Petersburg, Russia, and a PhD (2004) from Chalmers University of Technology, Gothenburg, Sweden. He is currently an Associate Professor at the Department of Mathematical Sciences, Aalborg University, Denmark. He specializes in optimization of physical and engineering systems.



**Shuai Zhao** (Member, IEEE) Shuai Zhao received B.S. (Hons.), M.S., and Ph.D. degrees in information and telecommunication engineering from Northwestern Polytechnical University, Xi'an, China, in 2011, 2014, and 2018, respectively. He is currently an Assistant Professor with AAU Energy, Aalborg University, Denmark.

From 2014 to 2016, he was a Visiting Ph.D. student at the University of Toronto, Canada. In August 2018, he was a Visiting Scholar with the University of Texas at Dallas, USA. From 2018 to 2022, he was a Postdoc researcher with AAU Energy, Aalborg University, Denmark. He is the Associate Editor of IEEE Transactions on Vehicle Technology and Guest Editor of IEEE Journal of Emerging and Selected Topics in Industrial Electronics and Elsevier e-Prime. His research interests include physics-informed machine learning, system informatics, condition monitoring, diagnostics & prognostics, and tailored AI tools for power electronic systems.



**Sven Kalker** (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from RWTH Aachen University, Aachen, Germany, in 2014 and 2017, respectively. He has been a research associate at the Institute for Power Electronics and Electrical Drives at RWTH Aachen University since December 2017, took over as head of the Reliable Power Electronics Group in September 2020 and became chief engineer in September 2022. His research interests include reliability, modeling, and sensing of power electronic systems.