Compact Nonlinear Thermal Modeling of Packaged Integrated Systems

Zao Liu †, Sheldon X.-D. Tan†, Hai Wang ‡, Sahana Swarup†, Ashish Gupta
† University of California, Riverside, US
‡ UESTC, Chengdu, China
Intel Corporation, Chandler, US
*

Abstract— This paper proposes a new thermal nonlinear modeling technique for packaged integrated systems. Thermal behavior of complicated systems like packaged electronic systems may exhibit nonlinear and temperature dependent properties. As a result, it is difficult to use a low order linear model to approximate the thermal behavior of the packaged integrated systems without accuracy loss. In this paper, we try to mitigate this problem by using piecewise linear (PWL) approach to characterizing the thermal behavior of those systems. The new method (called *ThermSubPWL*), which is the first proposed approach to nonlinear thermal modeling problem, identifies the linear local models for different temperature ranges using the subspace identification method. A linear transformation method is proposed to transform all the identified linear local models to the common state basis to build the continuous piecewise linear model. Experimental results validate the proposed method on a realistic packaged integrated system modeled via the multi-domain/physics commercial tool, COMSOL, under practical power signal inputs. The new piecewise models can lead to much smaller model order without accuracy loss, which translates to significant savings in both the simulation time and the time required to identify the reduced models compared to applying the high order models.

I. Introduction

Temperature has become a major concern and constraint for high performance packaged integrated system design as more devices are integrated on a chip [1, 2, 3]. Thermal management and related design problems continue to be identified by the Semiconductor Industries Association Roadmap [4] as one of the key challenges during the next decade for achieving the projected performance goals of the industry. Thus, accurate and efficient thermal modeling and analysis is vital for the thermal-aware circuit, chip and package designs to improve performance, reliability, power reduction, and online temperature regulation techniques [5, 6, 7].

For thermal modeling of packaged integrated systems, existing works on HotSpot [8, 7] attempts to solve this problem by generating the compact thermal model in a bottom-up manner based on processor and package structures. However, such compact models may suffer from accuracy loss, and have to be calibrated with hardware if more accurate models are required. Recently, top-down behavioral thermal modeling methods have been proposed using the matrix pencil method [9] and the subspace identification method [10]. All those methods assume that that packaged integrated systems are linear for thier thermal behaviors.

However, we show that thermal systems are fundamentally nonlinear as thermal conductivities of silicon and package materials are temperature dependent. To mitigate this problem, we apply the piecewise linear (PWL) scheme to characterize the thermal behavior of complicated thermal systems. Our experiments show that the nonlinear effects in the thermal systems are typically mild and weak but are still significant enough to warrant the PWL modeling. PWL method can lead to smaller models and reduced modeling costs compared to high order model approximation. This is important as the costs of identifying and simulating the reduced models will grow at least quadratically, it is very critical to reduce the model order to maintain the efficiency gain from the reduced order modeling. The new modeling algorithm, *ThermSubPWL*, partitions the temperature into a number of ranges and perform modeling to identify the linear local models for each temperature range using *ThermSubCP* method. A linear transformation method, which avoids the existing multi-transition requirement, is proposed to transform the identified linear local-models to the common state basis to build the continuous piecewise linear model. To the best knowledge of the authors, the proposed method is the first work addressing the nonlinear thermal modeling problem. Experimental results validate the proposed method on a realistic packaged integrated system modeled by the multi-domain/physics commercial tool, COMSOL, under practical power signal inputs. The new piecewise linear models can lead to much smaller model order without accuracy loss, which translates to significant savings in both the simulation time and the time required to identify the reduced models compared to the simple modeling method by using the high order models.

This paper is organized as follows. Section II reviews the thermal modeling problem for packaged electronic systems. Section III introduces the new piece-wise thermal modeling method, which shows how each local model is built with proper power inputs and how the transition matrices are generated to give a continuous linear model. The experiment results of *ThermSubPWL* are presented in section IV. Finally, Section **??** concludes the paper.

