Simulation of Solder Fatigue Effects on Typical BGA Package due to Material and Temperature Variations

James Petroski
Design by Analysis Technical Consulting
Parma, OH 44134 USA
James.Petroski@DesByAnalysis.com

Andy Carrasco
Flexetch
San Jose, CA 95136 USA
andres@flexetch.com

Abstract

The past two decades have seen many approaches to solder fatigue and solder joint life published. This subject has proved difficult as various failure mechanisms are proposed and examined. While these theoretical bases are discussed, it often leaves the end developer in a difficult situation as to how to apply an accurate simulation approach to fatigue failure for a particular package. In this paper, one approach advocated by Syed, Sharon, and Darveaux [1] is examined and used to evaluate a "typical" metal lid flip chip BGA (ball grid array) package under variation of a few key parameters under the control of a project via a 2-level DOE (Design of Experiment). The parameters examined in this paper are the in-plane thermal expansion coefficient of the PCB board (a material choice), the lid material (also a material choice), and the temperature range for the BGA environment (a boundary condition and a heat transfer effectiveness choice). Most other parameters that could also affect solder fatigue life are not within areas that are easily changed due to standards, manufacturing methods, cost, or regulations, and hence are not considered. The DOE results show that all three of these individual parameters are significant to the DOE models, but for solder bumps there is an interaction term of the lid CTE (coefficient of thermal expansion) and temperature ΔT, while the solder balls show a significant interaction of the PCB CTE and the temperature ΔT. Bump life cycles to failure ranged from approximately 35,000 to 24 million, while ball life cycles to failure ranged from 48,000 to 24 million.

.

Keywords
Keywords, separated, by commas

Nomenclature

BGA	Ball Grid Array
DOE	Design of Experiments
CTE	Coefficient of Thermal Expansion
PTH	Printed Through Hole
SI/PI	Signal Integrity/Power Integrity

1. Background

There have been numerous studies to understand and characterize solder used in electronic packaging. Unfortunately, solder's behavior in electronic systems is anything but simple and linear, which has led to a plethora of studies and literature contributions. That has made it quite difficult to use in standard product development, especially when trying to decide which model(s) should be used. This paper is intended to provide some guidance and clarity for such product development.

Many past papers have surveyed the methods of characterizing solder and solder failures due to fatigue (see Ref [1]-[6] for some examples). Generally, they are segregated into various categories. For example, [2] separates constitutive models into two types (with and without the definition of a yield surface). The extensive review in [5] divides fatigue based models into four categories (plastic strain-based, creep damage-based, energy-based, and damage accumulation-based). Ref [6] is one of the most recent and comprehensive surveys of all fatigue models examined to date.

Even when one characterization of solder is chosen, there are many factors found in the literature, such as strain rate, temperature, grain size, geometry chosen for testing (actual BGA ball vs. solder tensile specimen geometry), etc. These lead to models with widely different characteristics at the standard parameters, such as the Anand models plotted in Fig 1 below [7].

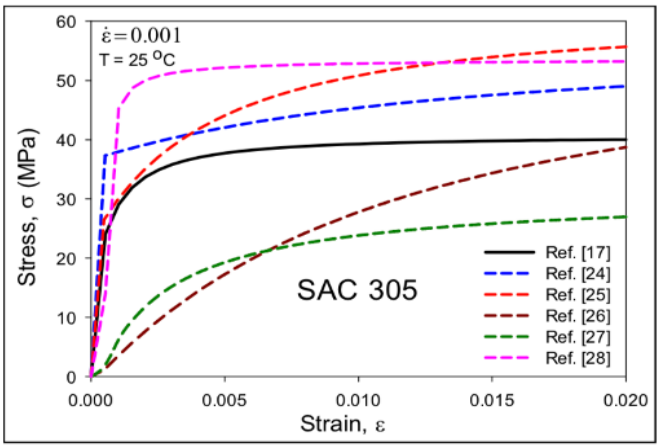

Figure 1: SAC305 Anand models [7]

Beyond this theoretical choice of models, there are choices in the geometry or material of the item to help reduce solder fatigue from thermal cycling. For a BGA, various approaches are taken to increase overall lifetime. For example, smaller die reduces the strain seen at the corner solder bump locations.

Copyright ©2024 by SEMI-THERM EDUCATIONAL FOUNDATION

ISBN 978-1-7355325-4-7
ISSN 2577-1000

Using a favorable CTE in the lid, substrate, or underlying PCB also help. Similarly, some lid geometries are more favorable under certain conditions. Applying a proper underfill to the die or using an appropriate lid adhesive can also be used to reduce solder fatigue. Each of these are variables one can use unless constrained by other environmental conditions or manufacturing methods.

Clearly, though many BGA packages are produced and used today, using them under extreme conditions and optimizing them for long lifetimes is a daunting task to examine with simulation.

2. Approach

Given this disparity of fatigue models and approaches to failure, this paper will examine one method. Choosing a method when numerous approaches are available means some criteria should be used. For this paper, the basic criteria used consists of the following items: (1) the theoretical background of failure is sound and examined in the literature, preferably by a number of papers; (2) test data is used to verify and/or tune the application of the theory; (3) solder material tests are used with proper environment (temperature, strain rates) and geometry (tensile samples, or actual part geometry such as solder balls) to create material models; and (4) effective and proven modeling approaches in FEA are published similar to the solder geometry of the actual problem.

Given this set of criteria, the BGA in this paper is examined using the approach enumerated in [8] by Syed, Sharon, and Darveaux. This paper fulfills the criteria of (1), (2), and (4). This paper is significantly more expansive and comprehensive in its approach than many other papers in the field by covering these three items in an integrated manner. The solder material models (SAC 305 for solder balls, Sn-2.5Ag for solder bumps) was compiled from various tests and conditions found in the literature. Syed et. al. in [8] use creep as the damage criteria for fatigue which is used in this study. Hence, elastic-plastic models are used at a low strain rate found in thermal cycling problems along with a creep model chosen for fatigue modeling and suited for the solder location (Garofalo model for SAC 305, and Combined Time Hardening for Sn-2.5Ag). The advantages and disadvantages of various fatigue models are well summarized in Table 1 of [6] for the interested reader.

3. Factors to limit solder fatigue

Since a BGA package is a mechanical system composed of different part geometries, materials, and material behavior, it is reasonable that the assembly can be optimized for solder fatigue life. Also, the temperature variation of the assembly contributes to the solder life outside of the mechanical specifications. But not every item or environment variable is under the control of the designer. Material or manufacturing limitations can restrict the variables or range of variables. And any cooling system capability (or limitation) will control the temperature variations.

Controlling the mismatch of CTEs within the package is the key to maximizing the solder fatigue life. Given how a typical flip chip BGA stack is assembled, the solder bumps between the die and the substrate are affected by the different CTEs of silicon and the substrate. The solder balls are affected by the different CTEs of the substrate and the underlying PCB. The lid, attached to the substrate via some adhesive or solder, affects both the bumps and balls due to the additional warpage it imposes on the package and its unique CTE. Fig 2 below from [9] shows a typical design of such a package.

