DESIGN OPTIMIZATION OF A BALL GRID ARRAY ELECTRONIC PACKAGE

Yatharth Vaishnani

Department of Mechanical Engineering
Texas A&M University

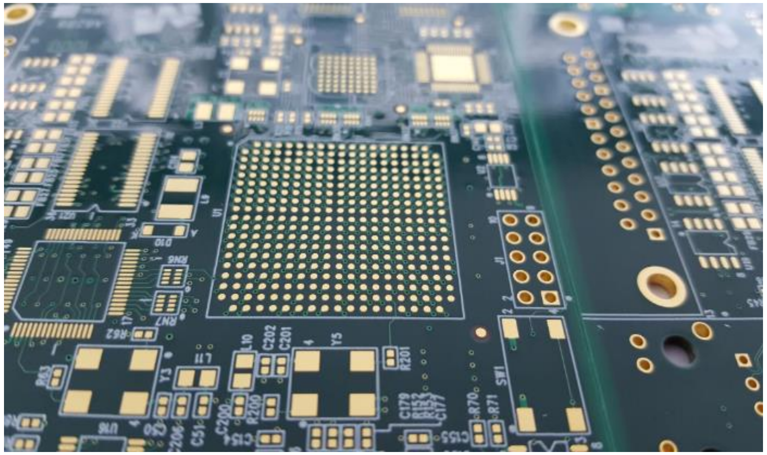
Motivation

- 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.

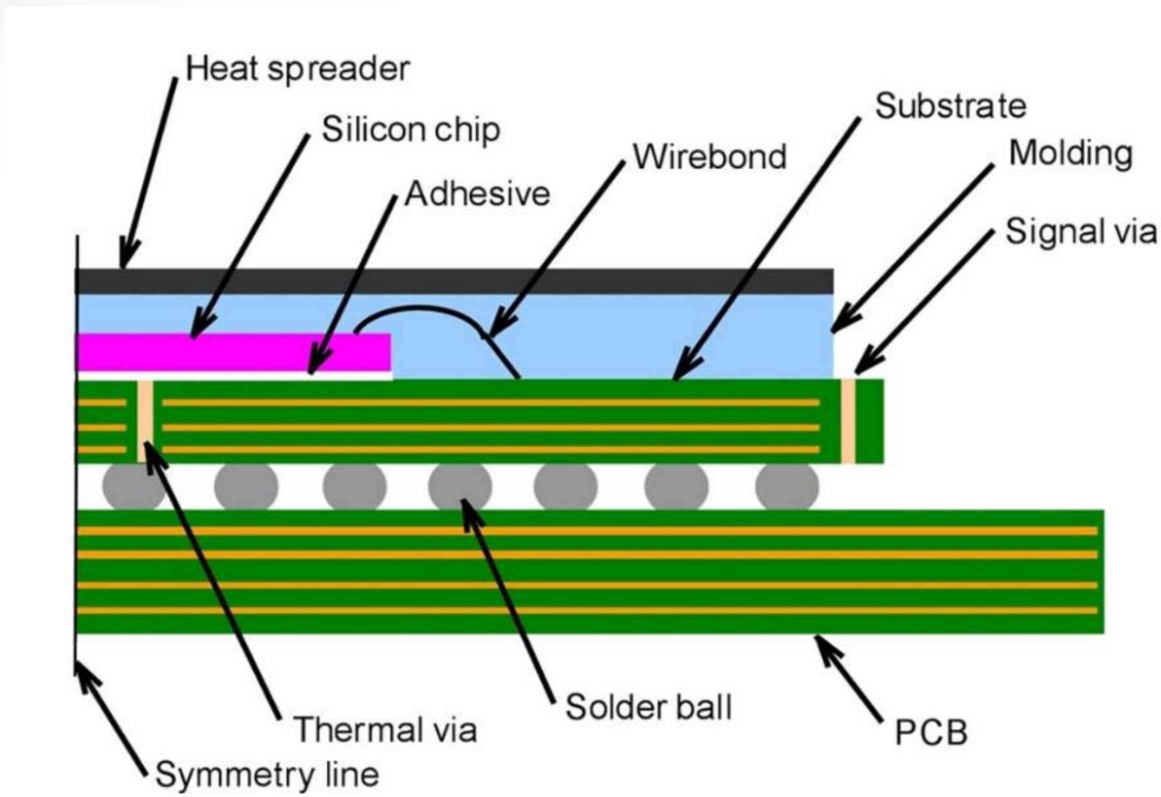
Failed IC in a laptop. Wrong input voltage has caused massive overheating of the chip and melted the plastic casing.

Ball Grid Array

- A type of surface mount packaging, a chip carrier used for Integrated Circuits.
- BGA packages are used to permanently mount devices such as Microprocessors.



