Multi-disciplinary Design Optimization of Electronic Packages

Yatharth Vaishnani[1]
Texas A&M University, College Station, Texas, 77840, USA

This report presents a comprehensive methodology for the design optimization of a Ball Grid Array Electronic Package, using the numerical methods for multi-objective design optimization. A multi-disciplinary design and multi-physics analysis tools are used to optimize key design parameters. To compute the conjugate heat transfer and thermal strain generated in BGA package, low fidelity models are used including resistor network with agreeable results with high fidelity models. The problem formulated with mixed-integer design variables along with the non-linear constraint functions and further optimized with gradient based as well as heuristic based method. A comprehensive post optimality study is also presented with scaling, sensitivity and multi-objective trade-off analysis using Pareto front. The present methodology can be applied to different electronic product design at various packaging levels.

Nomenclature

D_b	=	Diameter of solder balls (mm)
F	=	Input to module
G	=	Output from module
h	=	Heat transfer coefficient (W/m^{2}K)
H_b	=	Height of solder balls (mm)
J	=	Design objective
k	=	Thermal conductivity (W/mK)
l_s	=	Length of spreader (mm)
N_b	=	Number of solder balls
p	=	Pitch of solder ball array (mm)
R	=	Thermal Resistance (K/W)
t	=	Thickness (mm)
w_s	=	Width of spreader (mm)
x	=	Design variable

[1] Graduate Student, Mechanical Engineering, Texas A&M University, College Station, TX.

I. Introduction

THE majority of electronic parts failures are packaging-related. Packaging, as the barrier between electronic parts and the environment, is very susceptible to environmental factors. It has been observed that almost 85% of all failures in electronic packages today are caused by thermal effects and lack of improvement in designs which ensure a good heat dissipation [l]. To prevent this, a complete understanding of the heat dissipation characteristics from the package and into the surroundings is essential. This comprehension can lead to better designs of electronic packages which can be optimized for design and improve their thermal performance. There are mainly two approaches to proceed with the thermal modeling of these electronic packages. The first approach involves a complete numerical solution of the heat dissipation problem by means of FEA (Finite Element Analysis) using commercial packages like ANSYS Icepak or FloTherm XT. In fact, it is very much possible with modern computers and is widely applied to get highly accurate results, which can be incorporated in design optimization algorithms. But it is still relatively time-consuming and expensive to perform full package modeling from the chip to the system level.

A compact thermal model that captures the key thermal features of electronic packages but has reduced complexity is a useful and economic method to predict the junction temperature and analyze the heat dissipation in various types of electronic packages. This reduced complexity improves computational efficiency, allowing thermal simulations of the electronic system to be performed, using personal computers within a reasonable time. A typical electronic system comprises several printed circuit boards (PCBs), each of which may contain more than twenty, different types of packages. The mesh number and computational time will become prohibitive for simulation, and even the slightest design change will take several hours or even days to reevaluate its impact. A compact model may be simply a numerical model where the detailed features are simplified such that the meshes are not unusually skewed or have extremes in length scales, and the computation time vastly reduced. By using these low fidelity models, such as the resistance network models, the thermal performance can be gauged reasonably well, and the network can be implemented in a design optimization algorithm and obtain improved designs. The study is based on the design optimization of the solder ball array in a typical BGA package by using both a genetic algorithm and a gradient based – sequential quadratic programming algorithm to determine the global optimum of the design and where it occurs and validating the results with experimental and Finite Element Analysis results.

II. Problem Formulation

Current BGA Package designs are focused on minimizing the size of the package while designing the system to have good thermal dissipation effects by implementing high fidelity models which are computationally expensive and difficult to analyze. The aim of this work is to achieve reasonably good results by optimizing the package design using a resistance network model for objectives as described in the following section.

A. Objective

To design an electronic package to have optimal thermal performance and cost, the solder ball array has been chosen as the module which is to be optimized. To achieve this objective, different modules are considered which are illustrated with the help of a N^2 diagram. Therefore, the objectives can be defined as:

$$\begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = \begin{bmatrix} Temperature\ rise\ in\ Chip \\ Cost\ of\ the\ Package \end{bmatrix}$$

The objectives are determined as expressions in terms of the design variables and the module to be optimized is considered to be the solder ball array.