Examining this geometry, it is clear that the assembly works against itself to optimize the fatigue life of both the solder bumps and balls. Given that silicon has a CTE of approximately 2.6 ppm/K, a low CTE substrate is desirable for the solder bumps, something around 9-10 ppm/K with low CTE dielectrics. However, the solder balls see the substrate and the PCB that cause the CTE mismatch, so a higher CTE for the PCB will help the solder balls. The lid CTE and the lid adhesive likewise can be tailored to help either the bumps or balls, but not both at the same time.

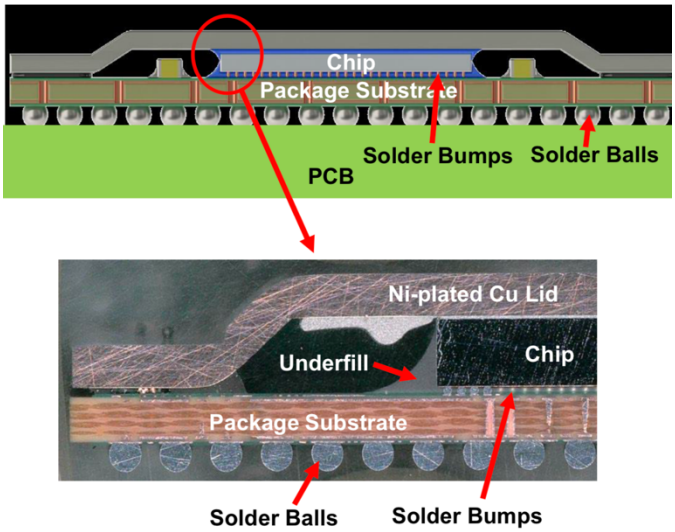

Figure 2 Typical lidded fcBGA cross section [9]

Since the CTE mismatches tend to produce a shearing load in the solder bumps and balls, another mitigation is to make them as tall as possible so the shear angle is smaller. Another factor is the size of the chip; if it is large, the amount of thermal expansion/contraction creates more shearing load than a smaller die on the solder bumps and possibly the balls. Finally, since all of the thermal loads is driven by the temperature range for the BGA, a lower ΔT will lower the thermal strains and increase fatigue life.

Since some of these items are beyond normal manufacturing options, or beyond the cooling system capacity of the system (hence the ΔT imposed on the BGA), this paper investigates a controllable set of factors for a typical BGA. Specifically, this paper examines a low and high CTE lid, a low and high CTE PCB, and a low and high temperature ΔT on the system (see section 5.1 below).

4. Substrate construction limitations

A number of interconnected drivers can result in packages that experience a high amount of warpage, and thus high solder fatigue. Some of the most common drivers are highlighted below.

(1) Signal propagation at high speeds typically requires use of low loss materials, which are usually not low CTE.

(2) High-speed signals need to be properly shielded and referenced by ground to avoid crosstalk. This is achieved vertically (through the stack) by alternating signal and gnd layers, and horizontally (in-plane) by including ground and

via shields between signals, or pairs of signals. Thus, high speed signals normally take up significant real estate on the substrate. And because Signal Integrity (SI) requirements usually dictate that signals are routed above the core of the substrate, the result is a substrate with many layers. For example, two high speed routing layers would require a 4-2-4 or 5-2-5 stackup, while three routing layers would require a 6-2-6 or 7-2-7 stackup. A typical stackup is shown in Fig 3 below, where a FC bump is routed from the Top layer, through 7 dielectric build-up layers using uVias, through a core using drilled PTH vias, through an additional 7 build-up layers, and finally exiting the substrate through a BGA ball.

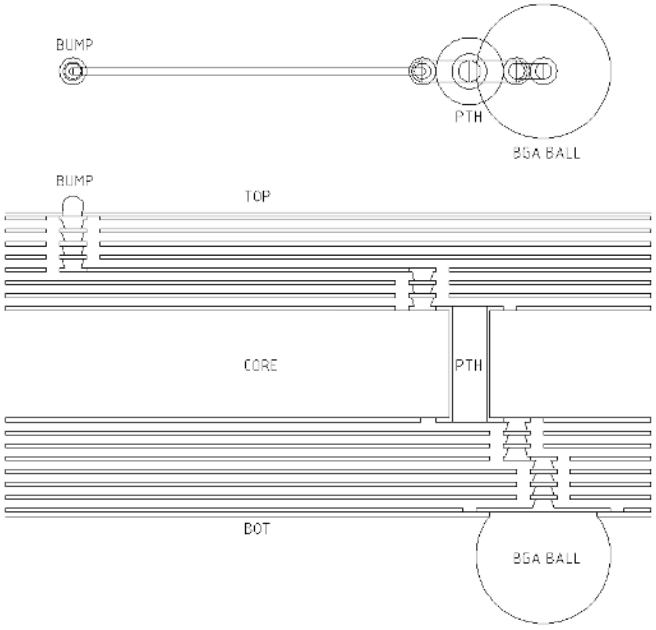

Figure 3 Diagram/stackup of 7-2-7 substrate

(3) Common fabrication methods call for a high amount of Cu coverage in each substrate layer, often 60-80%. This amount of Cu, together with more layers, drives up the effective CTE of the substrate. Managing the Cu balance involves careful control of metal balance from layer to layer and from above core layers to below core layers.

(4) The desire and pressure to increase bandwidth in devices leads to the inclusion of many high-speed signal lanes. To accommodate the number of lanes, die and package architects / engineers will increase the size of both silicon and package substrate, which grows the substrate size as well as the size of the bump field, exacerbating warpage.

(5) Larger devices with more high-speed signals/lanes will tend to consume larger amounts of power, leading directly to more heat and thermal expansion.

(6) The operational environment and cooling solution can result in additional unwanted warpage and solder fatigue if not optimized properly.

The aforementioned drivers will impose limitations on how the package is constructed, including material choices, thicknesses, and layout of geometry. As none of these considerations can necessarily be ignored or looked at independently, the package engineer is responsible for finding the best balance for each package substrate. It is crucial for the package engineer to be involved very early upstream during development of the chip architecture or floorplan, to ensure that decisions made early on will not have adverse effects on the assembly process. Similarly, the package engineer must also consider the downstream assembly processes to make all team members aware of the warpage-related requirements and methods to compensate for warpage. A good example is thermal/CTE compensation, which allows the substrate bump field to be pre-shrunk so that at assembly temperature, the bumps align, mate, and lock correctly with the silicon. The successful package engineer will carefully consider expert opinions from multiple disciplines and teams (physical design, mechanical, SI/PI, manufacturing partners/vendors), bring together team members to understand how each affect and are affected by the substrate, and drive the substrate to a successful execution through fabrication and assembly.

