*photonics*

MDPI

*Article*

# High-Precision Lumped-Parameter Thermal Modeling and CNN-LSTM-Based Temperature Prediction for Optical Interconnect Modules

**Tianxiang Lan, Chuanchuan Yang *, Hao Qin and Zhangyuan Chen**

State Key Laboratory of Photonics and Communications, Department of Electronics, Peking University, Beijing 100871, China; 2101111425@stu.pku.edu.cn (T.L.); qinhao0@pku.edu.cn (H.Q.); chenzhy@pku.edu.cn (Z.C.)
* Correspondence: yangchuanchuan@pku.edu.cn

**Abstract**

The increasing power consumption and intensified thermal coupling in next-generation high-speed optical interconnect modules present severe thermal management challenges. To address these issues, we propose a high-precision lumped-parameter thermal modeling method for optical interconnect modules. Building upon the constructed lumped-parameter thermal model (LPTM), a real-time transient temperature prediction framework utilizing a hybrid convolutional neural network and long short-term memory (CNN-LSTM) architecture is developed. The LPTM achieves high-precision modeling while reducing computational complexity. Compared with the simulation results from COMSOL Multiphysics 6.3, the proposed LPTM model exhibits an average mean absolute error (MAE) of only 0.52 °C and an average root mean square error (RMSE) of only 0.69 °C for the transient temperatures of key components. Furthermore, the CNN-LSTM neural network, trained on the dataset generated by the LPTM, achieves an MAE of 0.34 °C and an RMSE of 0.78 °C on an independent transient temperature testset generated by COMSOL Multiphysics. The proposed methodology successfully achieves both high-precision modeling and real-time, accurate transient temperature prediction, providing a solid theoretical foundation for thermal management in next-generation optical interconnect systems.

**Keywords:** optical interconnect modules; thermal management; lumped-parameter thermal model; temperature prediction; CNN-LSTM architecture

## 1. Introduction

Driven by the rapid growth of fifth-generation (5G) wireless communication, artificial intelligence, and cloud computing, data center traffic demands have increased sharply in recent years. Optical interconnect modules serve as the core transmission carriers for data center traffic. Thermal management is a critical factor in optical interconnect module design, as elevated temperatures can cause wavelength drift in lasers and degrade the performance of key devices such as lasers and modulators. Furthermore, high operating temperatures also accelerate device aging, ultimately degrading the overall performance and reliability of data center communication systems. Consequently, effective thermal design and management play an indispensable role in ensuring efficient and reliable communication across optical links.

Conventionally, the finite element analysis (FEA) method has been the primary solution for heat conduction analysis and thermal design in silicon photonics [1]. However,

Check for updates

Received: 15 September 2026
Revised: 3 October 2026
Accepted: 14 October 2026
Published: 5 November 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

as optical interconnect technologies evolve, pluggable optical modules are advancing toward higher transmission rates such as 1.6 T [2,3], leading to a substantial increase in component power consumption. Meanwhile, emerging architectures such as near-packaged optics (NPO) [4–6] and co-packaged optics (CPO) [7–9] highly integrate switch ASICs with optical engines, giving rise to complex internal thermal effects and severe thermal coupling issues. These advancements significantly increase the complexity of thermal design and management for next-generation optical interconnect modules [10]. Under these circumstances, FEA suffers from high computational complexity, making it difficult to satisfy the demand for rapid analysis and large-scale data generation in complex system design, thereby limiting further progress in thermal management. Therefore, developing a high-precision, low-complexity thermal modeling method has become a crucial research direction in optical interconnect thermal design. Two representative routes to such a low-complexity model are compact thermal networks [11], which describe a system by a small network of thermal resistances and capacitances whose parameters are identified from measurements or simulations, and reduced-order thermal models [12], which reduce a detailed finite-element model to a low-order subspace that retains the dominant thermal dynamics. Both approaches have been widely applied in electronics and optoelectronics, but their application to the heat conduction analysis of optical interconnect modules is still limited.

Concurrently, real-time temperature prediction based on high-precision, low-complexity thermal modeling demonstrates growing research significance. This technique overcomes measurement blind spots caused by physical constraints on sensor placement. It enables dynamic temperature alerts for key hotspots and provides real-time data support for proactive thermal management algorithms [13]. These capabilities help prevent thermal failures and improve the overall reliability of optical interconnect modules and data center optical links. In recent years, substantial research has been devoted to temperature prediction. For instance, the work in [14] employed a state-space-based Kalman predictor for real-time temperature prediction under noisy sensor conditions, and the works in [15,16] employed long short-term memory neural networks and other machine learning models to predict the temperature at the data center room level rather than at the module level of optical interconnects. In addition, the work in [17] performed a package-level thermal analysis of a photonic-electronic transceiver, but the transient temperature was not predicted. Due to the scarcity of detailed internal temperature data and the fast, complex power fluctuation in optical modules, module-scale temperature prediction remains largely unexplored.

To address these requirements, we propose a high-precision, low-complexity lumped-parameter thermal modeling method for optical interconnect modules and build an efficient temperature prediction framework using a hybrid convolutional neural network and long short-term memory (CNN-LSTM) architecture. By simplifying complex finite-element thermal fields into lumped nodes, the proposed lumped-parameter thermal model (LPTM) maintains high accuracy while significantly reducing computational complexity compared with FEA, enabling rapid heat conduction analysis and large-scale data generation. Furthermore, using the transient temperature dataset generated by the LPTM, a CNN-LSTM neural network is designed and trained for transient temperature prediction. A simulation platform was developed to validate the LPTM against COMSOL Multiphysics simulations. The results show that the average mean absolute error (MAE) and root mean square error (RMSE) for the transient temperatures of critical components are only 0.52 °C and 0.69 °C, respectively, confirming the accuracy of the LPTM. Additionally, numerical evaluations of the proposed temperature prediction neural network demonstrate an MAE of 0.34 °C and an RMSE of 0.78 °C on an independent transient temperature testset generated by COMSOL Multiphysics. These results demonstrate that the proposed CNN-LSTM network enables real-time and precise transient temperature prediction, laying a solid theoretical

foundation for the design and optimization of future thermal management algorithms in optical interconnect modules.

## 2. Lumped-Parameter Thermal Modeling of Optical Interconnect Modules

A lumped-parameter model simplifies a complex continuous system into discrete nodes. Building upon this general concept, the proposed LPTM is constructed based on the electro-thermal analogy. It performs system-level thermal analysis by mapping the heat transfer characteristics of physical entities to equivalent electrical components. Specifically, temperature difference corresponds to voltage, heat flow rate to electric current, thermal resistance to electrical resistance, and thermal capacitance to electrical capacitance. Through this approach, a complex heat conduction system is simplified into a thermal RC network composed of discrete nodes.