II. Outline of thermal modeling problem

We first present how the power inputs are modeled in our problem. A microprocessor chip is partitioned into $p = n \times m$ power grids as shown in Fig. 1, where each square power grid has a power source as an input and its measured temperature at its adjacent 4 corners as outputs. We can abstract this power grid model into a discrete linear system with $p = n \times m$ power inputs and q temperature outputs as shown in Fig. 2. The $n \times m$ power input distribution at one time instance is defined as a *power map*, which can be measured or computed practically.

In general, the abstracted p-input and q-output thermal system could be represented as

$$\begin{aligned} x(t+1) &= F(x(t), u(t)) \\ y(t) &= G(x(t), u(t)), \end{aligned} \quad (1)$$

where $F(x)$ and $G(x)$ both are nonlinear vector functions of state variable vector $x(t)$ and input signal vector $u(t)$. In our

*This work is supported in part by NSF grant under under No. CCF-0902885, in part by Semiconductor Research Corporation(SRC) Grant under No. 2009-TJ-1991.

978-1-4673-3030-5/13/$31.00 ©2013 IEEE

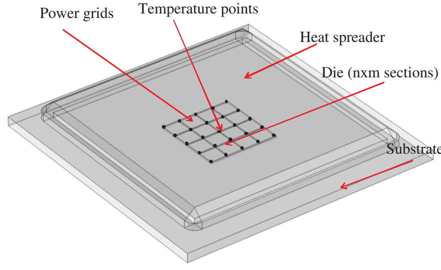


Fig. 1. Meshed chip and package

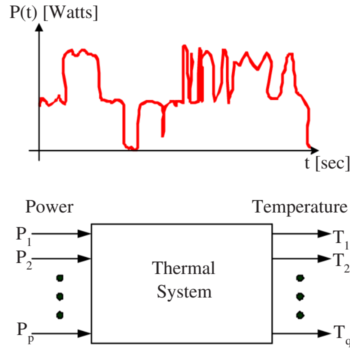


Fig. 2. The abstracted model system and correlated power inputs

problem, the input vectors $u(t) \in \mathbb{R}^{p\times 1}$ are the measured power input traces and output vectors $y(t) \in \mathbb{R}^{q\times 1}$ are the temperature responses.

Existing approaches typically assumes that thermal systems in Fig. 1 is linear. As a result, (1) can be rewritten as the standard linear state transition form:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{2}$$

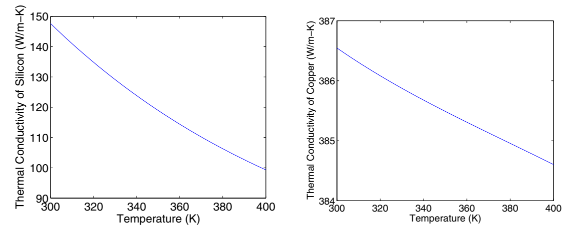
where $A \in \mathbb{R}^{l\times l}$ is a stable matrix, l is the number of states. $B \in \mathbb{R}^{l\times p}$, $C \in \mathbb{R}^{q\times l}$, and $D \in \mathbb{R}^{q\times p}$. With s input samples $u(t_i)$ and s output samples $y(t_i)$ where $i = 1, 2, \ldots, s$, the problem at hand is how to generate state matrices A, B, C, and D, where D is typically considered as a matrix of zeros.

Existing behavioral thermal modeling method works well when the system is linear and can be described by (2). However, thermal behavior of packaged electronic systems is typically nonlinear due to the temperature-dependent properties of the packaging materials [11]. Fig. 3 shows the temperature dependence of thermal conductivity of Si and Cu. Fig. 4 shows if we excite a chip package system shown in Fig. 1 with a sinusoid power input, we can clearly observe the harmonic components, which evidently indicates the nonlinearity of underlying thermal system although the nonlinear components are mild and weak. But such mild nonlinear behaviors, however, can still lead to significant loss of accuracy when low order is used as shown in Fig. 5.

To mitigate this problem, in this work, we propose to use linear models to represent the thermal behavior of the packaged electronic systems under different temperature ranges (piecewise linear model approach), thus this allows significant accuracy improvement by using just low order models.

