The electronic industry drive for miniaturization and increasing functional integration forces the development of feature sizes down to the nanometer range. Moreover, harsh environmental conditions and new porous or nano-particle filled materials introduced on both chip and package level - low-k and ultra low-k materials in Back-end of line (BEoL) layers of advanced CMOS technologies, in particular - cause new challenges for reliability analysis and prediction. The authors show a combined numerical/experimental way and results towards optimized fracture resistance of those structures under chip package interaction aspects utilizing integral bulk and interface fracture concepts, VCCT and cohesive zone models in multi-scale and multi-failure modeling approaches with several kinds of imperfections. As important preconditions for high-quality simulations, nano-indentation AFM, FIB and EBSD provide the desired properties, while FIB-based trench techniques using deformation analyses by grayscale correlation and numerical simulations provide the intrinsic stresses especially of thin films in BEoL layers.

1. Multiscale Modeling of BEoL structures under CPI
Two major developments in electronics – miniaturization down to the nanometer range and introduction of new high-tech, nano-particle filled or nano-porous materials demand for innovative simulation techniques. An evident example is approaching by the latest technology developments for Back-end of line (BEoL) layers of advanced Cu/Low-k 90, 45…22 nanometer CMOS technologies. Under those conditions it is the big challenge for packaging to bridge the wide gap between chip (nm and µm range) and application (mm and cm range). While the thermo-mechanical reliability is dependent on the layer/vias design and the materials used, it is additionally highly dependent on the interaction between the chip (incl. the BEoL layers) and the type, design, chosen materials, and manufacturing technology of the package – the so called chip package interaction (CPI) – see [1-2]. So, the wide range of structural dimensions – the nanometer range for the transistor and tiny BEoL-structures and centimeter range for the overall design space of a device is one important challenge for FEA-simulations. Staying with “conventional” Finite Element Analysis implies global-local modeling (also known as submodeling approach), multiple sub-structures (or superelements) as well as fracture and damage mechanics utilizing also cohesive zone models, visco-elasticity, plasticity (ratcheting under cyclic loading conditions) and creep of homogeneously constitutive behaving materials.
Assets and drawbacks of utilizing substructures and submodeling techniques are:

- Substructures as realized in commercial FE-codes do not reflect nonlinearities - constitutive behavior of materials, contacts, CZM (cohesive zone method) etc.
- The evaluation of local deformation, strain and stress fields in substructures is possible but, there is no way to calculate fracture parameters (energy release rate ERR, J-integrals or stress intensity factors SIF).
- Making use of substructures is often coupled with a lot of fancy work and therefore fault-prone and
- It is necessary to pay attention to stiffening effects at the substructure-global model boundary or interlocking of substructures among themselves.
- The submodeling approach has the advantage that it may contain nonlinearities, user written elements and materials and allows calculating fracture parameters etc. but, the results depend on the displacements at the boundary of the submodel taken from global model simulation results. If the global model is not precise enough, this can cause misleading results:

A BEoL-stack in a flip-chip assembly (Fig. 1) was taken to make use of the submodeling technique and to demonstrate possible results. Fig. 2 presents the submodel when deformed, with an initial crack in a bimaterial interface starting from underneath the chip corner with a length of up to 100 µm. The global model consists (for this demonstration purpose) of a substructure which is identical to the submodel with two modifications, so that the global model covers an initial crack with the same length in one case and no crack in the other one.

![Fig. 1. FC assembly with an initial crack (left), BEoL stack (middle) and a 3D submodel detail (right)](image1)

![Fig. 2. Submodel of the BEoL stack in a FC-assembly](image2)

![Fig. 3. Normalized ERR at the crack tip with and without an initial crack in the global model](image3)

As visible in Fig. 3, the crack driving force at the crack tip vs. the initial crack length shows a completely different behavior if the global model covers an initial crack.
crack of identical length or not. Of course, only short initial cracks show the right characteristics. This is determined by the close neighborhood of the submodel boundary compared to the crack length (the thickness of BEoL is below 10 µm) and by the relatively weak material behavior of the adjacent PI layer, the underfill and soldermask.

