Use of Lock-In Thermography for Non-Destructive 3D Defect Localization on System in Package and Stacked-Die Technology

Rudolf Schlangen, Shinobu Motegi, Toshi Nagatomo
DCG Systems, Fremont, CA, USA
Rudolf_Schlangen@DCGSystems.com

Christian Schmidt, Frank Altmann
Fraunhofer Institute for Mechanics of Materials Halle, Germany

Hiroaki Murakami
Toshiba, Yokohama, Japan

Stewart Hollingshead, John West
Texas Instruments, TX, USA

Abstract:
With the growing variety, complexity and market share of 3D packaged devices, package level FA is also facing new challenges and higher demand. This paper presents Lock-In Thermography (LIT) for fully non-destructive 3D defect localization of electrical active defects. After a short introduction of the basic LIT theory, two slightly different approaches of LIT based 3D localization will be discussed based on two case studies. The first approach relies on package internal reference heat sources (e.g. I/O-diodes) on different die levels. The second approach makes use of calibrated 3D simulation software to yield the differentiation between die levels in 8 die μSD technology.

Introduction
Increasing interconnect complexity, component density and number of stacked dies in 3 dimensional system in package (3D SIP) devices together with shrinking critical dimensions result in increasing challenges for isolation and root cause analysis of defects. In particular, new fault isolation methods with 3D defect localization capability are required. Standard methods to localize electrical defects like OBIRCH [1], emission microscopy or liquid crystal thermography [2] can not be applied, because optical access to the functional structures in 3D SIP devices is mostly not available.

Lock-In-Thermography (LIT) is an established technique for on-die defect localization, but based on its superior sensitivity it also allows for completely non-destructive localization of defects through covering package material. The first theoretical evaluation and proof of concept results for using LIT in general and the ELITE specific phase data in particular for non-destructive 3D localization of buried defects were already shown in [3] & [4]. The applicability of LIT for 3D defect localization on stacked die samples was first shown and discussed in [5].

This paper focuses on the practical implementation and optimization of the 3D defect localization based on two case studies, exemplary for the two most common product types/applications for today’s 3D packaging: first, the combination of one logic- and one memory die first and second a multi-chip memory stack for μSD technology.

Basic LIT theory
Key to successful localization of small heat sources, such as defect related hot-spots, is the real-time graphical lock-in methodology. Figure 1 shows the standard lock-in principal, used to detect a small signal buried in random noise of higher magnitude.

Figure 1: Illustration of basic thermal lock-in principle

The Elite delivers the same results – amplitude and phase. However, there are two main differences:

a) The input is not a continuous- / analogue-signal, but a stream of IR-images, hence discrete data points
b) The discrete lock-in algorithm has to be applied to every image pixel separately.

The SW performs all lock-in calculations (for 640 x 512 pixels) at real time, allowing to see immediate results and offering to accumulate data as long as necessary; there is no acquisition time limit. Acquisitions on hot-spots with moderate power dissipation (> 1mW) only require seconds up to a few minutes, but also extremely small heat sources can be localized by simply spending more time; the record so far is the clear detection of a 400nW source after 18 hours.

The real-time, pixel-wise lock-in algorithm is described in more detail in [1].

Figure 2: Illustration of thermal 3D defect localization on a stacked die memory product.

Figure 2 shows a scenario with two hot-spots, buried at different depths in a package. Applying the lock-in excitation voltage causes periodic heat generation at these two spots. The heat starts diffusing in all directions, also towards the package surface. Based on the individual depth of the heat source, it takes the heat different times to reach the surface, resulting in different phase shifts.

Since the ELITE directly measures the phase, $\phi$ can be used to calculate the depth of the defect. Calculation of the defect depth requires either precise knowledge of the thermal properties of the involved material (such information is usually not readily available) or a calibration via reference measurements, which will be subject of the first case study.

### LIT 3D defect localization, based on reference measurements

The first DUT is a fully encapsulated SIP with two dies stacked on each other. The lower die is mounted face-down in flip-chip technology. The upper die is glued onto the lower one and connected to the package lead-frame via bond-wires (see cross-section in Figure 3).

**Figure 3: Cross-sectional view of DUT**

The failing device shows a ~13Ω short at one IO pin and the aim is to localize this defect in x, y and z without removing the package.

The x, y localization can be obtained with just one LIT measurement and the results can be seen as the overlay between the LIT amplitude image (in colour) and the IR image of the same sample as background / topography information (in gray), shown on the left of Figure 4.