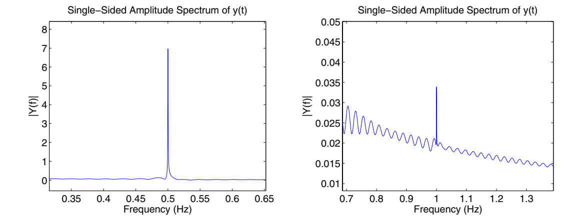
III. Piecewise linear thermal modeling approach – ThermSubPWL

In this section, we present the new thermal modeling technique (*ThermSubPWL*) to handle the nonlinearity of the thermal package.

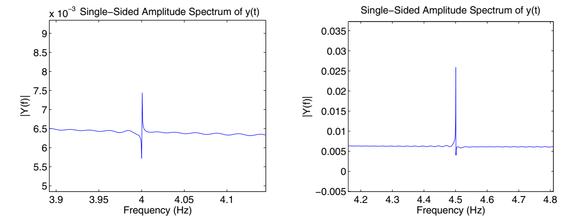


(a) (b)

Fig. 3. Temperature dependence of the thermal conductance (a) Silicon (b) Copper



(a) (b)



(c) (d)

Fig. 4. Frequency domain response of the thermal system under sinusoid input with frequency of 0.5HZ (a) baseband spectral (b) 1st order harmonics (c) 2nd order harmonics (d) 3rd order harmonics

A. Local models for partitioned temperature ranges

As shown in Fig. 4, thermal systems for packaged microprocessors show weakly nonlinearity. If we still use the linear models to characterize the system, we observed that we have to use higher order to get good approximation. Such approximate models of course will not show any nonlinear effects. But they will use more poles or states to emulate the effects of nonlinearity on thermal responses of those systems. So we will end up with much higher orders for the thermal models, which will hurt the performance of the thermal analysis. Such analysis also loses the nonlinear effects of the original systems.

To mitigate this problem and reduce the model order, in this section, we partition the temperature range into a number of sub-ranges and then we build the state space linear models for each temperature sub-ranges by the subspace identificatoin method [12] with sufficient training (to satisfy the persistently exciting condition) and these local models are then used to build piecewise linear thermal model for the whole thermal

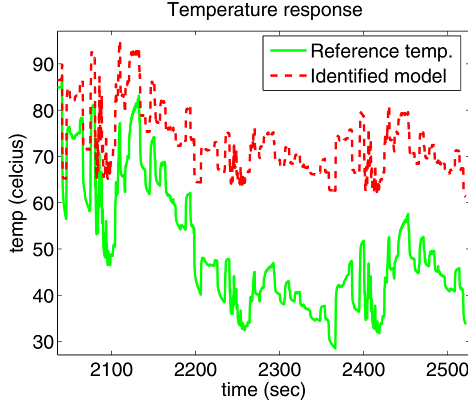


Fig. 5. Accuracy loss of the temperature response of the identified 4-th order linear model

system.

We notice that one issue with a such piecewise linear thermal modeling scheme is that temperature is location dependent across a whole package. There may existing temperature gradients among different locations. To mitigate this problem, we use the average temperature of any instance time to guide model switching. The thermal gradients in a well-designed chip is typically well managed and reduced by the online thermal management techniques [13, 14]. Even with some degrees of thermal gradients, the local models should be still valid as it is a localized model and should be valid for a temperature range.

In order to obtain the local models for different temperature ranges, we use stair-like input-output power-temperature data sets to identify these models as Fig. 6 shows. For example, the model M_i is identified during time interval $[t_{i-1}, t_{i+1}]$, which corresponds to the temperature range from $[T_{i-1}, T_{i+1}]$. Since model M_i is identified with the temperature data ranging from $[T_{i-1}, T_{i+1}]$, the correct using of the subspace identification method guarantees that the identified model is valid for this temperature range.

To avoid the predictability issue and improve the accuracy of the subspace identification method, we use independent power map configurations as given by sinusoid distribution discussed before to identify each local model for the corresponding temperature range.

By using the stair-like input-output data, the linear models of the subsystems in different temperature ranges could be accurately identified via the *ThermSubCP* method.

