# *ANALOGUE MIXED SIGNAL TEST DEVELOPMENT – A CONTINUOUS IMPROVEMENT EXERCISE IN QUALITY, RELIABILITY AND REDUCTION IN TEST COST OF SEMICONDUCTOR DEVICES*

*Peter Sarson*

*Division of Electronics and Informatics*

*Faculty of Science and Technology*

*Gunma University*

*This dissertation is submitted for the degree of Doctor of Philosophy*

*December 2018*

*To my wife, Bronwen, and my children, Penelope and Amelia.*

*There is no elevator to success; you have to take the stairs*

# DECLARATION

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

Signed:_____

Date:_____

Peter Sarson PH.D SMIEEE CMgr MCMI

Swindon, United Kingdom

# ABSTRACT

This Ph.D dissertation entitled "Analogue Mixed Signal Test Development – A Continuous Improvement Exercise in Quality, Reliability and Reduction in Test Cost of Semiconductor Devices" by Mr. Peter Garry Sarson, BEng (Hons) CMgr MCMI SMIEEE will discuss how an automotive grade semiconductor device production test can both be improved from a test coverage and cost standpoint. Therefore ensuring both reliability and quality of the device, but also guaranteeing all specifications that normally would not be possible with lower end automated test equipment.

It will be shown experimentally how the testing of nonvolatile memories were improved such that in 5 years' worth of high volume manufacturing, approximately 50 million devices, that not a single return was received from the field, due to a device failure, which was the result of the superior test solutions that were developed to test such devices.

It will also be shown that by exploiting the features of an automated test equipment instruments that by understanding the sampling theorem that a very efficient method can be developed for testing analog filters and similar types of circuitry such as amplifiers. Similarly by exploiting digital signal processing techniques, it will be shown how the performance of the ATE instruments can also be enhanced compared to how they are actually specified. Thus allowing difficult specification parameters that were not possible to be measured with a lower cost instrument would now become feasible to test, therefore reducing the need for more expensive test instruments and hence reducing cost of the overall test system.

# ACKNOWLEDGEMENTS

First, I would sincerely like to thank Prof. Haruo Kobayashi for his encouragement to actually apply for my Ph.D. with Gunma University based on the research I have conducted as part of my career over the last 18 years. Without his guidance in all matters relating to journal and conference papers, this Ph.D. dissertation would not have been possible.

I would really like to thank my wife for her support and understanding. For several months, I spent weekends and evenings writing and re-writing my papers and my dissertation and without her support, this would not have been possible,

Thanks must also go to friends and colleagues at ams AG. Specifically Markus Wuchse, Brian Davies, Gregor Schatzberger, Fritz Leisenberger, Thomas Riener, Rainer Minixhofer, Thomas Moerth, and Herwig Klimesch for their help and support. Without their understanding that I had undertaken this endeavor, I could not have completed the required work in the shortest amount of time.

I would like to thank Prof. Haruo Kobayashi for his help and patience in getting me through this endeavor. I would also like to thank the PHD committee, Prof. Kazumasa TAKADA, Prof. Kuniyuki MOTOJIMA, Prof. Hitoshi AOKI, Prof. Koji ASAMI for reviewing my work and helping me complete this dissertation, without their help this would not have been achievable.

Last but certainly not least, I would like to show my appreciation to both Lenny Leon, Ian Harrison and Derrick Dupre at Xcerra Corporation (formally LTX) who acted somewhat like my mentors when I first came into the ATE industry. Lenny Leon originally gave me the idea for the chirp that I have developed to what it is today.

Finally, I would like to thank the discoverer of coffee, without which nothing would be possible.

# CURRICULUM VITA

| | |
|---|---|
| 1998, July23rd | BEng (Hons) Electrical and Electronic Engineering with specialisation in Communications, University of Sheffield, United Kingdom. |
| 2000, Aug 1st | Senior RF Applications Engineer, LTX Corporation, United Kingdom, USA, Taiwan. |
| 2003, Mar 1st | Chartered Engineer, Institute of Electrical Engineers |
| 2006, Dec 1st | Senior Test Development Engineer, SiGe Semiconductor, United Kingdom |
| 2007, Jan 1st | Principal Test Development Engineer, Frontier Silicon, United Kingdom |
| 2008, Nov 1st | Test Development Manager, ams AG, Austria |
| 2014, May 1st | Chartered Manager, Chartered Manager Institute |
| 2016, Mar 1st | Senior Member, Institute of Electrical and Electronic Engineers |
| 2018, Feb 1st | Test Engineering Manager, Dialog Semiconductor, United Kingdom |

# PUBLICATIONS

Journal Papers

[1] Peter Sarson; Haruo Kobayashi, "Using Distortion Shaping Technique to Equalize ADC THD Performance Between ATEs", Journal of Electronic Testing: Theory and Applications, Springer (2017), DOI : 10.1007/s10836-016-5630-0

[2] Peter Sarson, G. Schatzberger, F. Leisenberger, "Fast Bit Screening of Automotive Grade EEPROMs - Continuous Improvement Exercise", IEEE Transactions on VLSI (2017), DOI : 10.1109/TVLSI.2016.2634589

[3] Peter Sarson, "An ATE Filter Characterization ToolKit Using a Discrete Chirped Excitation Signal as Stimulus", Journal of Electronic Testing: Theory and Applications, Springer (2017), DOI : 10.1007/s10836-016-5633-x

[4] Peter Sarson; Tomonori Yanagida; Shohei Shibuya; Kosuke Machida; Haruo Kobayashi, "A Distortion Shaping Technique to Equalize Intermodulation Distortion Performance of Interpolating Arbitrary Waveform Generators in Automated Test Equipment", Journal of Electronic Testing: Theory and Applications, Springer (2018), DOI : 10.1007/ s10836-017-5700-y

[5] Peter Sarson, "Automotive Grade Screening Procedure for Anti-Fuse Gate-Oxide Non-Volatile Memories", IEEE Design & Test.

[6] Peter Sarson; Tomonori Yanagida; Kosuke Machida, "Measuring Group Delay of Frequency Downconverter Devices using a Chirped RF Modulated Signal", Journal of Electronic Testing: Theory and Applications, Springer (2018), DOI : 10.1007/s10836-018-5708-y

International Conferences

[7] P. Sarson, G. Schatzberger, R. Seitz, "Automotive EEPROM Qualification and Cost Optimization", IEEE 22nd Asian Test Symposium, pp. 105-106, Yilan, Taiwan (2013).

[8] Peter Sarson, "Analog DfT & Test : implementing analog Iddq", 2014 19th IEEE European Test Symposium (ETS), ETS2 – Special Track on Emerging Test Solutions

[9] (Invited) Peter Sarson, "RF Filter Characterization Using a Chirp" , IEEE 9th International Design & Test Symposium, Algiers, Algeria (Dec. 2014).

[10] Peter Sarson, "ACR BER correlation to ATE for a COFDM VHF RX", 20th International Mixed-Signal Testing Workshop, pp.1 - 4, Paris, France (2015).

[11] Peter Sarson, "Very Low Supply Voltage Room Temperature Test to Screen Low Temperature Soft Blown Fuse Fails Which Result in a Resistive Bridge", 17th International Symposium on Quality Electronic Design, pp.135 - 139, Santa Clara, CA (May 2016).

[12] G. Schatzberger, F. P Leisenberger, P. Sarson, “Yield Improvement of an EEPROM for Automotive Applications While Maintaining High Reliability”, IEEE 34th VLSI Test Symposium, pp. 1-6, Las Vegas, NV (April 2016).

[13] Peter Sarson, Automotive Session: “Test Development : Getting to 0ppm,” presented at VLSI Test Symp., April 2016.

[14] Peter Sarson, “Group Delay Filter Measurement Using a Chirp”, IEEE 21st European Test Symposium, pp.1 – 2, Amsterdam, The Netherlands (May, 2016)

[15] (Invited) Peter Sarson, Haruo Kobayashi, “Using Distortion Shaping Technique to Equalize ADC THD Performance Between ATEs”, IEEE 21st International Mixed-Signal Testing Workshop, Pages: 1 - 6, Cataluya, Spain (July 2016).

[16] Peter Sarson, “Test Time Efficient Group Delay Filter Characterization Technique Using a Discrete Chirped Excitation Signal”, IEEE 47th International Test Conference, Fort Worth, TX (Nov. 2016).

[17] CK, Hsu, P. Sarson, G. Schatzberger, F. Leisenberger, J. Carulli, S. Siddhartha, KT Cheng, “Variation and Failure Characterization Through Pattern Classification of Test Data From Multiple Test Stages”, IEEE 47th International Test Conference, Fort Worth, TX (Nov. 2016)

[18] Peter Sarson, Shohei Shibuya, Tomonori Yanagida, Haruo Kobayashi, "A Technique for Dynamic Range Improvement of Intermodulation Distortion Products for an Interpolating DAC-based Arbitrary Waveform Generator Using a Phase Switching Algorithm", IEEE 35th VLSI Test Symposium, Las Vegas, NV (April 2017)

[19] Peter Sarson, “Automotive Alternative Test”, 2017 35th IEEE VLSI Test Symposium (VTS), IP Session: Automotive Test Solutions.

[20] Bagalini, E., Sini, J., Sonza Reorda, M., Violante, M., Klimesch, H., Sarson, P. “An automatic approach to perform the verification of hardware designs according to the ISO26262 functional safety standard”. In Test Symposium (LATS), 2017 18th Latin-American. IEEE.

[21] J. Mikulic, P. Sarson, G. Schatzberger, and A. Baric, “Fault-Based Test Methodology for Analog Amplifier Circuits,” IEEE East-West Design & Test Symposium (EWDTS), Oct 2017.

[22] P. Sarson, J Rearick, “Use Models for Extending IEEE 1687 to Analog Test”, IEEE 48th International Test Conference, Fort Worth, TX (Nov. 2017)

[23] S. Sunter, P. Sarson, “A/MS Benchmark Circuits for Comparing Fault Simulation, DFT, and Test Generation Methods”, IEEE 48th International Test Conference, Fort Worth, TX (Nov. 2017)

[24] C. Xanthopoulos, P. Sarson, H. Reiter, Y. Makris, “Automated Die Inking: A Pattern Recognition-Based Approach”, IEEE 48th International Test Conference, Fort Worth, TX (Nov. 2017)

[25] Peter Sarson; Tomonori Yanagida; Shohei Shibuya; Kosuke Machida, “Group Delay Measurement of Frequency Downconverter Devices Using Chirped RF Modulated Signal” , IEEE 36th VLSI Test Symposium, San Francisco, CA (April 2018)

[26] G. Schatzberger, F. P Leisenberger, P. Sarson A, Weisner, “High efficient low cost EEPROMs screening method in combination with an area optimized byte replacement startegy which enables high reliabilty EEPROMs”, IEEE 36th VLSI Test Symposium, San Francisco, CA (April 2018)

[27] Yanjun Li, Ender Yilmaz, Peter Sarson, Sule Ozev, “Online Information Utility Assessment for Per-Device Adaptive Test Flow”, IEEE 36th VLSI Test Symposium, San Francisco, CA (April 2018)

Filled Patents

[28] G. Schatzberger, F. P Leisenberger, P. Sarson, “High Efficient Low Cost EEPROM Screening Which Enables Area Optimized Dynamic Byte Error Correction”, European Patent Number, 17158340.4

Lectures Given

[29] Peter Sarson, Test Development, Arizona State University, April 2018

[30] Lunch & Learn, Test Engineering Overview, Dialog Semiconductor, Sept 2018

[31] Peter Sarson, Analogue Test Engineering, Politecnico Turino, Dec 2018

# CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS AND ACRONYMS

| | | |
|---|---|---|
| ADC | - | Analogue to Digital Converter |
| AIAG | - | Automotive Industry Action Group |
| AIMI | - | Analogue Infant Mortality Inducing |
| APQP | - | Advanced Product Quality Planning |
| ATB | - | Analogue Test Bus |
| ATE | - | Automated Test Equipment |
| ATPG | - | Automated Test Pattern Generation |
| AWG | - | Arbitrary Waveform Generator |
| BB | - | BaseBand |
| BIF | - | Built-In Function |
| BIST | - | Built-In Self-Test |
| BOST | - | Built-Out Self Test |
| BPF | - | Band Pass Filter |
| CMOS | - | Complementary Metal Oxide Semiconductor |
| CPK | - | Process Capability |
| CUT | - | Circuit Under Test |
| DAC | - | Digital to Analogue Converter |
| DIB | - | Device Interface Board |
| DAB | - | Digital Audio Broadcast |
| DC | - | Direct Current |
| DFT | - | Design For Test |
| DFT | - | Discrete Fourier Transform |
| DNL | - | Differential Non-Linearity |
| DPAT | - | Dynamic Part Averaging Testing |
| DPPB | - | Defect Parts Per Billion |

| | | |
|---|---|---|
| DPPM | - | Defect Parts Per Million |
| DOT | - | Defect Orientated Test |
| DSB-SC | - | Double Sideband, Suppressed Carrier |
| DSP | - | Digital Signal Processing |
| DUT | - | Device Under Test |
| EDA | - | Electronic Design Automation |
| EEPROM | - | Electrical Erasable Read Only Memory |
| EEHVOX | - | EEPROM High Voltage Oxide |
| ENOB | - | Effective Number of Bits |
| EVM | - | Error Vector Magnitude |
| FG | - | Floating Gate |
| FFT | - | Fast Fourier Transform |
| FMEA | - | Failure Mode Effect Analysis |
| FSR | - | Full Scale Range |
| GPS | - | Global Positioning by Satellite |
| GRR | - | Gauge Repeatability and Reproducibility |
| HF | - | High Frequency |
| HV | - | High Voltage |
| HPF | - | High Pass Filter |
| IATF | - | International Automotive Task Force |
| IC | - | Integrated Circuit |
| ICL | - | Instrument Control Language |
| IEEE | - | Institute of Electrical and Electronic Engineers |
| IF | - | Intermediate Frequency |
| IMD | - | Intermodulation Distortion |
| IO | - | Input Output |
| IoT | - | Internet of Things |

| | | |
|---|---|---|
| IDM | - | Integrated Device Manufacture |
| IP | - | Intellectual Property |
| IPOX | - | Inter Poly Oxide |
| ISI | - | Inter Symbol Interference |
| JTAG | - | Joint Test Action Group |
| LNA | - | Low Noise Amplifier |
| INL | - | Integral non-linearity |
| LPF | - | Low Pass Filter |
| IDM | - | Independent Device Manufacture |
| LO | - | Local Oscillator |
| LSB | - | Least Significant Bit |
| LSI | - | Large Scale Integration |
| LSL | - | Lower Specification Limit |
| MIMO | - | Multiple-Input Multiple-Output |
| MRB | - | Material Review Board |
| MSB | - | Most Significant Bit |
| NBTI | - | Negative Bias Temperature Instability |
| NCO | - | Numerically Controlled Oscillator |
| NFC | - | Near Field Communications |
| NMOS | - | N-type Metal Oxide Semiconductor |
| NVM | - | Non-Volatile Memory |
| OS | - | Operating System |
| OTP | - | One Time Programmable |
| PAT | - | Part Average Testing |
| PDL | - | Procedural Description Language |
| PMOS | - | P-type Metal Oxide Semiconductor |
| PPAP | - | Production part approval process |

| | | |
|---|---|---|
| PPB | - | Parts per Billion |
| PPM | - | Parts Per Million |
| PPT | - | Parts Per Trillion |
| PSW | - | Part Submission Warrant |
| QA | - | Quality Assurance |
| QAM | - | Quadrature Amplitude Modulation |
| QBD | - | Charge to Breakdown |
| QPSK | - | Quadrature Phase Shift Keying |
| RADAR | - | Radio Detection and Ranging |
| R&D | - | Research and Development |
| RAM | - | Random Access Memory |
| RMA | - | Return Material Advice |
| RMS | - | Root Mean Square |
| RF | - | Radio Frequency |
| RFSOC | - | RF System on a Chip |
| RX | - | Receiver |
| ROM | - | Read Only Memory |
| SAR | - | Synthetic Aperture RADAR |
| SAR | - | Successive Approximation Register |
| SFDR | - | Spurious Free Dynamic Range |
| SINAD | - | Signal to Noise and Distortion |
| SLM | - | Scribe Line Monitor |
| SNR | - | Signal to Noise Ratio |
| SPI | - | Serial Peripheral Interface |
| SOC | - | System on a Chip |
| SPST | - | Single Pole Single Throw |
| SPDT | - | Single Pole Double Throw |

| | | |
|---|---|---|
| TEM | - | Transmission Electron Microscope |
| TETRA | - | Terrestrial Trunked Radio |
| TCXO | - | Temperature Controlled Crystal Oscillator |
| TG | - | Transmission Gate |
| THD | - | Total Harmonic Distortion |
| ToF | - | Time of Flight |
| TSP | - | Time Stretched Pulse |
| TUX | - | Tunnel Oxide |
| TX | - | Transmitter |
| uP | - | Microprocessor |
| UHF | - | Ultra High Frequency |
| USL | - | Upper Specification Limits |
| UTP | - | Unit Test Period |
| UWB | - | Ultra Wideband |
| VGA | - | Variable Gain Amplifier |
| VHF | - | Very High Frequency |
| VLV | - | Very Low Voltage |
| WAT | - | Wafer Acceptance Test |
| WRT | - | With Respect to |