Overview of the Structure and Components

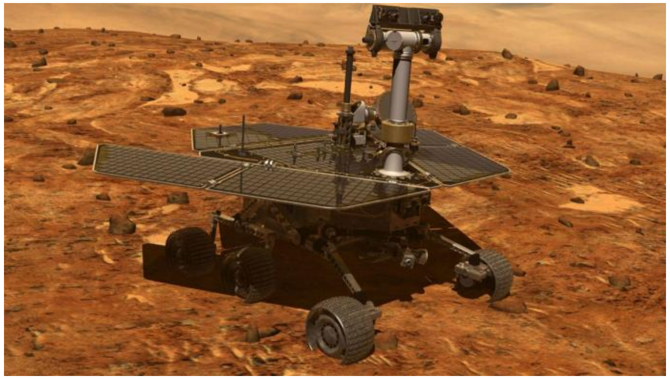


Problem Statement

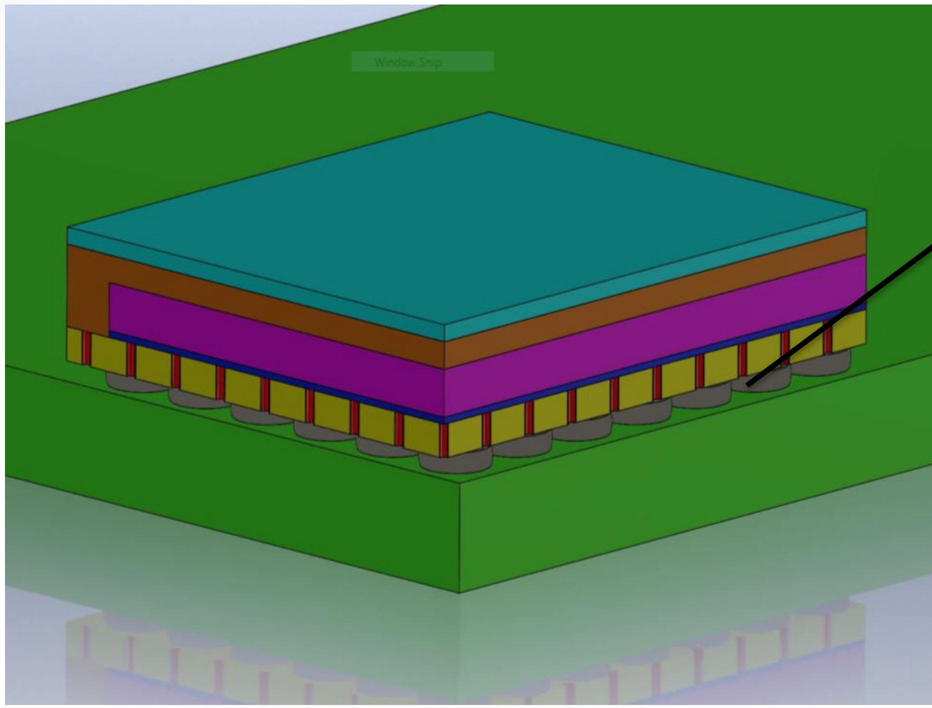
For a given ball grid array package, to optimize the dimensions of **solder ball array** , to have optimum thermal and structural performance while **minimizing** the **overall cost** of package design and **minimizing** the **silicon chip temperature**, with stress and geometrical constraints.

Applications

- Most common applications: Portable and telecommunication products
- *Space applications*: Ceramic packages were specifically tailored for high reliability, to provide processing power required for the Spirit and Opportunity Mars Rovers built by NASA-JPL