Note that, all the pairs of the two adjacent models, like M_i and M_{i+1}, are identified with a shared portion of data, which makes both models valid for the same temperature range, like $[T_i, T_{i+1}]$ shown in Fig. 6. The reason is that the transition from one thermal model to another thermal model is gradual, and this shared portion can facilitate determination of model transformation matrices as will be discussed below.

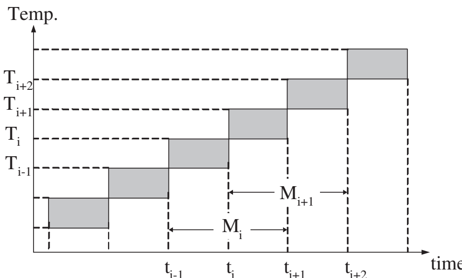


Fig. 6. Identification of linear subsystems for different temperature ranges

B. Determination of model transitions

The linear models for each subsystem could not be directly combined to build the piecewise linear model for the thermal system of the microprocessor package because these identified models are not built on the same state variable basis. Hence, linear transformations that transfer all these models to the same basis needs to be found. In [15], the transitions are assumed to be known at each time instance. Assuming the model transition is abrupt at the transition time instance t_k as shown in Fig. 7, it can be proved that the state of model M_a and state of model M_b is differed by a linear transformation T_{ba} as

$$x_{M_b}(t_k) = T_{ba} x_{M_a}(t_k) \tag{3}$$

where $x_{M_a}(t_k)$ is the state of model M_a at the transition time instance t_k, and $x_{M_b}(t_k)$ is the state of model M_b at the same transition time. Hence, to determine the linear transformation matrix T_{ba}, multiple transitions are required to solve the linear equations (4) in the sense of least squares as shown in [15].

$$[x_{M_b}(t_1), x_{M_b}(t_2), ...,] = T_{ba}[x_{M_a}(t_1), x_{M_a}(t_2), ...,] \tag{4}$$

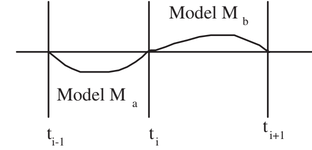


Fig. 7. Abrupt model transition at known time instance

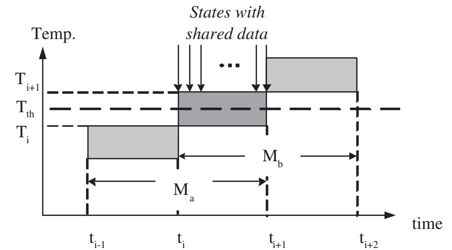


Fig. 8. Model transition from M_a to M_b

However, in our thermal system modeling, if we have specific temperature value for transitions between two models, we have to excite the states of the two models such that we have many independent states of two models and transitions happens between the two models with those states. This will lead to prolonged model identification process.

To mitigate this problem, we propose a *transition region* concept in this paper. We observe that the temperature transition from one model to another model is in general a gradual process as indicated in the time interval from t_i to t_{i+1} shown in Fig. 8, instead of an abrupt one that happens at a specific time instance. We define a transition region as shown in Fig. 8 in which both local models are valid (in other words, a state of model M_a will become a corresponding state of model M_b at any given time in this region).

As discussed before, the subspace identification method guarantees that any two adjacent models are valid for a portion of shared data sets from t_i to t_{i+1} (even though in the simulation we specify arbitrary T_{th} to determine the abrupt transition from M_a to M_b as shown in Fig. 8), thus, the relationship of the states for these two models within the range of the shared data set could be written as (5)

$$x_{M_b}(t_i : t_{i+1}) = T_{ba} x_{M_a}(t_i : t_{i+1}) \tag{5}$$

in which the matlab-like notation $x_{M_a}(t_i : t_{i+1})$ represents the states of model M_a during t_i to t_{i+1}, and $x_{M_b}(t_i : t_{i+1})$ represents the states of model M_b during t_i to t_{i+1} as Fig. 8 shows.