# List of Appendices

# 1 INTRODUCTION

The semiconductor business is at a junction where the required quality level of products is increasing at an exponential rate, whereas the test cost per device is being driven lower. More stringent quality requirements are the result of the concentration on market penetration for one product rather than multiple different products and hence the volume of any one-semiconductor device increases at the same rate as the acceptance of individual products into the market. A simple example of this is the cell phone market. 10 years ago, Nokia had several tens of mobile phones for sale at any one time and for each of these products there could have been several different suppliers of the same device function for each model of phone. In today's world, the old guard of manufacturers has failed to adapt to changing market conditions and the available choices to the consumer is a simple one between Apple and Samsung. As the market is growing at 5 % year on year, and as the market segment has shrunk to two models of phone it is obvious that the volume of the equivalent devices that once went into the Nokia has had a more than a 10-fold increase. Whilst quality levels have actually decreased in some cases, more returns have been received from end customers, for analysis by the semiconductor providers. Therefore, without significantly increasing quality assurance personnel, semiconductor providers would be at a loss to deal with this changing environment.

One way of dealing with the need to drive down the dppm in the field is by introducing more stringent testing at the backend of the semiconductor manufacturing flow. In the automotive business segment, it has long been the requirement to test every specification detailed in the datasheet at the test stage, together with the need to launch a

product into production in a safe launch mode where the product is tested at the temperature extremes with a characterisation style test program. However, this is extremely costly in terms of test time and the delivery times of such products are very long due to the throughput rates. In recent years, a trend has emerged where such automotive style backend test flows, are being demanded by the high volume device consumers simply due to the volume they are expecting. From an automotive point of view, the need is due to safety concerns. From a mobile phone manufacturer's viewpoint, if the product has a quality level of 100 ppm and 1 billion devices are expected per year, this would lead to a return rate of 100 thousand units per year. Simply put, this would be prohibitively expensive for the end consumer in rework and would require a Quality Assurance (QA) department of a company to be enormous to deal with so many returns per year.

In contrast with the above, there is the added pressure on test development to significantly decrease test times of products because of the overall cost of semiconductor devices. This is due to Moore's law which states that every 18 months there will be a reduction in process geometry size of 50 %, thus leading to the ability to produce twice as many parts one year compared to the last resulting in significantly cheaper semiconductor manufacturing year on year. In digital test development there has been somewhat of a drive to keep up with Moore's law in that more defects are being found as a result of more and more efficient ways of capturing them, using such test technique as scan, IDDq, cell aware etc. Unfortunately, in the analogue space, no such progression has been made and as such analogue mixed-signal test development is still in its infancy; therefore, with the pressures of the modern world cracks are becoming apparent. As process nodes shrink, it is becoming increasingly difficult to predict the type of random defects that have been standard within the semiconductor business using larger geometries. Hence increasing the pressure for the backend to find such issues efficiently and cost effectively. However, this is in stark contrast to reality.

As the world of IoT starts to develop, from 2018 to 2023, the number of analogue devices shipped worldwide is predicted to grow by some 11 %; this is estimated to be approximately 250 billion devices. This will further increase the pressure to decrease test time from a throughput point of view. There is only so much space for testers without building new facilities or outsourcing to external subcontractor manufacturers. However, unless a company wishes to invest heavily in quality assurance personnel, the

quality level of these parts needs to increase at the same percentage as the growth otherwise the level of returns coming from the field will simply be overwhelming. For a small company, this could be the difference between the success and failure of the venture.

The silicon foundries around the world are at a loss as to further decrease the number of defects per square millimetre, therefore the only option left to increase the quality level is somehow to force the latent defects to be 100 % detectable. There has been a lot of effort given to producing devices in the massive parallel testing space but this is only feasible for devices that are constantly running production, as the setup times of such systems are too high for medium volume products. Simply stated, it would take longer to setup the system; than it would take to test the amount of required devices.

What is needed is a way of testing devices that captures a specification based failure but at the same time is able to detect a latent defect that would fail in the field. For example, if such a device was in a safety critical system within an autonomous vehicle it could potentially kill the passenger.

This is a momentous task and it cannot be achieved solely from one element of the semiconductor process flow but needs a combined effort from all sides of the process. There has been much work done in both academia and industry to tackle these issues but usually only one aspect is addressed at any one time, and as a whole, the solutions that are produced do not work for similar companies, as the solutions may not take into account everything earlier or later in the process flow. This also leads to the lack of adoption of such solutions, or even, the will to look at these issues as the solution is simply not transparent.

Many areas can be addressed, in both design and the test world to help drive this process forward. However, a certain arrogance from design has been observed, it can be seen that it is believed that test is a simple discipline, and design is more important. This type of thinking is very common in the industry and can be observed simply by looking at the test time of a product. It can be seen that there is a strong correlation between a device that had a very strong collaboration and an efficient backend test process. However this requires to go one step further; the way a device is designed or even laid out can influence whether a certain defect can be detected or not. In many cases, analogue fault coverage can be as low as 40 % of the possible defects that can occur. However as the industry states that there is 100 % coverage because specifications are

measured, this has the effect of causing a false belief that the shipped devices exhibit a high quality level. The fact that all companies receive parts returned from the field at a certain defect parts per million rate indicates that there is no such thing as 100 % tested.

Any solution to any one part of the described problem above would only be temporary until some form of automated analogue fault simulators and analogue scan insertion tools are available. The author believes that specification based testing will, in one form or other, always be installed in an analogue mixed-signal test program; however at the moment, for almost 100 % of cases, these tests are handcrafted and take much time to code. Most handcrafted solutions do not use the ATE in the best possible way and hence there is much overhead that can be removed if some form of automation is used. In addition, the tester specifications can also be improved such that difficult specifications can be tested with a larger margin if the users would simply invest some time into improving the test solutions. However, it is in the ATE industries' interests to sell more expensive test options for those more difficult to test specifications. It is also not in the ATE industries interest to spend time to write software that produces automated test program generation. This would not enable them to sell more of their product but less as this would result in code that is more efficient, resulting in faster test times. The ATE industry gives the user the possibility to produce such code but does not make it an easy task. The ATE industry is concentrating more on the highly parallel testing methodologies as it enables them to sell higher end handling capabilities, however as described above this is only of interest to very high volume products.

During the authors career in the test development business he has strived to deliver superior device throughputs with the best yield and quality possible. There have been many instances that the author has been faced with poor design that resulted not only in poor yield but the high possibility of delivering devices that did not fulfil specifications to a customer simply due to design or process related issues. The author spent numerous hours and effort to deliver solutions that did not only guarantee that the delivered parts where 100 % in accordance with the devices specifications but also would be reliable in the field. The author has produced much work on reliability testing of semiconductors using high voltage and low voltage stress to create conditions within the semiconductor devices that would cause a latent defect that ordinarily would not be visible to show itself hence avoiding the shipment of the part into the field. The author has also looked at the system architectures of automated test equipment and identified several areas

where they are not used in the best way possible and as such found ways to produce superior throughputs and test to more stringent specifications by simple use of digital signal processing. This dissertation will build upon four topic areas he has worked on and published at international conferences and in journals. This dissertation will bring all four topics together and demonstrate at a system level how an analogue mixed signal device can be tested with the best throughput possible using automated test equipment but with the least amount of devices exhibiting latent defects hence ensuring the greatest quality and reliability shipped into the field.

## 1.1 Dissertation Overview

Appendix 1 shows a typical RF analogue mixed signal device, it has a digital communication block and the RF block. In this case, this RF receiver has an OTP and EEPROM. The end application of this device could very easily be a GPS or digital radio receiver that would be used in the automotive market.

The OTP would be used for trimming the device such that the performance of the device could be fine-tuned at the production test stage. This enables higher yields as the Gaussian distribution of the process is removed. Parameters such as the oscillator frequency, bandgap voltage, ADC reference voltage, and bias currents are some of a select few parameters that are subject to trimming. The trim codes would be blown into the OTP and hence the device performance is permanently changed. If programming of the OTP was such that it was possible that a bit could go from being seen as programmed to un-programmed could result in big issues for the device performance and potentially the safety of the system.

The EEPROM would be used for mission mode features of the device such as software code that needs to be loaded into the microprocessor at run time. This is usually referred to as firmware, an EEPROM is used so that the end supplier of the product can modify and update the software as and when they require doing so. If an EEPROM suddenly became un-programmable and no field update could be applied, such as deep space applications, would lead to the potential failure of the entire system if a critical system update could not be performed.

In general, the production test of a 16 bit ADC is pretty straightforward when testing integral and differential non-linearity however if noise style parameters are needed then problems can occur when trying to test parameters close to performance limit of a tester.

Usually the only option is to go to higher performance test instrument but this leads to expensive test costs.

On the other hand, testing of filters in production is a straightforward task for even the most stringent of parameters however; the challenge is do so in a test time efficient manner.

The dissertation is set out in the following way; Chapter 2 is a mini tutorial on testing of LSI devices in the 21st century, Chapter 3 & Chapter 4 will be dedicated to the efficient testing of analogue mixed signal devices such as BB and RF filters for such parametric tests such as 3 dB point and group delay. Chapter 5 will look at improving the performance of an AWG in terms of harmonics, and Chapter 6 will look at improving the two-tone intermodulation distortion performance of an AWG. Chapter 7 & Chapter 8 will concentrate on yield and quality improvements of NVM memories such an anti-fuse OTP and on EEPROM. Chapter 9 will discuss what the future holds for the testing of LSI devices in the next 10 years. Chapter 10 will conclude this dissertation.

## 1.2 Chapter 2 Abstract

This chapter gives a mini tutorial on testing of LSI devices. A description of an LSI tester is given as well as how the test equipment is used to interface a device to the tester. A description of how a device has design-for-test features added to the devices circuitry so that testing the part can be made more efficient and cost effective. An in-depth overview is given on how analogue-to-digital converters are tested.

## 1.3 Chapter 3 Abstract

This chapter describes how to measure a full suite of RF filter characteristics using reusable code to produce a chirp that sweeps the filters' frequency range. It shows how to measure 3 dB points, bandwidth and group delay of any filter and how accurately the technique correlates to measurements of the silicon performance made in the lab. The test time saving and stability of results are shown as well as the advantages of the technique with regard to having full characterisation data available in a production program. Historically test engineers have had the problem to be asked to characterize certain parts of the device during debug of initial silicon. This adds time pressure to an already time critical sections of a product delivery milestone because device

characterisation on ATE is never part of the original test quotation. This also adds considerable cost pressure since this takes extra software development and tester time to produce. The proposed method can alleviate a part of this problem.

## 1.4 Chapter 4 Abstract

This chapter discusses the implementation possibilities for making Group Delay measurements of RF frequency converting devices using a standard RF semiconductor Automated Test Equipment that cannot be done using the standard S-parameter measurement due to the difference in frequency from the input to the output of the device. We discuss how using a chirp waveform modulating an RF generator can be used to sweep the frequency response of a RF frequency-converting device and how to produce such a modulation waveform in digital signal processing. We discuss how to implement a group delay test based on our previous work in the baseband domain and how to understand the errors pertaining to measuring a Radio Frequency Receiver. The measurement of the Group Delay of an RF front-end filter and post down convert IF filter will be demonstrated. We discuss how to produce and maintain a stable frequency reference so that any down converted signal would be a true representation of the modulation signal applied to the RF Source and not corrupted by Phase Noise. We discuss how to implement a group delay test based on our previous work in the baseband domain and how to understand the errors pertaining to measuring a radio frequency receiver. The measurement of the group delay of an RF mixer and pre and post down convert RF/IF filters will be demonstrated.

## 1.5 Chapter 5 Abstract

This Chapter describes how using a phase switching technique can produce a low distortion signal from an Arbitrary Waveform Generator (AWG), and how this technique aligns the performance of the AWGs between LSI testers, to evaluate the total harmonic distortion (THD) performance of analogue-to-digital converters (ADCs). Once a device has been characterized, and correlated to the bench, the test engineer needs to start the release procedure in getting the device into a production ready state. One major issue that a test engineer faces is the difference in ADC THD performance test results using the same automated test equipment (ATE) manufacture testers. This paper will then show how the gauge repeatability and reproducibility (GRR) between testers can be produced more easily allowing less stringent guard-bands to guarantee the

performance of those devices that have performance criteria close to the device specification. This work will also go some way to proving previous papers' works on distortion shaping testing to enhance the spectral performance of arbitrary waveform generators.

## 1.6 Chapter 6 Abstract

This chapter describes a phase switching algorithm for Interpolating Digital-to-Analog Converter (DAC) based Arbitrary Waveform Generator (AWG). This confirms a previous exercise that was made by experiment with different Intermodulation Distortion (IMD) suppression techniques and starting phase shifts to suppress IMD tones of the AWG with the interpolating DAC. We show that the poor performance of the AWG can be improved by using the phase switching algorithm over the installed base of a company's tester platform. It is also shown that the IMD performance of AWGs across a company's tester installed base can be equalized, and how it can be achieved using the phase switching technique. We describe how the IMD specifications of the instrument are much worse than those actually measured, and by using phase switching, better performance can be achieved than what would be possible under normal conditions. We present how this technique allows the use of a low-cost tester resource to test IMD products with a higher dynamic range than what was previously possible.

## 1.7 Chapter 7 Abstract

This chapter presents a method for screening resistive bridges that are the side effect of an incomplete burning cycle of a gate oxide anti-fuse architecture one-time programmable (OTP) memory has been developed. This paper shows a correlation of OTP fuse failures that occur at low temperature, to very low supply voltage test at room temperature. Multiple wafer-maps are presented showing the effects of the low temperature and the effect of screening at very low voltage on both good material and material exhibiting a step field issue causing memory failures. This is further developed to demonstrate how the neighbour dice that would need manual inking are caught with the very low voltage test.

## 1.8 Chapter 8 Abstract

This chapter presents an optimization of an existing EEPROM production test flow by means of thorough analysis of the faulty dice and the test flow, which leads to an increase in the yield, a significant decrease in test time, and a decrease in the dppm (increase in quality) level leaving the factory. In order to manufacture high quality and cost effective EEPROMs suitable for automotive under-hood applications, several topics must be taken into account. As well as a high reliability EEPROM technology, the choice of an advanced memory architecture including error correction code and a highly sophisticated screening methodology in production test is necessary to achieve high quality in the field. The EEPROM production test flow must not only be able to screen out weaknesses of the process but must also be cost efficient. A majority of the tests executed in the EEPROM test flow are needed to check the quality of the processed oxides, which are the basic elements to realize the EEPROM function of the memory. Most of these tests are complex and time consuming.