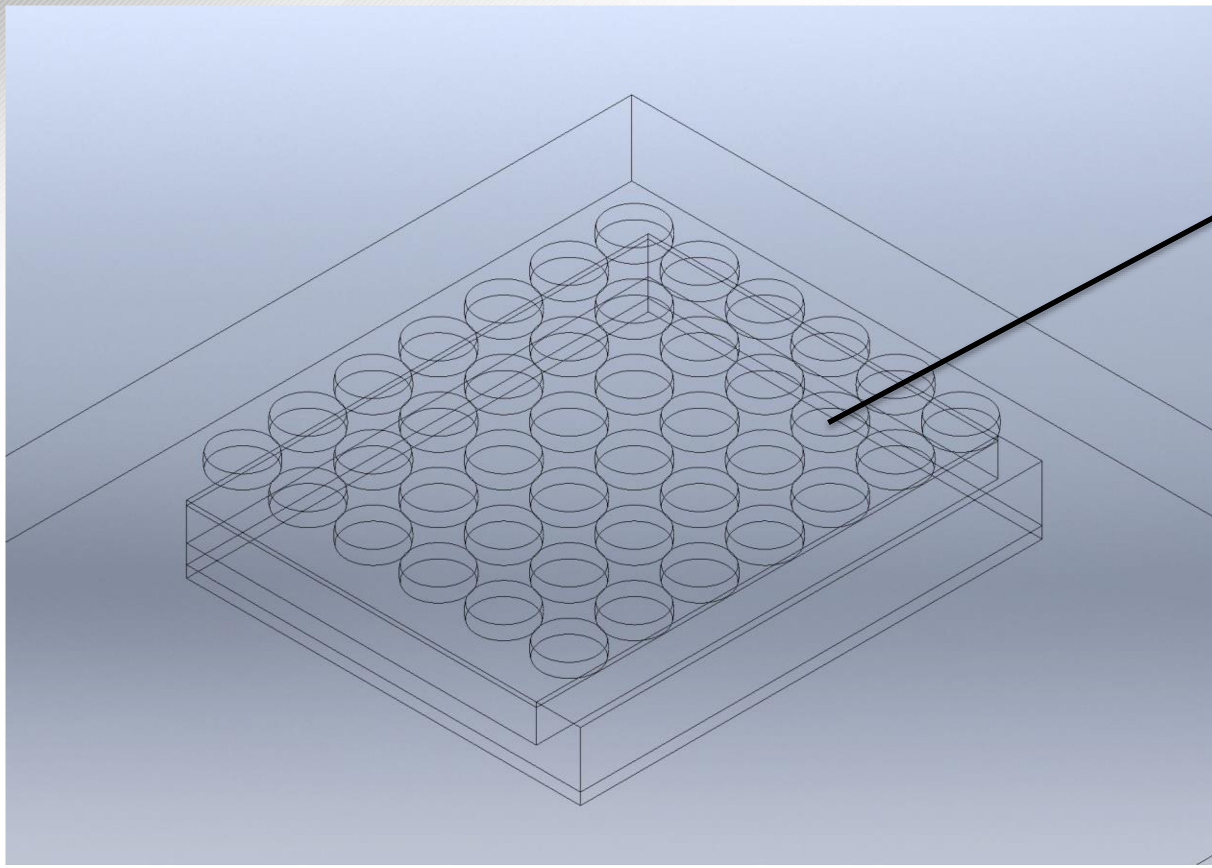


CAD Model of the BGA Package



Solder Ball Array

$$\begin{bmatrix} N_b \\ p \\ d_b \\ h_b \end{bmatrix}$$



Solder Ball Array

$$\begin{bmatrix} N_b \\ p \\ d_b \\ h_b \end{bmatrix} = \begin{bmatrix} 44 \\ 0.0014 \\ 0.0012 \\ 0.0003 \end{bmatrix}$$

Objectives

- Chip
  - Minimize the temperature
    - Improve thermal performance and heat dissipation by choosing an optimum combination of values of design variables
- Cost
  - Minimize total cost of the package
    - Reduce cost

Design Variables and Parameters

Objective Functions
min(chip temperature)
min(total cost of package)

Constraints
Geometrical Constraints:
- Clearance between the diameter of each solder ball and pitch is 0.2mm
- Diameter of each solder ball should be greater than its height

Stress Constraint:
- Shear stress in the solder joint should be less than its material shear strength

Temperature Constraint:
- Difference in temperatures of the chip and ambient should be less than 60^{o}C