B. Design Variables

The design vector is defined as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} N_b \\ p \\ D_b \\ H_b \end{bmatrix} = \begin{bmatrix} Number\ of\ solder\ balls \\ Pitch\ of\ solder\ ball\ array\ (mm) \\ Diameter\ of\ solder\ balls\ (mm) \\ Height\ of\ solder\ balls\ (mm) \end{bmatrix}$$

C. N^2 Diagram

Input		N_b, P, d_b, h_b						
	Constraints							**C_s, C_t**
		SBA	l_s, w_s		G_{sba}	G_{sba}	G_{sba}	
			Heat Spreader	G_{hs}	G_{hs}	G_{hs}	G_{hs}	
				Mold	G_m	G_m	G_m	
	F_t				**Thermal**	F_t		**F_t**
	F_s					**Strain**		
							Cost	**F_c**
								Output

Fig. 1 N^2 diagram with all the variables and modules

The N^2 diagram with all the design variables, input and output of different modules are shown in the Fig. (1). The modules shown in the figure are not completely decoupled from each other. The geometry module provides the dimensions and materials properties to the thermal resistance and thermal strain module. The temperature obtained from the thermal resistance module is feedforwarded to the thermal strain module and the constraint function is calculated. If the constraints are not violated, the cost module calculated for the output.

III. Thermal Resistance Model

A thermal resistor network has been used to model the mechanisms of steady state heat transfer within the BGA package. To account for different heat flows within the package, three types of thermal resistances have been considered: layer resistances have been used to model the various layers within the substrate of the package and the components that are layered on top of each other [2]. To simulate the approximation of the heat flow in three dimensions as changes in the cross-sectional area arise along the heat flow path, Spreading/Constriction resistances have been considered. To account for the interaction between device and surroundings, boundary resistances have been considered in this model. The formulation of the resistor network is shown as a schematic in the Fig. (2). The analytical formulation of the above-mentioned resistances will be described in detail as following.

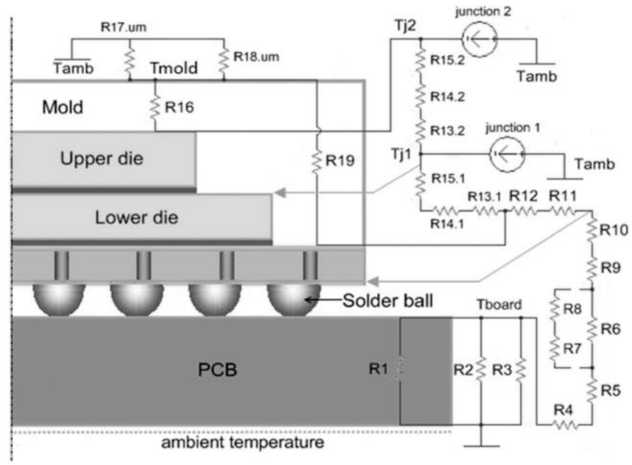

Fig. 2 Resistor network for temperature calculations

A. Layer Resistance

The BGA package is a superposition of several different structures named layers, which result from association of different parts, e.g., a substrate with copper traces. Material heterogeneity and anisotropy arise from the variety of material types and from the particular disposition of the different parts, respectively. Usually, a layer can be modeled as made of an orthotropic material, characterized by two different values of thermal conductivity, i.e., in-plane (k_{II}) and out-of-plan ($k_{\perp}$) conductivities, defined as follows:

$$k_{||} = \frac{\sum_{i=1}^{N} k_i t_i}{\sum_{i=1}^{N} t_i}$$

$$k_{||} = \frac{\sum_{i=1}^{N} t_i}{\sum_{i=1}^{N} t_i / k_i}$$

where t_i (m) is the thickness of the i^{th} layer, and k_i ($W\ m^{-1} K^{-1}$) is its thermal conductivity. According to the heat flow direction, the specific thermal resistance of a layer is given as follows:

$$R'' = \frac{t_{||}}{k_{||}} \ \ or\ \ R'' = \frac{t_{\perp}}{k_{\perp}}$$

where, t is the thickness of the structure (m) along the considered direction.

B. Spreading/Constriction Resistance

This resistance accounts for the three-dimensional propagation of the heat flow, arising in particular, where the heat transfer surface area changes abruptly. The term spreading refers to an increase in the area, whereas constriction is generally intended for a decrease. The spreading/constriction of heat flow occurs especially in the propagation, from the source to the substrate, from the substrate to the printed circuit board (PCB) through the ball array, within the PCB through thermal vias, and from the PCB to the heat spreader.

The following relations have been obtained analytically with high accuracy [3]. Kennedy et al. [3] considered a heat source placed on the top of an isotropic spreader with the bottom surface maintained at uniform temperature.