## 1.9 Chapter 9 Abstract

This chapter will describe the author's views on how the test landscape will change in the coming years. Overviews of many of the current research area will be discussed and how they will knit together to form the basis of a new test ecosystem for the future.

It is described how the author foresees the future of the test of analogue mixed/signal semiconductor devices. The author describes how they believe there is a lot of knowledge that was developed for the digital circuits that use to populate the larger geometries that today are mainly occupied by the analogue/mixed signal devices that used in the automotive industry. The author believes a great deal of this knowledge can be leveraged and used for analogue circuits to improve the quality and reliability of these devices to sub dppb levels. It is then discussed how the techniques developed above can be further improved by the use of dynamic part average testing to further drive down the dppb level. Further discussion is made on how there is a mountain of data that is collected across the entire semiconductor manufacturing ecosystem that is simply not exploited which results in a "data graveyard". Continued discussion on how the data from the fab can be used in the final test of a device that could reduce test time, increase quality levels by predict maverick events from the silicon foundry and adapting the test programs accordingly using adaptive test algorithms. A change of track is then made as the discussion diverts to the design for test of such devices, where a discussion

of the future of functional safety standard, ISO26262, in the whole semiconductor space is made, along with the new analogue test standard, 1687.2, that is being developed for standardized test access to a device. The chapter is finished by discussing the authors vision of what will come of the 2427 standard for analogue fault models. The period for this entire ecosystem to come into place is within the next 10 years.

# 2 LSI TESTING TECHNOLOGY OVERVIEW

## 2.1 Introduction

In this chapter, the basics of a production tester and the building blocks of a test program will be covered. A production tester is the main equipment with which all the research in this dissertation has been used to collect data, produce stimulus and overall test the functions of a device to ensure that the device is worthy of being sent into the field.

## 2.2 What is a production tester or ATE

Automatic or automated test equipment (ATE) is any apparatus that performs tests on a device, known as the device under test (DUT), using automation to quickly perform measurements and evaluate the test results. An ATE can be a system containing dozens of complex test instruments (real or simulated electronic test equipment) capable of automatically testing and diagnosing faults in sophisticated electronic packaged parts or on wafer testing, including system-on-chips and integrated circuits. Providers of ATE test platforms for the semiconductor market include, Teradyne, Advantest, and Xcerra. The production tester used in this research is the Xcerra Fusion MX tester, formally known as LTX and LTX-Credence. An example of a production tester is shown in Appendix 2.

### 2.2.1 Instruments

In previous years, instruments were housed in the tester mainframe. The raw signals were then sent via an umbilical cord to the test head. Once at the test head, the signal were formatted ready for use by the user. As the drive to produce better test instruments and smaller footprint testers has resulted in smaller mainframes and larger test heads, the test head now houses the instrument pin cards, see Figure 1. The mainframe now only houses parts of the tester that is still not possible to put in the test head such as RF generators and the tester workstation such as a PC.

### 2.2.2 Test Head

The test head is where the individual signals of the test come to the outside world. See Figure 2. These signals will be routed through the device interface board or loadboard to the device socket or probehead.

**Figure 1 –Xcerra Fusion MX instrument slots**

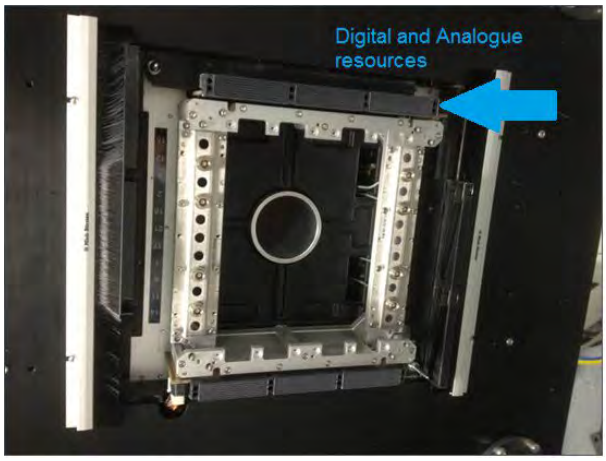



**Figure 2 –Xcerra Fusion MX test head**

### 2.2.3 Loadboards and sockets

For packaged parts, see Figure 3, a tester uses a loadboard or a device interface board, (Figure 4) to interface to the tester. On the loadboard there is a test socket, (Figure 5), the test socket connects the device pins to the instrument resources of the tester.


**Figure 3 –ams AG QFN package**

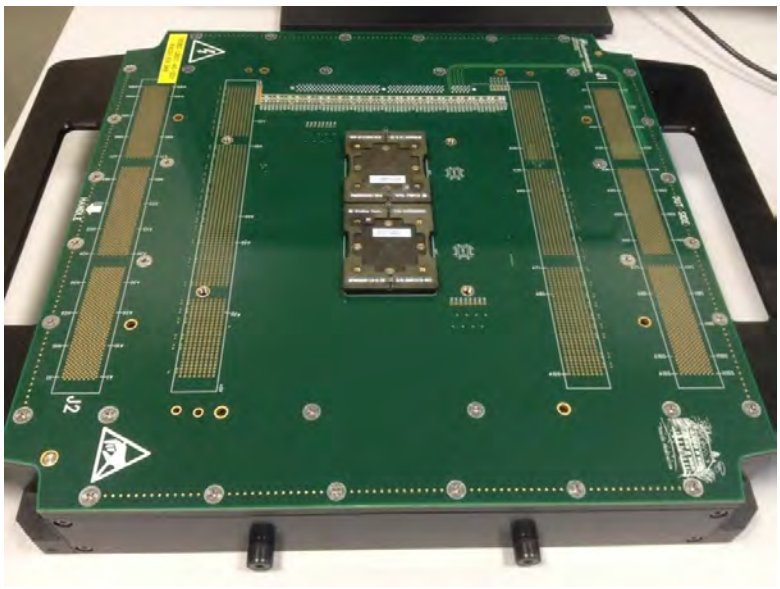


**Figure 4 –Xcerra Fusion MX loadboard**


**Figure 5 –Winway production socket**

### 2.2.4 Tester Resource Types

A modern semiconductor device has several functions (see Appendix 1), and as such requires a modern tester to be able to handle all these types of functions.

The basic building blocks of any modern tester incorporate the following:-

| | | |
|---|---|---|
| Digital pin cards | - | I2C commands, SCAN testing, device setup |
| VI Cards | - | source and measure of voltage and current |
| DSP cards | - | sourcing and measuring AC signals |

In summary, a VI would be used to force a voltage such as VDD or measure a bias current. A digital pincard would be used to setup a device so that the device is in a state that the function that requires to be tested can be tested. A DSP cards would be used to source or measure a signals such as a sinusoidal waveform.

### 2.2.5 Advantages and Disadvantages

A semiconductor production tester is very good at being just that, a semiconductor production tester. It is not a tool for characterizing devices as they would be done on the bench as some designers and product engineers believe. A product engineer would typically trigger an oscilloscope with a device pin going either high or low. Although this maybe a great feature for bench type equipment this type of approach for the production test would simply be classed as extremely bad design-for-test (DFT). When this type of testing is enforced on a production tester, it leads to asynchronous capturing of signal, a great deal of post processing of data and always leads to significant test times and bad yields due to unstable tests.

A well-developed production test procedure should be as simple as:-

- Apply VDD voltages
- Apply dc biases
- Send digital commands
- Wait a while
- Measure responses

Anything other than this is not a good production test setup.

## 2.3 Design for Test

### 2.3.1 Introduction

Design-for-test (DFT) is a name for design techniques that add certain testability features to a semiconductor chip design.

The premise of the added features is that they make it easier to develop and apply production tests for the chip. The purpose of production tests is to validate that the device contains no defects that could adversely affect the product's correct functioning.

### 2.3.2 Overview

During the course of the authors' career, he has been instrumental in driving DFT strategies during the development cycle of a semiconductor device. The author has advised and helped produce different DFT strategies to aid in the testability of a device that was going into mass production. He has helped reduce the cost of the required tester platform for testing a certain product by use of certain onboard circuitry or implementing test structures within a device to help test more simply. The following subchapters will give an overview of possible DFT schemes, below is a list of where the author has used these schemes in research and development (R&D).

Chapter 2.3.3 discusses potential issues surrounding bad DFT, the author has experienced many issues pertaining to clocking devices with a XTAL rather than with a tester resource. In all cases relating to such devices, the author strongly suggests to not use a XTAL or at least not to be solely reliable on a XTAL and to use the methodology developed in 2.3.12. The author has been implementing this test strategy for such devices for over 10 years.

Chapter 2.3.4 describes the DFT connections to a semiconductor IP block, in all cases a thorough review of the DFT of a device needs to be made before the product is released to production. If the connections to the test structures are not available at the periphery of the device then the DFT effort for the block was of no consequence. For all products being developed under the authors responsibility, a set of DFT meetings are carried out with the design team, during the development of a product, to ensure the most reliable and cost effective test solution is produced.

Chapter 2.3.5 details an example of how to use the testers' resources most efficiently, in this case, testing four DAC's in parallel. The author successfully deployed such a

DFT scheme in a very complex TETRA modem device where there were a bank of instrumentation DAC's.

Chapter 2.3.7 discusses using a DAC and an ADC that are contained within one device, to produce a digital test strategy to test mixed signal blocks. This allows a reduction in the cost of the tester platform that would test the device as an expensive digitizer could be avoided. The author has deployed such a technique in a custom ASIC that is used to measure radiation.

Chapter 2.3.8 discusses the issue of testing RF components with expensive RF Mixed Signal testers. The author has successfully deployed the onboard source and measurement techniques described in Chapter 2.3.9 & 2.3.10 in UHF NFC devices that allowed testing on standard configuration testers at ams AG. However, this is only possible if the design is centred to the process and testing for specification is not a priority although using IDDq Analogue, Chapter 2.3.14, to bridge this gap is very feasible. Reference [1], describes a DAB receiver that the author had to put into production, that had a severe design issues and only by testing with the actual modulation scheme used in the application was it possible to find accurately, the real specification based failures. In this case, what is described in Chapter 2.3.9 & 2.3.10 would not have been feasible.

Chapter 2.3.11 describes the RF equivalent of the mixed signal loopback described in chapter 2.3.7. As long as the RX and TX are in the same range, for the most part this is true, and then at least a mixed signal test program can be produced or at best purely digital. The author has deployed such a test strategy into a UWB MIMO device where all RF testing was carried out in loopback. The microprocessor onboard the device measured the test parameters. The device then flagged if the part passed or failed by simply sending a digital IO high.

Chapter 2.3.13 discusses the problem of testing an RF receiver or Transmitter and not having the reference clock phase locked to the tester. This chapter discusses how to get around this problem but without sacrificing phase noise performance by using a tester resource for a reference clock rather than a XTAL.

One of the coveted research topic from the author and something that is of great interest to him is the topic of defect coverage. In the automotive space this is a hot topic, however the author has been active in the area for many years regardless of current trends; Having focused on initially making simple current measurements during the test

of different blocks. This progressed with finally finding correlations, [2-4], that allowed the removal of complicated and expensive tests with simple current measurements. However, it was found that these current measurements could be more sensitive than the mixed signal equivalent, giving higher test coverage of defects that might go undiscovered. The author has deployed such techniques in almost every device he has been responsible for over the last 6 years and has found that this technique can even identify the root cause of a silicon failure. Chapter 2.3.14 gives details on the technique and a sample dlog is shown in Appendix 3.

### 2.3.3 Bad DFT

Bad DFT can lead to the following:

- Asynchronous captures resulting in > 10 times the needed samples
- Asynchronous captures resulting in instable measurements and poor repeatability
- Measurements being made in the lower regions of the instruments accuracy resulting in many averages, measurement instability and poor repeatability

However, for bad DFT, there is one thing clear; Bad DFT results in long test times and poor stability of measurements, which leads to much greater product costs and lower test efficiency.

### 2.3.4 Example of Good DFT

It does not matter how well a block of a semiconductor device is designed, if it is not possible to test it, then it is worthless as a production device. An IP block from a semiconductor design house can have as much DFT as possible but if the hooks are not placed such that the DFT can be accessed from the periphery of the device, then the DFT is also for nothing. It happens very often that the exact same design IP is placed in a device with completely differing modes of access, which can result in one placement of the IP having superior test time and yield compared to the next, simply due to access of certain nodes of the IP.

Figure 6 shows an SAR ADC from ams AG. On one implementation of this IP, the test engineer was required to read back the value of the ADC from a test register via a SPI bus. Not only is the SPI bus incredibly slow compared to the ADC IP but also now, there is the issue of multiple clocks. Imagine that this is an 12 bit ADC which means there are 4096 possible combinations of which all codes need testing with an

oversample of at least 16 times. This means that 65535 points need to be collected coherently using two clocks

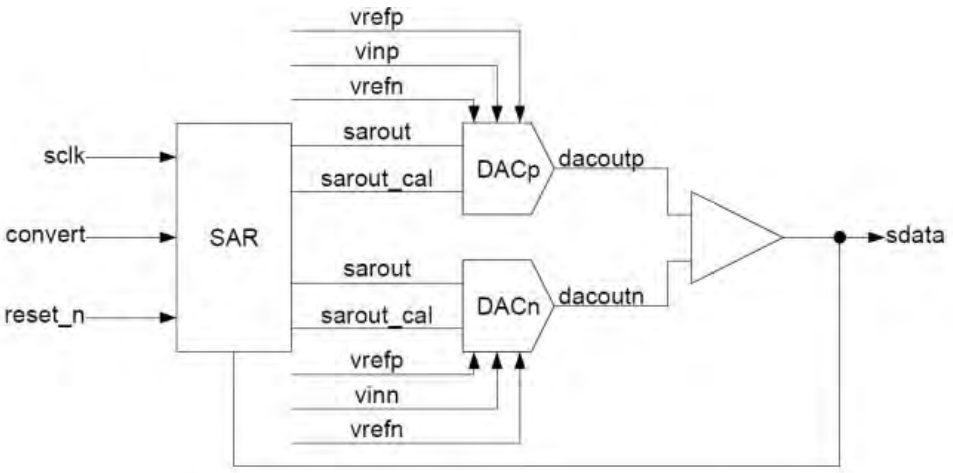



**Figure 6 –ams AG SAR ADC architecture IP**

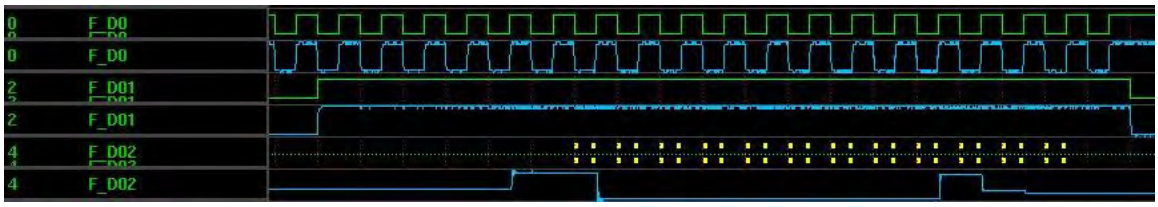



**Figure 7 –ams AG SAR ADC DFT clocking and data acquisition scheme**

that are not synchronized. It is easy to imagine that the test solution that was produced for this DFT implementation was sub-optimal to say the least.

Referring to Figure 7, it can be seen how the IP was designed to be used, here it is clear where the clocks to the ADC are being sent whilst at the same time strobing the data coming out of the ADC. This solution resulted in an extremely efficient test solution that gave high yields and optimum test time.