Conclusion: utilizing the submodeling technique requires also finding the appropriate submodel boundaries. The mechanical behavior of the underlying elements in the global model could be satisfied using the substructure technique. Other challenges arise for instance from the close neighborhood of structural dimensions in design and morphology of newly developed materials in BEoL layers of advanced Cu/Low-k 90, 45…22 nanometer CMOS technologies. For example, black diamond-I or black diamond-II used as new materials in low-k BEoL stacks are increasingly porous and interconnect materials or new functional layers come up as nano-particle filled high-tech compounds. Therefore, it is to be checked whether they can still be handled as homogeneous materials, anymore. To identify the constitutive behavior of those materials and the appropriate material properties is generally difficult, requires expensive techniques (SEM, AFM, AFAM, FIB, Raman spectrometry etc.) and preparations. Frequently, they can only be estimated with the help of additional assumptions – linearity, constant Poisson’s ratio, stiffness of substrates or fastenings etc. These difficulties and the urgent request for more knowledge about the physics determining the material behavior is the driving force for atomistic level simulations and molecular modeling. So, atomistic level simulations start to help explaining the physics of deformation and damage incl. size effects in the closest area of crack tips in nano systems [3-4], in MEMS devices [5] or underneath a nano-indenter [6], and support at the same time to close the gap to conventional FE-techniques with the help of different hybrid FE–MD–QM simulation algorithms [7]. Molecular dynamics (MD) techniques increase in popularity for polymeric materials, carbon nanotubes, -rings, -connectors etc. [8], to simulate and understand the moisture diffusion [9], the mechanical behavior and properties of certain bi-material interfaces [10] and to determine material properties [11]. Two major ways seem to satisfy the need for characterizing the underlying physics best and to close the gap between MD and FEM:

**Direct incorporation of homogenized MD-models:**
- Modeling and simulation of the molecular structure,
- Homogenization in a unit cell,
- Use it inside a macro-model, a FE-Model for instance.
- Such approaches could also base on FEM- or semi-analytical representations of the micro-structure – the field of meso-mechanics.

**Extraction of mechanical properties for use in FEM:**
- Modeling of the molecular structure,
- Simulations towards extraction of key-properties,
- (Young’s modulus, CTE, diffusion coefficients)
- Use these properties in a macro-model, a FE-Model for instance – see [12].
Because of the wide gap between MD and FEM and also because of the huge amount of computational resources necessary for appropriate MD simulations, the second way is preferred at the moment.

2. Multiple Failure Evaluation

Miniaturization, new materials and harsh environmental conditions cause new challenges for reliability analysis and prediction, i.e. the development of multiple failure criteria for combined loadings like residual stresses generated by several steps of the manufacturing process, various kinds of inhomogeneity, moisture diffusion and the well known thermal expansion mismatch problem. These circumstances, which can cause different failure modes like interface delamination, chip or encapsulation cracking, pop corn ing and/or fatigue of interconnects, have to be treated on a new qualitative level. Traditionally applied methodologies base on classical strength evaluations or/and life time estimations of solder interconnects by means of modified Coffin-Manson approaches, which rarely address multiple failure modes. Second, especially under cyclic loading conditions, fatigue of solder materials, fatigue crack propagation in polymers or at bimaterial interfaces and ratcheting of kinematic hardening metals take effect simultaneously. Therefore, approaches to evaluate such risks and damage propagation rates have to provide qualitatively comparable results which speak the same language in order to use it for design optimizations. Table 1 depicts some of the accepted different risk and fatigue evaluating approaches and their different evaluation parameters.

**Table 1. Package reliability evaluation issues**

<table>
<thead>
<tr>
<th>Overload Risk/Damage</th>
<th>Evaluation method</th>
<th>Evaluation Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip cracking under bending (overload)</td>
<td>Classical strength hypotheses: Weibull-plots</td>
<td>maximum tensile stress in the chip surface - $\sigma_{\text{max}}$ → failure-probability</td>
</tr>
<tr>
<td>Cracking risk of housings, encapsulations or substrates (overload)</td>
<td>Fracture mechanics: CTOD [15], SIF k, VCCT G [16-17], J-integral</td>
<td>K-factor, J-integral, Energy release rate $G$,</td>
</tr>
<tr>
<td>Bimaterial interface delamination (overload)</td>
<td>Bimaterial fracture mechanics: VCCT [16-17], ARE [22], M-Integral [18], CZM [19, 20], XFEM [21]</td>
<td>Energy release rate $G$, SIF k (phase angle weighted, both), crack propagation vs. load/time</td>
</tr>
</tbody>
</table>
Fatigue

| Material fatigue of solder joints or gen. metals for instance | Coffin-Manson like approaches based on creep or plastic strain, dissipated inelastic energy evaluation or models similar to Darveaux model | Mean cycles to failure (MCTF) \( N_f \) \( \Delta a/\Delta N \) |
| Fatigue crack propagation | Fracture mechanics: Paris-Erdogan approaches for K, J, CZM-approaches incl. damage propagation | \( \Delta a/\Delta N \) S-N curves |
| Fatigue bimaterial interface delamination | Bimaterial fracture mechanics: Paris-Erdogan like approaches for K, G, CZM-approaches incl. damage propagation | \( \Delta a/\Delta N \) S-N curves |