5. Sample BGA model and problem

A simple BGA was modeled for this analysis. This is a one quarter model where symmetric boundary conditions are applied to the cut surfaces. The model uses a standard JEDEC substrate board size (33x33 mm), 0.8 mm pitch solder balls, and 100 µm solder bumps between the single die and substrate. Note that the corner bump is removed under the die as often done for large die manufacturing. The die is 20x20 mm. Also note that only some of the individual solder bumps and balls are modeled. The majority are modeled as single parts, where orthotropic smeared material properties are used. The bump smear is a composite of solder and underfill, and the ball smear is a composite of solder and air. Figures 4-6 show the solid model geometry.

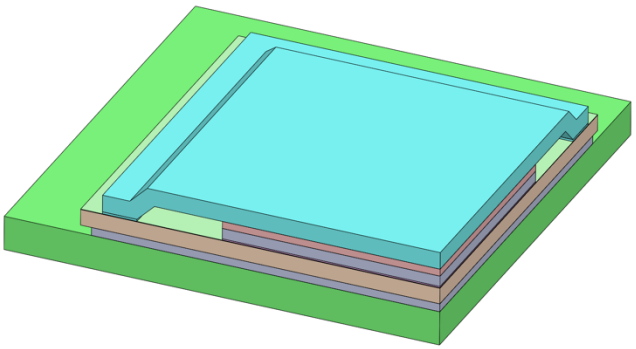

Figure 4 Quarter BGA model

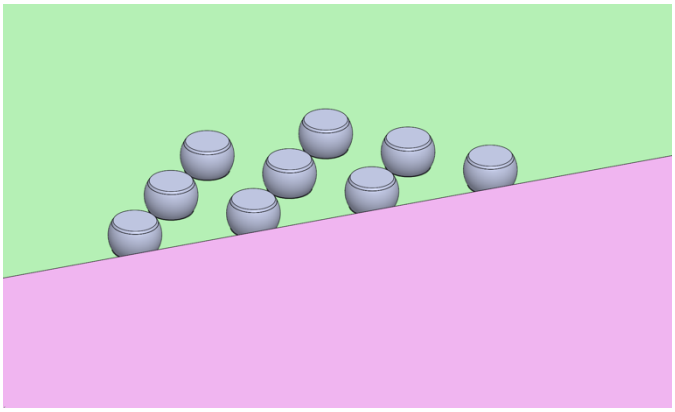

Figure 5 Solder bumps & bumps/underfill smear

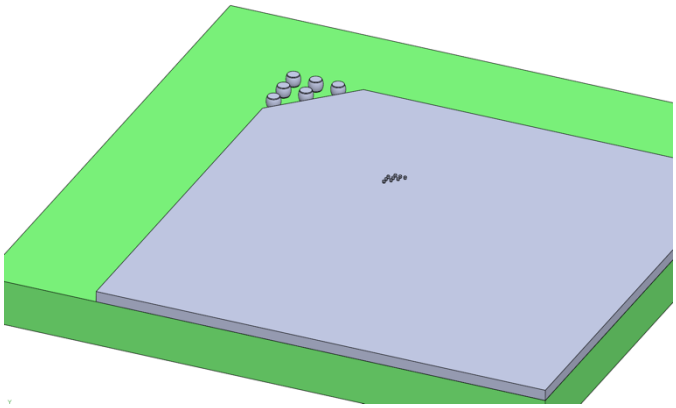

Figure 6 Solder balls & smear (bumps shown for reference)

The solder bumps and balls also need some particular geometry for the simulation. The shape of each type is correct but sliced so the program can calculate the amount of damage in the most critical areas, namely the top and the bottom of the bump/ball. This is where fatigue cracks occur, and averaging the damage over the volume of an entire bump/ball would not accurately predict the cycles to failure. For a bump, a thin 5 µm slice is used; the ball uses a 25 µm slice at the top and bottom. Each bump/ball is attached to a copper pad at the top and bottom, and this develops the correct CTE mismatch. The intermetallic layer (IMC) is not modeled, and is rarely done so in the literature. This is not seen in numerous studies as being significant and is left out, and this study follows that precedent. Figures 7 and 8 show the detailed geometry of the bumps and balls used in this study.

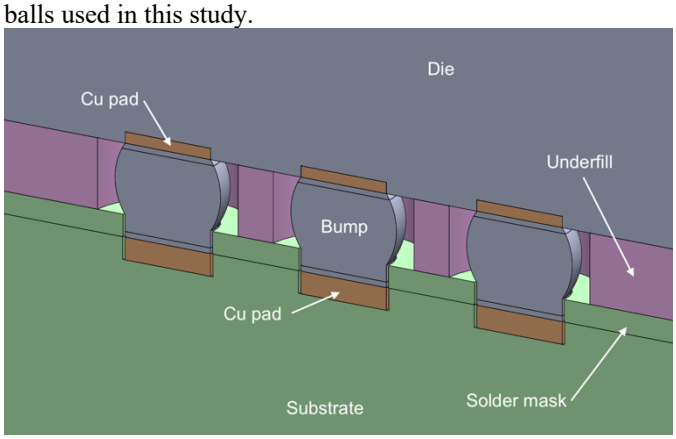

Figure 7 Solder bump cross section region

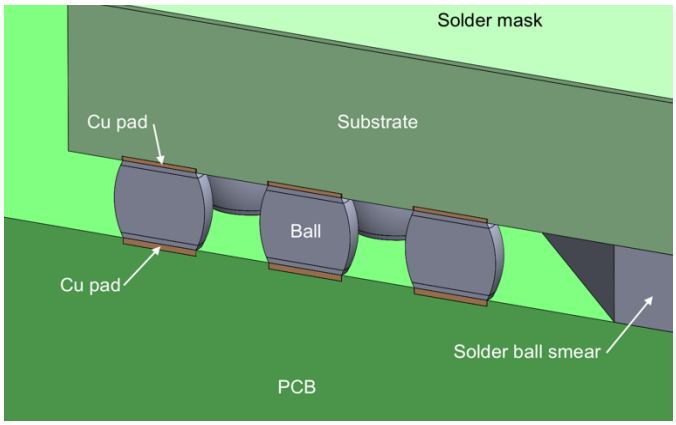

Figure 8 Solder ball cross section region

5.1. Design of Experiment setup

Since the simulations will examine three variables (lid CTE, PCB CTE, and the system ΔT) over a high/low range, a simple 2-level 3-factor DOE is used to create the simulation matrix. This results in 8 simulations runs (2^3), and one additional simulation run with a higher mesh density for checking. The DOE is particularly useful in finding the significant variables and any interactions between them, as well as the relative importance of any variable. Note this is a matrix used for two outputs – cycles to failure for the solder bumps and the solder balls. They are evaluated independently and so the significant variables and interactions may not be identical. Table 1 shows the DOE matrix and values.

The variable ranges were chosen to represent some typical BGA manufacturing operating conditions. With different material choices or operating environments, they easily could be larger/farther apart, but that is best covered with a 3-level DOE. Hence, here the PCB CTE is either 13 ppm/K (typical materials) or 9.5 ppm/K (a low CTE designed PCB). The lid CTE is either 17 ppm/K (copper) or an alloy with a lower value of 10 ppm/K. Finally, the environmental condition is either 20°C or 50°C for the entire BGA assembly. The die is not heated in this case (although it can be for actual problems). One thermal cycle is a two-hour period with 15-minute dwell times at each extreme, and 45 minutes to traverse from one temperature to the other.