Assuming a thermal RC network consisting of $N$ nodes, for the $i$-th node, the heat source power input equals the total outgoing heat flow, namely the sum of the heat absorbed and stored by the thermal capacitance, the heat conducted through adjacent thermal resistances, and the heat dissipated to the ambient environment. Consequently, the heat transfer differential equation for the $i$-th node is expressed as:

$$Q_i(t) = C_i \frac{dT_i(t)}{dt} + \sum_{j \neq i} \frac{T_i(t) - T_j(t)}{R_{ij}} + \frac{T_i(t) - T_d}{R_{id}}, \quad (1)$$

where $Q_i(t)$ denotes the heat source power of the $i$-th node, $C_i$ denotes the thermal capacitance of the $i$-th node, $T_i(t)$ denotes the transient temperature of the $i$-th node, $R_{ij}$ denotes the thermal resistance between the $i$-th and $j$-th nodes, $T_d$ denotes the temperature of ambient environment, and $R_{id}$ denotes the equivalent thermal resistance between the $i$-th node and the ambient environment. If a heat conduction path exists between two nodes, $R_{ij}$ takes a finite value. Otherwise, $R_{ij}$ is infinite. The same principle applies to the thermal resistance between a node and the ambient environment. An example of a LPTM consisting of two nodes is illustrated in Figure 1.

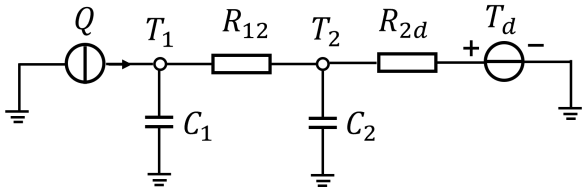



**Figure 1.** Schematic of a two-node lumped-parameter thermal model.

The proposed LPTM offers distinct advantages of low computational complexity and closed-form solvability, making it highly suitable for thermal modeling of optical interconnect modules. Moreover, the proposed modeling methodology is not restricted to conventional pluggable optical modules, but represents a highly scalable and generalized thermal modeling framework. Optical interconnect modules include numerous temperature-sensitive devices that require rigorous heat conduction analysis and thermal management such as lasers and modulators. While comprehensively considering all these constituent devices, we focus on the laser as the primary target device to elaborate our proposed thermal modeling methodology. For systematic analysis, the optical interconnect module is partitioned into four main constituent groups: the laser, thermal management components, other heat-generating components, and other non-heat-generating components. Specifically, thermal management components include elements required for thermal management, such as heat sinks, thermoelectric cooler (TEC), and module shells. Other heat-generating components consist of active components excluding the laser and TEC, such as driver chip

and photo detector (PD) chip. Other non-heat-generating components include passive and structural elements outside the aforementioned groups, such as lenses. Heat conduction is the dominant heat transfer mechanism within optical interconnect modules. According to this classification, the heat transfer paths in optical interconnect modules are as follows. Heat generated by the laser is conducted to the thermal management components. Heat from other heat-generating components flows to both the thermal management components and other non-heat-generating components. Finally, thermal management components and other non-heat-generating components exchange heat with the ambient environment via convection. The heat transfer paths in optical interconnect modules are illustrated in Figure 2.

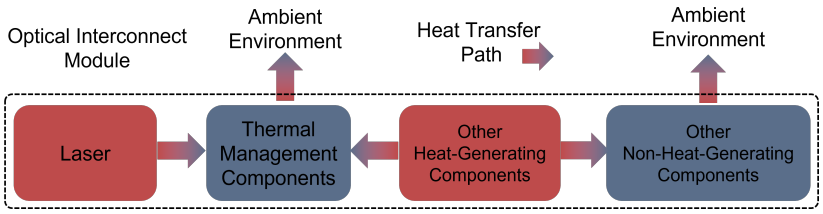



**Figure 2.** Schematic of heat transfer paths in an optical interconnect module.

Assuming that each of the aforementioned constituent groups corresponds to a single node, node 1 represents the laser, node 2 represents the thermal management components, node 3 represents other heat-generating components, and node 4 represents other non-heat-generating components. Based on this assumption, the generalized optical interconnect module depicted in Figure 2 can be modeled as the LPTM shown in Figure 3.

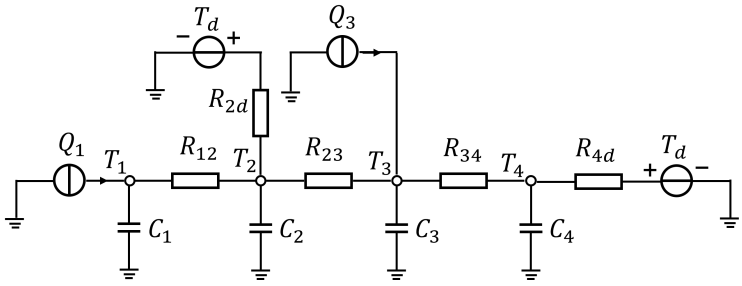



**Figure 3.** Lumped -parameter thermal model corresponding to the generalized optical interconnect module in Figure 2.

When modeling a practical optical interconnect module, the aforementioned constituent groups are further expanded by abstracting each component of the optical interconnect module into an individual node.

The topological connectivity among these nodes is established based on the module's internal physical structure, while the thermal capacitance of each node $C_i$ and the inter-node thermal resistances $R_{ij}$ are determined by structural and material parameters of the interconnect module. Specifically, the thermal capacitance of each node is directly calculated from the density $\rho_i$, the specific heat capacity $c_{p,i}$ and the volume $V_i$ of the corresponding component as $C_i = \rho_i c_{p,i} V_i$. The thermal resistance between two nodes is determined by the contact condition and the material parameters of the corresponding components. Although the thermal resistance can in principle be calculated directly, it is difficult to implement in practice. Therefore, a one-off finite element simulation is adopted in this work.

The heat flow rate $Q_{ij}$ from the $i$-th node to the $j$-th node is obtained by integrating the heat flux over the contact interface $A_{ij}$, and the inter-node thermal resistance is then obtained as

$$R_{ij} = \frac{\Delta T_{ij}}{Q_{ij}} = \frac{T_i - T_j}{\iint_{A_{ij}} \mathbf{q} \cdot \mathbf{n}_{ij}\, dA}, \tag{2}$$

where $\mathbf{q}$ denotes the heat flux density vector on the contact interface, $\mathbf{n}_{ij}$ denotes the unit normal vector pointing from the $i$-th node to the $j$-th node, and $\Delta T_{ij}$ denotes the temperature difference between the $i$-th node and $j$-th node. For the equivalent thermal resistance between a node and the ambient environment, the convective heat transfer is described by Newton's law of cooling $Q = hA\Delta T$, and thus the equivalent thermal resistance of the $i$-th node is obtained as $R_{id} = 1/(hA_{id})$, where $h$ denotes the convective heat transfer coefficient and $A_{id}$ denotes the area of the $i$-th component node in contact with the ambient environment. After the above parameters are obtained, the modeling process of the lumped-parameter thermal model is completed.

