Compact Nonlinear Thermal Modeling of Packaged Microprocessors

Zao Liu [†], Sheldon X. -D. Tan[†], Hai Wang[†], Yingbo Hua[‡] and Ashish Gupta[‡]
[†] Department of Electrical Engineering, University of California, Riverside, CA 92521
[‡] Intel Corporation, Chandler, AZ 85226 [*]

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 identifies the linear sub-models for different temperature ranges using a linear behavioral modeling method such as the subspace identification method. A linear transformation method is proposed to transform all the identified linear sub-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.

1. INTRODUCTION

Temperature has become a major concern and constraint for high performance packaged integrated system design as more devices are integrated on a chip. Thermal management and related design problems continue to be identified by the Semiconductor Industries Association Roadmap [1] as one of the five 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 [6, 2, 8].

For thermal modeling for packaged integrated systems, existing works on HotSpot [3, 8] 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 [4] and the subspace identification method [5]. The subspace identification method works well when the thermal systems are linear. But practical thermal systems may exhibit nonlinear and temperature dependent properties (due to temperature dependent materials). As a result, the subspace identification method may end up with the models with high orders to achieve the accuracy requirement, which will lead to high computing costs to identify and evaluate the models.

In this paper, we try to mitigate this problem by using piecewise linear (PWL) approach to characterizing the thermal behavior of the packaged microprocessors. The PWL nonlinear modeling is rationalized by the facts that the nonlinear effects in the thermal systems are mild and weak as shown in the paper but are still significant enough to justify the PWL modeling. Also 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 method identifies the linear sub-models for different temperature ranges using a linear behavioral modeling method such as the subspace identification method. A linear transformation method is proposed to transform all the identified linear sub-models to the same basis to build the continuous PWL model. We propose a new method to determine the transformation matrices based on the models we identified. Experimental results validate the proposed method on a realistic packaged microprocessor modeled via the multi-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 simulation and identification cost of reduced models compared to the high order models.

This paper is organized as follows. Section 2 reviews the thermal modeling problem for packaged electronic systems. Section 3 introduces the new piece-wise thermal modeling method, which shows how each sub-model is built with proper power inputs and how the transition matrices are generated to give a continuous linear model. The experiment results are demonstrated in section 4. Finally, Section 5 concludes the paper.

2. 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.

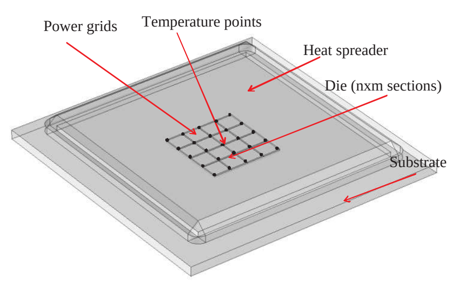


Figure 1: Meshed chip and package

Then we can abstract this $p = n \times m$ meshed grid into a discrete linear MIMO system with p inputs and q outputs as described by (1).

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

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 problem, the input vectors $u(t) \in \mathbb{R}^{p \times 1}$ are the measured

[*]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.

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

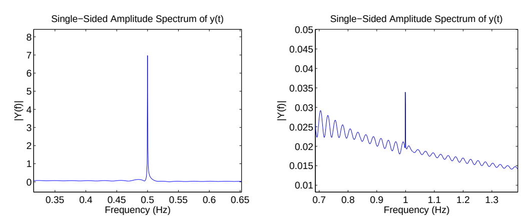
Existing approaches typically assumes that thermal systems in Fig. 1 is linear. As a result, (1) can be rewritten the standard linear state transition form:

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

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 methods like the subspace identification method [5] 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 (like Cu, and Si) [7]. Fig. 2 shows the baseband spectral and the harmonics of the package shown in Fig. 1 with a sinusoid power input. Although the nonlinear components are mild and weak, such mild nonlinear behaviors can still lead to significant loss of accuracy when low order is used as shown in Fig. 3.

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



(a) Baseband (b) Harmonics

Figure 2: Frequency domain response of the thermal system under sinusoid input with frequency of 0.5HZ

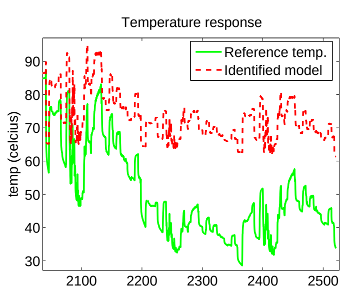


Figure 3: Accuracy loss of the temperature response of the identified 4-th order linear model

3. PWL THERMAL MODELING APPROACH

3.1 Identification of state space linear models

In this section, state space linear models for different temperature ranges are identified, and these models are to be used to build piecewise linear thermal model for the thermal systems.

As discussed before, we treat the thermal system of the microprocessor package as a MIMO system in which the number of correlated power inputs is p and the number of temperature output is q. Since the properties of the package materials are more or less temperature dependent as mentioned before in section 2, the state space models of the thermal behavior of the microprocessor package are different for different temperature ranges, which essentially lead to the nonlinearity. Thus, we use stair-like input-output data sets to identify these models as Fig. 4 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.