$Design\ Variables$

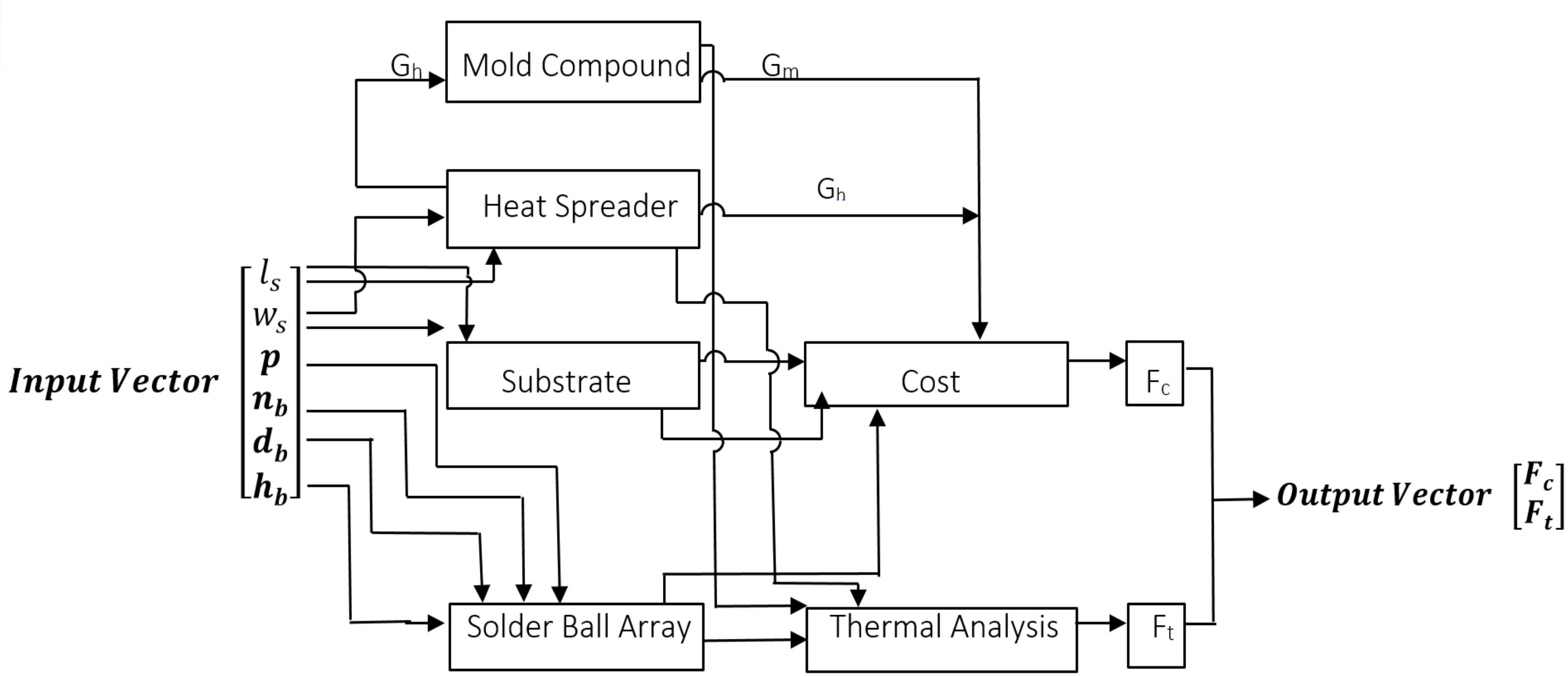
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} N_b \\ p \\ d_b \\ h_b \end{bmatrix}$$

Component	Attribute	Nomenclature	Units	Type
Substrate	Thickness	t_s	m	Parameter
Silicon Chip	Length	l_c	m	Parameter
	Width	w_c	m	Parameter
	Thickness	t_c	m	Parameter
	Heat flux	q_c	W	Parameter
Mold	Length	l_m	m	Dependent
	Width	w_m	m	Dependent
	Thickness	t_m	m	Parameter
	Geometry	G_m	m	Dependent
Heat Spreader	Length	l_h	m	Dependent
	Width	w_h	m	Dependent
	Thickness	t_h	m	Parameter
	Geometry	G_h	m	Dependent
Solder Ball Array	Geometry	G_b	m	Dependent
Adhesive	Thickness	t_a	m	Parameter
PCB	Length	l_p	m	Parameter
	Width	w_p	m	Parameter
	Thickness	t_p	m	Parameter
Air	Velocity	v	$m \cdot s^{-1}$	Parameter
	Temperature	T	K	Parameter

Parameters

Component	Material	Thermal Conductivity K (w m^{-1} K^{-1})
Solder Balls	Eutectic Solder (63Pb/37Sn)	50
PCB	FR4 Epoxy	0.343
Substrate	FR4 Epoxy + Copper Inlays	10
Silicon Chip	Silicon	148
Mold	G760Y Epoxy	1.3
Adhesive	Silver Filled Epoxy	2.45
Surrounding Environment	Air	0.027

Block Diagram



N^2 Diagram

Design Variables		N_b, P, d_b, h_b						
	Constraints							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	
					Thermal	F_t		$\mathbf{F_t}$
						Strain		$\mathbf{C_s}$
							Cost	$\mathbf{F_c}$
								Output