This was possible due to the following:-

Full control of test block and access to raw data is possible from

- SCLK
- SDATA
- CONVERT
- The signals just need to be available at defined pins using a test mode
- Using this test features allows the best testability currently available on the market

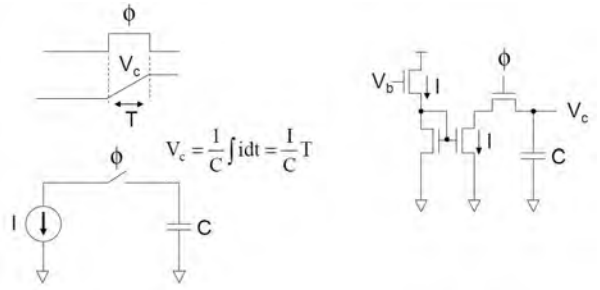



**Figure 8 –DFT circuit for generating an analogue ramp for ADC testing**

### 2.3.5 Ramp Generation with Charge Integration

In Figure 8, [5], a constant current charges a capacitor for time interval, T. This produces a linear ramp that can be used to test an ADC's integral non-linearity (INL) and differential non-linearity (DNL). However, it would be difficult test an ADC greater than 12 bits with this BIST architecture as the ramp linearity depends on the linearity of the current source and the output capacitor.

### 2.3.6 Another example of good DFT

Take the example of four identical 12-bit DAC's, within one device, to test these DAC's, one would apply a digital code ramp to the digital inputs and capturing the analogue output.

A medium to high-end mixed signal test would have four digitizers available so if possible, the tester could measure all four DAC's in parallel. Unfortunately, this is not generally the case. In a bad example, access would be gained to one DAC at a time through a test register. Again, in general, SPI would be a much slower interface than a directly accessing the DAC but the added issue here is only being able to capture one DAC at a time. This results in the possible tester usage efficiency being lower by a factor of four.

The most effective and cost effective way to test this bank of DAC's would be to parallel up the digital access (Figure 9), within in the device via a test connection so that each DAC data pin sees the same digital code allowing to ramp all DAC's at the same time. Now the tester could be used to its full potential and capture four individual signals using four digitizers in parallel. As the test time for an audio or instrument, DAC can be in the 100's of milliseconds then this type of thinking is necessary to ensure the profitability of a device going into mass production.

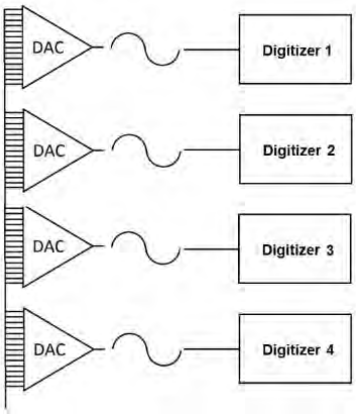



**Figure 9 –Test mode demonstrating the digital access to multiple DAC's**

## 2.3.7 Mixed Signal

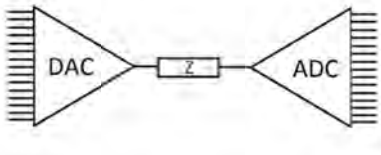



**Figure 10 –Test mode demonstrating mixed signal loopback test**

If there is a DAC and ADC within the same design with the same range and similar number of bits, a test mode with the analogue output of the DAC can be produced that connects the analogue input of the ADC via some filtering to the output of the DAC, Figure 10. By adding filtering to the output of the DAC smooths the transitions between the DAC codes and hence increases the effective number of bits of the DAC. A digital pattern then can be applied to the DAC input that would then stimulate the ADC. By implementing such a test mode, enables a purely digital test of complex analogue structures and can avoid using expensive analogue/mixed signal testers.

## 2.3.8 RF and High Speed Devices

As RF and high speed test instrumentation are very expensive and the availability on a production test floor is limited, then there is a distinct advantage of implementing in the chip design some "Design for Testability".

The advantages are the following:

- No tester access issues due to complex tester configuration
- Results can be better due tester setup issues avoided
- Test time and yield can be improved due to data transfer, tester noise etc is avoided

### 2.3.9 Receiver DFT

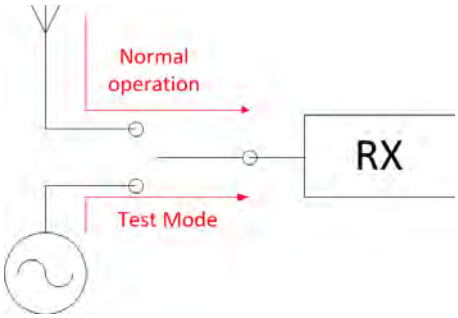



**Figure 11 –On-board oscillator for testing RF Receivers**

The test of RF receiver (RX) needs expensive RF test options such as the Xcerra MX RF16 test option. At ams AG, for example there are only two of these testers available and generally fully loaded. Bringing a new RF device into production will not trigger logistics to buy a new RF option however; it will cause longer lead times for all devices requiring the RF tester configurations. A way around this is to include a test mode where an on-board oscillator (Figure 11), that mimics the input power and frequency, can be switched in during the production test. This saves the need for an expensive and seldom available RF tester. Typical tests are defined in Chapter 4.

### 2.3.10 Transmitter DFT

The test of RF transmitter (TX) also needs the expensive RF test options to measure the output power of a RF Transmitter semiconductor device. A way around this is to include a test mode where a peak detector (Figure 12), regulates the RF output power so that a DC voltage can be measured then back correlated the actual output power of the transmitter.

### 2.3.11 RF Loopback - Transceiver

If a device has both a transmit and receive functionality a device (Figure 13), by the use of special test modes that switch the output of the transmitter to the input of the receiver, a device can effectively can be tested purely digitally if there are DAC's and ADC's in the basedband (BB) path, Figure 10. This is the ultimate analogue mixed signal test as the device test can be done highly multisite due to only requiring digital resources to do the test. However, that said a separate test of either the TX or RX should be done to be able to diagnose any yield issues seen in production. If either a RX or a TX tests are not done in combination with a loopback test, it is not possible to know which part RF loopback path of the device was failing.

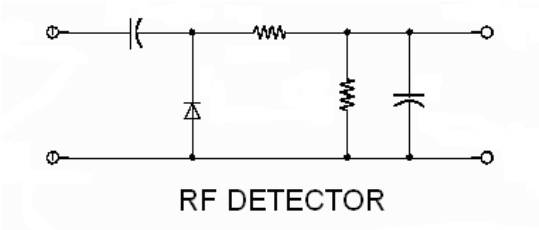



**Figure 12 –On-board peak detector**

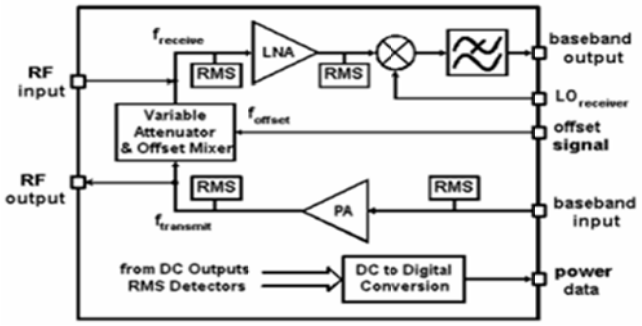



**Figure 13 –TX & RX loopback connections**

Ideally, the RX would be tested instead of the TX as the RX as an AWG can be shared easily across multiple sites whereas a TX needs a dedicated digitizer to measure the output thus limiting the possible multisite count or parallelism of the test solution to the number of digitizers in the test system. In addition, attention needs to be paid to the fractional synthesiser of such a device if it is shared between the RX and TX. If the phase error of the fractional synthesiser is critical to functioning of the device, which is usually the case with most modern wireless devices that use leading edge modulation schemes such as 128 level Quadrature Amplitude Modulation (QAM), the phase error will be observed by both the RX and TX and hence will be seen as no error.

If the transceiver is part of a radio frequency system on a chip (RFSOC) and has a microprocessor (uP) such as an ARM core, then the digital signals can be routed to and from the transceiver to the uP. Code can then be developed and executed once loaded into the uP such that a full suite of transceiver loopback tests are produced. If all tests pass, the uP can place a digital high on a pin that the tester checks for once that the test program in the uP was launched. This is a very exotic style of built in self-test, several companies such as Cambridge Silicon Radio used this approach to protect the IP of their devices from theft. If a test program is not compiled full access to the workings of a device can be discovered by interrogating a test program in detail.

## 2.3.12 XTAL Oscillator

Care should be taken with XTAL or clock circuitry as if the device can only be run with a XTAL then windowing will need to be any waveform capture as it will not be phase locked to the tester. This can be resolved by implementing a test mode that allows clocking the device from XTAL pins using a digital pin from the tester. However, by using a XTAL, a test of the gain of the XTAL oscillator, Figure 14, is done and the functionality of the device with the XTAL guaranteed. Unfortunately, XTAL oscillators have a start-up time, which sometimes can be up to 2 seconds. Hence using the actual XTAL in the test solution can lead to excessive test times just to guarantee the gain of the oscillator. This can be resolved by implementing a circuit to find the point at where the inverter would switch. By applying a small delta to either side of this bias the measured swing in comparison to the delta can be measured and hence the gain of the inverter.

The gain is measured by shorting the input to the output, using K3 and measuring the midpoint bias on the ppmu of the digital pin via the closed relay, K2. For ease of description, we will say this should be approximately VDD/2 or $V_{mid}$. Now opening the K3 and closing K1, we can apply a voltage of $V_{mid} + \Delta V$ to the inverter input via the digital pin ppmu via K1. The inverter should now be approximately 0V or $V_{low}$ and can

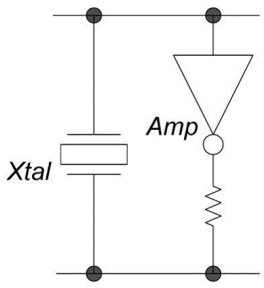



**Figure 14 –XTAL Oscillator circuit**

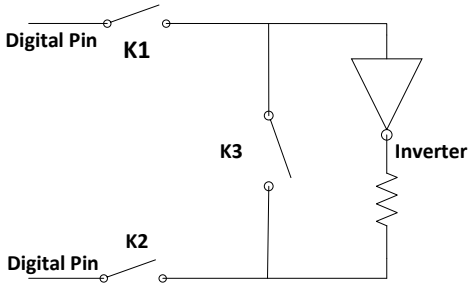



**Figure 15 –XTAL Oscillator gain measuring circuit**

be measured through K2, on the ppmu of the digital pin. We can now apply a voltage of $V_{mid} - \Delta V$ to the inverter input via the digital pin ppmu via K1. The inverter should now be approximately VDD or $V_{high}$ and can be measured through K2, on the ppmu of the digital pin. The gain can now be measured with equation (1a).

$$\text{Gain} = (V_{high} - V_{low})/2\Delta V \quad \textbf{(1a)}$$

### 2.3.13 Maintaining spectral purity – Phase Noise and Low Jitter reference clocks

As the drive to increase data rates; whilst maintaining the same spectral bandwidth, continues unabated results in ever increasing complex modulation schemes. In the past schemes, such as Quadrature Phase Shift Keying (QPSK) was at the forefront of digital modulation world now they are a confined to the history books. The latest digital protocols such 256 Quadrature Amplitude Modulation (QAM) have such small distances between vector positions, the control requirements needed to ensure adequate Error Vector Measurement (EVM) performance dictates that the phase noise, of the system clock or LO be greater than -120 dBc/Hz. The same requirements are needed for low noise systems, such as Global Positioning by Satellite (GPS), where the circuit is trying to detect very low signals. Using a XTAL, as described in 2.3.12, has great phase noise, $\theta_n$ , performance, but is impractical to use to test circuit parameters such as EVM as it is not phase locked to the test system. A few commonly used approaches are significantly flawed but are used none the less. These approaches result in a significantly degraded phase noise performance.

1) Digital Pin – A digital pin is used, as shown in Figure 15, to drive the reference clock, from a digital pattern. If the edge placement of the digital pin were accurate to a few picoseconds then the phase jitter of the pin would be sufficient. However, this is extremely unlikely in lower cost Mixed Signal testers where edge placement is usually specified to nanoseconds, which would result in a degraded phase noise performance due to the clock.

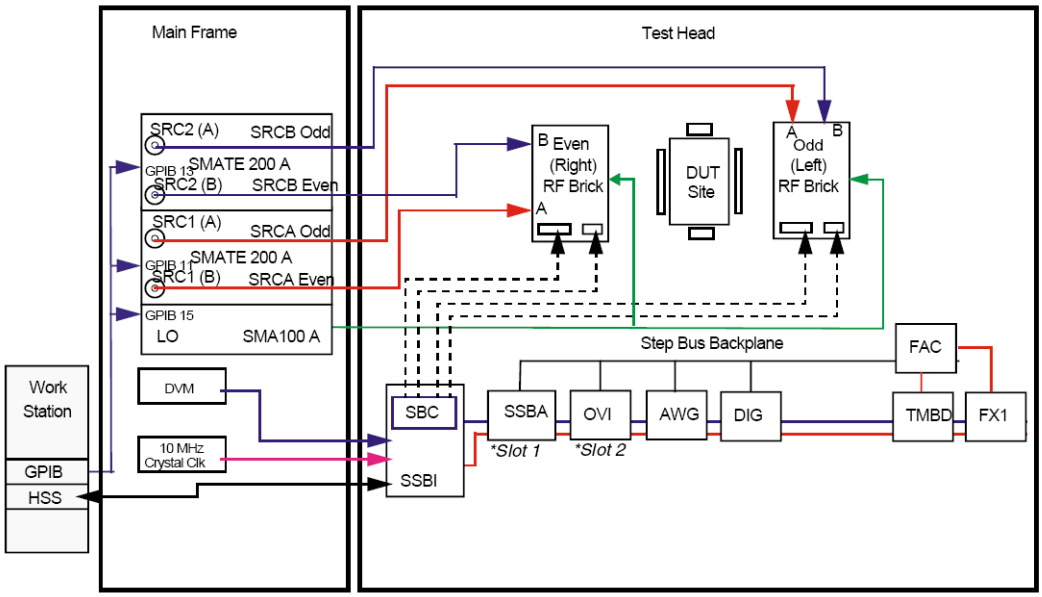



**Figure 16 –System overview of LTX MX RF tester**

2) AWG – An AWG can also be used to produce these signals, although more commonly as a sinewave than a square wave. In general, most high speed AWG's do not come with a specified phase noise performance, therefore cannot be relied upon to generate low phase noise clocks.
3) RF generator – Although RF generators can be used down to the MHz region, the best phase noise performance is exhibited at the higher frequencies, for a commonly used RF generator, at 20 MHz the phase noise specification is -116 dBc/Hz whereas at 1 GHz, the same generator has a phase noise specification of -126 dBc/Hz. If the requirement is a phase noise of -120 dBc/Hz then clearly using a source with a specified value of -116 dBc/Hz will not suffice.

To overcome this issue we can use high performance digital on-board circuitry to improve the phase noise of the RF generator to produce a significantly better phase noise clock that would be otherwise not be possible using the test system in its normal configuration to produce the required frequencies needed.

Using the RF subsystem to generate the Local Oscillator (LO) is to the DUT is a good choice as it is phase locked 10 MHz reference, Figure 16, that drives the whole test system. As such, the measured signals of any Transmitter or Receiver will be synchronised with the input signals, therefore allowing to measure complex waveforms such as 256 QAM, and thus calculations such as EVM are possible without complex windowing techniques.