Hence, in this way, instead of requiring multiple model transitions with independent states as in [15], we explore the states in the gradual transition region and compute the transformation matrix T_{ba} by solving (5) in a least square sense. By using T_{ba}, we could transfer model M_b to the basis of model M_a through

$$x_{M_b}(t_k) = T_{ba} x_{M_a}(t_k) \tag{6}$$

in which t_k is the transient time points.

Following this method, we could calculate the transformation matrices between any two adjacent thermal models by (5), and transform the model basis by (6). In this way, it is straightforward to transform all the identified local linear model to the common basis. We illustrate this by a 3-local-model system that has model M_a, M_b and M_c. It is straightforward to transform other model stats into common model basis of M_a by

$$\begin{aligned} x_{M_b}(t_k) &= T_{ba} x_{M_a}(t_k) \\ x_{M_c}(t_k) &= T_{cb} T_{ba} x_{M_a}(t_k) \end{aligned} \tag{7}$$

As a result, we can just use the common state x_{M_a} as the local model states for all the local models. With the linear local model built from different temperature range onto the same state basis through linear transformations, the piecewise linear model could smoothly switch from one model to another model , which benefits the simulation accuracy.

IV. Implementation and numerical results

A. Modeling and simulation environment setup

The packaged microprocessor design used in this study is shown in Fig. 9, where the convective boundary on the top of heat sink models the convective cooling from the fan placed above the processor. The aluminum heat sink is glued to the copper integrated heat spreader (IHS) that is attached to silicon die through a thin layer of thermal interface material. The materials and geometries of the major parts of the package are shown in Table. I, and we partition the die area into 4 × 4 power grids as shown in Fig. 10, and each grid represents a different function block.

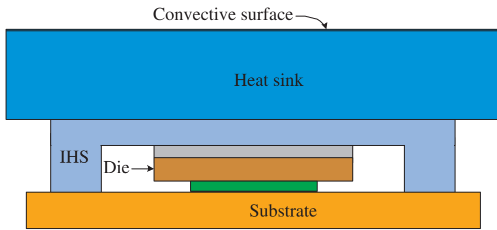


Fig. 9. Microprocessor chip package

To model the power consumption of these function blocks, the input power sources are placed in these power grids and we measure the temperature at the adjacent 4 corners of each square power grid. As a result, we end up with 16-input and 25-output thermal system. The convection coefficient of 450 ($W/(m^2 \cdot K)$) is used to model the convective air cooling effect from the cooling fan on top of the chip package.

To build a more realistic package with right dimension and materials, we applied COMSOL 4.1 [16] to build the package structures with on-chip power waveforms as inputs. The thermal response was obtained by COMSOL using the finite element method under the input power maps we generated. Fig. 11 shows the steady state temperature distribution under a given power input on the constructed package and the chip.

TABLE I
Material and geometry of the microprocessor package

Parts	Material	Dimensions (mm)
Die	Silicon	$10 \times 10 \times 0.7$
IHS	Copper	$31 \times 31 \times 1.5$
Heat sink	Aluminium	$64 \times 64 \times 6.3$
Substrate	FR4	$37.5 \times 37.5 \times 1.3$

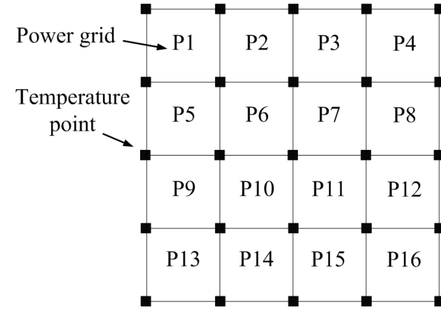


Fig. 10. Partitioned die area with power grids and temperature points

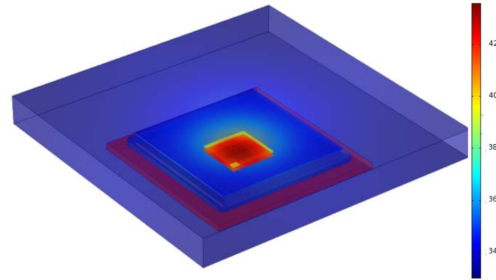
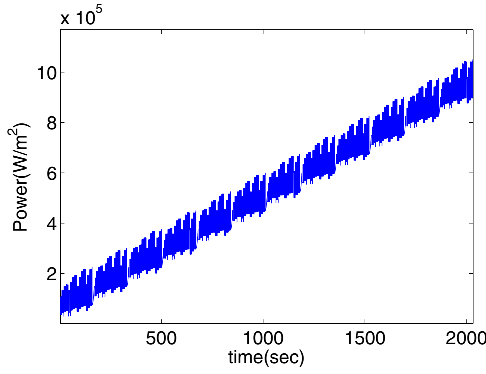