Run #	PCB CTE, ppm/K	Lid CTE, ppm/K	ΔT, °C
1	13	17	50
2	13	10	20
3	9.5	10	50
4	13	17	20
5	9.5	17	20
6	9.5	10	20
7	13	10	50
8	9.5	17	50

Table 1 DOE factors and runs

5.2. Solder material models

The material models for the solder are dependent upon what fatigue model is chosen to calculate the results. This paper follows the model type used in Syed et. al. [8].

The elastic-plastic portion of the Sn-2.5Ag bump behavior is modeled with a linear elastic component and then a plastic component. In the FEA code used here (Ansys 2022R2), the plasticity is modeled with a Multilinear Isotropic Hardening (MISO) model over various temperatures. The temperature dependence and higher stiffness of the solder at lower temperatures is shown in the MISO plot, Fig 9.

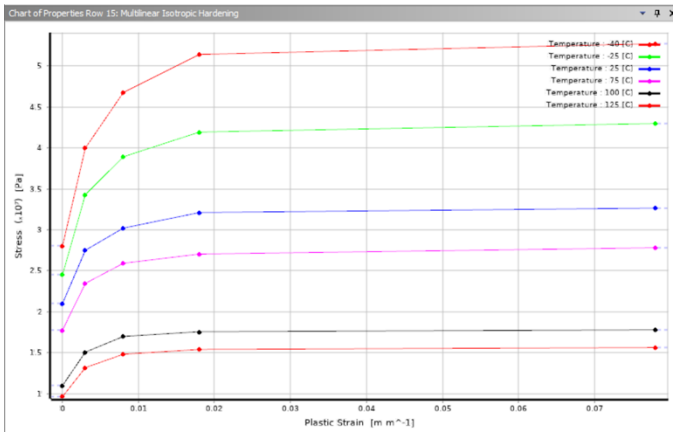

Figure 9 MISO plot Sn-2.5Ag solder

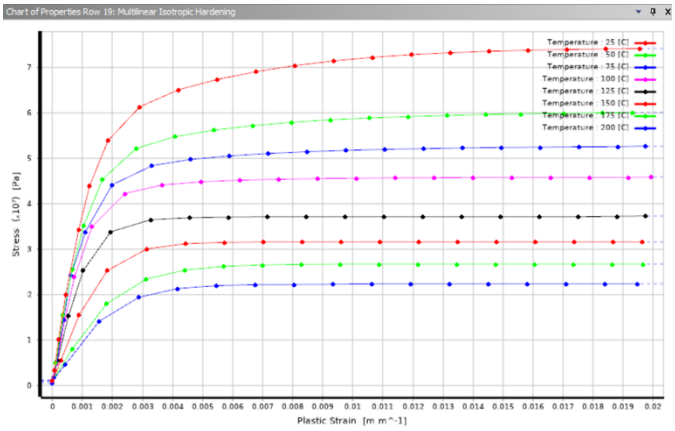

Figure 10 MISO plot SAC305 solder

22	Generalized Garofalo	
23	Reference Units (Length, Time, Temperature, Force)	mm...
24	Creep Constant 1	2630
25	Creep Constant 2	0.0453
26	Creep Constant 3	5
27	Creep Constant 4	6302

Figure 11 Garofalo (creep) model SAC305

As described in [8], the damage calculated for the bumps is best handled differently than the larger balls evaluated in typical literature. The steady state creep behavior is described by sinh-based creep behavior shown in equation 1. However, the key difference between this and a standard creep model often used (such as Garofalo) is that there is stress-based portioning of the damage calculation for both a low stress region and a high stress region. Hence the fatigue lifetime is calculated with equation 2.

$$\dot{\varepsilon}_{ss} = C_{ss}[\sinh(\alpha\sigma)]^n \exp\left(\frac{-Q_a}{kT}\right) \quad (1)$$

$$\frac{1}{N_f} = (C_1 SED_{GBS} + C_2 SED_{MC}) \quad (2)$$

In equation 2, N_f is the number of cycles to failure for the individual bump, SED_{GBS} and SED_{MC} are the accumulated strain energy densities (SED) for the for the low stress region (GBS) or the high stress region (MC). GBS occurs in solder when the grain boundaries slide (GBS). Higher stress regions occur with matrix creep (MC) and cracks happen inside the grains and through many of them. C_1 and C_2 are coefficients used to weight each type of damage, and this is done through experimental testing. Since Syed et. al. [8] used Sn-2.3Ag, this was assumed close enough to the Sn-2.5Ag used in this paper and the constants used were 0.00011 and 0.00028 for C_1 and C_2 respectively.

Computing the SEDs for the Sn-2.5Ag solder bumps requires examining the SED in the 5 µm slices in the top and bottom of each bump. That necessitates a proper creep model, and for this work (performed in the commercial code Ansys 2022R2) the creep model is in the form of Combined Time Hardening, different than what is used for the solder balls (discussed next).

The solder ball lifetimes can be found with a simpler method so a somewhat simpler model is used. The solder used here is SAC305, and correlation published by Syed [10] uses a creep energy density correlation to find the lifetime. As with the solder bumps, the solder is modeled in the FEA material model with a linear elastic section (temperature dependent as well), a plasticity MISO approach, and the sinh-based creep model (equation 1) with Garofalo constants. Figures 10 and 11 show the plasticity plots (MISO) and the Garofalo constants used.

Reference [10] provides the correlation for solder lifetime for the solder balls. As also done in [8], the paper correlates the theory to tests for creep failure in repeated cycling. The accumulated creep strain and dissipated creep strain energy is used in [10]. As used in [8], both low and high stress regions for damage accumulation are discussed in [10] but the SnAgCu solder damage is nearly all found in the high stress prediction for accelerated temperature ranges. Thus, total accumulated creep strain can be used for the solder ball lifetime. Equation 3 is the correlation to cycles to failure.

$$N_f = (0.0019\, w_{acc})^{-1} \quad (3)$$

6. Analysis

With the background presented to this point, the BGA geometry and the various material properties were set up as a finite element model (FEM) in the Ansys 2022R2 program. A total of eight analyses were run for the DOE matrix in Table 1, and one additional analysis of a higher density mesh was solved for run #1 to examine mesh density effects.

The eight run analyses were conducted with a fairly coarse mesh. In [8], generally finer meshes were used, but here a coarse mesh was used and as a reference run 1 as the higher density mesh (similar to what is used in [8]) is run for comparison. If the coarse mesh is successful, then it is a useful tool for mapping out the experiment space when larger numbers of analyses are run. Figures 12, 13, and 14 show the coarser meshing used for the overall model, the solder bumps, and the solder balls.