On this basis, the heat source power $Q_i(t)$ is either set to a fixed value or dynamically modeled according to the actual operating conditions of the optical interconnect module. By substituting the derived thermal RC parameters and heat source excitations into the single-node heat transfer Equation (1) and coupling the equations across all nodes, a complete system of differential equations is constructed to describe the overall thermal dynamic behavior of the module. Finally, time discretization is applied to this equation system, which is solved iteratively using the backward Euler method. This method achieves accurate computation of the transient temperature evolution for all internal nodes with low computational complexity. The LPTM effectively models the heat conduction processes both within the optical interconnect module and with the ambient environment. By adjusting parameters such as node topology and heat source power, the model not only captures the operational temperature dynamics of individual components, but also flexibly accommodates thermal variations resulting from diverse packaging architectures and operating conditions. Consequently, it enables low-complexity, high-precision transient temperature simulations, establishing a solid theoretical foundation for subsequent real-time temperature prediction.

To demonstrate the practical implementation of the proposed modeling methodology, a corresponding LPTM was constructed with reference to an optical module architecture from InnoLight Technology [18], a leading commercial transceiver vendor. The structural configuration of the optical module is illustrated in Figure 4. The module includes two submounts. The laser chip and driver chip are positioned on submount 1 while the PD chip is positioned on submount 2, and both submounts are positioned on a common heat sink. The resulting LPTM is depicted in Figure 5, where the node label definitions are provided on the right-hand side of the schematic. The notations for thermal capacitances, thermal resistances, and related parameters remain consistent with the definitions established above. The accuracy evaluation and validation of this model will be presented in Section 4 through comparisons with COMSOL Multiphysics simulation results.

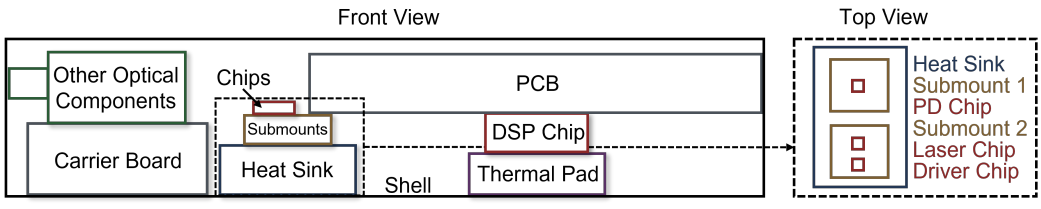



**Figure 4.** Structural schematic of the example optical module.

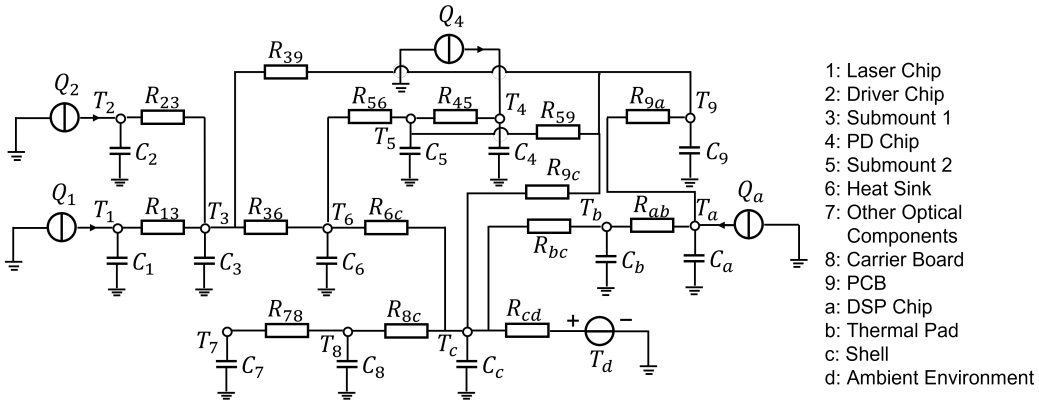



**Figure 5.** Lumped-parameter thermal model of the example optical module.

## 3. Temperature Prediction Neural Network Based on Lumped-Parameter Thermal Model

The proposed LPTM efficiently generates a large amount of high-precision transient temperature data for optical interconnect modules, providing high-quality dataset support for training temperature prediction neural networks. In this paper, we not only propose a high-accuracy LPTM framework for optical interconnect modules, but also design and train a high-accuracy neural network for temperature prediction which can achieve accurate real-time temperature prediction, providing theoretical guidance for the design and optimization of thermal management algorithms.

To achieve real-time and accurate temperature prediction in optical interconnect modules, we propose a hybrid deep learning architecture integrating convolutional neural network (CNN) and LSTM network. This architecture leverages the capabilities of CNN in local spatial feature extraction and LSTM in modeling long-term temporal dependencies, thereby efficiently handling the thermal responses in optical interconnect modules.

### *3.1. Heat Source Power Modeling and Dataset Generation*

To train and test the temperature prediction neural network, a large amount of transient temperature data reflecting the actual operating conditions of the optical interconnect module needs to be generated using the LPTM. Once the thermal parameters such as thermal capacitances and thermal resistances are determined, it is necessary to model the heat source power in the optical interconnect module.

We first perform time dicretization with a time step of $\Delta t$. Let $D_s$ denote the reference number of packets processed per second by the optical interconnect module, and $D(k)$ denote the actual number of packets processed during the $k$-th time step. We define the load intensity of the optical interconnect module at the $k$-th time step as $U(k) = \frac{D(k)}{D_s \Delta t}$. The load intensity modeling methodology and an example load sequence are illustrated in Figure 6, where $q_{ij}$ denotes the transition rate from state $i$ to state $j$, and $\mathcal{P}(a)$ denotes a Poisson distribution with parameter $a$. To effectively simulate the fluctuations of data center traffic, we model $D(k)$ using a Poisson process modulated by a two-state (busy–idle) Markov chain. The mean state durations for the busy and idle states are denoted by $t_1$ and $t_2$, respectively. And the Markov state corresponding to the $k$-th time step is denoted by $S(k)$. Setting the average load intensities for the two states to $u_1$ and $u_2$, the corresponding intensity parameters of the Poisson process are obtained as $u_1 D_s$ and $u_2 D_s$, respectively. Under practical conditions, since the number of packets per time step is large, the packet counts in the two states are approximated by Poisson distributions with parameters $u_1 D_s \Delta t$ and $u_2 D_s \Delta t$, respectively. After determining the state for each time step according to the

Markov chain definition, $D(k)$ is generated based on the corresponding Poisson parameter, from which $U(k)$ is subsequently calculated.