Song et al. [4] generalized the solution of the heat transfer problem by considering a convective boundary condition at the bottom of the spreader. Using these analytical models, highly accurate analytical expressions for the resistance has been obtained, shown as follows:

$$R_{ave} = \frac{0.5 \cdot \alpha \cdot (1-e)^{\frac{3}{2}} . \left(\frac{tanh\,(\lambda_c . {}^{\tau}/_{\alpha}) + \frac{\lambda_c}{B_i}}{1 + \left({}^{\lambda_c}/_{B_i} \right) \cdot tanh(\lambda_c . {}^{\tau}/_{\alpha})} \right)}{\sqrt{\pi} \, . \, k \, . \, a}$$

where,

$$e = \frac{a}{b}, \tau = \pi + \frac{1}{\sqrt{\pi} \, . \, e} \quad and \;\; Bi = \frac{1}{\pi . k . b . R_o}$$

C. Boundary Resistance

To take into account the ambient convection and radiation effects or dissipation, the boundary condition resistances are considered where the package is located. These resistances can be calculated with well-known analytical relations.

$$R_c = \frac{1}{hA}$$

Here, h is the convective heat transfer coefficient (W$m^{-2}K^{-1}$), and A is the heat transfer area (m^2).

D. Constraints and Bounds

The constraints imposed on the design of the BGA package can be classified into geometric limit in solder ball array, thermal strain limit by solder ball and operating temperature limit for the chip. The mathematical formulation of the geometric constraint can be written as,

$$x_3 - x_2 < 0.2\, mm$$
$$x_4 < x_3$$

By applying the definitions of the resistances defined in the section above, thermal resistances corresponding to different stacked layers of the BGA are calculated. Depending on the direction of heat flow, equivalent thermal resistances have been calculated and an equivalent thermal circuit with the chip effectively being the heat source (with no backflow of heat, however) is drawn and solved to obtain the heat flows through the various layers of the BGA and most importantly, the temperature of the chip is determined which is an objective to be optimized for this study and used in subsequent calculations to determine the shear stresses in the solder joints.

For the stresses in the assembly components, the design is assumed to be one with an organic lid and material properties are taken into consideration to determine the axial compliances and subsequently, the normal and shear stresses in the assembly components are computed using the strain module explained by E. Suhir [5].

$$Shear\ Stress\ in\ Solder\ Joints < Shear\ Modulus\ of\ Solder\ material\ (37e06\ {}^{N}/_{m^2})$$

Also, the operating temperature limit for the chip is considered as a constraint for the thermal resistance module. It can be given by,

$$\Delta T_{chip} < 60℃$$

The plastic BGA package geometry is selected from the JEDEC MO-192 [6] and 216 [7] standards. These standards define substrate sizes and solder ball footprints. The bounds for the geometrical design variables are provided with the same considerations as given below:

$$\begin{bmatrix} 25 \\ 0.0008 \\ 0.0006 \\ 0.0002 \end{bmatrix} \leq \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \leq \begin{bmatrix} 2500 \\ 0.0020 \\ 0.0020 \\ 0.0020 \end{bmatrix}$$

IV. Design Space Exploration

Initially the design space exploration is done by performing the Design of Experiments with the 4 variables conducted with 3 different levels for each using orthogonal array method. As the problem has the non-linear constraints, there were a least chances for the design to be feasible. From the calculation of the mean effects of each variable, there could not be any definite conclusion drafted from the orthogonal DOE search. And therefore, for the optimization of the BGA package cooling, it becomes necessary to approach using the numerical methods.

V. Multi-Objective Problem

For a comprehensive multi-disciplinary design optimization with multiple objective function that includes all of the pertinent design aspects is defined. In the present study, multi-objective optimization is explored, which consists of minimizing a positively weighted convex sum of the objectives [8]. For BGA design, the objective function that incorporates the important design criteria along with the thermal strain constraints is given by,

$$J = w_1 J_1 + w_2 J_2$$

where, J_1 and J_2 are thermal and cost objective function respectively and w_1 and w_2 are the weight for each objective function. The weights are defined as positive constants which are to be selected by the designer based on the design priorities. Here, for the convenience, more weightage is decided for the temperature value and thus, the value of w_1 and w_2 are kept as 0.67 and 0.33 respectively. Here, the scaling of both the objective functions is generally done when the units of the objective values are different. But the shadow minima found for individual objective functions are in the similar range and which depicted that the scaling of these objective functions are not required in this case.