The phase noise generated by a DUT is simply a function of the XTAL oscillator circuitry that would drive the device, Figure 14, which can usually be modelled by a simple inverter circuit. The major phase noise contribution is where the inverter is switching states, as defined in 2.3.12. Therefore, the quicker the signal passes through this switching region, the better the phase noise performance of the XTAL oscillator will be. This can be expressed mathematically simply as the rate of change of the signal, equation (1b). One misconception that is common is that the XTAL input voltage has to go from 0 → VDD. This is not true and actually reduces the rate of change of voltage, equation (1b), if produced, as the signal takes longer to get to and from VDD. As long as the voltage swing of the reference clock is greater than the switch points of the XTAL oscillator the DUT will function as expected.

$$\theta_n \propto \frac{\Delta V}{\Delta t} \qquad \textbf{(1b)}$$

A technology that has an exceptional rate of change performance and high frequency switching characteristics is the Emitter Coupled Logic (ECL). If this technologies performance could somehow be extracted and used with a loadboard it could be possible to generate a superior phase noise signal than with just the tester alone. It is possible with the RF subsystem of most testers to produce frequencies up to 6 GHz, if it was possible to both divide this frequency down and produce a square wave with an excellent rate of change, it could be possible to generate a superior phase noise signal.

Using the MC100EL33 ECL divider from OnSemi, it is possible to divide down signals at frequencies of up to 3.8 GHz, with a clock-to-clock jitter of 1 ps and rise and fall times of 225ps. By daisy chaining either three or four of these devices together, it is possible to generate the frequencies that are usually required for a XTAL, between 10 – 40 MHz, with exceptional phase noise properties. However, ECL has strange voltage levels in that they switch between 3.3 and 4 V i.e. a 0.7 V voltage swing that are generally not useful. However, if the DC component is filtered off with AC coupling, it will be possible to generate a 0.7 Vpk-pk signal centred on 0V with exceptional phase noise properties. If the LO requires a DC offset, it is possible to use one of the testers DC instruments, with a lot of filtering, to shift the generated reference clock to the required voltage offset.

Using a real world GPS device as an example that was recently put into production, the reference clock required a 10 MHz reference that would be supplied by a Temperature Controlled Crystal Oscillator (TCXO) in the application. The phase noise spec of the device was -86.5 dBc/Hz at 1.590 GHz at 100 kHz offset and -116.2 dBc/Hz at 1.590 GHz at 1 MHz offset. By using the ECL divider circuity, Appendix 4, with four divide by 4 dividers and using the RF generator with a frequency of 2.56 GHz, a pure 10 MHz signal was produced. This resulted in the ability to measure the phase noise contribution of device repeatedly, Figure 17 & Figure 18, rather than measuring the random noise added to the device by the test system that results in poor repeatability and non-correlating results.

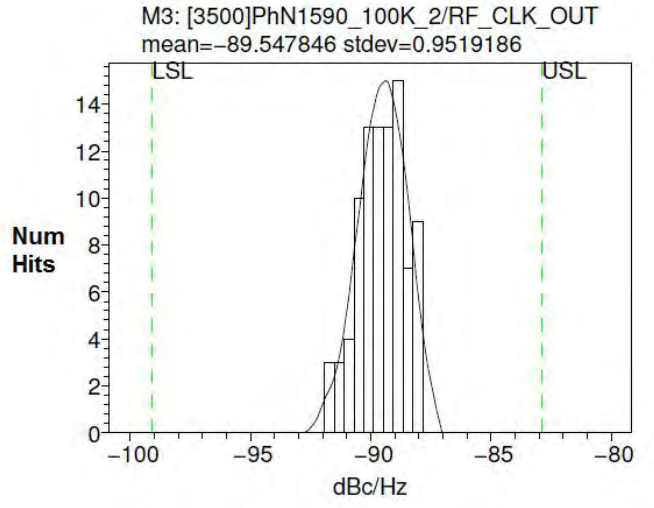



**Figure 17 –Relatability histogram of phase noise of a GPS receiver at 1.59GHz at 100 kHz offset**

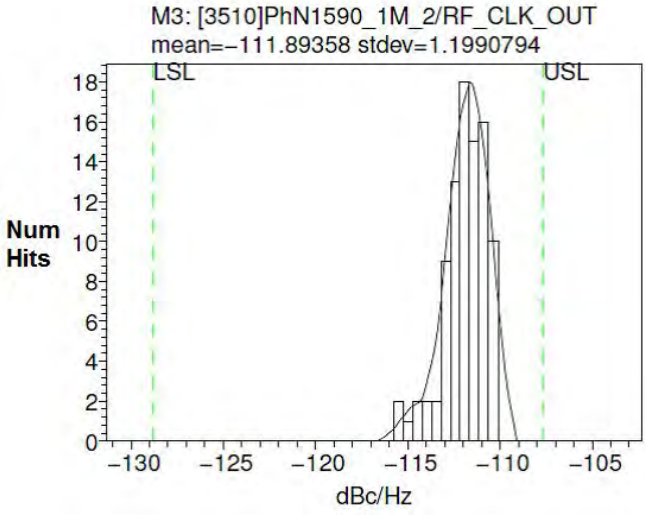



**Figure 18 –Relatability histogram of phase noise of a GPS receiver at 1.59GHz at 1 MHz offset**

### 2.3.14 IDDq Analogue

IDDq in the digital world is simply a set of digital scan vectors that place the digital in a power down mode such that leakage of the transistors in that mode can be measured on the supply, VDD. For an LSI device, we can replicate in several ways.

If each block of a device is clearly identified, and the associated external circuitry to each block designed in such a manner that each block can be powered independently then it becomes possible to measure both the powered down and powered up biased current of each block with no influence from surrounding circuitry. Although this in itself would be excellent coverage compared to the norm, if an extra step is taken to check the difference between the current of each powered block then excellent fault coverage can be attained. One of the excellent side effects of this is the ability to detect a specification failure based on the current consumption. This is simply because there has to be a failure in the circuity that would cause a failure with the bias current that would lead to the specification failure. See Appendix 3 for an example.

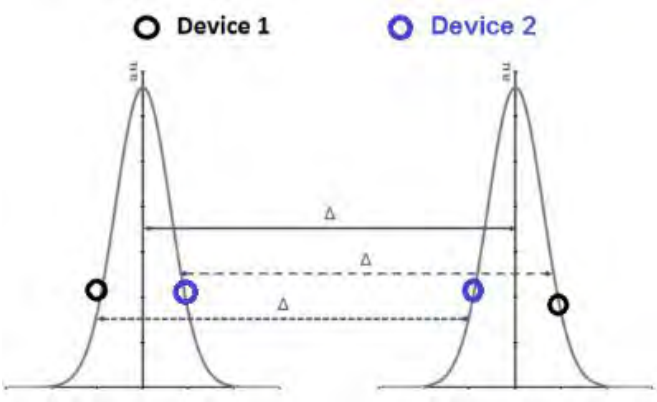



**Figure 19 – Distribution of two test showing a difference between two devices**

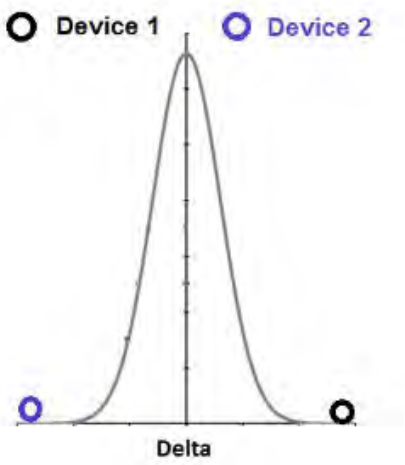



**Figure 20 – Delta test showing the devices as outliers**

## 2.4 Built in Self-Test

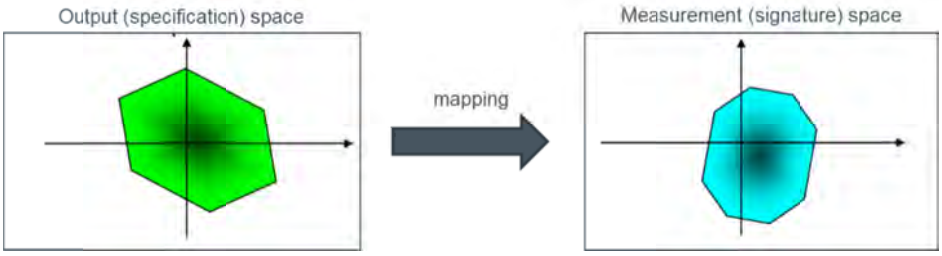



**Figure 21 –BIST measurement accuracy to specification space diagram**

The principle of BIST is simply to add additional circuitry so that the DUT is able to test itself

Requirements:

- On-chip test pattern generation
- On-chip test response analysis
- BIST control

Desired features:

- Allow an independent test of each block in an analogue chain
- Limit the need of complex functional tests
- Facilitate the reuse of test techniques
- Use techniques for test generation and response analysis highly
- Portable, mostly digital
- Transparency with respect to DUT functionality
- Low area overhead

Advantages:

- IP is protected as test program can be embedded in the chip memory

However, the issue with BIST is that there is not a one to one comparison with an actual device test and hence test escapes are highly likely and greater than 100ppm, [5]. Any BIST technique relies on mapping the output parameter space of specifications into the measurement signature space, considering analogue tolerances (Figure 21), [5].

For this reason, BIST cannot be used in automotive applications, as it could be a potential quality risk.

## 2.5 Analogue Test Bus

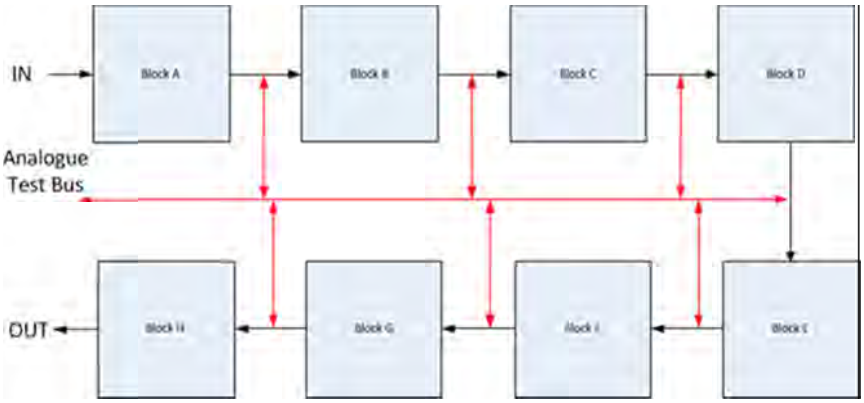



**Figure 22 –A typical analogue test bus**

As part of any BIST or DFT scheme requires that access can be gained to certain blocks so that they can be stimulated and measured either by on-board or tester instruments. This can be as simple a switch within the device so that a path can selected during test or a very complex multiplexer, Analogue test bus (ATB), which can switch signal all over a complicated RFSOC (Figure 22). Either way the solutions that are made are adhoc and at best are company specific and at worst, every designer of a company designs a test bus.

Therefore, it is clear that there is no standardisation in this market segment in terms of test access of analogue mixed signal blocks. What is standard is that these devices all use either a SPI or and I2C digital communications protocol to set test registers and put the devices into certain operating modes. If these two key ideas could be leveraged into a standard, then there could be the potential to automate, using tools from electronic design automation (EDA) vendors, this part of the design process. This would produce a standard model that all designers from around the globe could follow and hence a designer could easily understand a design from another engineer. This would also have the benefit of speeding up the design flow significantly. There would also be the benefit that if there was a standardised test access that the test development time would be improved with the side effect that test metrics could be improved due to better layout of loadboards due to the standard access. This is of course of high interest to the business for time to market reasons, and as such, the time to market of a device could be decreased significantly.

## 2.6 IEEE test standards

The following IEEE standards have been around for time and for digital devices the standards are very well adopted, however there has been some attempt to standardise analogue test access but not with much adoption from industry.

IEEE 1149.1 Standard digital boundary-scan test (1990) – Very well adopted

- Facilitates observability and controllability of signals in digital
- ICs, providing an structured solution to PCB testing

IEEE 1149.4 Standard mixed-signal test bus (1999) – Seldom used

- Facilitates observability and controllability of signals in analog
- ICs, providing also support for BIST techniques

IEEE 1500 Standard test method for embedded cores – Seldom used

- Facilitates access to embedded cores in SoCs

IEEE 1687 Standard for access and control of instrumentation embedded within a semiconductor device. – Gaining traction

- Facilitates standardized access to embedded cores in SoCs

From an analogue mixed signal point of view, the IEEE 1149.4 adds much overhead to a chip that may only have 8 pins, this is because the standard reuses the 1149.1 access that requires 4 extra JTAG pins. For a 1024 pin uP sharing 4 pins for test access is no problem however if the device is an 8 pin analogue mixed signal serial peripheral interface (SPI) driven device that only requires 3 digital pins then sharing 4 pins is simply not possible.

## 2.7 1687.2 – A new beginning

In 2014, a group of motivated industrial test veterans came together in Paderborn, Germany for the European Test Symposium and discussed as part of the emerging test strategies issues pertaining to the lack of standardised DFT and test access were causing in the analogue mixed signal test arena. The author was a key contributor to these discussions. The result and conclusion to the meeting of people was to kick off an industrial working group that originally consisted of six people from industry to talk about how to come to a standard that EDA vendors can use to produce automation tools that can help in the design flow of analogue mixed signal devices. This team has now

expanded to over 10 companies and is at a stage where a draft standard is ready. In 2017, an IEEE study group was formed so that the process of officially creating the standard could be started. The author is the organiser and instigator of this IEEE study group.

## 2.7.1 Scope and Purpose

### 2.7.1.1 Scope

This new standard formalises how to describe circuitry within an integrated circuit (IC) that is relevant to testing for fabrication defects in analogue functions of the IC. The description is also relevant to automating insertion of analogue test access circuitry, and to automating generation of the manufacturing tests. Analogue test access includes access to analogue signals and to digital signals in which the edge positions contain analogue information; analogue test control includes control of circuit under test (CUTs), control of analogue test instruments within the IC, and control of the access paths between CUTs and analogue test instruments, the latter including on- and off-chip instruments.

In this context, an “analogue” function means a function that has input, internal, or output signals with meaningful values in a defined continuous range, and the function has at least one tested performance that is sufficiently non-deterministic that its test has upper and/or lower limits (the limits may be real numbers or quantised digital equivalents). An analogue instrument is one that measures analogue properties of signals.

### 2.7.1.2 Purpose

The primary purpose of this new standard is to facilitate automating test generation for analogue functions within ICs.

The second purpose of this new standard is to facilitate automating insertion of analogue test access circuitry during the IC’s design to permit manufacturing tests to be better, for instance by detecting more defects, or by testing in less time, or using a lower-cost tester.

This new standard focuses on describing, in a machine-readable way, analogue test access techniques already used widely in the industry. When a new alternative offers a significant advantage, this new standard will describe specific analogue test access

circuitry that is better in some way, for instance by enabling faster test times, more reliable test access, or automation of more tasks.

This new standard will describe circuitry using existing languages and DFT standards, wherever possible, to minimize the amount of new automation needed. When that is not possible, this new standard will minimise the changes to the existing languages or test access standards to accommodate analogue functions. Specifically, this new standard will use the instrument connectivity language (ICL) and procedural description language (PDL) defined in the IEEE 1687 Standard for access and control of instrumentation embedded within a semiconductor device.

### 2.7.2 Expected outcome and benefits

Many needed benefits motivated development of this standard. The overall benefit should be a reduction in DFT and test development effort, and implementation time, as a result of an analogue test access description language that facilitates easier and more precise communication between IC designers, test engineers, and EDA tool developers. Specific benefits are listed below.