In order to avoid the predictability issue and improve the accuracy of the subspace identification method, we use independent power map configurations as input to identify each state space model [5]. We arbitrarily choose orthogonal power map configurations generated by the 2-D function set

$$g_{mn}(x, y) = sin(m\pi x/L_x) sin(n\pi y/L_y) \quad (3)$$

in which m and n are the indices starting from 1; x and y are the position variables; L_x and L_y are the size of the chip in the direction of x and y axis respectively.

By using the stair-like input-output data, the linear models of the subsystems in different temperature ranges could be accurately identified via the subspace identification 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. 4. 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 soon.

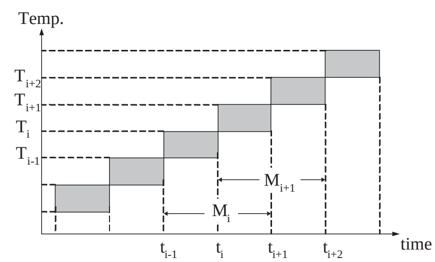


Figure 4: Identification of linear subsystems for different temperature ranges

3.2 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 [9], the transitions are assumed to be known at a certain time instance t_k, and it can be proved that the state of model M_p and state of model M_q is differed by a linear transformation at t_k. T_{pq} as

$$x_{Mp}(t_k) = T_{pq} x_{Mq}(t_k) \quad (4)$$

where $x_{Mp}(t_k)$ is the state of model M_p at the transition time instance t_k, and $x_{Mq}(t_k)$ is the state of model M_q at the same transition time. Hence, to determine the linear transformation matrix T_{pq}, multiple transitions are required to solve the linear equations (5) in the sense of least squares.

$$[x_{Mp}(t_1), x_{Mp}(t_2), ...,] = T_{pq}[x_{Mq}(t_1), x_{Mq}(t_2), ...,] \quad (5)$$

However, in our thermal system, transition from one model to another model is a gradual migration as indicated in the time interval from t_i to t_{i+1} shown in Fig. 5, instead of an abrupt one that happens at a specific time instance. We could arbitrarily specify a transition threshold temperature T_{th} as shown in Fig. 5 inside the transition region to mimic abrupt model switching from M_p to M_q during simulation. However, in the process of building piecewise linear model, our method makes use of this gradual transition region to determine the state transition matrix T_{pq}. As discussed before, the subspace identification method guarantees that any two adjacent models are valid for a portion of shared data

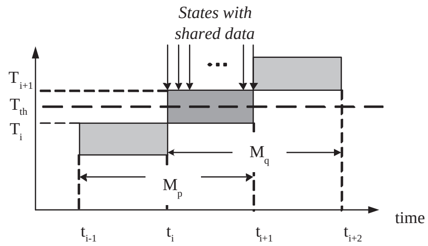


Figure 5: Model transition from M_p to M_q

sets from t_i to t_{i+1} (even though the in simulation we specify arbitrary T_{th} to determine the abrupt transition from M_p to M_q), thus, the relationship of the states for these two models within the range of the shared data set could be written as (6)

$$x_{M_p}(t_i : t_{i+1}) = T_{pq} x_{M_q}(t_i : t_{i+1}) \tag{6}$$

in which the matlab-like notation $x_{M_p}(t_i : t_{i+1})$ represents the states of model M_p during t_i to t_{i+1}, and $x_{M_q}(t_i : t_{i+1})$ represents the states of model M_q during t_i to t_{i+1} as Fig. 5 shows. Hence, in this way, instead of requiring multiple model transitions as in [9], we make use of the states in the gradual transition region and find the transformation matrix T_{pq} by solving (6) in a least square sense. By using T_{pq}, we could transfer model M_p to the basis of model M_q.

Following this method, we could transfer all the identified linear models to the same basis with the corresponding transformation matrices, thus, it is straightforward to build the piecewise linear model by partitioning the overall temperature range into the sub-ranges and attributing them to the corresponding linear models. In this way, different linear models are chosen in simulation based on the temperature ranges of the chip package, improving the accuracy of the prediction of the on-chip temperature response.

4. IMPLEMENTATION AND NUMERICAL RESULTS

4.1 Modeling and simulation environment setup

The packaged microprocessor design used in this study is shown in Fig. 6, 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. 1, and we partition the die area into 4×4 mesh grid as shown in the previous section.

Table 1: Material and geometry of the 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	$37.5 \times 37.5 \times 1.3$

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 [10] 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.

The transient power input for each power grid (its magnitudes will be determined by the specific power map) is shown in Fig. 7. At the model identification stage, PRBS (Pseudo Random Binary Sequence) signals with stair-like envelops are used as inputs to characterize the system parameters; and in the validation phase, the input signals are from our industry partner. PRBS signal has the white-noise like spectrum so that it can excite all the thermal system states.