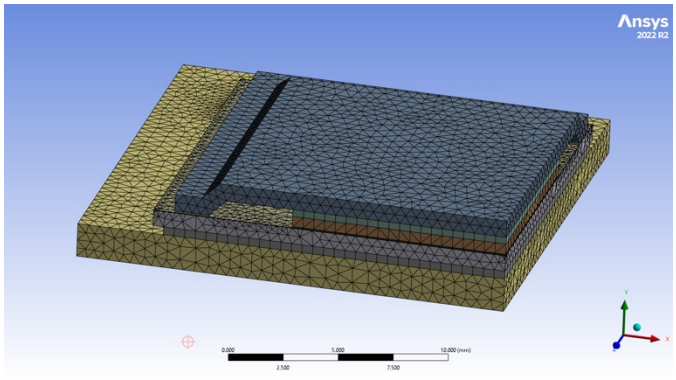

Figure 12 Full BGA model mesh

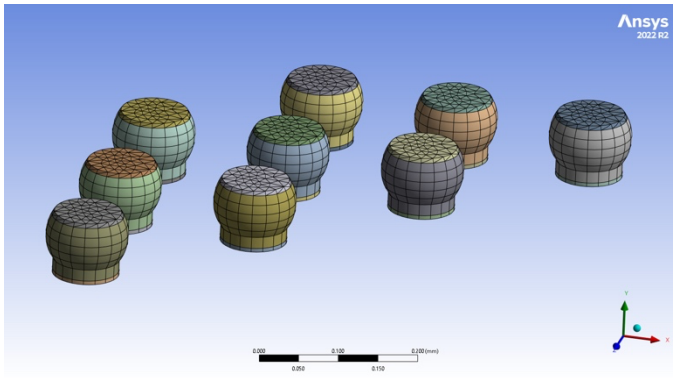

Figure 13 Solder bumps mesh

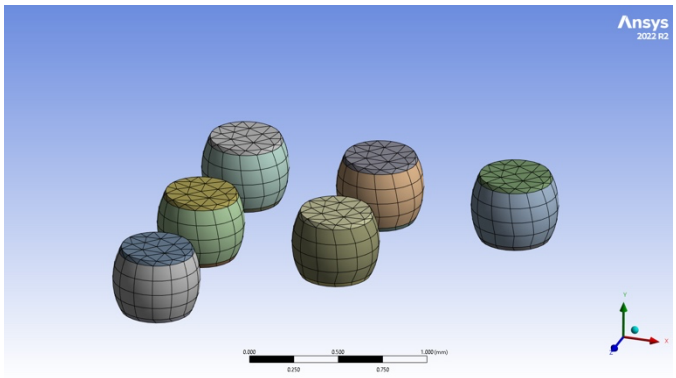

Figure 14 Solder balls mesh

A total of three thermal cycles are applied to the model (either a ΔT of 20°C or 50°C), with each cycle lasting 3000 seconds, as shown in Fig 15.

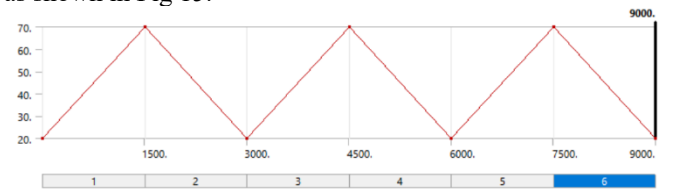

Figure 15 Thermal cycle of 20 to 70°C

In all of the DOE runs, it is important to check the base analysis results and ensure the solution is working properly. For these static structural analyses, base checks of the deformation (the primary solution of the FEA, with strains and stresses derived afterwards) should be made to ensure all elements are deforming in a reasonable manner. This can be a problem with thin or small elements, especially if contact elements are used between parts. Fig 16 shows a sample total deformation plot for run 8. All nine analyses run passed these deformation checks.

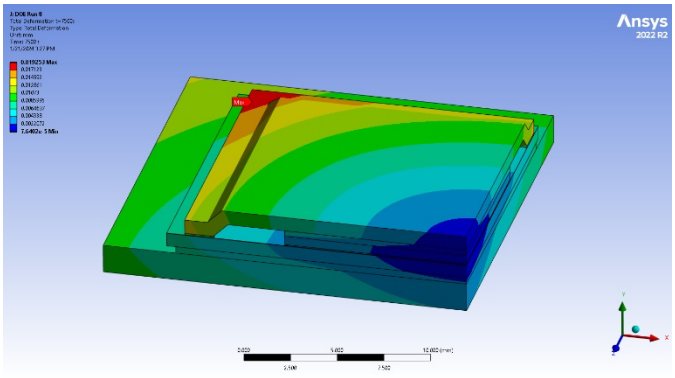

Figure 16 Run 8 total deformation plot

The plots of total strain (summation of elastic, plastic, and creep) for the solder bumps and balls are shown below for the worst strain point in the third thermal cycle (7500s) for run 8. As expected, the worst strains occur in the regions contacting either pad at the top and bottom extremes of the solder.

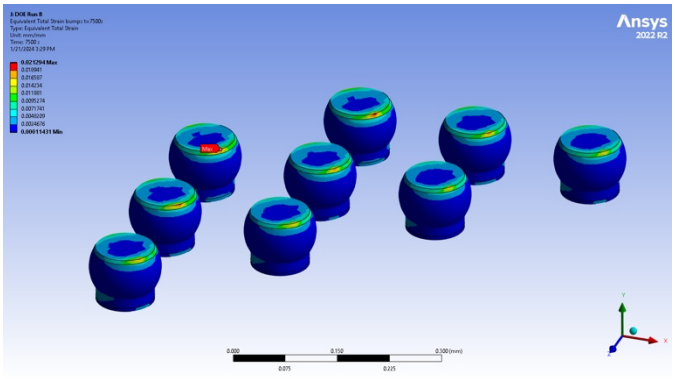

Figure 17 Solder bumps total strain, 7500s

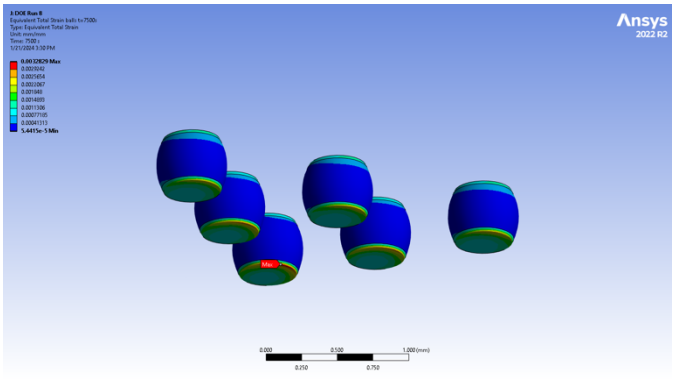

Figure 18 Solder balls total strain, 7500s

Looking at how the total strain progresses during the three thermal cycles, one can see the total strain increasing in Fig 19 for each cycle of the solder balls. This difference in strain from one cycle to the next usually stabilizes between cycles 2 and 3, though some solders are stable between cycles 1 and 2. Fig 19 shows the difference in total strain from cycle 1 to 2 is larger than 2 to 3, thus requiring three cycles for this solder.