Benefits of using 1687 digital automation for mixed-signal DFT

- Control ATB using PDL procedures instead of bit-bashing
- Create test patterns that are portable from design to design
- Create test patterns that are re-targetable to higher levels in the design

Benefits of using this analogue test access and control description standard

- Standardized description of analogue test access, instruments, tests, block control, signals, …
- Clear, unambiguous communication of test intent
    - ICL, PDL that could be used to create tests automatically using existing 1687 tools
- A systematic DFT flow that facilitates automation (existing and future automation)

Expected benefits of using future automation based on this DFT standard:

- Automatically access an on-chip signal by simply naming the signal to be probed

  - Automated checking that access does not cause conflicts, e.g., accessing two outputs, driving an output, failing to release clamp, failing to apply clamp after test, etc.
  - Automated setup of circuit blocks for testing, independent of surrounding blocks
  - Automated checking that a design allows all intended test accesses, without conflicts
- Extend ATE OS-based commands to on-chip signals
  - Create on-chip signals by driving in digital pattern to on-chip DAC
  - Enable ATE analogue signal or on-chip DAC to drive an ATB
  - Convey ATB signal to a named node, and from that or another node
  - Enable ATE to measure analogue signal directly via ATB or via on-chip ADC
- Automatically access multiple ADCs/DACs serially in minimal test time, or less time
- Automatically minimize the test time to access a set of DC signals, by optimal sequencing
- ATPG that calibrates on-chip ADC/DAC, via ATE and ATB, to facilitate BIST of other functions
- Create a test suite for an analogue function block that can go with that block to a next IC design
- Automatically retarget analogue tests from block-level to chip-level
  - Definition of generic measurements would allow mapping to ATE instruments.
- Automatically retarget analogue tests to system consisting of DUT and on-chip instruments.
  - Mapping of generic measurements to on-chip instruments.

Benefits needed but not expected:

- Better DFT
- ATPG that automatically reduces analogue test time
- Test methods that allow use of lower-cost ATE
- Automated insertion of ATB
- Resolution to the debate about specification testing *vs*. DOT

## 2.8 Design & Test Quality Flows – Automotive vs Consumer

The design and test of devices for the automotive market are considerably different from the consumer flow. The following sections will give a brief overview of the automotive process, how it differs to the consumer flow and how the consumer flow is adapting to the current changes in product volumes.

### 2.8.1 Production part approval process

Production part approval process (PPAP) is used in the automotive supply chain for establishing confidence in component suppliers and their production processes. Actual measurements are taken of the parts produced and are used to complete the various test sheets of PPAP.

Although individual manufacturers have their own particular requirements, the Automotive Industry Action Group (AIAG) has developed a common PPAP standard as part of the Advanced Product Quality Planning (APQP) – and encouraging the use of common terminology and standard forms to document project status.

The PPAP process is designed to demonstrate that the component supplier has developed their design and production process to meet the client's requirements, minimizing the risk of failure by effective use of APQP. Requests for part approval must therefore be supported in official PPAP format and with documented results when needed.

*The purpose of any Production Part Approval Process (PPAP) is*

- *to ensure that a supplier can meet the manufacturability and quality requirements of the parts supplied to the customer*
- *to provide evidence that the customer engineering design record and specification requirements are clearly understood and fulfilled by the supplier*
- *to demonstrate that the established manufacturing process has the potential to produce the part that consistently meets all requirements during the actual production run at the quoted production rate*

The result of this process is a series of documents gathered in one specific location called the "PPAP Package". The PPAP package is a series of documents that need a formal certification by the supplier and approval by the customer. The form that summarises this package is called PSW (Part Submission Warrant). The signature in the supplier certification area of the PSW indicates that the supplier responsible person has

reviewed this package and that the customer responsible person has not identified any issues that would prevent its approbation.

The documentation on the PPAP package is closely related to the advanced product quality planning process used during the design and development of new vehicles and component systems to reduce the risk of unexpected failure due to errors in design and manufacture. The PPAP manual is published by the AIAG, and specifies generic requirements for obtaining PPAP approvals. Additional customer specific requirements may be imposed by particular clients and incorporated in the purchasing contracts. Details of 'customer specific' requirements may be found on the International Automotive Task Force (IATF) website or supplier portals provided by the vehicle manufacturers.

Suppliers are required to obtain PPAP approval from the vehicle manufacturers whenever a new or modified component is introduced to production, or the manufacturing process is changed. Obtaining approval requires the supplier to provide sample parts and documentary evidence showing that,

*1) The client's requirements have been understood*

*2) The product supplied meets those requirements*

*3) The process (including sub suppliers) is capable of producing conforming product*

*4) The production control plan and quality management system will prevent non-conforming product reaching the client or compromising the safety and reliability of finished vehicles.*

PPAP is the confirmation that the product meets the customer requirements for high volume production. The PPAP will be considered signed when the customer approves a full PSW and added to the PPAP folder.

Although there is no specific PPAP process for consumer devices, certain mobile phone manufactures are producing their own versions of this process with more stringent requirements.

### 2.8.2 Design Considerations

During the design of an automotive device, the design would follow certain design rules such that the CPK (process capability) of device parameters meet a six sigma criteria, CPK = 1.66, i.e. there should be six sigma distance between the design distribution and the specification limit. This should ensure that the product is designed robustly i.e. it is a six sigma design.

It is mandatory for all automotive devices to be able to store a unique identification into the device. This is to enable full traceability of the device, during the wafer sort of the product, x and y co-ordinates of the silicon die on the wafer will be stored along with the wafer number from the lot. Although not mandatory, some manufactures store the lot number of the device also. This is done by using an on-chip non-volatile memory to store the data that is programmed during the wafer-sort test flow of the silicon wafer. An example is shown in chapter 7.

An automotive device has to have all the specifications that are listed within the datasheet with a minimum and maximum value tested directly with the ATE. No BIST or BOST is allowed. This is strictly forbidden as the BIST or BOST may have a fault and hence cause an error in the test of the device.

For a consumer device, no additional area, for whatever reason, would be used unless absolutely necessary as any silicon area overhead increases the die area and reduces the number of possible die per wafer. Therefore, increasing the cost per die and reducing the potential profit of the device. As consumer devices sell for extremely low cost; any saving, however small will be taken to increase the profit, however small.

### 2.8.3 Failure mode and effects analysis

Failure modes and effects analysis (FMEA) is an approach for identifying all possible failures in a design, a manufacturing or assembly process, or a product or service. “Failure modes” means the ways, or modes, in which something might fail. Failures are any errors or defects, especially ones that affect the customer, and can be potential or actual. “Effects analysis” refers to studying the consequences of those failures. Failures are prioritized according to how serious their consequences are, how frequently they occur and how easily they can be detected. The purpose of the FMEA is to take actions to eliminate or reduce failures, starting with the priority.

FMEA also documents current knowledge and actions about the risks of failures, for use in continuous improvement. FMEA is used during design to prevent failures.

Design and Test FMEA’s are mandatory for automotive devices and if done correctly can consume several days of time for the whole design and test team of a device, as such doing a thorough FMEA is extremely expensive in human time. As such, unless totally mandatory, a FMEA will be avoided. The results of the FMEA study for an

automotive customer will be a part of a PPAP package delivered to the end customer for sign off.

An FMEA for a consumer device is done very seldom, however, for some mobile phone manufactures; a FMEA is becoming a mandatory part of the acceptance process.

ISO26262 is an extended form of an FMEA, which is gaining traction in the automotive market, chapter 9.7 explains some future expectations from this standard.

## 2.8.4 Test Considerations

Before an Automotive device can be released into production, full traceability back to the bench characterisation needs to be performed for all parts of the production. This comes in the following forms.

### 2.8.4.1 Test List

A consumer device can be tested in anyway seen fit with BIST, BOST, using digital only test techniques etc. As long as the test escapes are not so high that the customer cannot handle them, the test solution is usually accepted.

Alternatively, for an Automotive Test solution, any minimum and maximum specifications detailed in the datasheet have to be tested directly with the ATE test system. This makes and automotive test solution much more expensive to produce, as these specifications need to be correlated over all the specified temperatures detailed in the datasheet.

### 2.8.4.2 Golden Sample Correlation

A set of 30 reference samples that were tested on the bench need to be tested and correlated on the ATE and a correlation report produced. Any non-correlating items need to be resolved before the following release to production activities can be executed.

For a consumer device, if there are some non-correlating items it is ok as long as a failure can be caught with that means of test.

### 2.8.4.3 Repeatability Analysis

A repeatability analysis needs to be undertaken and the distribution compared to the limits. If the board is multisite then the repeatability study needs to be completed for each site of the loadboard. The CPK for each parameter is then calculated, for an

automotive device the CPK has to be a minimum of 1.66 or 6 sigma. For a consumer device, the CPK requirement is more relaxed at 1.33 or 3 sigma.

#### 2.8.4.4 Gauge Repeatability and Reproducibility

Gauge Repeatability and Reproducibility (GRR) measures the amount of variability induced in measurements by the measurement system itself, and compares it to the total variability observed to determine the viability of the measurement system. There are several factors affecting a measurement system, including:

In systems making electrical measurements, sources of variation include electrical noise and analogue-to-digital converter resolution.

Specification, the measurement is reported against a specification or a reference value. The range or the engineering tolerance does not affect the measurement, but is an important factor in evaluating the viability of the measurement system.

There are two important aspects of a Gauge R&R:

*Repeatability: The variation in measurements taken by instruments, on the same, or replicated item and under the same conditions.*

*Reproducibility: the variation induced when different operators, instruments, or laboratories measure the same or replicate specimen.*

It is important to understand the difference between accuracy and precision to understand the purpose of Gauge R&R. Gauge R&R addresses only the precision of a measurement system. It is common to examine the P/T ratio that is the ratio of the precision of a measurement system, to the tolerance of the manufacturing process, of which it is a part. If the P/T ratio is low, the impact on product quality of variation due to the measurement system is small. If the P/T ratio is larger, it means the measurement system is taking a large fraction of the tolerance, in that the parts that do not have sufficient tolerance may be measured as acceptable by the measurement system. Generally, a P/T ratio less than 0.1 indicates that the measurement system can reliably determine whether any given part meets the tolerance specification. A P/T ratio greater than 0.3 suggests that unacceptable parts will be measured as acceptable (or vice versa) by the measurement system, making the system inappropriate for the process for which it is being used.

For the test solution, Gauge R&R should be done between loadboards to ensure that any replicated boards of the reference board will operate identically to the reference test

loadboard. This should be done using the same devices and ATE tester so that the all variables are the same other than the loadboard being compared. Five devices tested 10 times, in each site, 2 times, for both loadboards. For a quad site solution, this would result in 200 test runs, 4 times, twice for each board.

Gauge R&R should also be executed between several ATE systems of the same type. This should be done to ensure that the results across the installed tester base of the company are guaranteed to be the same no matter what tester was selected to run production. This should be done using the same devices and loadboard so that all the variables are the same other than the ATE systems being compared. The complete test setup including handler should be used to prove the full system. Depending on the hander, a special setting maybe needed to ensure the devices are always tested in the same sequence. However, as there is always a die ID, the results per device can always be sorted accordingly if the die ID is read back and dlogged.

As the handler is a required part of the tester GRR study, a repeatability analysis is not feasible as with the loadboard Gauge R&R. Therefore, 50 parts tested twice on two test systems can be used to make the GRR study. Gauge R&R is a requirement for a PPAP documentation package.

Although GRR is not a requirement for consumer products, by doing such a study on all products ensures a stable production process resulting in a low probability of yield issues due to bad loadboards or incorrect usage of ATE instruments. As such, most semiconductor companies perform some sort of Gauge R&R analysis on their products to ensure a smooth production flow.

### 2.8.5 Corner Lot

For an automotive device, it is very important to understand how the device will behave over the process corners of the silicon foundry. As such, a corner lot is a mandatory part of the release to production flow.

A corner lot is where the silicon foundry purposely non-standard wafers based on the probability of process shifts. By producing such wafers, the silicon supplier can guarantee that the automotive design is robust across the natural variations the silicon foundry would experience during production.

The corner lot test results need to be documented and shared with the customer as part of the PPAP procedure.

Generally a consumer customer will not request a corner lot as it is not mandatory, however mobile phone manufacturers are starting to request corners of a process that historically Automotive customer would not have asked for hence, in some case, creating more stringent requirements for a consumer device than an automotive device.

### 2.8.6 Safe Launch

Safe launch is the process of testing all devices at the maximum and minimum specification temperatures. This is a very expensive procedure as a production lot is tested 3 times. To be able to remove the safe launch stage of an automotive device, all failures at the minimum and maximum temperatures have to be shown to have failed at the standard test temperature what would normally be the standard production stage of the device. This has to be done with either 3 lots or 100,000 devices, whichever is larger, i.e. if one lot has 10,000 devices then 10 lots would be needed, if another device has 34,000 devices per lot then only three lots are needed.

If the silicon supplier cannot prove that failures at temperature can be caught with the standard flow at nominal temperature then the safe launch stage of production cannot be exited.

Safe launch is required so that it can be guaranteed to an automotive customer that no devices failing temperature margins will be delivered to the end user.

The data for safe launch stage would need to be shared with the customer as part of the PPAP before removal from the production flow.

Historically, no consumer would request such a flow; however, the trend is changing, consumer customers are currently requesting a safe launch style rampup with 1 Million units. All failures at the corner temperatures are required to be analysed to ensure all failures are captured at the standard test temperature and if not appropriate changes made to the test program to ensure that the failure would be caught in the future either with tighter gaurdbands or additional tests.

## 2.9 ADC Testing Background

### 2.9.1 INL and DNL

There are several key design parameters an ADC is designed towards, the most critical and most common that are tested for are the linearity specifications. These are defined as integral non-linearity (INL) and differential non-linearity (DNL).

DNL reveals how far a code is from a neighbouring code. The distance is measured as a change in input-voltage magnitude and then converted to LSBs. Note that INL is the integral of the DNL errors. The key for good performance for an ADC is to ensure that there are no missing codes. This means that, as the input voltage is swept over its range, all output code combinations will appear at the converter output. A DNL error of less ±1 LSB guarantees no missing codes.

A perfect ADC is shown in Figure 23 with a DNL error of 0 LSB. In Figure 24, three DNL error values are shown. One code has a DNL error of -0.5 LSB, one code has a value equal to 1 LSB, and one code is missing giving a DNL of -1 LSB. As most ADCs are specified maximum DNL error of less than +/-1 LSB and because production test limits are tighter than the data-sheet limits, no missing codes are usually guaranteed. With a DNL value greater than -1, the device has missing codes.

INL is defined as the integral of the DNL errors, so good INL guarantees good DNL. The INL error tells how far away from the ideal transfer-function value the measured converter result is.

Figure 25 shows an Ideal ADC transfer function vs two real world transfer curves. If there is a positive offset, a part of the ADC's range is lost. If the offset is negative, the converter will just display zeros. Thus, with a negative offset error, you must increase the input voltage slowly to determine where the first ADC transition occurs; again, a part of the ADC range is lost.

As with offset error, you lose dynamic range with gain error. For example, if a full-scale input voltage is applied and the code obtained is 4050 instead of the ideal 4096 (for a 12-bit converter), this is defined as negative gain error, and in this case, the upper 46 codes will not be used. Similarly, if the full-scale code of 4096 appears with an input voltage less than full-scale, the ADC's dynamic range is again reduced (Figure 26).