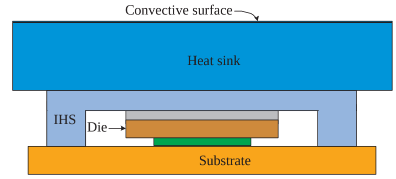
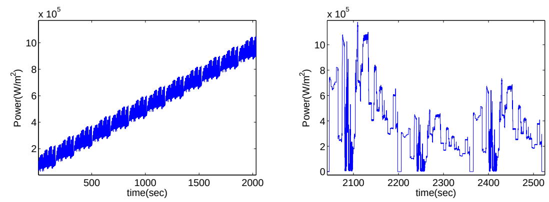


Figure 6: Microprocessor chip package



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

(b) The transient input signal for model validation

Figure 7: Input power trace for model identification and validation

4.2 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. 8. 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. 8 (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 sub-models. In order to avoid the predictability issue as discussed before, for each range of the input power, 16 orthogonal configurations are generated by $g_{mn}(x, y)$ as defined in (3).

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 sub-models, it is reasonable to partition the overall temperature range into the sub-ranges that the sub-models correspond to. The simulation result in Fig. 9(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 sub-models and 4 sub-models respectively. From the simulation, it clearly shows the performance improvement as more linear sub-models are used as shown in Fig. 9, and the output error information of the identified system is summarized in Table 2, 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 3, 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.

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.

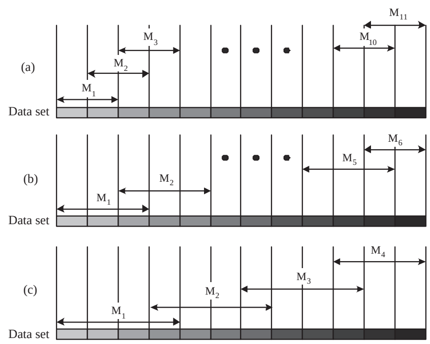


Figure 8: Data partition schemes for model identification (a) scheme-1(11 sub-models) (b) scheme-2(6 sub-models) (c) scheme-3 (4 sub-models)

Table 2: Errors with PWL models (order: 4)

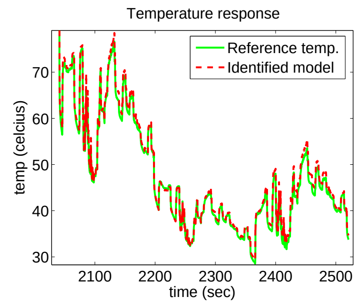
Num of linear models in use	11	6	4
Max mean error	2.1%	3.9%	5.9%

Table 3: Comparison of model accuracy and cost

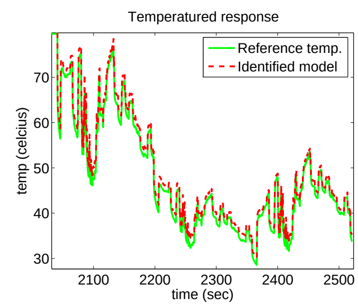
Comparison Items	Error	ID time	simulation time
PLM (order:4)	2.1%	63.8 sec	7.88 sec
LM (order:15)	2.3%	627.1 sec	22.2 sec

5. CONCLUSION

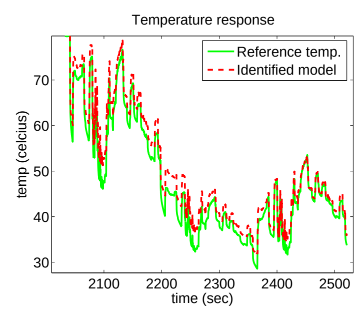
This paper has proposed a new thermal nonlinear modeling technique for packaged microprocessor systems. The contributions of this paper are three folds as summarized below: (1) piecewise linear model scheme has been proposed to model the thermal behavior of a microprocessor package system with nonlinear effects. (2) A novel procedure has been proposed to identify the linear sub-models for different temperature ranges. (3) A new method has been applied to find the linear transformation matrices that transform all the identified linear models to the same basis. 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 means significant savings in model identification and simulation time.



(a) PLM built with 11 sub-models



(b) PLM built with 6 sub-models



(c) PLM built with 4 sub-models

Figure 9: Transient view of one on-chip temperature response of the piecewise linear models (PLM) built with different number of sub-models

6. REFERENCES

[1] "International technology roadmap for semiconductors(ITRS)," 2009 update, http://public.itrs.net.

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

[3] 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)*, 2004, pp. 878–883.

[4] 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)*, 2008, pp. 456–461.

[5] Z. Liu, S. X.-D. Tan, H. Wang, R. Quintanilla, and A. Gupta, "Compact thermal modeling for package design with practical power maps," in *Proc. 1st International IEEE Workshop on Thermal Modeling and Management: Chips to Data Centers (TEMM)*, 2011, pp. 1–5.

[6] M. Pedram and S. Nazarian, "Thermal modeling, analysis, and management in VLSI circuits: Principles and methods," *Proc. of the IEEE*, vol. 94, no. 8, pp. 1487–1501, Aug. 2006.

[7] 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, no. 1, pp. 124–130, March 2004.

[8] 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)*, 2003, pp. 2–13.

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

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