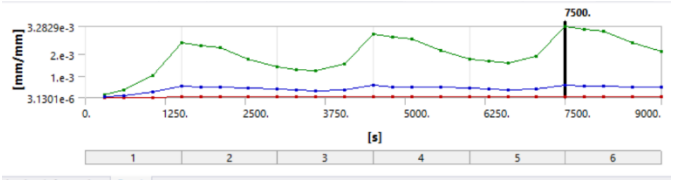

Figure 19 Solder ball total strain plot

From this point, the solder lifetime calculations are performed. To solve equations (2) and (3) for the solder bumps and balls, the formulas must be solved for the thin regions at the top and bottom sections. Depending upon the finite element software, this will likely be a script to find the combination of the strain energy determination and finding it for each element and summing it over those sections. Here, the nine solder bumps and six solder balls each need two scripts to account for the tops and bottom sections, or a total of 30 scripts are required to be run. Additionally, each element in the solder bumps must be evaluated to determine if the strain energy density is in the SED_{GBS} or SED_{MC} category, and then partitioned and summed appropriately.

Tables 2 and 3 below show the shortest solder life (fewest number of cycles to failure) for each run case, for bumps and balls.

DOE Run #	Bump #	Location	N_f
1	3	Top	34944
2	2	Top	10380100
3	2	Top	126020
4	2	Top	24196400
5	2	Top	19594400
6	2	Top	8945040
7	3	Top	126977
8	2	Top	198963

Table 2 Minimum solder bump life

DOE Run #	Ball #	Location	N_f
1	6	Bottom	77924
2	1	Top	7896440
3	5	Bottom	55755
4	1	Top	7906040
5	5	Bottom	23935800
6	5	Bottom	23417300
7	5	Bottom	47874
8	5	Bottom	55992

Table 3 Minimum solder ball life

Some basic observations are useful here. First, the smaller ΔT of 20°C has the longest life for both the bumps and solder balls (runs 2, 4, 5, and 6). This is expected as the strains between the two temperatures have a smaller difference, and hence a smaller SED. Second, something the DOE analysis will show better, there are cases where the lowest cycle life may be either the bumps or balls. For example, in run 1, the worst-case bump has half the life cycles that the worst-case ball does. For run 3, the worst-case ball has less than half the life of the worst-case bump.

Clearly there are conditions in a BGA design that can favor one solder location (bump vs. ball) over the other for the same ΔT cycle swing; this is a useful point to understand if one needs to "tune" the BGA to shift N_f in a particular direction. This is one of the key observations in this study.

6.1. DOE analysis

Given these results, an analysis via DOE software draws out additional useful information. For this analysis, the commercial software Design-Expert 13 was used to analyze the solder life outputs for the design matrix with an ANOVA for each bump and ball matrix. For each ANOVA, it was determined a square root of N_f best allowed for creating suitable transfer functions (due to large differences of lifetimes at low vs. high ΔT). The solder ball ANOVA result is shown in Fig 20.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	2.774E+07	4	6.935E+06	134.57	0.0010	significant
A-PCB CTE	27040.66	1	27040.66	0.5248	0.5212	
B-Lid CTE	1.166E+06	1	1.166E+06	22.63	0.0176	
C-dT cycle	2.526E+07	1	2.526E+07	490.12	0.0002	
BC	1.289E+06	1	1.289E+06	25.02	0.0154	
Residual	1.546E+05	3	51530.44			
Cor Total	2.789E+07	7				

Figure 20 Solder bump ANOVA

The equation for bump life using coded coefficients (-1 and +1 for variable ranges) is shown in equation 4:

$$\sqrt{(N_f)} = 2112.75 + 58.14\,A + 381.75\,B - 1776.8\,C - 401.5\,BC \quad (4)$$

where A is PCB CTE, B is the lid CTE, C is the ΔT, and BC is the interaction term of B and C. With a predicted R^2 of over 96%, this is a good fit to the analysis runs, and this shows in the plot of predicted vs. actual data points for the DOE (Fig 21):

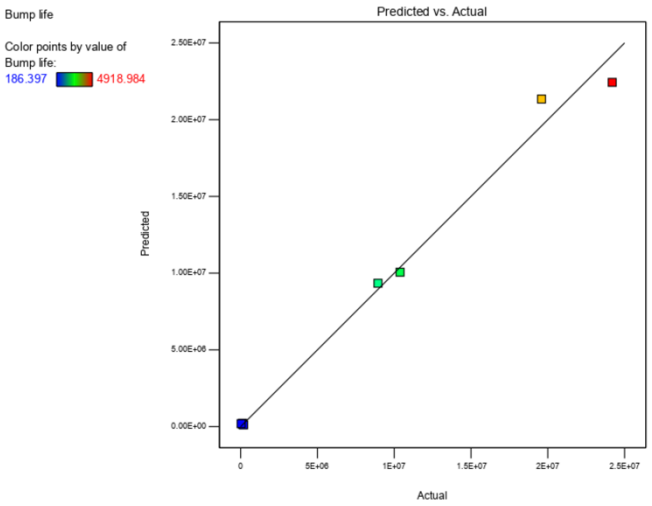

Figure 21 Bump DOE predicted vs. actual

The bumps show one interaction present, the lid and PCB CTEs. For this analysis, the effect is more present at lower temperature ΔT (black line) than the higher ΔT value (red line) as seen in Fig 22, and shown in the 3D plot in Fig 23:

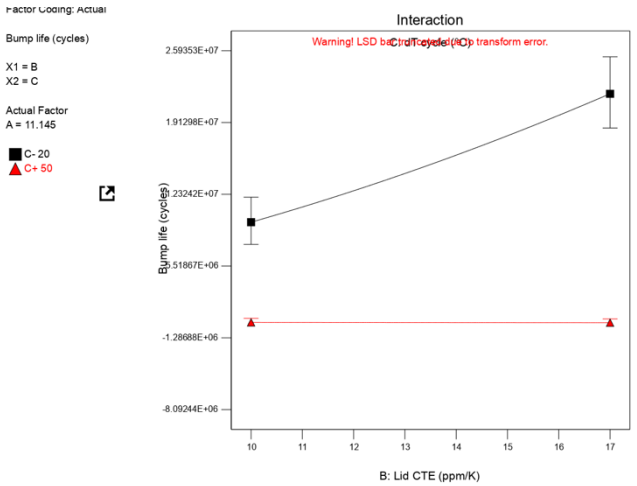

Figure 22 Bump interaction plot

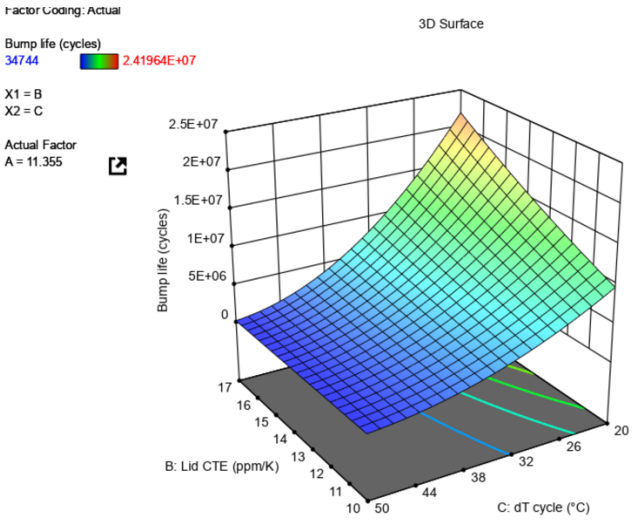

Figure 23 Bump interaction 3D plot

Similarly, the solder ball ANOVA (Fig 24) produces a coded transfer function with a square root for the N_f term as shown in equation 5.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	3.009E+07	4	7.522E+06	13998.03	< 0.0001	significant
A-PCB CTE	2.084E+06	1	2.084E+06	3877.79	< 0.0001	
B-Lid CTE	1630.32	1	1630.32	3.03	0.1799	
C-dT cycle	2.586E+07	1	2.586E+07	48130.93	< 0.0001	
AC	2.139E+06	1	2.139E+06	3980.39	< 0.0001	
Residual	1612.08	3	537.36			
Cor Total	3.009E+07	7				

Figure 24 Solder ball ANOVA

$$\sqrt{(N_f)} = 2040.31 - 510.36\,A + 14.28\,B - 1798.04\,C - 517.1\,AC \quad (5)$$

where now the interaction term is AC is significant vs. BC for the bumps. This equation is also a good fit as the adjusted R^2 term is over 99%, and is seen in the predicted vs. actual plot in Fig 25:

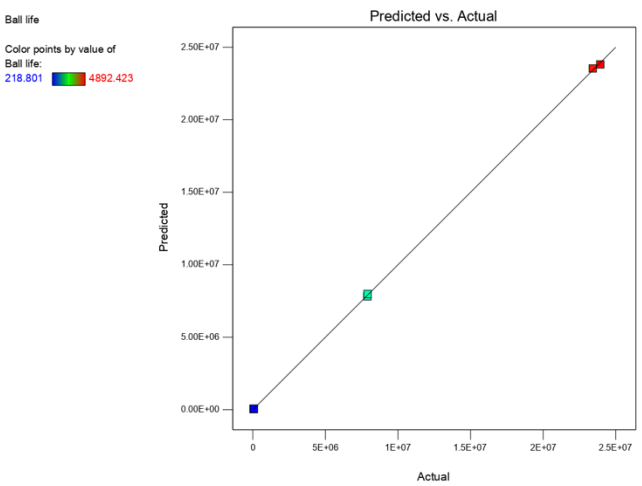

Figure 25 Ball DOE predicted vs. actual

The interaction for the ball transfer function is between the PCB CTE and the temperature difference ΔT as shown in Fig 26 and 27. As with the one solder bump interaction, the effect is more noticeable at the lower ΔT than at the upper one.

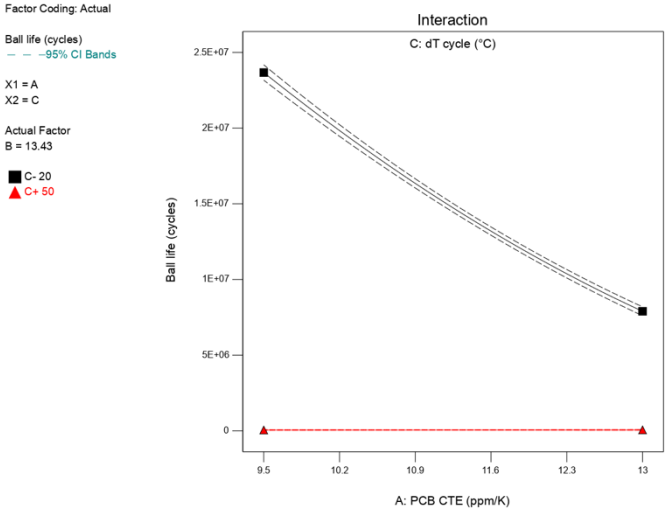

Figure 26 Ball interaction plot

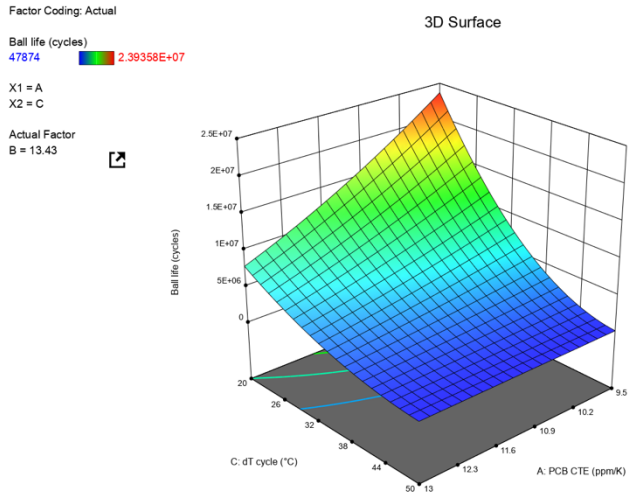

Figure 27 Ball interaction 3D plot

6.2. DOE discussion

In examining the ANOVA results, it is clear that different parts of the BGA affect the solder lifetimes differently, and this is due to the proximity of BGA parts to the solder. The solder bumps show little affect from the PCB CTE as the A coded coefficient in the transfer function is only 58 compared to 510 in the solder ball function. Similarly, the lid CTE affects the solder bumps far more than the balls, with coded coefficients of 382 versus 14 for the bumps and balls, respectively. Therefore, one could "tune" a solder lifetime by tuning the parts closest to the desired solder region.

As a side note, the lid CTE would have far more effect if it was adhered to the substrate with a greater stiffness adhesive. The relatively soft silicone RTV allows more compliance than an epoxy or a solder does. This is another variable to consider if the manufacturing process allows.

Another result is that the design space shows that under certain conditions the bumps have the minimum life, and under others the balls have the minimum life. The solder choices affect this but the factors in this DOE also affect it. In run 1 the bumps have the shortest lifetime, and in run 3 (lowering the lid and PCB CTEs) the balls now show the shortest life. However, run 3 provides a longer solder life compared to run 1. One can see that changing the factors may shift the shortest lifetime solder joint from one area to another (balls vs. bumps) but the overall affect may be a longer lifetime for the mounted BGA. Also noteworthy is that interaction terms in the transfer functions are present, but may be unimportant in an actual application. In both cases the interactions only matter at the lower end of the temperature range (ΔT=20°C), but the minimum solder lifetimes are in millions or tens of million cycles; this is likely far beyond the needed lifetime. In the higher temperature range (ΔT=50°C) the interactions are negligible.