Fig. 11. Steady state temperature distribution simulated by COMSOL 4.1

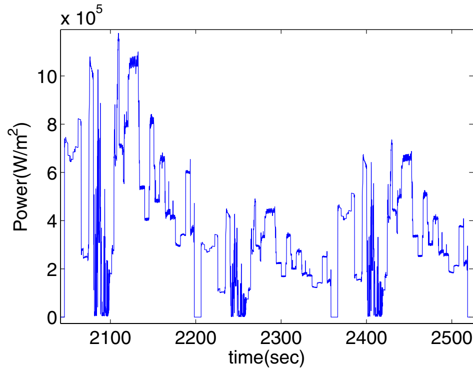
The transient power input for each power grid (its magnitudes will be determined by the specific power map) is shown in Fig. 12. At the model identification stage, PRBS (Pseudo Random Binary Sequence) signals with stair-like shown in Fig. 12(a) envelops are used as inputs to characterize the system parameters; and in the validation phase, the input signals are from our industry partner as shown in Fig. 12(b). PRBS signal has the white-noise like spectrum so that it can excite all the thermal system states.

B. Piecewise linear model identification and validation

In this case, the stair-like envelop contains 12 'steps' that corresponds to 12 different ranges of input power intensity as shown in Fig. 13. We could arbitrarily partition the data and attribute them to different linear models that need to be identified with these data. At the beginning, we use 'scheme-1' shown in Fig. 13 (a) to identify the linear models. In this scheme, each model is identified based on two consecutive data sets, and the adjacent models are built with one shared data set. In this way, 11 models will be identified in total given the 12 data sets, and the piecewise linear model is to be built with these 11 local models. In order to avoid the predictability



(a) The stair-like input signal used for model identification



(b) The transient input signal used for model validation

Fig. 12. Input power trace for model identification and validation

issue as discussed before, for each range of the input power, 16 orthogonal configurations or power maps are generated.

By choosing 4-th order model, and using the subspace identification method, all the 11 linear models could be identified. Applying the proposed method to determine the transformation matrices, all the linear models could be transformed to the same basis. Since the piecewise linear model built up in this way contains multiple local models, it is reasonable to partition the overall temperature range into the sub-ranges that the local models correspond to. The simulation result in Fig. 14(a) confirms that the temperature value predicted by the output of the identified piecewise model (dash line) closely matches the reference data (solid line).

In comparison, we also use different schemes of data partition. By using 'scheme-2' and 'scheme-3', we end up building the piecewise linear models with 6 local models and 4 local models respectively. From the simulation, it clearly shows the performance improvement as more linear local models are used as shown in Fig. 14, and the output error information of the identified system is summarized in Table II, where we list the maximum of the mean errors (*Max Mean error*) among all the ports over the entire transient simulation period. We can clearly observe that, for the same order, the error reduces as number of the linear models in use increases, which shows a compelling evidence of using piecewise linear model for compact thermal behavior modeling and simulation.

On the other hand, to make the linear model achieve comparable accuracy with the piecewise linear model, high order model needs to be chosen. In our experiment, we used 20412 transient time points to identify the model. As summarized in Table III, the time required to identify (ID time) the high order linear model (LM) is 627.1 seconds, while on the other hand, the time required to identify all the piecewise linear models (PLM) is 63.8 seconds. Hence, the speedup factor for model identification is 9.8 comparing with linear model. Also, we used 25412 time points in transient simulation, and the high order linear model uses 22.2 seconds to conclude the simulation, and the piecewise linear model uses only 7.88 second to conclude the simulation, which is approximately 35% of the simulation time of the high order linear model.