SIF (stress intensity factor), CTOD (crack tip opening displacement), VCCT (virtual crack closure technique), ARE (area release technique), CZM (cohesive zone/surface modeling), XFEM (extended FEM), \( \Delta a/\Delta N \) (crack extension vs. number of loading cycles)

It is to be noted that some of the approaches are restricted to LEFM (SIF, mode separation for VCCT), some have to be checked for path independence of the results (J- or M-integrals) and some seem less applicable for unloading situations (J-integral, VCCT, ARE etc.). While CZM and XFEM have the potential of incorporating micromechanical conditions and processes with some difficulties - mesh dependence, time integration instability and the huge number of model-parameters that have to be measured prior to the simulations. However, in sensitivity analysis and especially in optimization studies the overload preventing parameters can act as boundary conditions, nonlinearly restricting the allowable parameter window. All fatigue evaluating approaches have to provide a consistent kind of measure in order to build a common objective function for multi objective optimizations. In as much, \( \Delta a/\Delta N \) as provided by most of the approaches discussed above seems to constitute a good basis for that universal failure criterion.

3. FEA-based DoE/RSM and Optimization

A great variety of parameters can influence the reliability of advanced electronic assemblies.

**Design:** Type of package itself (FC, CSP, WLCSP, BGA, SBGA etc), interconnect technique, embedding/housing

**Geometry:** thicknesses of substrates, silicon dies, die sizes, pitch, diameter of UBM, height and volume of solder bumps, height and shape of the fillet of an underfill a.s.o.

**Material:** Young’s modulus, Poisson’s ratio, CTE, \( T_g \), hardening behavior, yield stress, creep/relaxation coefficients, compressibility and the temperature dependence of all these parameters,

**Manufacturing:** Cure conditions of polymeric materials, soldering and reflow thermal conditions, sequential build up of the several thin layer and assembly processes.
The set up of a parametric study rapidly leads to an exponential growing amount of test combinations. Advanced DoE (design of experiments) approaches as realized in appropriate software tools help to design such parametric studies in order to obtain the best information on the influence of parameters on desired goals with a minimum of tests. This is vital especially for highly nonlinear, transient and/or spatial FE-simulations. Such software tools offer graphical representations about all relations and provide response surfaces (RSM) - mathematical models best fitting the simulation results.

Such response surfaces give an excellent basis for optimizations of the design and material properties with the goal of best thermo-mechanical reliability, lower costs etc. – see also [23]. On the other hand, all model parameters naturally show scattering. As a consequence all results depend on this scattering and show also a stochastic distribution. Optimal solutions found for determined model parameters often reside in a region near or directly on a boundary established by the risk of fracture, for instance. But, because of this stochastic distribution of the results, some of the potential solutions reside now in the forbidden region of the solution space and will probably fail. In order to reach a robust design, probabilistic design methods like the First Order Reliability Method (FORM) or Second Order Reliability Method (SORM) utilize the so called reliability index which is the distance of the nearest failure point from the mean value. This distance is a direct measure for reliability. It denotes how many variance steps lie between the parameter mean and the nearest failure point. In other words, if we have a distance of six standard deviations – that is 6-Sigma – practical no item exceeds the specifications – see also the expressions in [14].

In summary, a design for reliability bases on physics of failure and turns out to reach a probabilistic level, such as 6-sigma, which means 3.4 failures in one million parts. That’s why; such a design is called a “robust design”.

Fig. 5. Coefficients of a polynomial fit of the response surface shown in Fig. 6

Fig. 6. Dependence of accumulated equivalent creep strains in solder balls on the CTE of underfill and soldermask

Fig. 6 illustrates the dependence of accumulated equivalent creep strains on the CTE of underfill and soldermask of a flip-chip assembly under thermal cycling conditions. Fig. 5 gives an impression about how much influence the coefficients of the polynomial representation of a 2-parametric model have on the accumulated equivalent creep strains in solder balls. It is easy to see that the CTE
of the underfill (alpha_UF) is the most important parameter with the constant part and with the 3rd power part but, also the same parts of the CTE of the soldermask (alpha_SM) and the bilinear combinations of both of them play an important role.