VI. Numerical Approach for Optimization

In order to find out the optimal design vector for the proposed multi-objective function for BGA model, different numerical methods are opted. From the BGA model, it can be directly concluded that heuristic based methods would be a good option as the problem is having mixed-integer design space with non-linear constraints. But the heuristic based methods such as Genetic Algorithm are computationally very costly compared to the gradient based methods. Therefore, first the BGA package design is optimize with the gradient based method and the results were compared with the heuristic based method. Here, for the numerical optimization, optimization toolbox that is available in MATLAB [9] is utilized.

A. Gradient based method

The mixed-integer design space is assumed to be continuous and the sequential quadratic programming algorithm is used to proceed with the optimization. Here, MATLAB based fmincon (non-linear programming solver) is used to handle non-linear constraints effectively. To ensure that the best local optimal point is achieved, multi-start is performed with this algorithm. The Fig. (3) shows the convergence history of gradient based optimization with multi-start. It can be seen that the algorithm gets trapped in local optima a few times, but a best optimal point is also obtained in several times. The optimal design vector and objective value is given by,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} N_b \\ p \\ d_b \\ h_b \end{bmatrix} = \begin{bmatrix} 44 \\ 0.0014 \\ 0.0012 \\ 0.0003 \end{bmatrix}$$

$$J = 25.92$$

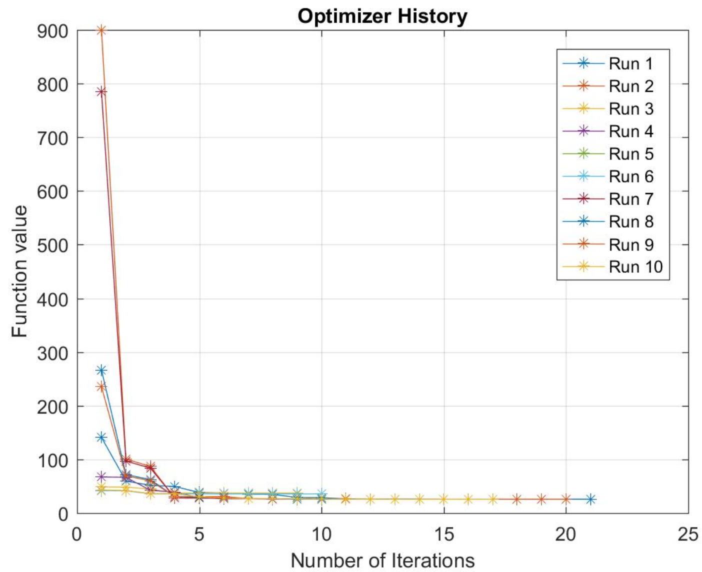

Fig. 3 Multi-start with Sequential Quadratic Programming algorithm

B. Heuristic based method

Further, the optimal point obtained from the gradient based method with multi-start is verified with heuristic based method. As discussed before, Genetic Algorithm is a good option for optimization of a system with mixed-integer design space with non-linear constraints. The Genetic Algorithm is tuned to converge and obtain an optimal point well within 100 generations. The tuning criteria, optimal design vector, objective value and the optimizer history is shown in the Fig. (4).

- Initial population = 100
- Initial penalty = 1100
- Cross-over fraction = 0.9
- Elite count = 1

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} N_b \\ p \\ d_b \\ h_b \end{bmatrix} = \begin{bmatrix} 44 \\ 0.0014 \\ 0.0012 \\ 0.0003 \end{bmatrix}$$

$$J = 25.92$$

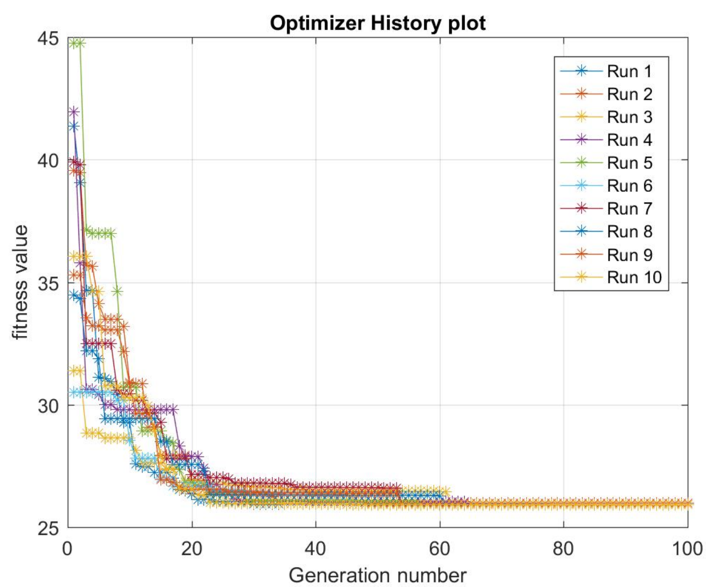

Fig. 4 Optimizer history with Genetic Algorithm

From the Fig. (3) and (4), it can be concluded that the numerical optimization with gradient based method as well as with the heuristic based method is same. This suggests that the obtained optimal point may be a global optimal point.