We remark that although in our case, order 15 of the linear model is not significantly higher than the order 4 in a sense, yet, the time required to identify the state space model through subspace identification method increases significantly because a large amount of input and output data are required to identify the state space model accurately. As a result, choosing low order model to identify the targeted dynamic system leads to substantial savings in subspace identification method, which is important in the process of building and calibration a dynamic model in a dynamically changing environment. Also, piecewise linear model achieves substantial savings in simulation time because the lower order model is used in simulation.

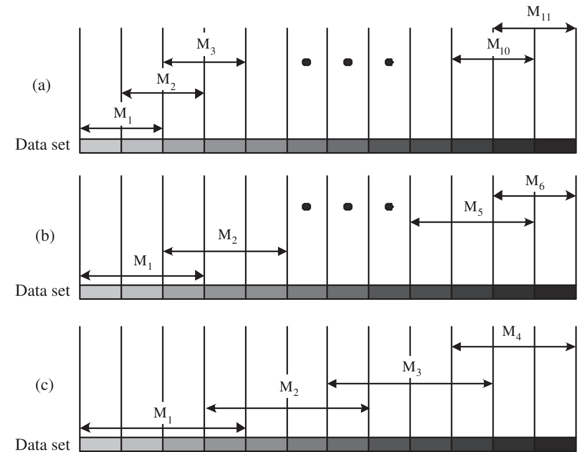


Fig. 13. Data partition schemes for model identification (a) scheme-1(11 local models) (b) scheme-2(6 local models) (c) scheme-3 (4 local models)

TABLE II
Errors with different identified models (order: 4)

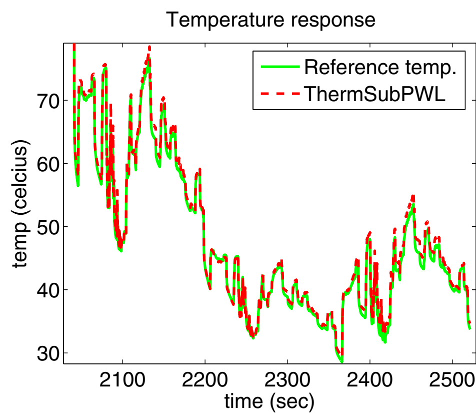
Num of linear models in use	11	6	4
Max mean error	2.1%	3.9%	5.9%

TABLE III
Comparison of model accuracy and cost

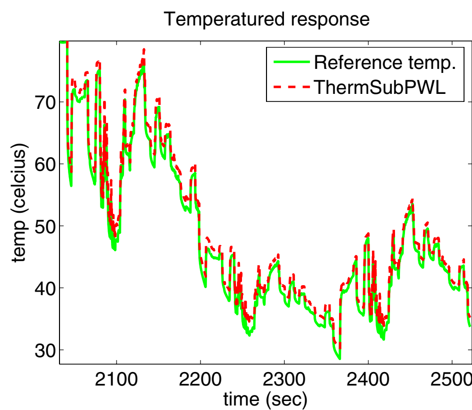
Comparison Items	Error	ID time	simulation time
PLM (order:4)	2.1%	63.8 sec	7.88 sec
LM (order:15)	2.3%	627.1 sec	22.2 sec

V. Conclusion

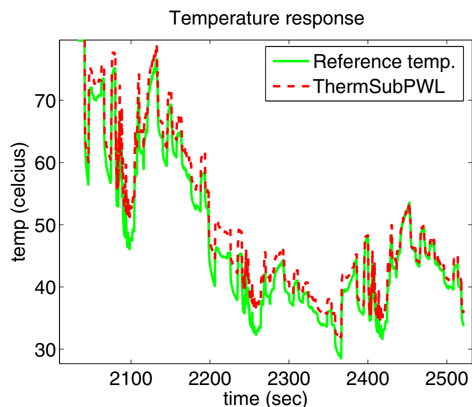
This paper has proposed a new thermal nonlinear modeling technique (*ThermSubPWL*) for packaged microprocessor systems. The new modeling algorithm, *ThermSubPWL*,