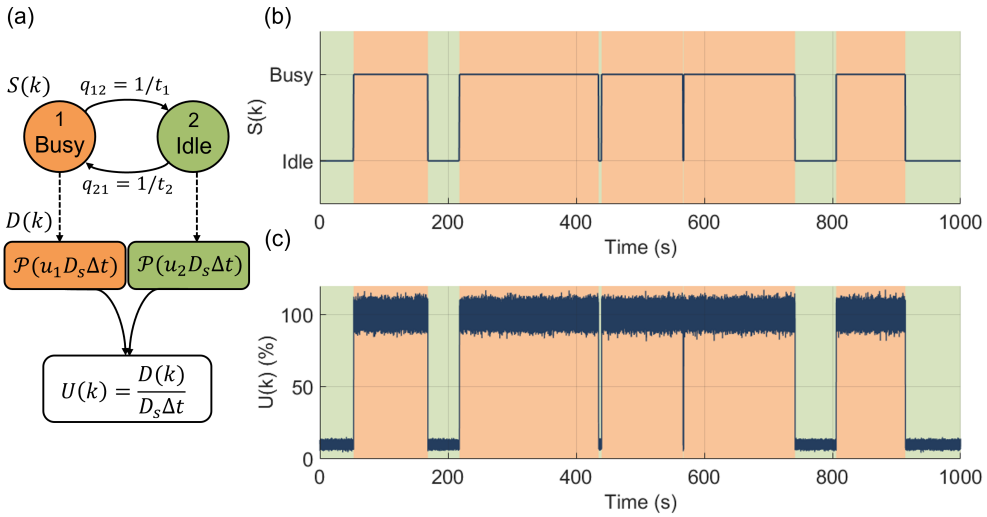



**Figure 6.** (**a**) Schematic of the load intensity modeling methodology. (**b**) Example of a state transition sequence. (**c**) Example of a load intensity sequence.

The heat source power of each component in the optical interconnect module is correlated with the load intensity. We model the heat source power of the module's components into three parts: static heat source power, dynamic heat source power, and noise power. Among these, the dynamic heat source power is directly proportional to the load intensity $U(k)$. The heat source power modeling is expressed as

$$\begin{aligned} Q_i(k) &= P_{i,static} + P_{i,dynamic}(k) + N_i(k) \\ &= P_{i,static} + P_{i,d}U(k) + N_i(k), \end{aligned} \tag{3}$$

where $Q_i(k)$ represents the heat source power of the $i$-th component at the $k$-th time step, $P_{i,static}$ denotes the static heat source power of the $i$-th component, $P_{i,dynamic}(k)$ represents the dynamic heat source power of the $i$-th component at the $k$-th time step, $N_i(k)$ denotes the noise power of the $i$-th component at the $k$-th time step, and $P_{i,d}$ denotes the dynamic heat source power of the $i$-th component when the load intensity is 1.

Once the heat source power is obtained, it is applied as the excitation for the LPTM and substituted into the system of differential equations. Following time discretization, the backward Euler method is employed to iteratively solve the equation system. After obtaining the transient temperature data, a high-accuracy dataset satisfying the training and testing requirements of the temperature prediction neural network is generated.

### *3.2. Temperature Prediction Neural Network*

The designed CNN-LSTM neural network for temperature prediction can predict the future temperature evolution of individual component nodes in the optical interconnect module based on historical temperature information. The architecture of this neural network is illustrated in Figure 7, where $x_t$, $h_t$, and $c_t$ denote the input vector, hidden state vector, and cell state vector of the LSTM layer at the $t$-th time step, respectively, and $L$ represents the sequence length used for training. The neural network includes four components: a one-dimensional convolutional (1D-CNN) layer, a ReLU nonlinear activation layer, an LSTM layer, and a fully connected (FC) layer. The inputs of the neural network consist of the component temperatures and the Markov chain states of the heat source power over a

past time window, while the outputs are the predicted component temperatures over a future time window. The 1D-CNN layer is utilized to extract local temporal features of the temperature variations while expanding the feature dimension of the data. The LSTM layer then processes the dimensionally expanded time series and outputs the hidden state at the final time step, thereby realizing the mapping of time-series information to a high-dimensional feature vector. Finally, the FC layer maps this high-dimensional feature vector into the temperature prediction outputs.

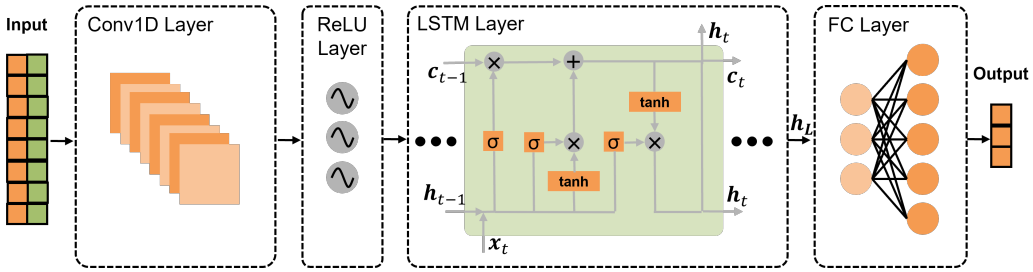



**Figure 7.** Structure of the CNN-LSTM temperature prediction neural network.

Because state transition is difficult to predict, the high-order penalty of the traditional mean squared error (MSE) loss function easily causes the neural network to overcompensate for temperature variations at the transition edges, thereby degrading the prediction accuracy in non-transition intervals. To address this issue, a piecewise weighting mechanism is introduced in this paper to modify the standard MSE. The standard MSE loss is maintained in the non-transition intervals, while an attenuation factor $\alpha$ is introduced in the transition intervals to suppress the excessive interference of state transition on the overall training process. The modified loss function is expressed as

$$\mathcal{L}_{MMSE} = \frac{1-(1-\alpha)s}{N}\|\mathbf{y}-\hat{\mathbf{y}}\|_2^2, \tag{4}$$

where $N$ represents the prediction sequence length, $\mathbf{y}$ denotes the ground-truth temperature sequence, $\hat{\mathbf{y}}$ denotes the predicted temperature sequence. And $s$ is the state transition indicator variable, which takes the value of 0 if there are no state transitions within the prediction interval, and 1 if state transitions occur. Finally, $\|\mathbf{a}\|_p$ represents the $L_p$-norm of the vector $\mathbf{a}$.