C. Hessian at optimal point

Further, to identify whether the optimal point is validating the optimality conditions or not, the hessian is calculated at the same. From the calculated eigen values of the hessian, the matrix can be said as a positive definite and the optimality conditions are validated.

$$H = \begin{bmatrix} 0.002 & -481.7 & 641.97 & 15.71 \\ -481.7 & 9.5e07 & -1.3e08 & 2.5e07 \\ 641.97 & -1.3e08 & 2.1e08 & -4.4e07 \\ 15.71 & 2.5e07 & -4.4e07 & 1.2e08 \end{bmatrix} \qquad \Lambda = \begin{bmatrix} 1.75e-11 \\ 6.55e06 \\ 1.04e08 \\ 3.09e08 \end{bmatrix}$$

D. Sensitivity analysis

The sensitivity of the design variables are calculated by the first order derivative of the objective function with respect to the design variables. For all the objective functions J, J_1 and J_2, the normalized sensitivity of each design variables are calculated and shown in the Fig. (5).

$$\nabla \bar{J} = \frac{x^o}{J(x^o)} \nabla J = \begin{bmatrix} 0.15 & 0.91 & 0.14 & -0.23 \\ -0.22 & 2.89 & -3.78 & 0.32 \\ -1.68 & 1.73 & -1.49 & -0.01 \end{bmatrix}$$

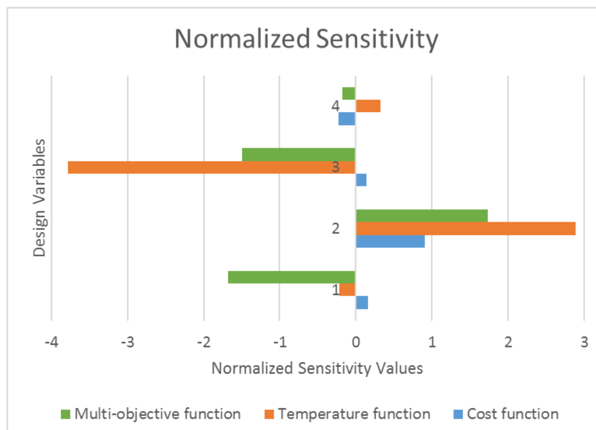

Fig. 5 Normalized sensitivity analysis with design variables x_1, x_2, x_3 and x_4

E. Scaling of design variables

From the sensitivity analysis, it can be seen that the design variables are having different effects when the normalized sensitivity is calculated. This may affect the optimization results for both the gradient based method as well as the heuristic based method. To visualize the effect of scaling, the hessian at the optimal point has been reduced to the order of 1 and again optimized with the gradient based method. The table (1) shows the results obtained by scaling. It can be concluded that there is not much difference in the optimized value of the objective function as well as the design variables, but the computational time can be seen reduced by about 50% with gradient based sequential quadratic programming algorithm.

Table 1 Results after scaling of design variables

		Before Scaling	After Scaling
Computational Time (s)		51.64	25.19
Optimized Output		25.92	25.94
Design Variables	**x_1 (N_b)**	44	45
	x_2 (p)	0.0014	0.0015
	x_3 (D_b)	0.0012	0.0013
	x_4 (h_b)	0.0003	0.0003

VII. Post Optimality Analysis

For the post-optimality analysis, the multi-objective trade-off is analyzed by capturing the Pareto front. The Fig. (6) shows the Pareto front obtained with the weighted-sum method. From the Pareto front the optimal design for different weightages for the objective function J_1 and J_2 can be decided by calculating the trade-off for either of the objective function. As discussed before, this solely depends upon the design engineer to make the decision and proceed further with the optimization.

Also, there are some discrepancies at the end of the Pareto front, which can be avoided by using Normalized Boundary Intersection method. But for this study, the critical area is very well captured and thus the Pareto front from the weighted-sum method is accepted.