Resistor Network

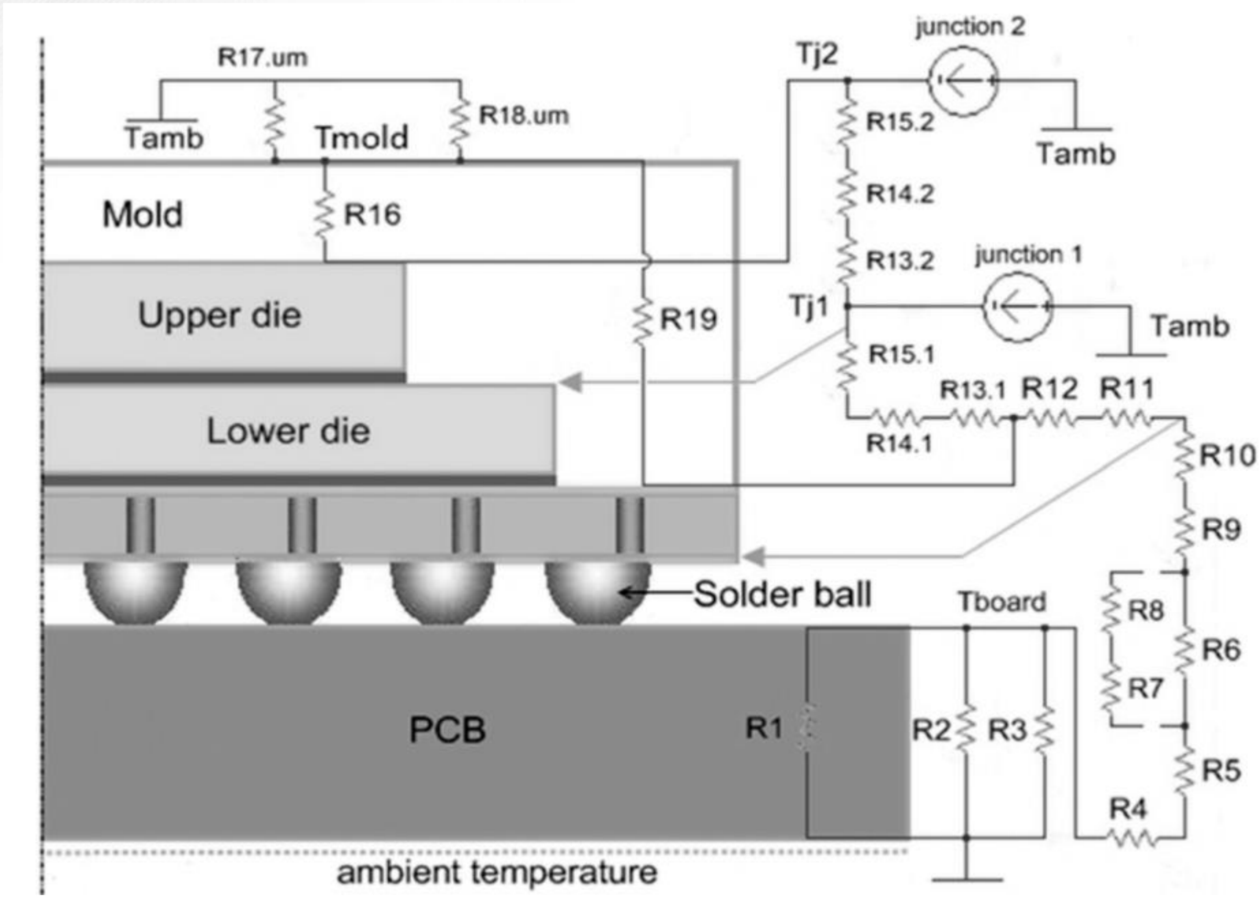
Steady state heat transfer within a BGA package has been modeled by means of a thermal resistor network.

Three types of thermal resistances have been considered:

- ***Layer resistances***: To model various layers within the substrate of the package that are layered on top of each other.
- ***Spreading/Constriction resistances***: To mimic approximation of the heat flow in 3D as changes in the cross-sectional area arise along the heat flow path.
- ***Boundary Condition resistances***: To account for the interaction between device and surroundings.

Resistor Network

Schematic



Thermal Analysis

Layer Resistance

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

$$k_{\parallel} = \frac{\sum_{i=1}^{N} k_i t_i}{\sum_{i=1}^{N} t_i} \qquad k_{\perp} = \frac{\sum_{i=1}^{N} t_i}{\sum_{i=1}^{N} \frac{t_i}{k_i}}$$

Spreading/Constriction Resistance

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

$$e = \frac{a}{b}, \tau = \frac{t}{b}, \lambda_c = \pi + \frac{1}{\sqrt{\pi} \cdot e}, \text{ and } Bi = \frac{1}{\pi \cdot k \cdot b \cdot R_0} \qquad \alpha = \sqrt{\frac{k_{\perp}}{k_{\parallel}}}$$

Boundary Condition Resistance

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

Assumptions

- Homogeneous and temperature-independent thermal properties
- Low fidelity, simplified models considered instead of high fidelity Finite Element Analysis models, to reduce computational time
- Electrical and electro-magnetic effects on the thermal analysis not considered
- Fatigue loads on solder balls have not been considered

Validation of Results

T_{max}	Result		
Component	Model (°C)	Simulation (°C)	Reference (°C)
Chip	41.9	43.8	43.3
Solder Ball	28.7	30.3	31.2
Substrate	25.4	25.9	24.7

Results obtained from resistance network model are compared with reference model and a high fidelity FEA model. The obtained results are in reasonable agreement with both, the reference and FEA model.