## 4. Performance Validation of the Lumped-Parameter Thermal Model and Temperature Prediction Neural Network

In this section, we conducted numerical simulations to validate the accuracy of the LPTM for the optical interconnect module and to test the performance of the CNN-LSTM temperature prediction neural network.

### *4.1. Validation of the Lumped-Parameter Thermal Model*

To validate the accuracy of the LPTM, we adopted the same optical module structure as presented in Section 2 (illustrated in Figure 4) and established a simulation project in the COMSOL Multiphysics platform. The constructed simulation model of the optical module is depicted in Figure 8. After completing the geometry and material modeling, the laser chip, driver chip, PD chip, and digital signal processing (DSP) chip were configured as fixed power heat sources, while the boundary condition of the optical module shell was set to convective heat flux. A transient simulation of the temperature rise process from 0 to 400 s was conducted. Subsequently, the transient temperature of each component was volume-averaged to obtain the variation curves of the volume-averaged temperatures for

each component. These curves serve as the benchmark for the subsequent comparative validation of the LPTM's accuracy. The key parameters of the COMSOL simulation are summarized in Table 1.

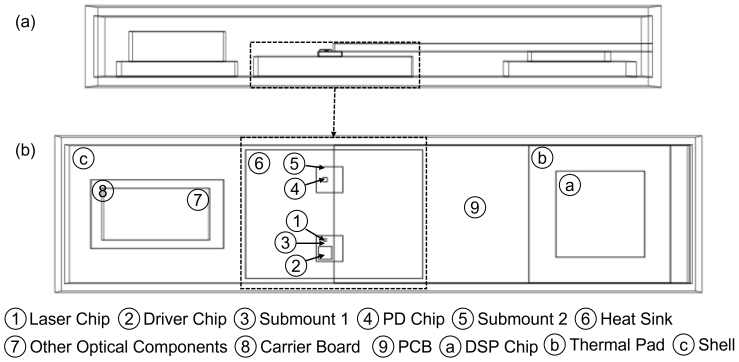



**Figure 8.** COMSOL simulation model of the optical module. (**a**) Front view. (**b**) Top view.

**Table 1.** Key parameters of the COMSOL simulation.

| Parameter | Value |
|---|---|
| Heat source power of laser chip | 0.033 W |
| Heat source power of driver chip | 1.0 W |
| Heat source power of PD chip | 0.5 W |
| Heat source power of DSP chip | 2.5 W |
| Shell-to-ambient convective heat transfer coefficient | 50 W/(m$^2$·K) |

According to the thermal capacitance and thermal resistance acquisition method described above, the parameter values of the LPTM are obtained, as listed in Table 2. It should be noted that, in the COMSOL simulation, the contact thermal resistances were set at the key contact interfaces of the module with the layer resistance parameter set to a realistic value of $10^{-5}$–$10^{-6}$ K·m$^2$/W. For the LPTM, the contact thermal resistances are lumped into the equivalent inter-node thermal resistances which means the thermal resistance values listed in Table 2 are effective thermal resistances that already incorporate the contact thermal resistances.

**Table 2.** Thermal capacitances and thermal resistances of the LPTM.

| Thermal Capacitance | Value (J/K) | Thermal Resistance | Value (K/W) |
|---|---|---|---|
| $C_1$ | $1.34 \times 10^{-5}$ | $R_{13}$ | 10.27 |
| $C_2$ | $7.34 \times 10^{-4}$ | $R_{23}$ | 6.45 |
| $C_3$ | $1.13 \times 10^{-2}$ | $R_{36}$ | 6.33 |
| $C_4$ | $4.08 \times 10^{-5}$ | $R_{45}$ | 47.56 |
| $C_5$ | $1.13 \times 10^{-2}$ | $R_{56}$ | 6.19 |
| $C_6$ | 2.07 | $R_{6c}$ | 0.38 |
| $C_9$ | 1.33 | $R_{9c}$ | 45.26 |
| $C_a$ | 0.16 | $R_{9a}$ | 85.90 |
| $C_b$ | 0.82 | $R_{ab}$ | 4.31 |
| $C_c$ | 8.86 | $R_{bc}$ | 2.47 |
| | | $R_{cd}$ | 4.96 |

We conducted a transient simulation of the temperature rise process from 0 to 400 s for the identical optical interconnect module using the proposed LPTM employing the same parameters as listed in Table 1. We compared the temperature variation curves of the

volume-averaged component temperatures obtained from the COMSOL simulation with the corresponding node temperatures simulated by the LPTM for four key components: the laser chip, driver chip, submount 1, and heat sink. The comparison results are illustrated in Figure 9, where the solid blue curves represent the COMSOL simulation results and the dashed orange curves represent the LPTM simulation results. As observed from the figure, the temperatures of all components exhibit a consistent temporal trend. During the initial heating phase, the temperature rises rapidly. Subsequently, as internal heat generation and external heat dissipation gradually reach a dynamic equilibrium, the temperature flattens and eventually stabilizes at a steady-state value. Throughout both the transient heating phase and the steady-state period, the temperature curves calculated by the LPTM closely match those from the COMSOL simulation. To quantitatively evaluate the accuracy of the LPTM, we further calculated the MAE and RMSE between the transient temperature sequences obtained from the COMSOL and LPTM simulations. The MAE and RMSE are defined as

$$\begin{aligned} E_{MAE} &= \frac{1}{N}\|\mathbf{T}_1 - \mathbf{T}_2\|_1 \\ E_{RMSE} &= \frac{1}{\sqrt{N}}\|\mathbf{T}_1 - \mathbf{T}_2\|_2, \end{aligned} \tag{5}$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ denote the transient temperature sequences of the COMSOL and LPTM simulations, respectively, and $N$ represents the sequence length.

The transient temperature sequence errors for the four key components are illustrated in Figure 10. The average MAE and RMSE of the transient temperature sequences across four components are merely 0.52 °C and 0.69 °C, respectively. These results thoroughly validate the accuracy and effectiveness of the LPTM for heat conduction analysis in optical interconnect modules.

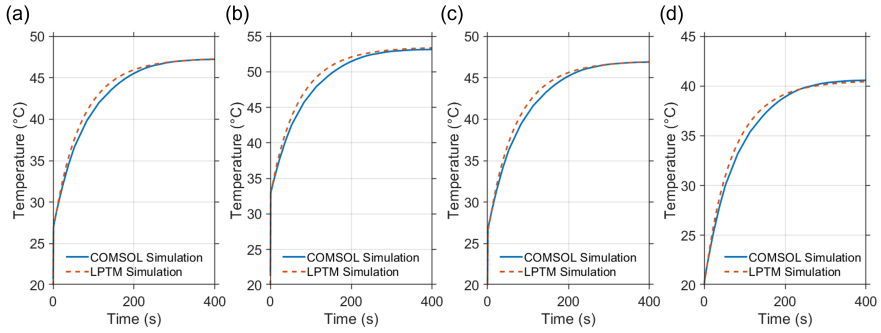



**Figure 9.** Comparison of simulation results between COMSOL and the LPTM. (**a**) Laser chip. (**b**) Driver chip. (**c**) Submount 1. (**d**) Heat sink.