In a second step, we want to use the ELITE phase result to determine the defect depth. The upper most row of images in Figure 5 shows multiple such phase results of the failing pin, measured at different lock-in frequencies. The extracted phase values are plotted in Figure 6 (labelled fail). Since the thermal properties of most involved materials (e.g. mould compound and die attach tape) are not known precisely enough, the use of a readily available 3D simulation SW is not practical with such samples. The much faster approach is to compare the thermal signature of the fail to reference heat sources on known levels in the package. In many cases, IO diodes can serve as such reference heat sources by simply biasing them in forward direction.

The used reference points are indicated in the cross-sectional view of the DUT (labelled A & B) and the results are shown in the Figures 4, 5 & 6 in comparison to the fail signature.

**Figure 4: Elite results of failing pin and two reference points, all at 1x magnification**
The fail signature shows consistently higher phase shift than the reference data of the lower die, clearly indicating that the defect must be below the lower die. Since the difference is small, this allows concluding that the short must be located in the uppermost layer of the lead-frame, indicated by point C.

This result matches the theory and prior experience of the customer and we are awaiting confirmation by 3D X-ray and/or PFA.

(The combination of 3D LIT for initial defect localization followed by high resolution 3D x-ray for final defect analysis has proven to be a very fast and completely non-destructive solution for package level FA, with multiple application examples already published in [6].)

Figure 6 contains an additional set of reference measurements (dashed lines), measured on a good / reference sample, showing very close match to the reference results of the failing DUT, proving the consistency of the phase data. This consistency allows using once measured reference data for any sample of the same technology simplifying and increasing the through-put of the proposed technique. Furthermore, such reference data may be used to extract the thermal properties of the involved materials allowing the use of dedicated 3D SW, which will be discussed in more detail with the second case study.

Advanced 3D defect localization based on phase-simulation SW

In contrast to the DUT of the first case study, most commercially available stacked memory devices consist of 8 stacked memory dies and most recently even up to 16, with ever shrinking die thicknesses of 30\(\mu\)m and below.

Initial experimental results and simulations regarding such die stacks (published in [5]) could show that the glue layers in-between the stacked dies are poor thermal conductors and therefore induce a clear difference in thermal time delay (measured as phase) going from one Si-die to the next. Consequently, the same measurement principle as with the first case study (compression to a local reference, on the same die) would yield clear layer detection, but since all memory dies are bonded in parallel with the predominant device architectures, using IO diodes as layer specific reference heat sources is mostly not applicable.

Fortunately, one fundamental requirement for such aggressive 3D packaging is a very tight process control, ensuring minimum variation of layer dimensions and therefore also providing very reproducible heat flow properties within different samples of the same technology. This allows yielding reliable identification of the defective layer by comparing to the reference or calibration data, previously measured on known defective devices.
Use of 3D SW

Besides the high level of process control, the limited number of utilized materials in memory stacked-die packages makes the use of dedicated phase simulation SW advantageous. With the shown example (8+1 die μ-SD), the full stack consists of the covering mould compound and a strictly repetitive stack of dies and die-attach-tape (glue layer). Even-though the individual layer thicknesses may vary, there are only three different materials which need to be characterized to correctly simulate the thermal behaviour of the whole stack; mould compound, Si-die and die-attach-tape.

The SW is based on an analytical solution of the 1-dimensional heat spread function. The procedure begins with generating a layer model of the used sample including all layers, their individual thicknesses and materials. The simulation SW calculates and plots the expected phase values vs. lock-in frequency for all relevant layers.

In case of dealing with unknown (not thermally characterized) materials, the user starts with a set of default parameters and can fine-tune / calibrate the model by adjusting the used material parameters to match reference measurements.

Figure 9 shows the results; the points are reference measurement results and the solid lines are the simulated phase shift based on a calibrated simulation model, both plotted versus the lock-in frequency, showing excellent agreement.

Due to the small number of materials to be accounted for and the repetitive nature of the layer structure, 3-4 reference devices with known heat-source in different layers are considered sufficient for such model alignment on a given 8-stacked-die technology.

For the same samples shown in Figure 8, new data was acquired now measuring at 2Hz and 5Hz and comparing to the simulation model. The changed procedure allowed correctly identifying 86% of all samples, and still 100% with +/- 1 die.

Optimized Measurement Procedure

The lock-in frequency is the most important measurement parameter. If chosen too high, the resulting temperature change visible on the surface decreases to a level where sufficient SNR cannot be achieved in a reasonable time. This is especially critical for a defect in the lower layers, where there is more material for the heat to penetrate before affecting the surface temperature. The other way around, if the defect is located on one of the upper layers (close to the surface), it’s advantageous to measure with increased lock-in frequencies to achieve higher phase values, allowing better separation between neighbouring layers.