Although this was a simple 3 factor DOE, a larger number of factors would likely show some significant factors to tune in a BGA. The lid adhesive was already mentioned. Another significant contributor would be the substrate CTE, and this can also be tuned. This factor would show some interesting results as the substrate is in between the solder bumps and balls. Shifting it lower would likely help the solder bump life, and higher would likely help the solder ball life. A variation of this paper's DOE would be to use a single ΔT value (no longer a factor) and add the substrate CTE, and use the shortest lifetime, bump or ball, as the output. This would likely produce an output model showing the best CTE choices for the substrate given the variations in the lid and PCB CTE values.

6.3. Higher mesh density results

Run 1 was meshed with a higher density and helped evaluate the mesh dependency of the model (and the suitability of using a lower quality mesh). The overall mesh was increased from 378,844 elements to 963,712 elements with significant refinements in and near the solder bumps and balls. Fig 28 and 29 show the solder refinements.

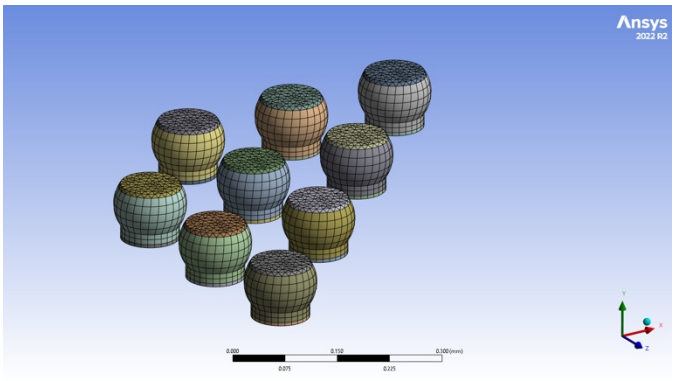

Figure 28 Solder bump mesh refinement

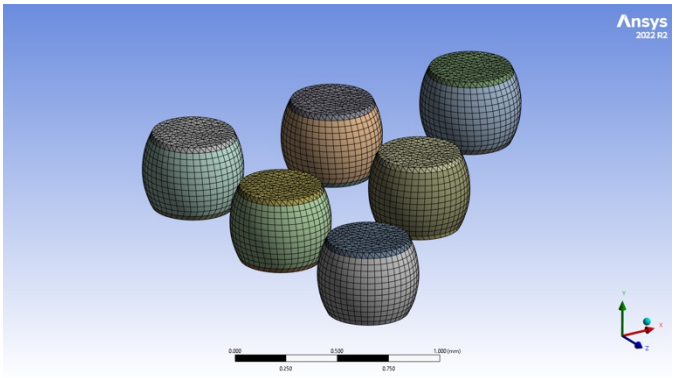

Figure 29 Solder ball refined mesh

The key output is the shortest lifetime of the two solder regions compared to the coarser mesh used otherwise. This comparison is given in Table 4.

Solder	Coarse mesh		Refined mesh	
Bump	3, Top	34944	3, Top	32217
Ball	6, Bottom	77924	1, Bottom	81814

Table 4 Solder lifetime coarse vs refined mesh

The lifetime numbers are similar (< 10% difference) and have the same location for the bump but a different one for the solder ball. Not shown in this table is that the solder balls have similar lifetimes for the solder ball bottoms and only the lowest one is shown. The solder bump lifetime is lower with the refined mesh, indicating one should run a refined mesh to derive final lifetime numbers. The coarse mesh is reasonably close and can be used in exploratory DOE calculations to arrive at designs more quickly (the author's computer needed 2.3 hours to solve the coarse mesh problem, and 13 hours to solve the refined mesh).

7. Conclusions

Performing a three-factor, two-level DOE of a typical BGA, a suitable FEA model using previously demonstrated methods for determining solder ball and bump life (see [8] and [10]) was conducted and analyzed. This study found:

- The balls and bumps are affected differently by the factors used in this analysis; this is affected by the distance from the solder location to the part (the longer the distance, the less the effect)
- Different factors can create the shortest life in either the bumps or the balls; the factor combinations in this model can cause either location to fail first

- Using these observations about the factor effects, it is possible that a BGA can be tuned for maximum solder life
- There are factor interactions but these only occur at low temperature cycles; at upper ranges (50°C) the interaction disappears
- Other factors not evaluated in the DOE will have significant effects, and may need to be considered
- A coarse mesh is feasible for an exploratory DOE and acceptable results are possible, but the final design should have a refined mesh to improve lifetime calculations

Acknowledgments

The work in this paper was partially supported by Rand Simulation via their provision of Ansys licenses for all the models and solutions.

References

1. Lee, W. W., Nguyen, L. T., & Selvaduray, G. S. (2000). Solder joint fatigue models: review and applicability to chip scale packages. Microelectronics reliability, 40(2), 231-244.
2. Chen, G., Zhao, X., & Wu, H. (2017). A critical review of constitutive models for solders in electronic packaging. Advances in Mechanical Engineering, 9(8), 1687814017714976.
3. Mukherjee, S., Nuhi, M., Dasgupta, A., & Modarres, M. (2016). Creep constitutive models suitable for solder alloys in electronic assemblies. Journal of electronic packaging, 138(3), 030801.
4. Depiver, J. A., Mallik, S., & Amalu, E. H. (2020, September). Comparing and benchmarking fatigue behaviours of various SAC solders under thermo-mechanical loading. In 2020 IEEE 8th Electronics System-Integration Technology Conference (ESTC) (pp. 1-11). IEEE.
5. Su, S., Akkara, F. J., Thaper, R., Alkhazali, A., Hamasha, M., & Hamasha, S. D. (2019). A state-of-the-art review of fatigue life prediction models for solder joint. Journal of Electronic Packaging, 141(4), 040802.
6. Gabriel, O. E., and Huitink, D. R. (October 22, 2022). "Failure Mechanisms Driven Reliability Models for Power Electronics: A Review." ASME. J. Electron. Packag. June 2023; 145(2): 020801.
7. Basit, M., Motalab, M., Suhling, J. C., & Lall, P. (2015, July). Viscoplastic constitutive model for Lead-free solder including effects of silver content, solidification profile, and severe aging. In International Electronic Packaging Technical Conference and Exhibition (Vol. 56895, p. V002T01A002). American Society of Mechanical Engineers.
8. Syed, A., Sharon, G., & Darveaux, R. (2012, May). Factors affecting Pb-free flip chip bump reliability modeling for life prediction. In 2012 IEEE 62nd Electronic Components and Technology Conference (pp. 1715-1725). IEEE.
9. Lau, John H., (April 13, 2022). Recent Advances and Trends in Advanced Packaging. Presented at IEEE EPS Binghamton Chapter (slide 11)
10. Syed, A. (2004, June). Accumulated creep strain and energy density based thermal fatigue life prediction models for SnAgCu solder joints. In 2004 Proceedings. 54th electronic components and technology conference (IEEE Cat. No. 04CH37546) (Vol. 1, pp. 737-746). IEEE.