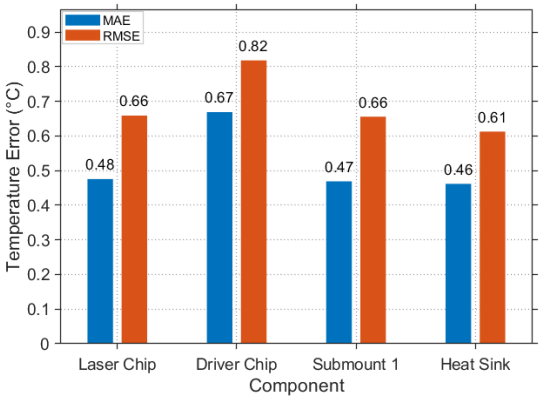



**Figure 10.** Transient temperature sequence errors between COMSOL and LPTM simulations for different components.

### *4.2. Performance Testing of the Temperature Prediction Neural Network*

The temperature prediction neural network constructed in this paper possesses comprehensive prediction capabilities for the node temperatures of all components in the optical interconnect module. The laser is the core device of the optical interconnect module. Its performance is highly sensitive to temperature variations and directly determines the overall performance of the module. Therefore, this sub-section primarily presents the performance of the temperature prediction neural network at the laser node. For the proposed CNN-LSTM neural network, the input window length of historical temperatures and states is set to 30 s, and the output window length for predicted temperatures is set to 10 s. Its structural hyper-parameters are summarized in Table 3.

**Table 3.** Structural hyper-parameters of the temperature prediction neural network.

| Layer | Parameters |
|---|---|
| 1D-CNN Layer | Kernel size: 3; input channels: 2; output channels: 8 |
| ReLU Layer | None |
| LSTM Layer | Input size: 8; hidden size: 16 |
| FC Layer | Input features: 16; output features: 100 |

To train and test the CNN-LSTM temperature prediction neural network, we adjusted the modeling parameters of the heat source power and established three different scenarios, generating three transient temperature sequence datasets for the laser chip with a duration of 86,400 s each dataset. The parameter settings for these three scenarios are summarized in Table 4. The datasets were partitioned into training, validation, and test sets at a ratio of 8:1:1, and the hyper-parameters employed during the neural network training are listed in Table 5. In addition to the three LPTM datasets, we have also generated an independent transient temperature dataset with COMSOL Multiphysics for the same optical interconnect module and used it to further test the performance of the proposed temperature prediction neural network. The heat source power of this COMSOL dataset was modeled with the parameters of Scenario 1 in Table 4.

**Table 4.** Modeling parameters of the heat source power under different scenarios.

| Parameter | $t_1$ (s) | $t_2$ (s) | $u_1$ | $u_2$ |
|---|---|---|---|---|
| Scenario 1 | 120 | 120 | 1.0 | 0.1 |
| Scenario 2 | 100 | 140 | 1.0 | 0.1 |
| Scenario 3 | 140 | 100 | 1.0 | 0.1 |

**Table 5.** Training hyper-parameters of the temperature prediction neural network.

| Parameter | Value |
|---|---|
| Batch size | 256 |
| Learning rate | 0.001 |
| Weight decay | 0.0001 |
| Epoch | 20 |
| Attenuation factor $\alpha$ | 0.1 |
| Optimizer | AdamW |

The trained temperature prediction neural network was first tested on three LPTM testsets. The complete predicted temperature sequence is obtained by concatenating the individual output sequences. We compared the ground-truth temperature sequence with

the predicted temperature sequence for each scenario, and the comparison results are illustrated in Figure 11. The MAE and RMSE between the ground-truth and predicted temperature sequences under each scenario are presented in Figure 12.

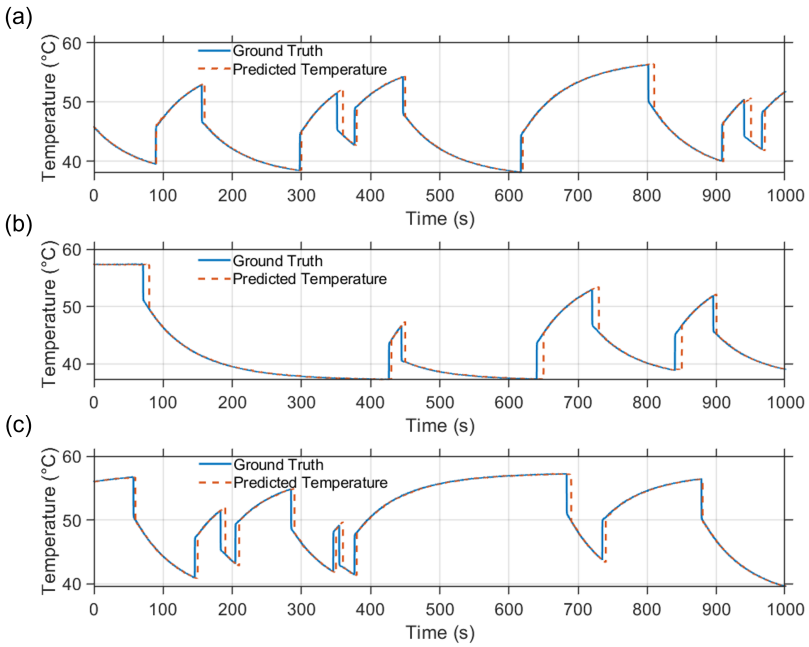



**Figure 11.** Comparison between the ground-truth and predicted temperature sequences under different scenarios. (**a**) Scenario 1; (**b**) Scenario 2; (**c**) Scenario 3.