4. Material Properties and Initial Stress State Determination

Complementary to the simulation side of reliability estimations, serious issues are connected with the collection of appropriate material properties in the miniaturized range addressed – Young’s modulus, initial yield stress, hardening, in particular. In addition, the evaluation of residual stresses, especially of thin films, inherent to many of the thin film deposition and treatment processes [24], is an important precondition for high-quality simulations. Accordingly, the estimation of intrinsic stresses of thin films on substrates or layered structures by the help of a FIB-based trench technique and numerical simulations will be exemplified.

The authors developed a new technique based on stress release by FIB milling. Digital image correlation (DIC) algorithms are utilized to determine stress release deformations from SEM micrographs, captured before and after ion milling. As a result residual stresses can be computed [25-28].

In the past it was demonstrated that the extraction of in-plane stress values of thin membranes can be easily carried out measuring local stress relaxations. Tiny through slots milled into the membrane inside a FIB equipment lead to stress relief deformations causing some tens of nanometers slot opening. Because the mechanical problem for the displacement field can be solved analytically, a relatively simple parameter fitting allows computing stress values [25].

Stress measurements in thin layers or layer structures are a more general problem of concern. First attempts to tackle this problem were made by the authors of [27, 28], milling tiny trenches into the layers of interest. Afterwards, trench opening/closing is used as a measure of released stresses. The applied approach suffers from the fact that available analytical solutions of the corresponding mechanical problem do not describe stress relaxation fields nearby ion milled trenches well enough.

Consequently, more accurate finite element analysis has to be applied in combination with displacement field measurements around the milled trench. Both displacement fields from FEA and DIC measurement, respectively, have to be compared by some kind of least square fit in order to determine the stress values describing the experimental data in a best way.

A measurement routine developed for this purpose is illustrated by the following stress determination example. A highly stressed Si3N4 layer with compressive stress has been deposited on a Si substrate. Fig. 7 shows a typical trench milled into the thin layer and partially into the Si substrate, in order to release stress and to cause a trench closing. High resolution SEM images are picked up before and after trench milling in a FIB cross beam equipment. Between these captures the sample is rotated 54° towards the FIB ion column to carry out trench milling.

Subsequent images, with and without trench respectively, are processed by an appropriate DIC software. The resultant displacement field for the trench in Fig. 8 is given in Fig. 9.
Obviously the left trench rim moves towards the right one, which can be seen from the sign of the displacements, i.e. the trench is closing after milling. This situation refers to compressive stress in the layer.

Fig. 9 represents a finite element simulation carried out for the experimental trench size. Stresses at the trench site before milling were set arbitrarily to 587 MPa. Because of the linear elastic behavior of stress release, numerical displacement fields can be scaled until they fit experimental data. Residual stress in the vicinity of the trench milling can be found multiplying the originally assumed stress of 587 MPa by the scaling factor of the best fit.

5. Conclusions

Design for thermo-mechanical reliability of electronic components and microsystems on the basis of parameterized Finite Element Models and DoE/RSM-approaches is more and more used for optimizations at early phases of the product development process. Besides this, feature sizes at the nanometer range and the introduction of new high-tech, nano-particle filled or nano-porous materials cause novel challenges for reliability analysis and prediction, i.e. the development of multiple failure criteria for combined loadings including residual stresses, interface delamination, cracking and fatigue of interconnects simultaneously. The authors face up to multiscale modeling approaches, damage and fracture mechanics approaches on the basis of continuum mechanics and measurement techniques of material properties in the miniaturized range addressed. Reliability predictions of these miniaturized multi-material systems frequently require considering not only the variety of environmental loadings and combined multiple failure criteria, but also intrinsic stress situations from previous technological steps. A new analysis technique based on stress release by FIB milling and high-resolution displacement measurement has been proposed.

6. Acknowledgments

The authors express their thanks to Emmanouella Dermitzaki and Bernhard Wunderle from Micro Materials Center at the Fraunhofer IZM Berlin for their help and support with molecular modeling results. The authors acknowledge also the German Federal Ministry of Education and Research for funding and promoting substantial parts of the related work (registered under funding number 13N9228).
References