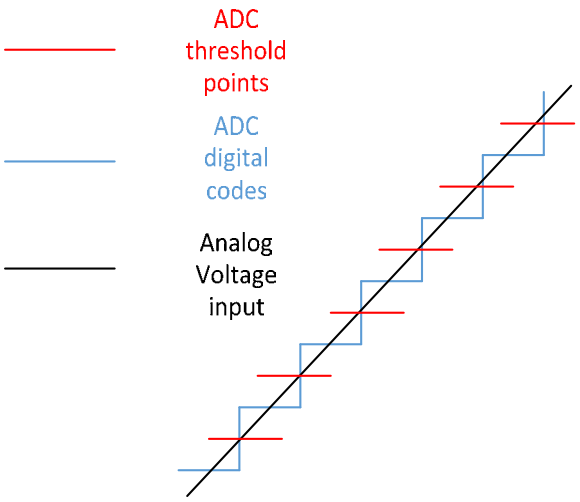



**Figure 23 –A perfect ADC response**

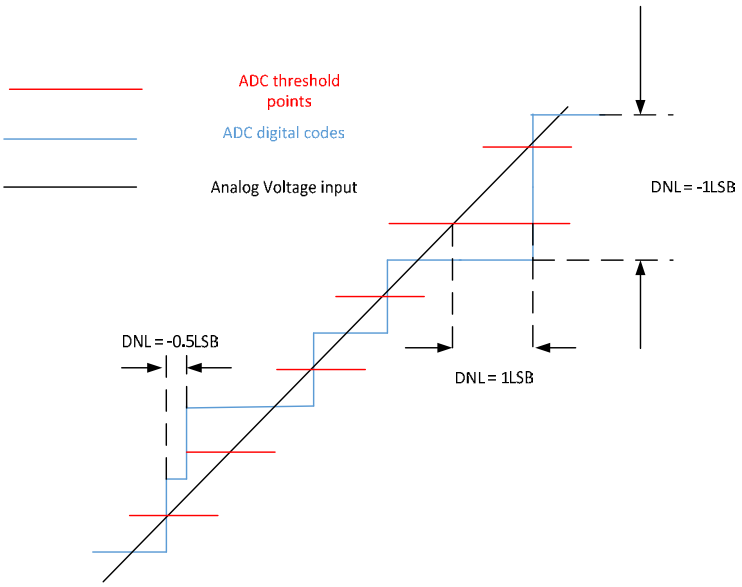



**Figure 24 –An ADC response with errors**

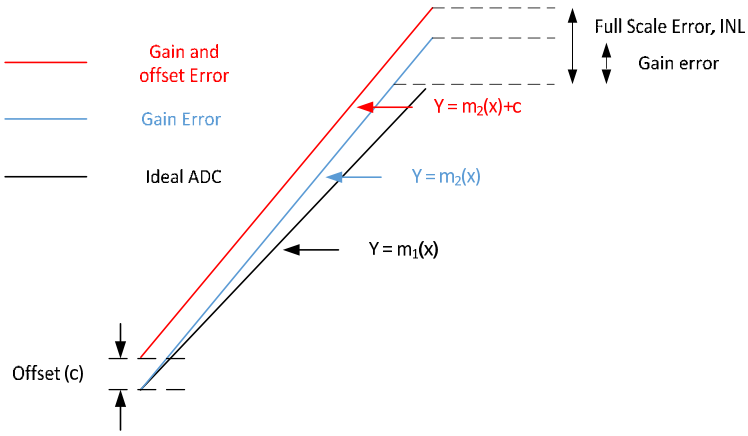



**Figure 25 –An ADC response with offset and gain error**

INL error is described as the deviation, in LSB or percent of full-scale range (FSR), of an actual transfer function from a straight line. The INL-error magnitude then depends directly on the position chosen for this straight line. At least two definitions are common: "best straight-line INL" and "end-point INL" (see Figure 27).

Best straight-line INL provides information about offset (intercept) and gain (slope) error, and the position of the transfer function. It determines, in the form of a straight line, the closest approximation to the ADC's actual transfer function. The exact position of the line is not clearly defined, but this approach yields the best repeatability, and it serves as a true representation of linearity.

End-point INL passes the straight line through end-points of the converter's transfer function, thereby defining a precise position for the line. Thus, the straight line for an N-bit ADC is defined by its zero (all zeros) and its full-scale (all ones) outputs.

The best straight-line approach is generally preferred, because it produces better results. The INL specification is measured after both static offset and gain errors have been nullified, and can be described as Equation (2).

$$INL = \frac{V_D - V_{ZERO}}{V_{LSB\_IDEAL}} - D, where\ 0 < D < 2^N - 1 \quad \textbf{(2)}$$

$V_D$ is the analogue value represented by the digital output code D, N is the ADC's resolution, VZERO is the minimum analogue input corresponding to an all-zero output code, and VLSB-IDEAL is the ideal spacing for two adjacent output codes.

INL and DNL can be measured with either a DC voltage ramp or a low-frequency sine wave as the input. A simple DC (ramp) test can incorporate a logic analyser and a high-precision DC source for sweeping the input range of the device under test (DUT).

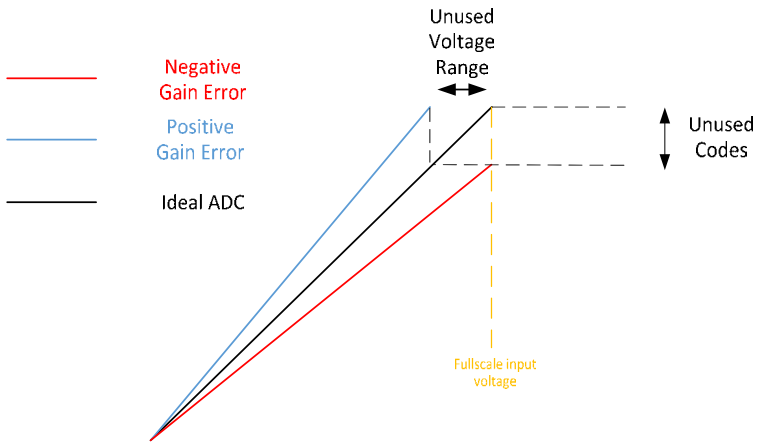



**Figure 26 –An ADC response with loss of range and codes**

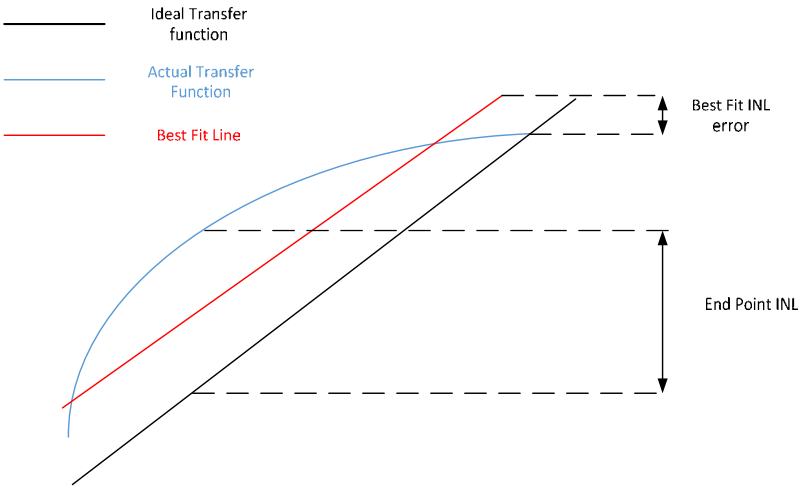



**Figure 27 –An ADC response showing different INL comparisons**

For an ATE system, synchronization is very important to know exactly where the ramp has started in real time, in addition, when to start sampling the output of the ADC. To judge the DNL to an accuracy of so many percentages of a bit, it is required to sample the output multiple times per code and hence the slope of the DC instrument needs to be controlled such that the resolution of the source is stepping up between the code segments. Therefore, if a 10 hits per codes are used, a DC ramp is produced that has a slope of 10:1 and the digital subsystem of the tester samples the ADC output 10 times more than the ADC resolution. Therefore, for an 8-bit ADC, 2560 points would need to be collected. Also as 10 samples are collected per code would result in theoretically being able to guarantee a DNL of 0.1LSB. However, in practice this is not possible, as guardbands and safe guards are needed in production. In this example, it would be possible to guarantee a DNL of around 0.25LSB as in general 4 times the accuracy of the number of codes are needed to be collected to guarantee the specified value i.e. to be able to guarantee a DNL of 0.1LSB, 40 codes per sample would need to be collected. Therefore, for this 8-bit example, to guarantee a DNL of 0.1LSB, 10240 codes would need to be sampled. In addition, the voltage per bit needs to be 40 times more accurate than the code bit width i.e. for an 8-bit ADC with a 700mV FSR, a source with an accuracy of better than 68uV is needed to be able to guarantee the DNL of 0.1LSB.

### 2.9.2 SINAD, SNR & ENOB

Some ADCs perform well only with input signals at or near DC. Others perform well with input signals from DC up to Nyquist frequency. Just because DNL and INL meet the system requirements does not mean the converter will give that same performance

when AC signals are considered. DNL and INL are DC tests. We must look to the AC specs to get a good feeling for AC performance. The key specs to review are signal-to-noise ratio (SNR), signal-to-noise and distortion ratio (SINAD), total harmonic distortion (THD), spurious-free dynamic range (SFDR) and intermodulation distortion (IMD). SINAD is defined as the rms (root mean square) value of an input sine wave to the rms value of the noise of the converter (from DC to the Nyquist frequency, including harmonic content, harmonics that occur at multiples of the input frequency (Figure 28)). SNR is similar to SINAD, except that it does not include the harmonic content. Thus, the SNR should always be better than the SINAD. Equation (3) describes SINAD and SNR and they are typically expressed in dB.

$$SINAD\ (dB) = 6.02 * N + 1.76 \quad \textbf{(3)}$$

where N is the number of bits (resolution). For an ideal 12-bit converter, the SINAD is 74dB. Rearranging equation (3) to be rewritten in terms of N, equation (4), it would reveal how many bits of information are obtained as a function of the RMS noise.

$$N = (SINAD - 1.76)/6.02 \quad \textbf{(4)}$$

Equation (4) is the definition for the effective number of bits, or ENOB

It should be noted that SINAD is a function of the input frequency. As frequency increases toward the Nyquist limit, SINAD decreases. If the specification in the data sheet is tested at lower frequencies compared to the Nyquist frequency, it can be guaranteed that the performance will be much as the input frequency approaches Nyquist. ENOB degrades with frequency primarily because THD gets increasingly worse as the input frequency increases. For example, with a SINAD minimum value of 68dB will result in an ENOB 11 bits. Therefore, 1 bit of information has been lost due to the converter's noise and distortion performance. This means that a 12-bit converter can provide a maximum accuracy of 0.05%.

SNR is the signal to noise and distortion ratio with the distortion components removed. SNR reveals where the noise floor of the converter is located.

### 2.9.3 Total Harmonic Distortion (THD)

Dynamic errors and integral nonlinearities contribute to harmonic distortion whenever an ADC samples a periodic signal. For pure sine wave inputs, the output harmonic-distortion components are found at spectral values whose non-aliased

frequencies are integer multiples of the applied sinusoidal input tone. The amplitudes of the non-aliased frequencies, which depend on the amplitude and the frequency of the applied input sine wave, are generally given as a dB-ratio with respect to the amplitude of the applied sine-wave input. Their frequencies are usually expressed as a multiple of the frequency of the applied sinusoidal input signal.

THD is the rms sum of all harmonics in the output signal's Fast Fourier Transform (FFT) spectrum. All harmonics are included by definition, but the first three represent the major contribution to output distortion in a given converter. THD is described by equation (5).

$$THD\ (dBc) = 20\log_{10}\left(\frac{\sqrt{[{V_{HD2}}^2+{V_{HD3}}^2+{V_{HD4}}^2+\cdots{V_{HDx}}^2]}}{V_{Fund}}\right) \quad \textbf{(5)}$$

where $V_{fund}$ is the rms fundamental amplitude, and $V_{HD2}$ through $V_{HD5}$ represent the rms amplitudes of the 2nd to Nth-order harmonics.

THD values are usually specified in either decibels (dB) or decibels with reference to the carrier frequency or fundamental (dBc).

### 2.9.4 Spurious-Free Dynamic Range (SFDR)

SFDR is defined as the ratio of the RMS value of an input sine wave to the RMS value of the largest spur observed in the frequency domain using an FFT plot. It is typically expressed in dB. SFDR is important in certain communication applications that require maximizing the dynamic range of the ADC. Spurs prevent the ADC from converting small input signals, because the distortion component can be much larger than the signal of interest. This limits the dynamic range of the ADC. It should be noted that a large spur in the frequency domain might not significantly affect the SNR, but will significantly affect the SFDR, see Figure 29.

The term spurious-free dynamic range is usually applied for cases in which the harmonic distortion and spurious signals are regarded as undesirable spurs in the output spectrum of a sampled pure-sinusoidal input tone. SFDR indicates the usable dynamic range of an ADC, beyond which a spectral analysis poses special detection and thresholding problems. Though similar to THD, SFDR addresses the converter's in-band harmonic characteristics.

Spurious-free dynamic range is the ratio of rms amplitude of the fundamental to the rms value of the largest distortion component in a specified frequency range. SFDR is important because noise and harmonics restrict a data converter's dynamic range.

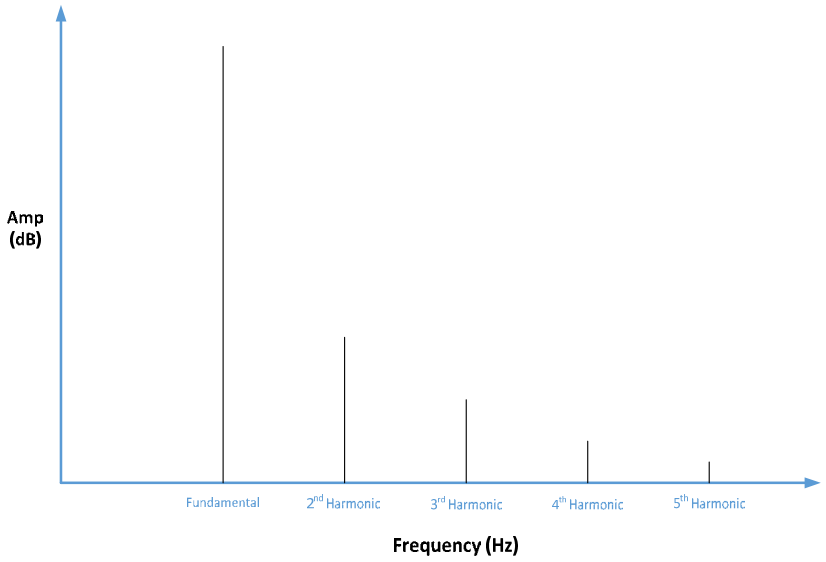



**Figure 28 –An ADC harmonic spectrum with a single tone input**

For spectrally pure sine-wave inputs, SFDR is the ratio of the amplitude of the averaged FFT value at the fundamental frequency, $V_{Fundamental}$, to the amplitude of the averaged FFT value of the largest-amplitude harmonic $V_{HD\_MAX}$, equation (6), or spurious signal component $V_{SPUR\_MAX}$, equation (7), observed over the entire Nyquist band.

In general, SFDR is a function of the amplitude and the frequency of the analogue input tone and, in some cases, even the sampling frequency of the converter under test. Therefore, when testing an ADC for its spurious-free dynamic range, the sampling frequency, input frequency, and amplitude should be defined.

$$SFDR\ (dBc) = 20\log_{10}\frac{V_{Fundamental}}{V_{HD\_MAX}} \quad \textbf{(6)}$$

or

$$SFDR\ (dBc) = 20\log_{10}\frac{V_{Fundamental}}{V_{SPUR\_MAX}} \quad \textbf{(7)}$$

## 2.9.5 Two-Tone Intermodulation Distortion (Two-Tone IMD)

IMD is generally caused by modulation, and it can occur when an ADC samples a signal composed of two or more sinewave signals. IMD spectral components can occur at both the sum ($f_{IMF\_SUM}$) and the difference ($f_{IMF\_DIFF}$) frequencies for all possible integer multiples of the fundamental frequencies. For the two-tone IMD test, the input test frequencies $f_{IN1}$and $f_{IN2}$ ($f_{IN2} > f_{IN1}$) are required to be much lower than the Nyquist frequency ($f_{SAMPLE}/2$).The effect that the two input signals have on the ADC causes

intermodulation distortion. These intermodulation distortion products decrease in amplitude by the power of the actual intermodulation distortion product i.e. the 2nd-order products would have the largest contribution with the 3rd being the 2nd largest and so on. See Figure 29.