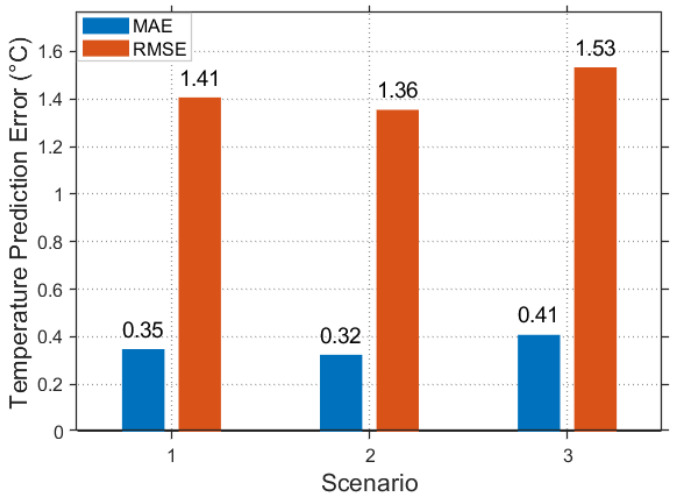



**Figure 12.** Temperature prediction errors under different scenarios.

As illustrated in the figures, the temperature sequences predicted by the proposed CNN-LSTM neural network show high agreement with the ground-truth temperature sequences. Moreover, across different scenarios, the MAE and RMSE between the predicted and ground-truth temperature sequences average only 0.36 °C and 1.43 °C, respectively. The RMSE is dominated by the state transition intervals, which cannot be predicted from the network input. In the steady intervals, which contain no state transition, the proposed temperature prediction neural network achieves a high prediction accuracy, with an MAE of 0.065 °C and an RMSE of 0.198 °C. These results clearly demonstrate that the proposed CNN-LSTM neural network achieves highly accurate temperature prediction, confirming

the accuracy of this architecture for real-time transient temperature prediction in optical interconnect modules.

Subsequently, the trained temperature prediction neural network was further tested on the COMSOL transient temperature dataset. We also compared the ground-truth temperature sequence with the predicted temperature sequence, and the comparison results are illustrated in Figure 13. The predicted temperature sequence is in close agreement with the ground truth temperature sequence, with an MAE of 0.34 °C and an RMSE of 0.78 °C, which is close to the results obtained on LPTM testsets, where the average MAE is 0.36 °C and the average RMSE is 1.43 °C. These results further verify the accuracy of the proposed temperature prediction network in an independent temperature prediction task.

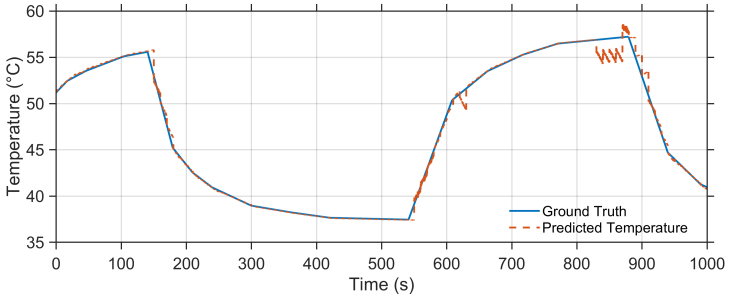



**Figure 13.** Comparison between the ground-truth and predicted temperature sequences of COMSOL dataset.

## 5. Discussion

In this section, the robustness of the proposed temperature prediction neural network was examined with two additional experiments. In the first one the temperature sequence of the input was perturbed by additive Gaussian noise, which represents the sensor noise of a module monitor, and three levels were compared, the original dataset and standard deviations of 0.10 and 0.20 °C. In the second one the state sequence was shifted in time with respect to the temperature sequence to simulate a misalignment of the load state estimation, with five conditions, the aligned case and shifts of +0.2 s, −0.2 s, +0.5 s and −0.5 s. The results are shown in Figures 14 and 15. Under the sensor noise the overall MAE increases from 0.360 °C to 0.412 and 0.498 °C, while under the temporal shifts it varies only between 0.349 and 0.371 °C, which is within about 3% of the aligned value. The temperature prediction neural network is therefore strongly robust to both the sensor noise of the temperature sequence and the misalignment of the load state sequence.

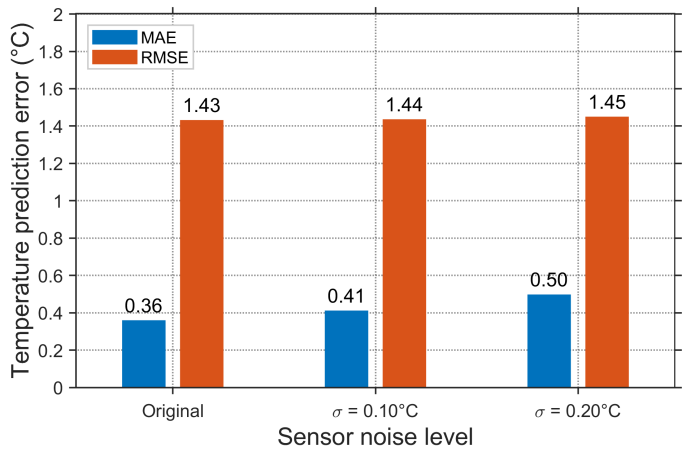



**Figure 14.** Temperature prediction errors under different sensor noise levels.

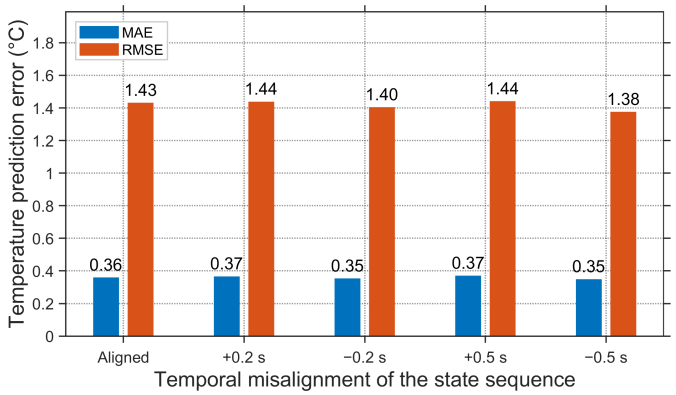



**Figure 15.** Temperature prediction errors under different misalignments of the state sequence.

## 6. Conclusions

To improve the thermal management of optical interconnect modules, we propose a high-precision, low-complexity LPTM method. Building upon the LPTM, a CNN-LSTM neural network architecture is further proposed for temperature prediction in optical interconnect modules. By applying the LPTM to heat conduction analysis, the complex system is simplified into a node network, significantly reducing computational complexity while maintaining essential thermal characteristics to ensure accurate analysis. The accuracy of the LPTM is validated by comparing its transient temperature simulation results with those from COMSOL simulations. Furthermore, a hybrid CNN-LSTM neural network combining CNN and LSTM architectures is developed. The LPTM is leveraged to generate high-quality, large-scale transient temperature datasets, enabling efficient training and rapid deployment of the neural network. The proposed CNN-LSTM neural network achieves an MAE of 0.34 °C and an RMSE of 0.78 °C on the independent testset generated by COMSOL Multiphysics, which demonstrates its effectiveness and accuracy in temperature prediction for optical interconnect modules. The LPTM facilitates rapid heat conduction analysis for future complex optical interconnect modules, and its capacity for large-scale data generation provides a theoretical foundation and dataset support for designing and optimizing thermal management algorithms. Meanwhile, the real-time prediction capability of the neural network offers vital model support for thermal management and resource allocation in optical interconnect modules. In summary, the proposed low-complexity, high-precision LPTM and CNN-LSTM prediction architecture improve thermal management technologies for next-generation optical interconnect modules, ultimately driving performance enhancements and technological evolution in data center optical interconnects.

**Author Contributions:** Conceptualization, T.L. and C.Y.; methodology, T.L.; validation, T.L. and H.Q.; formal analysis, T.L.; data curation, T.L.; writing—original draft preparation, T.L.; writing—review and editing, C.Y., H.Q. and Z.C.; visualization, T.L.; supervision, C.Y. and Z.C.; funding acquisition, C.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China, grant number 2023YFB2905501.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References


1. Oprins, H.; Ban, Y.; Cherman, V.; Van Campenhout, J. Thermal aspects of silicon photonic interposer packages. In *Proceedings of the 2020 21st International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*; IEEE: New York, NY, USA, 2020; pp. 1–6.
2. Yu, H.; Xudong, G.; Su, M.; Li, J.; Chen, J.; Zhang, R.; Gao, P.; Tao, M.; Wu, Y.; Li, Q.; et al. 1.6 T (8× 200Gb/s) 2× FR4 Silicon Photonic IMDD Transceiver with Monolithically Integrated Ultra-Low Crosstalk and Wideband Multiplexer. In *Proceedings of the Optical Fiber Communication Conference*; Optica Publishing Group: Washington, DC, USA, 2026; p. Th4A.7.
3. Isono, H. Latest standardization trend and future prospects for 800G/1.6 T optical transceivers. In *Proceedings of the Next-Generation Optical Communication: Components, Sub-Systems, and Systems XII*; SPIE: Bellingham, WA, USA, 2023; Volume 12429, pp. 45–51.
4. Fan, Y.; Luo, M.; Liu, C.; Zhang, S. Thermal solution study of 51.2 T Near-packaged optics switch. In *Proceedings of the 2023 22nd IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*; IEEE: New York, NY, USA, 2023; pp. 1–8.
5. Akahoshi, T.; Yamamoto, T.; Tsujino, S.; Takahashi, M.; Matsubara, T. Near-packaged optics module with 32Gbps x 16 lanes for short reach interconnect. In *Proceedings of the Next-Generation Optical Communication: Components, Sub-Systems, and Systems XII*; SPIE: Bellingham, WA, USA, 2023; Volume 12429, pp. 219–221.
6. Masuyama, K.; Shirao, M.; Araki, S.; Hasegawa, K.; Okuda, S.; Higashide, S.; Ohata, N. High-Density Integration of III-V Devices and EICs using Vertical-Coupled Photonic Packages with Glass-IP and Redistribution Layers. In *Proceedings of the 2026 IEEE 76th Electronic Components and Technology Conference (ECTC)*; IEEE: New York, NY, USA, 2026; pp. 68–72.
7. Muth, K.; Raghuraman, V.; Kannan, S.; Potluri, H. High density integration technologies for SiPh based optical I/Os. In *Proceedings of the 2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)*; IEEE: New York, NY, USA, 2023; pp. 212–215.
8. Zheng, Q.; Yang, P.; Xue, H.; He, H.; Cao, R.; Dai, F.; Sun, S.; Liu, F.; Wang, Q.; Cao, L.; et al. Research on 3D optical module integrating edge coupler and TSV. *J. Light. Technol.* **2022**, *40*, 6190–6200. [CrossRef]
9. He, H.; Xue, H.; Sun, Y.; Liu, F.; Cao, L. Design and realization of multi-channel and high-bandwidth 2.5 D transmitter integrated with silicon photonic MZM. *J. Light. Technol.* **2022**, *40*, 5201–5215. [CrossRef]
10. Tan, M.; Xu, J.; Liu, S.; Feng, J.; Zhang, H.; Yao, C.; Chen, S.; Guo, H.; Han, G.; Wen, Z.; et al. Co-packaged optics (CPO): Status, challenges, and solutions. *Front. Optoelectron.* **2023**, *16*, 1. [CrossRef] [PubMed]
11. Hu, Z.; Cui, M.; Wu, X. Real-Time Temperature Prediction of Power Devices Using an Improved Thermal Equivalent Circuit Model and Application in Power Electronics. *Micromachines* **2024**, *15*, 63. [CrossRef]
12. Li, D.; Lan, T.; Wang, Z.; Ye, Z. Fast Thermal Resistance Distribution Analysis in High-Power VCSEL Array Module. *Materials* **2025**, *18*, 5210. [CrossRef]
13. Lee, E.K.; Viswanathan, H.; Pompili, D. Vmap: Proactive thermal-aware virtual machine allocation in hpc cloud datacenters. In *Proceedings of the 2012 19th International Conference on High Performance Computing*; IEEE: New York, NY, USA, 2012; pp. 1–10.
14. Fu, Y.; Li, L.; Wang, K.; Zhang, C. Kalman Predictor-Based Proactive Dynamic Thermal Management for 3-D NoC Systems With Noisy Thermal Sensors. *IEEE Trans.-Comput.-Aided Des. Integr. Circuits Syst.* **2017**, *36*, 1869–1882. [CrossRef]
15. Liu, J.; Fan, R.; Li, Z.; Harnpornchai, N.; Qian, J. Proactive Cooling Control Algorithm for Data Centers Based on LSTM-Driven Predictive Thermal Analysis. *Appl. Syst. Innov.* **2026**, *9*, 21. [CrossRef]
16. Lin, J.; Lin, W.; Lin, W.; Wang, J.; Jiang, H. Thermal prediction for air-cooled data center using data driven-based model. *Appl. Therm. Eng.* **2022**, *217*, 119207. [CrossRef]
17. Coenen, D.; Kim, M.; Oprins, H.; Ban, Y.; Velenis, D.; Van Campenhout, J.; De Wolf, I. Thermal modeling of hybrid three-dimensional integrated, ring-based silicon photonic–electronic transceivers. *J. Opt. Microsyst.* **2023**, *4*, 011004. Correction in *J. Opt. Microsyst.* **2024**, *4*, 019801. [CrossRef]
18. Chen, L.; Zhai, X.; Sun, Y.; Wang, D.; Yu, D.; Li, A.; Qian, C. Optical Module and Communications System. WIPO Patent WO2023077903A1, 11 May 2023.


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.