(a) PLM built with 11 local models



(b) PLM built with 6 local models



(c) PLM built with 4 local models

Fig. 14. Transient view of one on-chip temperature response of the piecewise linear models (PLM) built with different number of local models

partitions the temperature into a number of ranges and perform modeling using *ThermSubCP* to identify the linear submodels for each temperature range. A linear transformation method, which is based on newly proposed transition region concept, is proposed to transform the identified linear local-models to the common state basis to build the continuous piecewise linear model. Experimental results have validated the proposed method on a practical microprocessor package modeled by commercial multi-domain/physics tool, COMSOL V4.1, under practical power signal inputs. The new piecewise models can lead to much lower model order without accuracy loss, which translates to significant savings in model identification time and simulation time compared to the time required to identify the high order models as shown in our experiments.

VI. REFERENCES

[1] N. Allec, Z. Hassan, L. Shang, R. P. Dick, and R. Yang, "ThermalScope: Multi-scale thermal analysis for nanometer-scale integrated circuits," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, (Piscataway, NJ, USA), pp. 603–610, IEEE Press, 2008.

[2] W. Huang, E. Humenay, K. Skadron, and M. R. Stan, "The need for a full-chip and package thermal model for thermally optimized IC designs," in *Proceedings of the 2005 international symposium on Low power electronics and design*, (New York, NY, USA), pp. 245–250, ACM, 2005.

[3] Y.-K. Cheng, C.-C. Teng, S.-M. Kang, and C.-H. Tsai, *Electrothermal analysis of VLSI systems*. New York, NY, USA: Cambridge University Press, 2000.

[4] "International technology roadmap for semiconductors (ITRS), 2011," 2011. http://public.itrs.net.

[5] M. Pedram and S. Nazarian, "Thermal modeling, analysis, and management in VLSI circuits: Principles and methods," *Proc. of the IEEE*, vol. 94, pp. 1487–1501, Aug. 2006.

[6] D. Brooks and M. Martonosi, "Dynamic thermal management for high-performance microprocessors," in *Proc. of Intl. Symp. on High-Performance Comp. Architecture*, pp. 171–182, 2001.

[7] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan, "Temperature-aware microarchitecture," in *Proc. Int. Symp. on Computer Architecture (ISCA)*, pp. 2–13, 2003.

[8] W. Huang, M. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S. Velusamy, "Compact thermal modeling for temperature-aware design," in *Proc. Design Automation Conf. (DAC)*, pp. 878–883, 2004.

[9] D. Li, S. X.-D. Tan, and M. Tirumala, "Architecture-level thermal behavioral characterization for multi-core microprocessors," in *Proc. Asia South Pacific Design Automation Conf. (ASPDAC)*, pp. 456–461, 2008.

[10] T. Eguia, S. X.-D. Tan, R. Shen, E. H. Pacheco, and M. Tirumala, "General behavioral thermal modeling and characterization for multi-core microprocessor design," in *Proc. Design, Automation and Test In Europe. (DATE)*, pp. 1136–1141, March 2010.

[11] M. Rencz and V. Szekely, "Studies on the nonlinearity effects in dynamic compact model generation of packages," *IEEE Transactions on Components, and Packaging Technologies*, vol. 27, pp. 124–130, March 2004.

[12] T. Katayama, *Subspace Methods for System Identification*. Springer, 2005.

[13] I. Yeo, C. C. Liu, and E. J. Kim, "Predictive dynamic thermal management for multicore systems," in *Proc. Design Automation Conf. (DAC)*, DAC '08, (New York, NY, USA), pp. 734–739, ACM, 2008.

[14] A. K. Coskun, T. S. Rosing, and K. C. Gross, "Utilizing predictors for efficient thermal management in multiprocessor SoCs," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 28, pp. 1503–1516, October 2009.

[15] V. Verdult and M. Verhaegen, "Subspace identification of piecewise linear systems," in *Proc. 43rd IEEE Conference on Decision and Control (CDC)*, pp. 3838–3843, 2004.

[16] www.comsol.com, "Comsol mutiphysics: User guide," *Version 4.1*.