Additionally, measuring at multiple frequencies improves the accuracy and confidence level of the 3D localization. To accommodate for the above, the proposed procedure starts with one acquisition at a moderate frequency (at best one where all layers deliver acceptable results, e.g. 5Hz as indicated in Figure 9) and depending on this first phase result the next measurements will be performed at either higher or lower frequencies. In the shown example, for phase below 150° at 5Hz, the next measurements would be done at 7 and 10Hz, whereas for the lowest layer (die 1), with more than 200° at 5Hz, the next acquisition should be at a lower frequency (e.g. 3Hz).

Depending on the desired level of confidence vs. the required acquisition time, the procedure may end after 1, 2 or even more matching data-points between measured- and simulated phase values.

One additional benefit of using multiple phase vs. frequency data-points is that unforeseen influences like delamination between layers or the presence of multiple defects in the same x, y area but on different layers are likely to influence the resulting phase values to not follow the expected square root trend of the simulation data. This would allow excluding those DUTs and not getting false localization.

Increased Throughput

The necessary acquisition time to achieve a certain SNR is linearly dependent to the camera frame rate (number of images captures per second). In standard operation, the utilized 640x512 pixel focal plane array has 100Hz frame-rate.

Initially, the LIT acquisition should cover the full DUT to localize the defect in x and y and also give the first (rough) z-localization to determine which frequencies to be
measured next. This often requires the use of the full pixel-array and a de-magnifying (wide-angle) lens. After this point, the 3D localization only requires the phase information on the centre of the hot-spot at different lock-in frequencies, so continuing to measure with the full pixel array is unnecessary.

To ensure stable and reproducible camera / system operation we propose to use a pre-defined sub-array of 128x128 pixels for the 3D (only) acquisition, allowing the camera to operate at almost 9x higher frame rate, reducing the necessary acquisition time by almost 9x.

**Electrical Setup**

Since the resistivity of shorts to be detected can vary by orders of magnitude, it’s important to set the right voltage levels for defect excitation; high enough to get sufficient signal (fast SNR) but still below power levels where the temperature dependence of the thermal diffusion properties of materials would have to be taken into consideration.

For example, with the used 8+1 μSD devices excitation power between 1 to 5mW is expected to be best and the lock-in voltage should be selected accordingly.

Additionally, any kind of SIP is likely to show non-linear IV-characteristics which may include current spikes due to device initialization as shown in Figure 10 between 1 and 2.5V. Since such initialization procedures may cause non-defect related thermal activity it is beneficial to avoid this voltage range, initialize the DUT before the LIT acquisition and switch between two voltages above the initialization limit (e.g. form \( V_{B \text{ low}} \) to \( V_{B \text{ high}} \)) during the LIT measurement.

**Figure 10: IV plot of DUT**

Especially with SIP, an increasing number of products requires such initialization and is equipped with on-chip / in-package voltage regulators, which often make simple switching or toggling of Vdd inapplicable since the local power dissipation of the defect does not change with the outer modulation, therefore can’t be detected based on the LIT principle anymore.

For such cases, a tester (ATE) may be used for controlled defect activation whereby the LIT setup synchronizes to the ATE operation (see Figure 11). A more detailed explanation and multiple application examples were published in [7].

**Conclusion:**

The 3D integration of multiple dies into one package as for SIP and high density memory products like μSD has proven to be a clear trend and a fast growing market. The resulting complexity of such 3D architectures and the shrinking geometries of today’s packages with emerging Si-interposers, TSVs and micro-bump technology offer great performance increase and cost reduction for various applications, but also require new techniques and methodologies for effective debug and FA.

Lock-In-Thermography is one of the most promising techniques to deliver non-destructive 3D defect localization through the still covering packaging material.

Based on the first case study it could be shown how 3D LIT can distinguish between different die- and packaging layers in a two-die SIP, using local reference points (e.g. IO-diodes) on each die for direct comparison.

This technique is fast and reliable, given access to the individual die levels.

The second case study focused on more aggressive μSD technology, where the first approach fails since individual reference points are not available due to the parallel connection of all memory dies.

Using an additional SW to simulate the thermal properties of the DUT in combination with a limited number of reference measurements on known defective samples delivered very promising results, with 86% success rate of localizing the failing die out of the 8+1 die-stack and 100% with ±1 die accuracy.

Furthermore, the influences of various measurement parameters as lock-in-frequency and excitation voltage have been discussed and an optimized procedure could be proposed where multiple LIT results are used to improve localization accuracy whereby the overall acquisition time can be reduced due to sub-array operation of the infra-red camera.

The proposed procedure is currently evaluated on 16+1 stacked die memory products.

Additionally, combining 3D LIT for initial localization with the high spatial resolution of 3D x-ray has proven to be extremely useful for fast and completely non-destructive package and even board-level FA.
Acknowledgement:

Special thanks to the Infineon FA team in Munich, Germany, for the fruitful discussions and for triggering the idea of using IO-diodes as DUT-internal reference points.

References:


