

Latent defect screening for high-reliability glass-ceramic multichip module copper interconnects

E. J. Yarmchuk
C. W. Cline
D. C. Bruen

Nonlinear electrical conduction effects arising from local heating provide a means for detecting the presence of latent defects that are at risk of becoming open electronic circuits under stress. A special choice of ac drive current results in a dc intermodulation signal produced by nonlinear conduction, and offers several practical advantages over previous techniques involving ac harmonics. A key feature is the use of digital signal processing to provide speed, accuracy, and flexibility of measurement. A manufacturing screening system involving an automated probe integrated with the defect-detection tool is used to test high-performance glass-ceramic substrate interconnects used in advanced microelectronic packaging.

Introduction

Production defects, often involving contamination, can result in narrow constrictions in the conducting lines of electronic packages [1]. These are mechanically weak and can fail from the stresses of assembly, testing, and thermal cycling during usage, and thus are of particular concern in environments with very high reliability requirements. In complex glass-ceramic packages, conductors are formed from a metal paste that is patterned onto multiple dielectric surfaces and squeezed into punched via holes between layers. The package is then fired in a sintering oven. These processes create conditions under which line constriction defects can occur. **Figure 1** is a micrograph of such a defect in a copper conducting line in a multilayer glass-ceramic interconnect package.

A very narrow constriction, as shown, creates a small amount of extra resistance in the line so that the electrical performance of the conductor is virtually unaffected until the constriction actually breaks. The ability to screen parts for these latent defects can significantly improve interconnect reliability.

The low resistance of constriction latent defects makes them impossible to detect by the straightforward approach of measuring for increased electrical resistance. According to Holm [2], the resistance associated with a

small circular contact spot between much larger conducting members equals the material resistivity divided by the spot diameter. Thus, a 1- μm -diameter constriction in a larger-diameter cylindrical copper conductor adds a resistance of only 17 m Ω . With typical line resistances of several ohms or more, the defect contribution would be hidden within the normal process tolerances for resistance. However, techniques based on nonlinear conduction [3–8] can highlight the presence of constrictions by detecting harmonic generation caused by local resistive heating at the constriction. This paper describes a production screening system based on nonlinear conduction that combines features of several of these approaches to achieve speed, sensitivity, and accuracy in latent defect testing of copper conductors in glass-ceramic chip carriers.

Nonlinear conductivity techniques

Figure 2 illustrates the basic principle behind nonlinear conductivity measurements. In the figure, a current I is applied to a conductor having a constricted region near the center. Ohmic losses in the constriction result in a local hot spot with a temperature rise δT that is proportional to $I^2 R_c$, where R_c is the extra resistance associated with the constriction. This raises the resistance by an amount $\delta R = \alpha R_c \delta T$, where α is the temperature

©Copyright 2005 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

0018-8646/05/\$5.00 © 2005 IBM

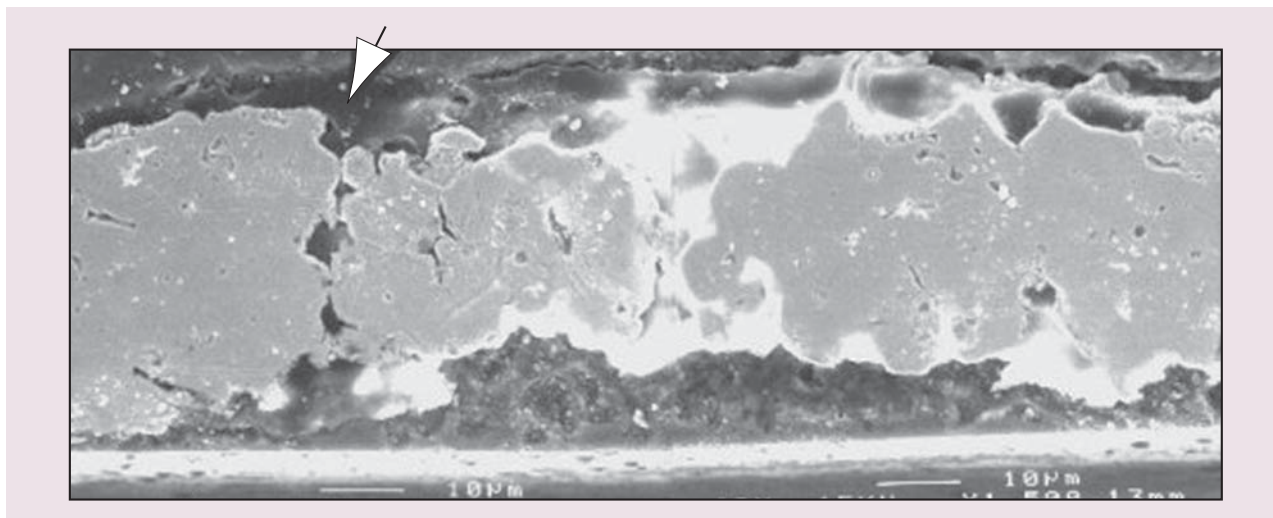


Figure 1

Micrograph showing a defective conducting line. The overall height of the image is approximately 50 μm .

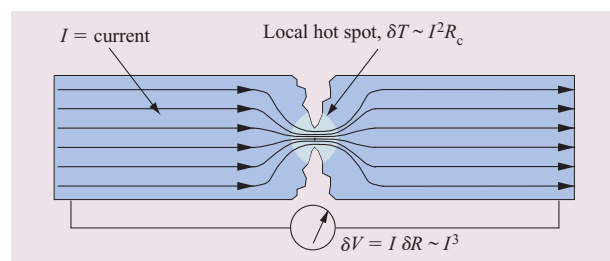


Figure 2

Diagram illustrating the principles of nonlinear conduction at a constriction.

coefficient of resistance. Thus, there is a constriction-related component in the voltage across the conductor that is proportional to the cube of the current. In principle, one could detect this by looking for departures from linearity in an I - V curve. However, at reasonable current levels (~ 1 A), the signal of interest can be as low as a few microvolts, while the overall voltage drop $V = IR_{\text{total}}$ can be tens of volts, making this impractical. Successful methods of detection [3–8] utilize an ac excitation current so that the cubic nonlinear relationship results in harmonic generation of signals at frequencies other than the drive frequency itself. Through filtering and narrow-band detection, discrimination of defect signals that are millions of times smaller than the fundamental signal becomes possible.

With the use of nonlinear conductivity, the effect of the constriction resistance becomes distinguishable from the

overall line resistance, making latent defects detectable even in the presence of resistance variations in a part due to process or line length differences from net to net. However, there is still a residual dependence on the overall line due to the fact that regions of good line also undergo small temperature excursions due to joule heating. Therefore, to achieve the greatest sensitivity, further discrimination between line and defect signals is desirable. Halperin et al. [8] describe in detail how the phase of the harmonic signal can be used to provide additional discrimination. The concept is based on the fact that the thermal response time of a narrow constriction differs dramatically from that of a good line. The response time affects the phase of the harmonic signal, allowing phase-sensitive detection of the voltage (e.g., via lock-in amplifier) to block the unwanted portion of the signal arising from good lines while passing the desired constriction-related part.

In a constriction, the localized hot spot has both a small heat capacity and a low thermal resistance to the surrounding metal line. This makes the thermal response time short compared with the period of the power input, so the phase lag between temperature and power is very small. In normal line regions, however, heat flows rather poorly to the surrounding insulator materials and the heat capacity is much greater, so the response time is longer than the period of the driving power. Here, the line acts like a thermal integrator, and its temperature lags the power by 90° . For specific assumed geometries, formulae can be derived for the magnitude and phase of normal line and defect signals. These are considered as a function of drive frequency and the electrical and thermal

properties of the conductors and surroundings [as in Reference [7], Equations (1) and (2)]. These are useful for determining the appropriate operating frequency of a tester. For example, a frequency f of about 1 kHz is good for printed circuit boards with large-cross-section conductors in low-thermal-conductivity dielectric substrates, while the thinner conductors and higher thermal conductivity of glass-ceramic dielectric substrates require a much higher frequency, typically 1 MHz.

In the high-frequency tester described by Halperin et al. [8], the drive consists of a combination of three currents; dc, 1.145 MHz, and 1.6 MHz. The nonlinear voltage, which is proportional to the current cubed, contains terms at a variety of frequencies consisting of multiples of the two ac frequencies as well as sums and differences of them. One of these, 455 kHz, is a standard intermediate frequency used in RF electronics, and was chosen as the signal of interest. This is detected using a lock-in amplifier referenced to a phase-adjustable 455-kHz signal derived within the tester by mixing of the two ac drive components. This permits selective detection of the voltage component at the thermally induced phase shift corresponding to the defect response.

This technique works well, but requires very sophisticated analog electronics with low intermodulation distortion on the drive circuits. Since the fundamental drive signals are only three to five times higher than the detection frequency, very-narrow-band filtering with passive low-distortion components must be used in the detector. In some cases, testing involves passing currents through or near magnetic materials which respond nonlinearly due to curvature of their B-H loop response. Very large apparent "defect" signals appear at the 455-kHz intermodulation frequency when testing circuits through reed relays or on lines close to ferromagnetic substrates such as Invar, making this tester unsuitable for such applications.

An alternative approach, outlined by Whitley [3] for contact resistance measurements and discussed by Kanno and Minowa [6] in the context of device and contact nonlinearity, involves the application of a drive current consisting of two frequencies, one of which is twice the other. In this case, one of the cubic nonlinear terms appears at dc. (Although not described in terms of two frequencies, the pulse nonlinearity technique described by English et al. [4] is related to this.)

Using dc as a detection signal eliminates the stringent requirement for low intermodulation distortion on the current drive circuitry, since the drive signal can be made completely free of dc current by passing it through a capacitor. Also, the filtering requirements are greatly simplified because the drive frequency is widely separated from the detection frequency. Interestingly, this technique shows virtually no sensitivity to the presence of

ferromagnetic materials. The lack of inductive coupling at dc may be the reason for this.

Difficulties usually associated with measurements at low frequencies, such as $1/f$ noise, thermal emfs, and line-frequency interference, can be minimized using modern digital signal processing techniques. Another apparent drawback of detecting at dc is that there is no opportunity to implement the desired thermal phase selection in the detector circuit. However, as described below, the equivalent selectivity can be achieved by controlling the relative phases of the two drive components [9].

Theory of operation

The present test system uses a phase-adjustable $1f + 2f$ technique combined with digital signal processing to achieve the extra selectivity provided by thermal phase discrimination with greatly simplified electronics. The applied current I is given by

$$I = I_1 \sin \omega t + I_2 \sin(2\omega t + \phi), \quad (1)$$

where $\omega = 2\pi f$ is the angular frequency of the fundamental component of the drive current, and ϕ is the adjustable phase that is applied to the $2f$ component. Typical values for the current amplitudes I_1 and I_2 are 0.8 A and 0.4 A, respectively. From experience, the drive frequency f is 1 MHz, since this has been shown to provide good sensitivity for copper line defects in glass-ceramic substrates.

The role of the phase ϕ becomes clear by considering the two limiting cases for the thermal response of the device under test. In one limit, the temperature follows the power input with no phase lag, while in the other, the temperature lags the power by 90° . Roughly speaking, these limits correspond respectively to the behavior of constriction defects and good lines. In the following discussion, the limiting case of zero phase lag between temperature and power is described in terms of an idealized defect, while that of a 90° phase lag is described as an idealized good line. As is described later, this tester measures both components. Because the phase ϕ is determined absolutely, the reported results correspond directly to these limiting cases.

Physically, an idealized defect is a body with an electrical resistance R_c , a thermal resistance R_{th} to the surrounding conductor, and a negligibly small heat capacity. The temperature excursions of the body relative to the surrounding line are directly in phase with the power and have a peak value equal to the power times the thermal resistance, or $\delta T = I^2 R_c R_{th}$. The overall temperature-driven resistance modulation is given by the average temperature over the constriction, but for the purposes of this discussion the peak value is used. The nonlinearity signal is then $V_d \approx I \alpha R_c \delta T = \alpha I^3 R_c^2 R_{th}$.

Substituting the current given by Equation (1) and keeping only the dc terms, the result is

$$V_d \approx - \left[\frac{3}{4} \alpha R_c^2 R_{th} \right] I_1^2 I_2 \sin \phi. \quad (2)$$

Generally speaking, metals follow the Wiedemann–Franz law, which implies that R_{th} is proportional to R_c . Holm (Reference [2], p. 193) solved the case of self heating of a small spot contact which, in the limit of small peak temperatures, gives $R_{th} = R_c/8LT$, where L is the Lorentz constant ($2.45 \times 10^{-8} \text{ V}^2/\text{K}^2$) and T is the absolute temperature. Thus, the defect signal scales as the cube of the constriction resistance. The peak temperature rise is quite small in most cases. A 1- μm diameter constriction having a resistance of 17 m Ω has a peak temperature swing of only 7°C with 1.2 A of applied current (I_1 and I_2 both at their peak values).

The idealized good line has a finite heat capacity C_{th} and negligible heat loss to the environment (infinite thermal resistance), so a line with electrical resistance R_l undergoes temperature swings of $\delta T = (R_l/C_{th}) \int I^2 dt$, and the nonlinearity signal is $V_l = I \alpha R_l \delta T = (\alpha R_l^2/C_{th}) \int I^2 dt$. Substituting from Equation (1) and keeping only dc terms gives

$$V_l = \left[\frac{3}{8\omega C_{th}} \alpha R_l^2 \right] I_1^2 I_2 \cos \phi. \quad (3)$$

Thus, using $\phi = 270^\circ$ provides maximum positive signal from an idealized defect and nothing from an idealized good line. Conversely, $\phi = 0^\circ$ provides maximum signal from an idealized good line and nothing from an idealized defect. In these equations, the terms in brackets are the quantities that are intrinsic to the device under test, or DUT. Letting D equal the intrinsic defect term from Equation (2) and L equal the intrinsic line term from Equation (3), the observed dc voltage normalized by the drive currents is given by

$$\frac{V}{I_1^2 I_2} = L \cos \phi - D \sin \phi. \quad (4)$$

In real circuitry, good lines do have finite thermal resistances to the substrate and would have both L and D components. Equation (4) shows that one could adjust the phase ϕ to produce zero detected signal for good lines just as Halperin et al. [8] describe, thereby achieving maximum discrimination of latent defect signals in the presence of large line signals. However, for maximum flexibility, the present tester is implemented with a dual-phase measurement capability that provides readout of both the L and D values for each net tested. Phase adjustment for line signal rejection can be done after the fact by mathematically rotating the orthogonal components L and D , using individualized settings for

each net. Using an absolute phase reference based on idealized line and defect signals also eliminates the possibility of drift or misadjustment in an analog phase control circuit.

Signal interpretation

In practice, the geometry of a line constriction comprising a latent defect is not known in advance. In fact, there is no way to uniquely determine the nature of a constriction from the size of its nonlinearity signal. For example, a latent defect such as that shown in Figure 1 could comprise a collection of parallel current paths with various constriction sizes. Latent defect screening, however, is a matter of probabilities, balancing yield loss against potential future failure. High nonlinear conductivity signals are generally indicative of short, narrow constrictions which are mechanically weak and represent a future reliability exposure. A precise physical interpretation is not required, but it is valuable to understand at least roughly what the signals L and D imply about the conductor geometry.

As a first simplification, assume the signal to be due to a single constriction. Next, assume that the signal from the good line is negligible, or that it is known and can be subtracted. Now L and D represent the constriction alone. On the basis of the previous discussion, if D is much greater than L , the constriction must be short, since it appears to be an ideal defect which is thermally shorted to the nearby conductor. In this case the constriction resistance can be estimated from the magnitude of D using Equation (2) together with the expression given above for thermal resistance in terms of R_c . The diameter of the spot contact is then given by $d = \rho/R_c$, where ρ is the electrical resistivity of the metal. If, on the other hand, L is greater than D , the constriction must be relatively long. In this case, not much can be inferred about the line cross section without making additional assumptions.

Thus, the length of the constriction determines the relative magnitudes of L and D . The crossover in behavior from ideal defect to ideal line occurs at a characteristic length that depends on the thermal parameters of the metal. By applying dimensional analysis to the 1D heat-conduction equation, one can show that this characteristic length (where $L = D$) is equal to $\sqrt{\kappa/\omega C}$, where κ is the thermal conductivity of the metal, ω is the angular frequency of the power input, and C is the specific heat. This characteristic length is about 4.2 μm for copper at a frequency of 1 MHz. For glass-ceramic substrates, constrictions that represent a reliability exposure have a cross-section diameter of about a micron or less. The effective length of a spot contact is about the same as its diameter, so relevant constrictions will have $D \gg L$ at 1 MHz, which is desirable.

Test procedure

Capitalizing on the simplicity of the analog circuitry offered by the $1f + 2f$ technique, the tester uses a digital signal processor to control the drive currents and process the detected signals to extract L and D while rejecting low-frequency amplifier drift and thermal electromotive forces. The electronics provide digitally synthesized drive currents which are applied to the DUT through a blocking capacitor and detect the nonlinearity signal using a low-pass filter and amplifier connected to an A/D converter. Another key element is a current probe, also connected to an A/D converter, to sense the applied current at the DUT. Details of the circuitry are described later in this paper.

Each test measurement consists of four electrical bursts, as illustrated in **Figure 3**. The bursts actually contain 150 cycles of the 1-MHz fundamental frequency, but the figure shows only a few cycles in expanded form to illustrate the shape of the waveforms. The current probe signal is sampled by the digital signal processor (DSP) during the first half of each burst, and the filtered nonlinearity signal is sampled and averaged over the second half. A discrete Fourier transform (DFT) is applied to the current probe waveform data to extract the amplitude and phase of the $1f$ and $2f$ components of the current. Dividing the averaged nonlinearity signal by $I_1^2 I_2$ yields the normalized signal [left side of Equation (4)] V_n , while the relative phase ϕ is given by the DFT phase at $2f$ minus twice the DFT phase at $1f$. Accounting for an undetermined offset voltage V_{off} (assumed constant over the set of bursts), the combined readings provide four equations in three unknowns (L , D , V_{off}) where $i = 0, 1, 2, 3$:

$$V_n(i) = L \cos \phi_i - D \sin \phi_i + V_{\text{off}}. \quad (5)$$

Such an over-determined set of equations provides an opportunity to find L , D , and V_{off} by least-squares fitting. This yields the following set of three simultaneous linear equations:

$$\begin{pmatrix} \sum \cos^2 \phi_i & -\sum \cos \phi_i \sin \phi_i & \sum \cos \phi_i \\ -\sum \cos \phi_i \sin \phi_i & \sum \sin^2 \phi_i & -\sum \sin \phi_i \\ \sum \cos \phi_i & -\sum \sin \phi_i & 4 \end{pmatrix} \begin{pmatrix} L \\ D \\ V_{\text{off}} \end{pmatrix} = \begin{pmatrix} \sum V_n(i) \cos \phi_i \\ -\sum V_n(i) \sin \phi_i \\ \sum V_n(i) \end{pmatrix}, \quad (6)$$

where the sums are over the four bursts. These are solved by inverting the 3×3 matrix, or by applying Cramer's rule [10].

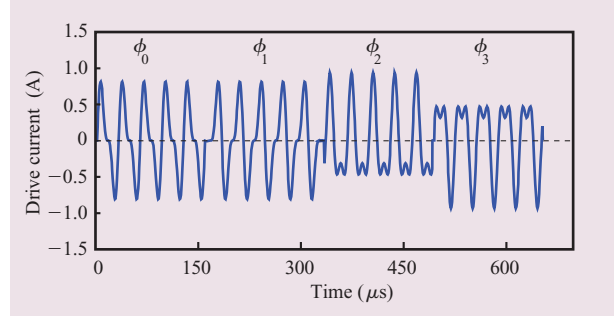


Figure 3

Diagram illustrating the current waveforms used in the four-burst measurement technique. Each burst actually contains 150 cycles of the 1-MHz fundamental frequency. Only a few cycles are included in expanded form to illustrate the shape of the waveforms.

In an early version of this tester, the drive current was applied with phases set to $\phi = 0^\circ, 180^\circ, 270^\circ$, and 90° . The first two bursts resulted in normalized signals of $V_{\text{off}} \pm L$, and the second two gave $V_{\text{off}} \pm D$, so taking half the difference of the pairs of readings provided offset cancellation. (The waveforms shown in Figure 3 actually correspond to these phase choices.) However, the phase of the 2-MHz component relative to 1 MHz is not easy to control over a wide range of load resistance and is also affected by cable inductance that can vary from tester to tester. Therefore, a more general approach utilizing a current probe to determine the actual currents was adopted. For convenience, the same burst generation code was used as in the early-version tester, resulting in digitally synthesized signals that are nominally set up for $0^\circ, 180^\circ, 270^\circ$, and 90° at low load resistance. Now, however, phase shifts at higher loads have no effect on the measurement, since the algorithm directly accounts for the actual phase using the current measurements. The relative burst phases also shift slightly because of differences in the operating characteristics of the output amplifier in different quadrants (note the asymmetry of the waveforms in Figure 3), but these are also automatically compensated. Applying four roughly orthogonal bursts provides approximately equal weighting in the data for L and D and ensures that the determinant of the matrix in Equation (6) will not be zero.

The inclusion of a current probe also allows the drive current to be controlled over a wide range of load resistances. The procedure involves applying a single burst (the first of the usual four bursts, as shown in Figure 3) using a nominal amplitude setting and then computing the I_1 amplitude using the DFT of the current probe signal. The amplitude setting is multiplied by the ratio of the desired I_1 value to the measured I_1 , and another burst is applied. This process is repeated once

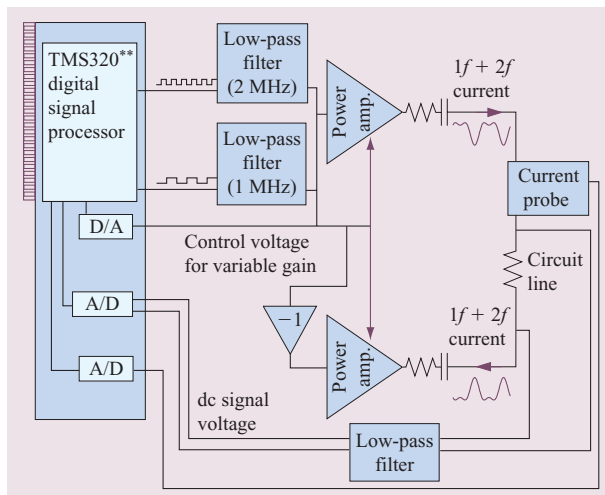


Figure 4

Block diagram of the electronics.



Figure 5

Photograph of the automated test system for latent defect screening.

more to obtain a final amplitude setting, which is then used during the application of the four bursts comprising the nonlinearity signal measurement. The relative amplitude of the I_2 component is set by the hardware circuitry to be about half of I_1 . This ratio can change with load resistance and cable inductance, but the signal from each burst is always normalized by the measured $I_1^2 I_2$, so variations in I_2 have no effect on the absolute scale factor for the L and D readings. In this tester, controlling I_1 keeps the signal-to-noise (S/N) ratio roughly constant and keeps the power dissipation in the DUT at acceptable

levels. In contrast, analog test systems [8] relying on fixed source resistances in the drive circuits to maintain constant current have a rapid drop-off in drive current as the load resistance approaches the source resistance. With a cubic dependency on drive current and no measurement feedback to provide compensation, the normalization factor causes large-scale factor variations in the defect signal.

A significant source of noise in low-frequency electronics is line frequency interference. In this design, the total duration of the bursts is only 0.6 ms, which is much shorter than the 16.7-ms period of 60-Hz interference. As a result, the line frequency signal is almost constant over each group of bursts, which modifies V_{off} but has very little effect on the signals of interest.

The use of short bursts also makes possible extremely rapid testing, an important feature for gang-probed and relay-switched test systems. Including computation and other overhead, a single measurement can be completed in approximately 1 ms. This is much shorter than the move and settle time of the flying probe system, so averages are taken over groups of repeated measurements to improve S/N ratio without significantly affecting test time. The drive-level adjustment is done only once per test site, and the cycle time for just the bursts and computation overhead is about 0.8 ms, so averaging the typical set of 20 measurements takes about 16 ms. Statistical analysis of the set of readings also provides a measure of the variability that can be useful when initial results are suspect and as a diagnostic tool when tracing noise sources, which can include nearby video monitors.

Electrical design

Figure 4 is a block diagram of the combined analog and digital electronics for this tester. The digital portion, shown as the circuit card, is located in a PC which also contains the probe control system. The analog circuitry is contained in a separate box located on the probe tool to minimize the length of cable required to connect to the circuit line being tested (see **Figure 5**).

The drive current is applied using a differential pair of power-operational amplifiers. This configuration cuts the drive voltage in half relative to ground as compared with a single-ended drive, which improves safety and eases the requirements for the voltage rating of the power amplifiers. Peak currents of 1.2 A [$(1f \times 0.8 \text{ A}) + (2f \times 0.4 \text{ A})$] can be driven in lines of up to 40 Ω resistance using a ± 40 -V supply voltage. Series capacitors are included to block dc current through the DUT. The drive circuit has a voltage-controlled gain that is set by the DSP using a D/A converter.

The $1f$ and $2f$ sinusoidal voltage signals are synthesized from transistor-transistor logic (TTL) signals generated

by the DSP. In earlier, lower-frequency prototypes these were generated using computed waveforms and a D/A converter. Modern, high-speed D/A converters are capable of generating the waveforms needed for this 1-MHz tester, but the signal processor card used here did not have that capability. Taking advantage of the insensitivity to harmonic content in the drive current afforded by the $1/f + 2f$ technique, simple low-pass filters are sufficient to convert square waves to sinusoids in this test system. The requisite phase control is very straightforward using TTL signals. The 1-MHz and 2-MHz signals are derived by dividing down from a common 8-MHz clock with variable delays from a shift register providing 90 degrees of relative phase per delay count.

The detected signal is obtained by low-pass filtering the voltage across the DUT. It is critical that the initial stages of filtering use passive, linear components to avoid generating false signals. In this design, four stages of RC filtering (three μs each) are used to provide rejection of the 1-MHz voltage by a factor of about 10^{-5} prior to the first active device which amplifies the signal before digitization at the signal processor card.

Figure 4 also shows the probe which senses the current at the DUT and its connection to the A/D converter at the signal processor card. A Tektronix CT-2 current probe which has a sensitivity of 1 V/A, a very low insertion impedance, good accuracy (3%), and excellent frequency response is used. Accurate current measurement is important, since the current readings are used to normalize the signals. As an example, a 3% inaccuracy causes a 9% error in L and D after the cubic current relationship is considered. Current probes could be individually calibrated and the readings corrected by software, but variations have been found to be substantially less than 3%, so they are currently left uncompensated. Current probe location also affects accuracy. First attempts with probes mounted inside the analog electronics box were inaccurate because the capacitance of the cables provided a path for current that was parallel to the DUT. Locating the current sensor very near one of the contacting probes significantly improved the accuracy of the DUT current measurement.

System integration

Automated testing for latent defects in advanced multichip module (MCM) high-performance glass-ceramic chip carriers is performed as part of the manufacturing process at the IBM East Fishkill, New York, facility. Some MCM designs for high-end microprocessors can support up to 16 chips and require 100% latent defect testing on approximately 9,000 critical signal nets. Latent defect testing is performed using high-speed single-sided two-point flying probe automated test

equipment (ATE), as shown in Figure 5. An industrial computer is used to control the ATE through Industry Standard Architecture (ISA) and peripheral component interconnect (PCI) bus interfaces to sensors, actuators, and mechanical positioning servo controllers, as well as the latent defect measurement electronics (ISA card). Flying probes have the advantage of being data-driven and thus flexible and adaptable to different product designs. The testers are networked to permit access to design-dependent tester data, and storage of test results to a common database. This allows cross-tool verification of defects and simplifies logistics and tracking through the test sector.

The test cycle for a single conductive line consists of commanding the probes to the x - y positions over the network endpoints, engaging the probes, and triggering and acquiring the latent defect measurement signal. The nonlinearity signal is compared to a threshold established to detect signal anomalies which pose a reliability risk. This threshold has been empirically determined on the basis of the occurrence of latent defects found in reliability testing.

A significant fraction of the tested nets connect from the top to the bottom surface of the carrier. Conductive cloth is used to short the surface I/O contacts on the bottom side during tests, which permits testing any two of these top-to-bottom nets together in series from the top surface alone. Defects detected using this series measurement technique are isolated to a specific defective net using different net pairings during subsequent retesting.

The probes are actuated using a closed-loop servomechanism to provide precise programmable probe pressure and to limit impact forces. The vertical probe suspensions are beryllium copper in a double cantilever arrangement designed to provide compliance while minimizing lateral displacement when contact is made. The tips are made from tungsten that has been rounded to a 25- μm diameter, which gives excellent wear resistance and limits damage to the contact pads.

The constricted electrical path associated with imperfect pad-to-probe contact (due to pad contamination, asperities, or other causes) can produce signals which are indistinguishable from signals produced by an actual defect embedded in a conductor. To verify that a detected signal originates from an actual defect in the conductor, re-probing is performed in place or with small x - y displacements from the pad center to validate the signal while changing the location of the probe-to-pad contact.

Contamination due to oxidation on the probe tips or product pads can present formidable contact problems not easily overcome with repeated probing. Applying a 2-MHz ac bias voltage across the probes as they are

brought into contact with the net endpoints (i.e., completing the circuit with the net under test) was found to be an effective means for cleaning the probes. Fortunately, this requires a voltage of only about 7 V rms, which results in no detectable probe tip erosion. The touchdown order of the probes is alternated during actual testing of successive nets to ensure that both probes benefit equally from this cleaning effect.

Other applications

Although flying probe testing offers flexibility, gang probing can be attractive for high-volume applications involving relatively low numbers of circuit lines. In such cases, matrix switching with magnetically activated reed relays is used to select individual nets for testing. As mentioned earlier, the $1f + 2f$ technique has minimal sensitivity to the presence of ferromagnetic materials, making latent defect screening possible in these applications. The use of very brief pulsed measurements is also important for matrix-switched systems, since the overall throughput is determined mainly by test time instead of probe movement time. By using just a single quad burst measurement, test time is reduced to less than a millisecond, which is comparable to the switching times of typical relays.

The use of short bursts also takes advantage of a beneficial scaling relation for S/N ratio, pulse duration, and conductor temperature rise as functions of drive current. Note that in the example above, the single measurement is noisier by a factor of $\sqrt{20}$ compared with the averages of 20 readings that are used today. If the drive current is increased by a factor of 2, the nonlinearity signal increases eightfold, which more than compensates for the loss of averaging (net S/N is 1.8 times better). Unfortunately, raising the drive current leads to higher power dissipation in the circuit line, which may unintentionally damage good lines or those with minor defects that would otherwise pass the screening. However, it is not the instantaneous power that determines whether a line will overheat, but instead the integrated power or total energy dissipated at the defect by the measurement (as mentioned earlier, instantaneous temperature swings are quite small even in very small constrictions). In this example, the power is higher by a factor of 4, but the total duration of the measurement is cut by a factor of 20. The net energy dissipated at the defect is therefore *lower* by a factor of 5. Thus, by raising drive current and cutting burst time, one can simultaneously achieve faster, better, and safer testing! The limit to this scaling comes as the instantaneous temperature at a constriction approaches the melting point of the metal. For a 50-m Ω defect, a temperature of 1000°C requires about 8 A of drive current (see Equation (13.13) on p. 64 of Reference [2]), so there is considerable improvement yet available with this approach.

The simplicity of the $1f + 2f$ technique makes it possible to operate the analog circuitry over a wide range of frequencies. A single system capable of printed circuit board testing in the kHz range and complex ceramic MCM testing at MHz frequencies can be envisioned as an extension of this technique. Prototypes have been built that are capable of operating over a wide range of frequencies (one design spanned the range of 10-kHz to 270-kHz testing) by using D/A-converter-generated waveforms for the $1f$ and $2f$ signals. The low-pass filter for signal detection need only be designed to reject the lowest desired drive frequency. The same filter suffices for all higher drive frequencies. The filter must be implemented with passive components to avoid nonlinearity, so the ability to cover a range of frequencies without modification is critical. The extremely narrow-band filters required for ac harmonic detection systems make them unsuitable for multi-frequency operation.

Unfortunately, working at frequencies much less than 1 MHz prevents one from using the very short bursts that are so desirable for rapid testing (see above). One limitation comes from the fact that a low-pass filter time constant must increase as the corner frequency is lowered; hence, the bursts must be longer to allow the signal to reach full amplitude. Another limitation comes from the need to attenuate line frequency interference. In the 10-kHz to 270-kHz test system described above, the solution to line frequency noise was to integrate each measurement over exactly one line cycle (16.67 ms) to cancel 60-Hz interference and all its harmonics. This is effective, but it increases the quad burst measurement time to nearly 100 ms after allowing for filter settle on each burst and considering other signal integrity overhead. This is comparable to the 300-ms settle time of the 1-kHz tester described in Reference [8], which is limited by similar filtering issues.

Summary

A latent-defect-detection system has been developed for routing production testing of glass-ceramic chip carriers based on the $1f + 2f$ nonlinear conduction technique with the addition of phase selectivity using control of the drive current phase. Digital control of the test process and advanced signal processing algorithms were used to provide significant practical advantages such as greatly simplified analog circuitry, accurate and stable calibration with no adjustments, no load-resistance dependence, and very short measurement times.

Acknowledgments

The authors thank A. Halperin for sharing his expertise regarding nonlinear conductivity measurements and for contributing to the development of the production-level test tool described here. The authors gratefully

acknowledge Tony Sucato for his assistance in the system integration and debug of the automated latent defect tester.

References

1. H. Liu and T. A. Wassick, "Identification of Latent Defects in Advanced Glass Ceramic MCM Packaging," *Microelectronic Failure Analysis*, Electronic Device Failure Analysis Society (EDFAS), an affiliate of ASM International, Materials Park, OH, November 2004, p. 163.
2. R. Holm, *Electrical Contacts, Theory and Applications*, Fourth Edition, Springer-Verlag, Berlin, 1967, p. 16.
3. J. A. Whitley, "Method and Means for Measuring Constriction Resistance Based on Nonlinearity," U.S. Patent 3,500,188, March 10, 1970.
4. A. T. English, G. L. Miller, D. A. H. Robinson, L. V. Dodd, and T. Chynoweth, "Pulse Nonlinearity Measurements on Thin Conducting Films," *J. Appl. Phys.* **49**, No. 2, 717–722 (1978).
5. T. H. DiStefano and A. Halperin, "Nonlinearity Detection Using Fault-Generated Second Harmonic," U.S. Patent 4,496,900, January 29, 1985.
6. M. Kanno and I. Minowa, "Application of Nonlinearity Measuring Method Using Two Frequencies to Electrical Components," *IEEE Trans. Instrum. & Meas.* **IM-34**, 590–593 (1985).
7. T. H. DiStefano, A. E. Falls, A. Halperin, and J. D. Mackay, "Defect Detection Using Intermodulation Signals," U.S. Patent 4,868,506, September 19, 1989.
8. A. Halperin, T. H. DiStefano, and S. Chiang, "Latent Open Defect Detection Using Phase-Sensitive Nonlinearity Detection Technique," *IEEE Trans. Comp., Pkg., & Manuf. Tech. B* **18**, No. 2, 358–365 (1995).
9. E. J. Yarmchuk, "Electrical Defector Detector for Circuit Lines," U.S. Patent 5,469,051, 1995.
10. D. Zwillinger, Ed., *CRC Standard Mathematical Tables and Formulae*, Thirtieth Edition, CRC Press, Boca Raton, FL, 1996, p. 135.

Received September 16, 2004; accepted for publication February 16, 2005; Internet publication August 12, 2005

Edward J. Yarmchuk *IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (yarmchuk@us.ibm.com)*. Dr. Yarmchuk is a Research Staff Member at the Thomas J. Watson Research Center. He received B.A. and Ph.D. degrees in physics from Rutgers University in 1974 and 1978, respectively. Following a postdoctoral appointment at the University of California in Berkeley working in low-temperature physics, in 1980 he joined IBM at the Thomas J. Watson Research Center, where he has worked on printing, storage, and applications of measurement techniques to test and process control. He has received two IBM Outstanding Innovation Awards and three IBM Outstanding Technical Achievement Awards for his contributions to these various fields. Dr. Yarmchuk is an author or coauthor of 51 patents and 13 technical papers. He is a Senior Member of the Institute of Electrical and Electronics Engineers and a member of the American Physical Society.

Christopher W. Cline *IBM Systems and Technology Group, Mid Hudson Research Park, 1580 Route 52, East Fishkill, New York 12533 (cline@us.ibm.com)*. Mr. Cline is an Advisory Engineer in the MLC Test Engineering Department at the IBM facility in East Fishkill, New York. He received a B.S. degree in 1983 and an M.S. degree in 1984, both in mechanical engineering, from Columbia University. Mr. Cline is the author or coauthor of two patents and four technical papers.

Dominic C. Bruen *IBM Systems and Technology Group, Mid Hudson Research Park, 1580 Route 52, East Fishkill, New York 12533 (bruen@us.ibm.com)*. Mr. Bruen is an Associate Engineer in the MLC Test Engineering Department at the IBM facility in East Fishkill, New York. He received an A.A.S. degree in electrical technology from DeVry Institute of Technology in 1990, a B.A. degree in computer science from Mount Saint Mary College in 1995, and an M.S. degree in computer science from Marist College in 2000.