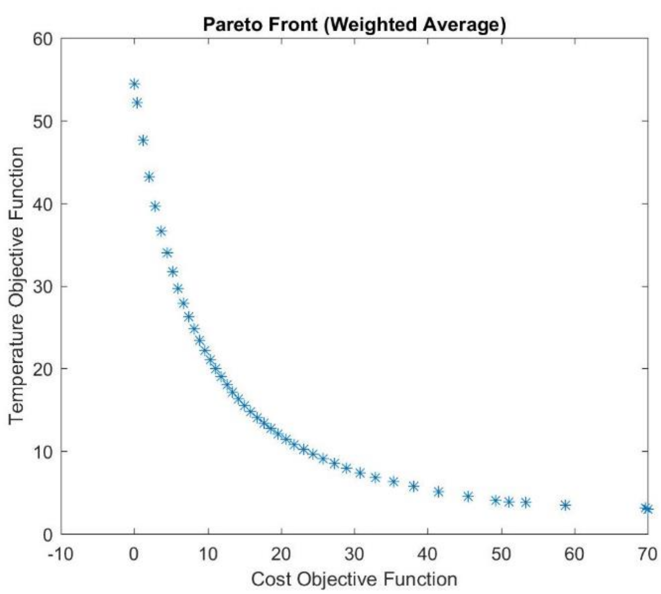

Fig. 6 Multi-objective trade-off analysis with Pareto front

VIII. Summary and Conclusion

A multidisciplinary design optimization comprising of cost and thermal performance objectives, using both heuristic and gradient based optimization techniques have been applied on a Ball Grid Array electronic package to find an optimal design configuration. The solder ball array module is chosen as the subject of design optimization, with the number of solder balls, the pitch of the array, diameter and height of solder balls were considered to be the design variables comprising the design vector. To model the system, a low fidelity thermal resistance network model which agreeably mimics the heat transfer in three dimensions in the package has been extensively studied and applied to the chosen system. The optimal design is obtained by performing gradient based optimization using Sequential Quadratic Programming algorithm and results were compared with heuristic based Genetic Algorithm.

The Hessian was calculated at the optimal point to verify the optimization criteria. Sensitivity and scaling analysis has been conducted to observe the effect of design variables over the objectives. Finally, a multi objective trade off analysis is performed by constructing a Pareto front using the weighted-sum approach.

From this study, the following are the conclusions derived:

- The gradient based methods are quite efficacious compared to the heuristic based methods with the given assumption of continuity of the design variables.
- There is not much difference reported while assuming the number of solder balls as a continuous variable for the gradient based optimization.
- Multi-start approach with Sequential Quadratic Programming algorithm in this case provides a best-local minima that can also be achieved by computationally costly Genetic Algorithm.
- Though the scaling of design variables do not change the optimal solution obtained, it decreases the computational cost almost by 50%.

Future work involves implementing models in high fidelity commercial packages such as ANSYS for higher accuracy and accounting for more complexity of the system in consideration.

References

[1] Brombacher, A. C., "Reliabilitv of Design - CAE Techniques for Electronic Components and Systems", *Wiley*, 1992.

[2] Luigi P. M. Colombo, Alexey Petrushin, Davide Paleari, "Simplified Thermal Model of a Stacked Ball Grid Array Package", *Journal of Electronic Packaging*, 2011.

[3] Kennedy, D. P., "Spreading Resistance in Cylindrical Semiconductor Devices," *Journal of Applied Physics*, 31, 1960, pp. 1490–1497.

[4] Song, S., Lee, S., and Au, V., "Closed-Form Equation for Thermal Constriction/Spreading Resistances with Variable Resistance Boundary Condition," *IEPS Conference*, Atlanta, GA, 1994, pp. 111–121.

[5] E. Suhir, "Analytical Thermal Stress Model for a Typical Flip-Chip (FC) Package Design", *Journal of Materials Science: Materials in Electronics*, 2018.

[6] Solid State Product Outline, Low Profile Square Ball Grid Array Family, MO-192-E: *JEDEC*, 2001.

[7] Solid State Product Outline, Thin Profile, Square and Rectangular Ball Grid Array Family, 1.00 & 0.80 Pitches, MO-216-D, *JEDEC*, 2002.

[8] M. Balachandran and J. S. Gero, "A comparison of three methods for generating the Pareto optimal set," *Eng. Optimization*, vol. 7, pp. 319–336, 1984.

[9] Optimization Toolbox, MATLAB, Software Package, Ver. R2016b, MathWorks, Natick, MA.