Design of Experiments

Exp. No.	No. of Solder Balls	Pitch of Solder Balls (m)	Diameter of Solder Balls (m)	Height of Solder Balls (m)	Rise in Chip Temperature (°C)	BGA Price ($)	Feasible?
1	100	0.0008	0.0006	0.0002	28.3488	40.65	Yes
2	100	0.0014	0.0012	0.0010	43.9910	59.34	Yes
3	100	0.0020	0.0018	0.0016	38.4141	139.19	Yes
4	500	0.0008	0.0012	0.0016	-15.1026	239.54	No
5	500	0.0014	0.0018	0.0002	10.7273	150.15	No
6	500	0.0020	0.0006	0.0010	116.3582	149.92	No
7	1000	0.0008	0.0018	0.0010	37.6249	618.39	No
8	1000	0.0014	0.0006	0.0002	12.6138	172.10	Yes
9	1000	0.0020	0.0012	0.0016	12.6973	557.37	No

Observations:

- 5 out of 9 experiments result in infeasible solutions
- Nothing conclusive can be determined from DOE

Numerical Approach for Optimization

Gradient-based

- fmincon used with initial point from DOE
- Multi-start is performed with gradient-based method

Heuristic-based

- GA with mixed-integer optimization is used
- Fine-tuned for convergence

Post-optimality

- Hessian value checked (Positive Definite)
- Scaling and Sensitivity of Design Variables

Multi-objective

- Trade-off analysis using Pareto-front

Gradient-based Method

Algorithm Selection

- With gradient-based approach, SQP algorithm used
- Multi-start with fmincon

Results

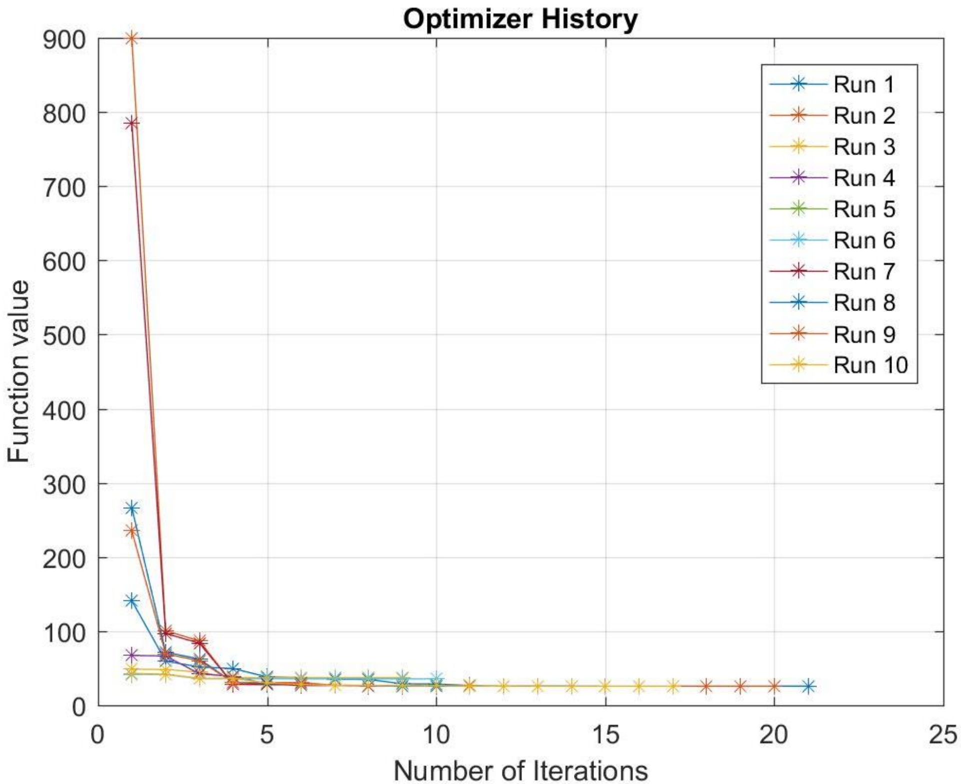
- With multi-start, fmincon converges to optimal solution
- Results depends upon the initial starting point
  - Gets trapped in local optima

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

- Hessian and KKT conditions checked at the "best-local optimal solution

Assumption

- Discrete variable (N_b) considered as continuous



Heuristic-based Method

Algorithm Selection

- With heuristic-based approach, GA is used
- Multiple runs for GA

Results

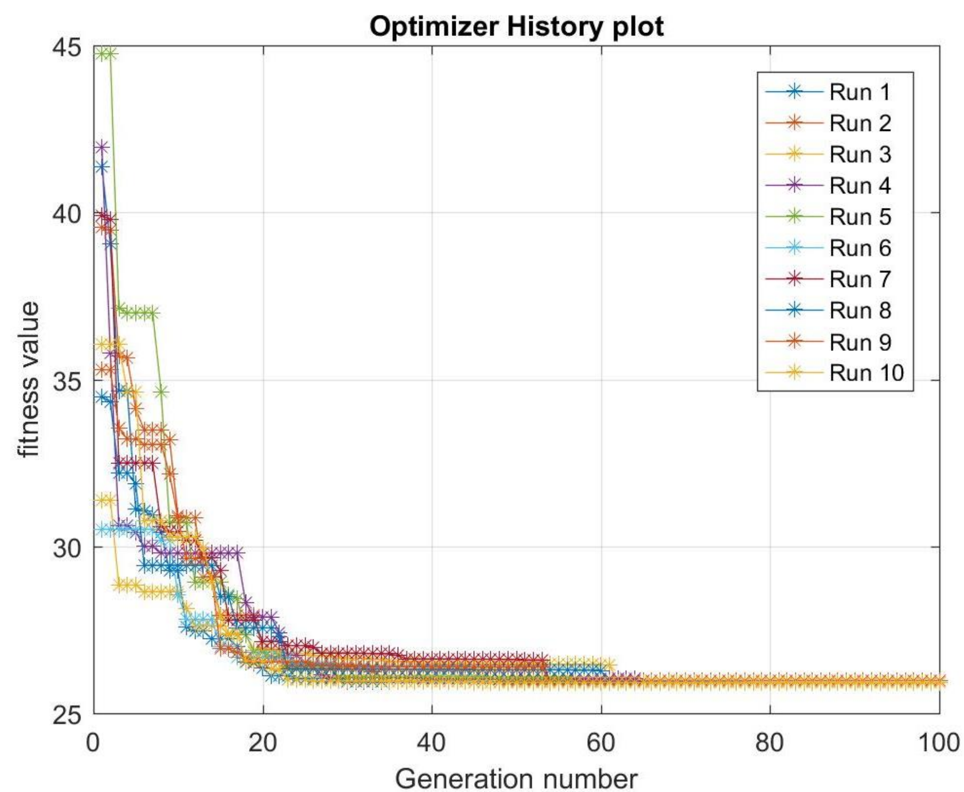
- Compared to gradient-based method, GA is computationally costly
- After fine-tuning, GA converges to optimal solution

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

- Both the optimal design vectors obtained from Gradient-based and Heuristic-based method is same

Fine-tuning

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



Post Optimality Analysis

- Optimality Conditions
  - Hessian and KKT Conditions
- Scaling Analysis
- Multi-Objective Trade-off Analysis

Hessian & KKT-Conditions

Hessian Calculation

- The calculated Hessian at the "best local optimal" solution is

$$H = \begin{bmatrix} 0.002 & -481.7 & 641.97 & 15.71 \\ -481.7 & 9.5e07 & -1.3e08 & 2.5e07 \\ 641.97 & -1.3e08 & 2.1e08 & -4.4e07 \\ 15.71 & 2.5e07 & -4.4e07 & 1.2e08 \end{bmatrix}$$

- The Eigen values of the Hessian matrix are calculated as

$$\Lambda = \begin{bmatrix} 1.75e-11 \\ 6.55e06 \\ 1.04e08 \\ 3.09e08 \end{bmatrix}$$

- Which makes the Hessian matrix a Positive Definite

Sensitivity Analysis

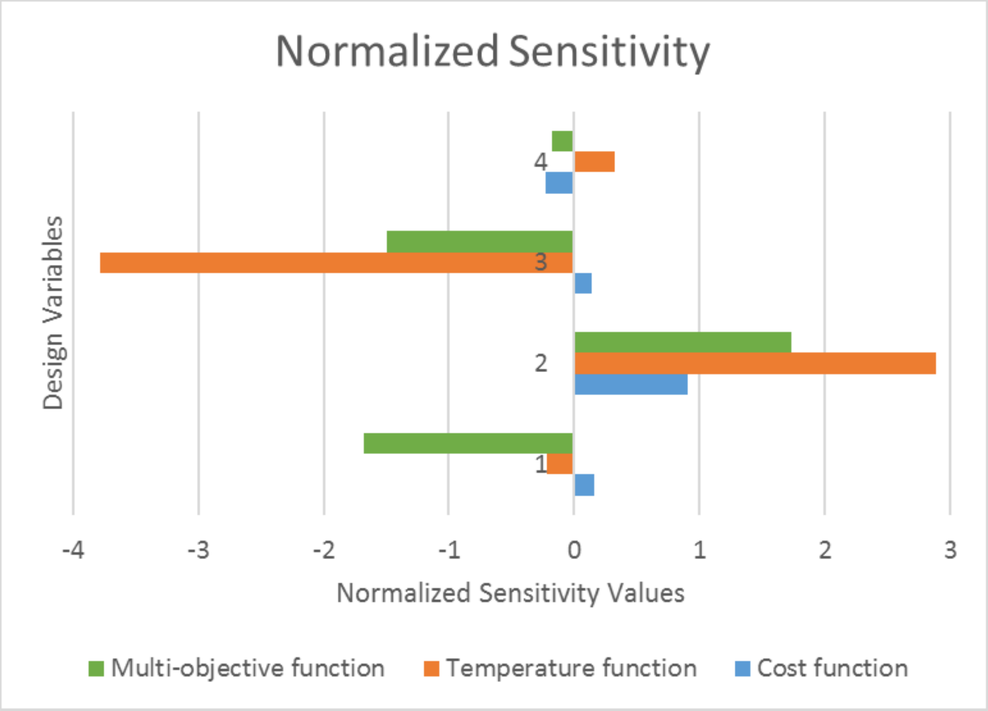
Sensitivity Analysis Approach

- First Order Derivative calculated with central-difference Finite Difference Method
- Further, normalized to calculate "relative" sensitivity

Results

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

- Both the optimal design vectors obtained from Gradient-based and Heuristic-based method is same



Scaling of Design Variables

Scaling Approach

- At the optimal point, the Hessian is calculated
- The scaling factors decided by trial-and-error approach
- Diagonal values of Hessian is scaled to the order of 1

		Scaling Analysis	
		Before Scaling	**After Scaling**
Computational Time (s)		51.64	25.19
Optimized Output		25.92	25.94
Design Variables	**Nb**	44	45
	p	0.0014	0.0015
	db	0.0012	0.0013
	hb	0.0003	0.0003

Results

- No significant difference in the optimal value obtained
- fmincon converges to the “best-local optimal” point
- Time required to converge reduces by 50%

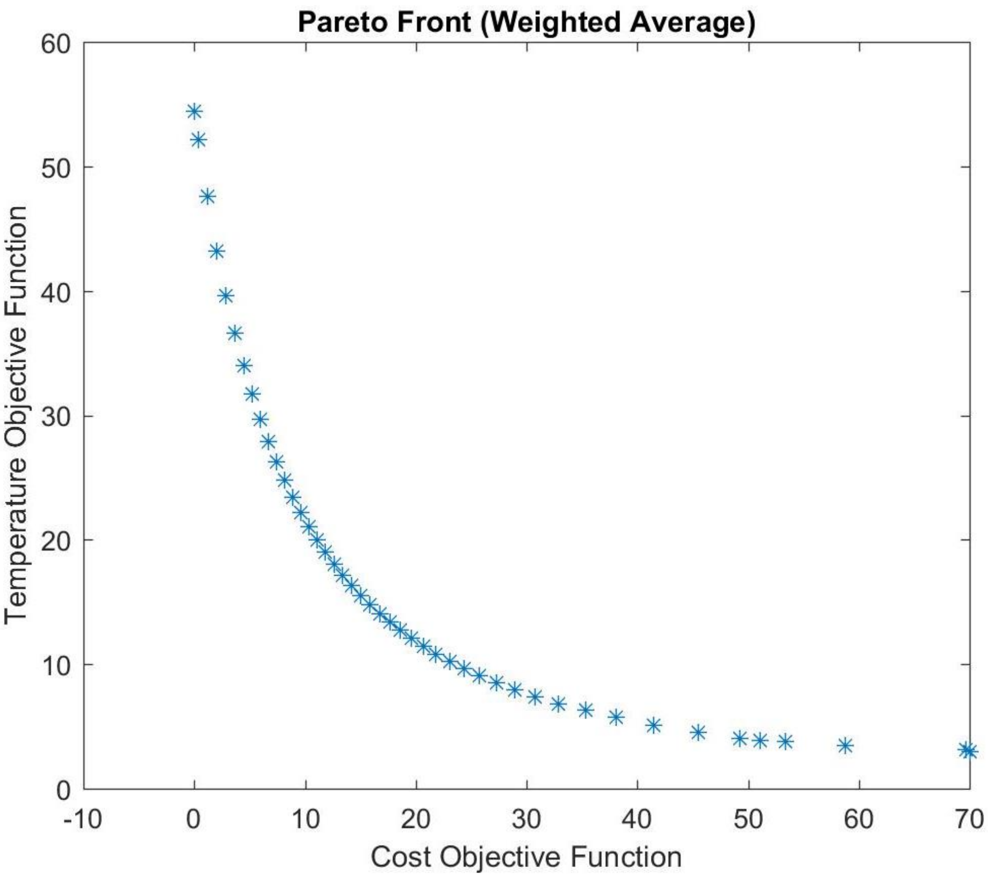
Multi-objective Trade-off

Pareto-Front Approach

- Weighted Sum Approach is used for finding out the Pareto-Front

Results

- Pareto-Front is well distributed with the Weighted Sum Approach method
- Further study needs to be done with NBI method for verification.



Conclusions

- fmincon with multi-start is able to give the optimal solution
- Assuming the discrete variable into continuous variable doesn't change the solution
- Scaling of the design variable accelerates the gradient-based approach
- Performance can be sacrificed for cost reduction

Future Work

- Observe different materials
- Perform NBI method for Pareto-Front
- Different design standards for manufacturing can be included for the optimization

Thank you!!

Questions?