In general, each of these frequency components will have a different amplitude and phase, which depends on the specific non-linear function being used, and on the amplitudes and phases of the original input components. This effect can lead to unbalanced intermodulation tones as observed in the characterisation and measurement of the AWG in Chapter 5, 6

2nd-order intermodulation product components are located at

$f_{IN1} + f_{IN2}$, $f_{IN2} - f_{IN1}$ and $2f_{IN1}$ and $2f_{IN2}$

and 3rd-order intermodulation product components are located at

$2f_{IN1} + f_{IN2}$, $2f_{IN1} - f_{IN2}$, $2f_{IN2} - f_{IN1}$, $f_{IN1} + 2f_{IN2}$, and $3f_{IN1}$ and $3f_{IN2}$

and 4th-order intermodulation product components are located at

$3f_{IN1} + f_{IN2}$, $3f_{IN1} - f_{IN2}$, $3f_{IN2} - f_{IN1}$, $f_{IN1} + 3f_{IN2}$, $2f_{IN1} + 2f_{IN2}$ and $4f_{IN1}$ and $4f_{IN2}$

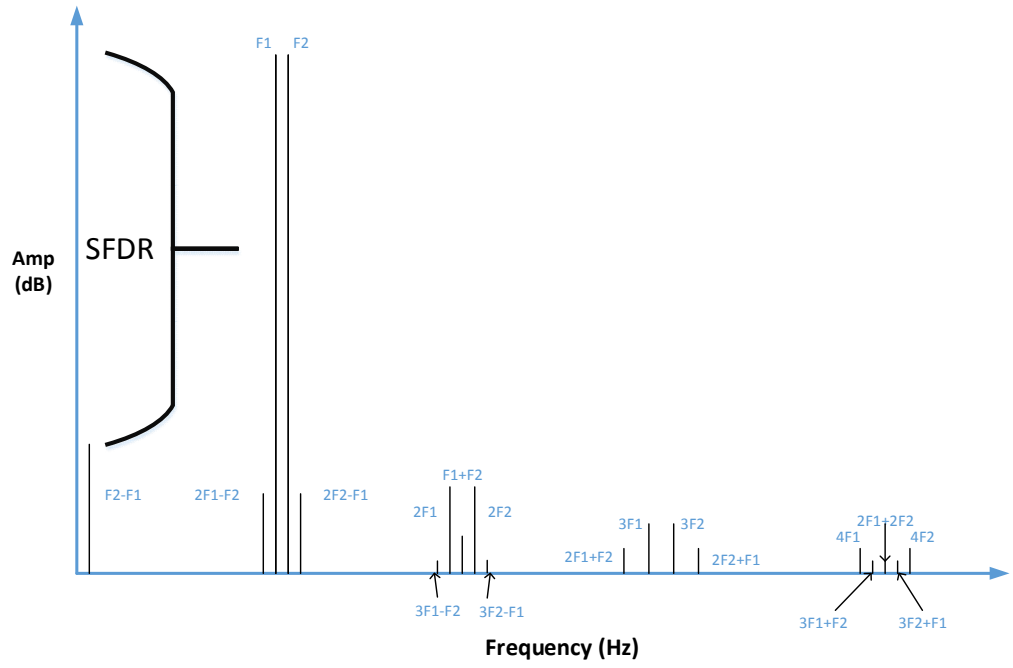



**Figure 29 –An ADC harmonic spectrum with a two tone input**

# 3 AN ATE FILTER CHARACTERISATION TOOLKIT USING CHIRPED EXCITATION SIGNALS AS STIMULI

## 3.1 Introduction

The focus of this chapter is to discuss an innovative ways to characterize filters using different chirp methodologies. The two applications described are a Time of Flight (ToF) device, that has Low Pass Filter (LPF) with a frequency range of 150 kHz and a UHF Near Field Communications (NFC) device, containing several Band Pass Filters, with bandwidths ranging from 100 kHz to 1 MHz. These methods can be used to save test time, save waveform memory space and improve time to market compared to using the standard setup and measurement of each individual frequency component shown in Figure 30. Using a frequency sweep would allow the user to test all the defined test frequency tones in one shot allowing the user to build a characterisation signature of the device in a production test. This would therefore result in a much more test time efficient solution with the added benefit of having additional data available when needed.

The decision to use a chirp as the main engine to solve this problem was due to several reasons; one could easily make a multi-tone as a test stimulus and characterise a filters' response using this approach. Unfortunately, for devices that have a very low input power sensitivity, such as UHF NFC devices with an input level of 60 mV pk-pk (-20 dBm), a multi-tone can cause some issues:

1. The individual frequencies interact, [6], and cause intermodulation distortion products that can affect the overall results in the frequency sweep, refer to Chapter 6.
2. Each individual frequency component has to be divided by the sum of the total number of frequencies so that the overall crest factor, [8-9], or peak to average value is below the maximum input power.
3. As the nonlinearity of a device is proportional to the input value raised to power of the term used, generated by the Taylor series, See chapter 6. Therefore, the higher the input amplitude used will result in a larger distortion value at specified harmonic or IMD component, therefore using a chirp with low input values will also reduce enable a more pure spectral measurement of an RF amplifier compared to using a multitone signal.

Intermodulation is the amplitude modulation of signals containing two or more different frequencies, caused by nonlinearities in a system. The intermodulation between each frequency component will form additional signals at frequencies, [6], that are not just at harmonic frequencies, integer multiples, of either, like harmonic distortion, but also at the sum and difference frequencies of the original frequencies and at multiples of those sum and difference frequencies. The theoretical outcome of these nonlinearities can be calculated by generating a Taylor series, See Chapter 6. Reference [6], shows co-channel distortion caused a multitone signal on a RF device due to 3$^{rd}$ order non-linearity within the IC. By using a chirp, only one frequency is stimulating the device at any one time and hence by using a chirp, the test equipment, and device under test avoids the effects of nonlinearities of the system. Reference [7] shows the effect of a discrete linear chirp on a Power Amplifier (PA), it can be seen that the effect of the chirp has no effect on the passband of the device compared to the ideal even though there is 3$^{rd}$ order non linearity associated with the IC. Therefore, a chirp can be used to produce and error free measurement.

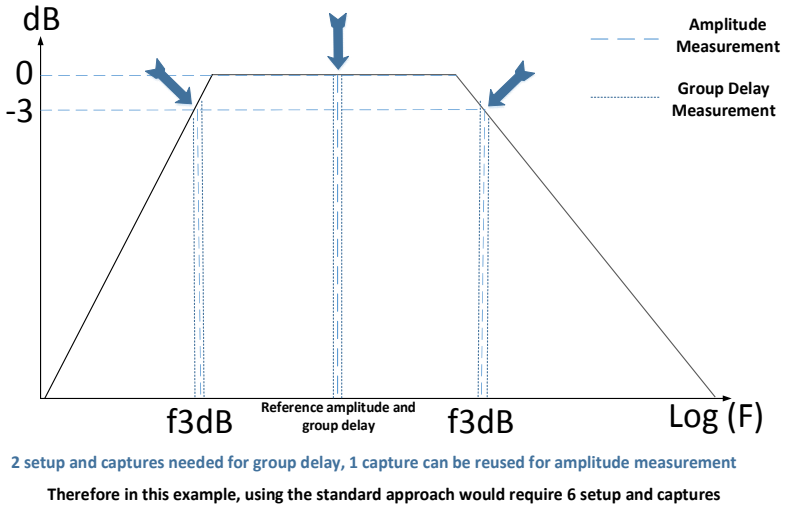



**Figure 30 – Filter frequency response showing the standard measurement analysis**

There are some key parameters of the filters that all designers are interested in, mainly the magnitude response, i.e. 3dB, 10dB, and bandwidth of the filter. Other parameters that are often requested are the phase response and group delay of the filter. In signal processing, the group delay is a measure of the time delay for the amplitude envelopes of the various sinusoidal components of a signal through a device under test, and is a function of frequency for each component.

All frequency components of a signal are delayed when they pass through a device such as a filter. This signal delays are different for the various frequencies unless the device as a linear phase response characteristic.

The delay variation means that signals consisting of multiple frequency components suffer from time-phase distortion because these components are not delayed by the same amount of time at the output of the device. This changes the shape of the signal waveform in addition to any constant delay or scale change. A sufficiently large delay variation can cause problems such as poor fidelity in audio or inter-symbol interference (ISI) in the demodulation of digital information from an analogue carrier signal. High-speed modems use adaptive equalizers to compensate for non-constant group delay, [10].

It is therefore critical for applications where no phase change is a key factor (such as audio and RADAR) to measure the group delay of the circuit if it cannot be guaranteed by design or measurement by other means.

Researching the topic of chirps, one would find several hundred references to use of this method in the simulation of different applications and hence how to generate such signals, [11-14]. Therefore, this concept was investigated further for automated test equipment (ATE) testing, even though its fundamental ideas were described in [13].

## 3.1.1 Used waveforms

### 3.1.1.1 Continuous Linear and Discrete Chirp

What a Linear Chirp consists of is quite simple: in a continuous system, the instantaneous frequency of the signal linearly increases (i.e. increasing from 100Hz to 1MHz) with no frequency jumps. Equation (8) describes a chirp mathematically:

$$F(t) = f_0 + kt \quad \textbf{(8)}$$

where $k = \frac{f_1 - f_0}{T}$

$f_0$ is the starting frequency (at time t = 0), $f_1$ is the final frequency and T is the time taken to get from $f_0$ to $f_1$. k is called the rate of frequency increase or chirp rate.

The corresponding time domain representation for a continuous linear chirp and discrete linear chirp can be observed in Figure 31 & Figure 32. The frequency spectrum of a discrete linear chirp can be observed in Figure 33.

It is also required to understand the phase response of the chirp to be able understand how to process the data at a later stage. Equation (9) describes how the phase of the chirp changes over time as the frequency increases. It can be seen that there is a $t^2$ element and this will be later observed in some of the DSP calculations of phase.

$$x(t) = \sin[\varphi_0 + 2\pi(f_0 t + kt^2/2)] \quad \textbf{(9)}$$

where $\varphi_0$ is the initial phase (at time t = 0). Thus, this is also called quadratic-phase signal.

### 3.1.1.2 Stepped Chirp

Before a chirp can be used to measure filters' amplitude characteristics, the amplitude response in the frequency domain has to be guaranteed to be equal for all frequency components. If this were not true, then a distorted signal would already be present at the source i.e. each frequency of the chirp could have different amplitudes, this can be caused by parasitics present on the loadboard. Therefore, unless some forms of compensation for each frequency were possible at the output, this exercise would not be

achievable without significantly intensive post DSP calculations that could lead to excessive test time. Hence, for accurate amplitude measurements, a signal with equal amplitudes is required. This can be achieved by using a stepped chirp. (Figure 35). Figure 35 shows the FFT of a generated stepped chirp waveform, the specification being from 1 kHz to 100 kHz in 1 kHz steps. What Figure 35 intends to show is the gain flatness of the waveform, a stepped chirp is flat in the frequency domain and hence has no ripple i.e. no peak to average variation or a crest factor equal to 1.4.

#### 3.1.1.3 Application of Waveforms

To save memory, a discrete version of the Continuous Linear Chirp can be created, equation (8), with the same frequency resolution, $F_f$, as the Stepped Chirp waveform. However, this comes at the cost of gain flatness, Figure 33; but still has superior intermodulation distortion characteristics compared to a multione but exhibits similar crest factors when the technique in [13] is used to generate a multitone .

As each frequency arrives at the destination (i.e., the filter), at one Unit Test Period (UTP) at a time and each frequency is coherently sampled, and then the captured signal can be treated as set of individually captured signals and hence a standard FFT can be used for DSP purposes.

Figure 35 shows frequency components at the end of the waveform after 100 kHz. This is spectral leakage of the FFT process as the chirp is not a repeating frequency but an increasing frequency over time. This leakage occurs at the frequency resolution of the chirp, $F_f$, as defined later in equation (11).

For the amplitude response tests, a bandpass filter (BPF) was used as the test vehicle with a discrete stepped chirp, Figure 34. However, for the group delay tests a low pass filter (LPF) was used with a discrete linear chirp. The test setup is shown in Figure 44. This was done in this manner as it was simply due what was investigated on a project-by-project basis within ams, with several years spanning the different projects where the different techniques were developed and implemented.

For the sake of completeness, a comparison of the chirp to a multitone should be drawn. Figure 36 shows how the each individual frequency component of the multitone can be seen as individual signals, in this case only 4 sinewaves. Figure 37 shows how the four sinewaves add and subtract during a period of the mutitone signal, a problem can be seen due to the crest factor of the signal being 3x higher than the individual

frequency component amplitudes. Figure 38, shows an FFT of the composite waveform shown in Figure 37, the 4 individual frequencies can now be seen in the frequency bins. One method to solve the crest factor issue can be to use Schroder, Newman, or Kitayoshi method to adjust the phase of the combined signal to produce a minimal crest factor but maintaining the individual relative signal amplitudes, Figure 39 & Figure 40. We show here some simulation results of employing the Schroder, Newman and Kitayoshi method defined in [13]. Referring to Figure 41, Figure 42 & Figure 43, it can be seen that the amplitudes of all the signals have been sufficiently equalised to produce a crest factor reduced waveform for a signal with 4, 32 and 512 tones respectively. This is achieved by adjusting the phase of each signal proportionally to the maximum number of tones in the signal. For these examples, the crest factor is approximately equal to 1.9 for all.

Another side effect of using a multitone is intermodulation of tones within the signal; this effect could cause errors in the measurement of the device under test. This can be overcome with using lower signal amplitudes to avoid the distortion, however, if a certain spec is needed then lowering the value outside of the required range is not an option and hence distortion will be present. Using a stepped chirp can overcome these issues but at the cost of longer test time.

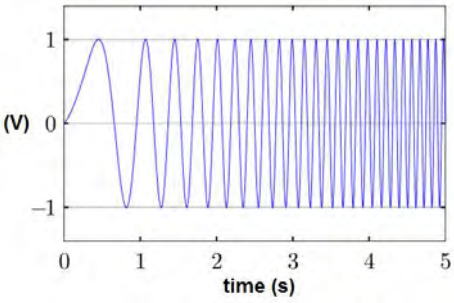



**Figure 31 – A Continuous Linear chirp**

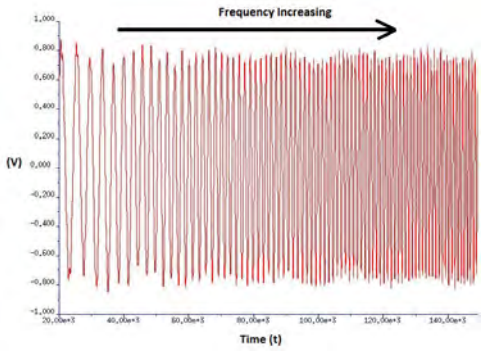



**Figure 32 – Time Domain of Discrete Linear chirp**

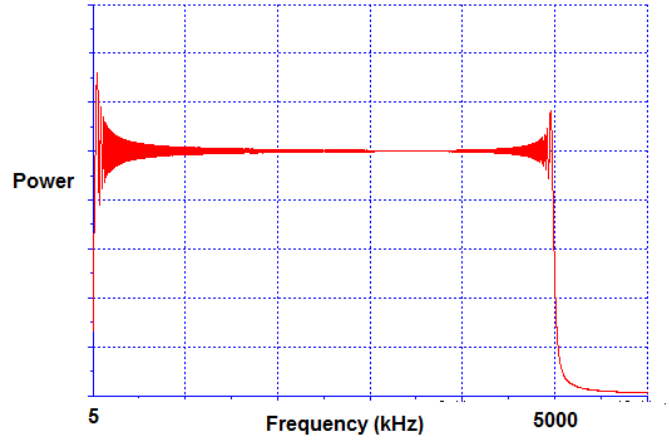



**Figure 33 – Frequency Domain of Discrete Linear chirp**

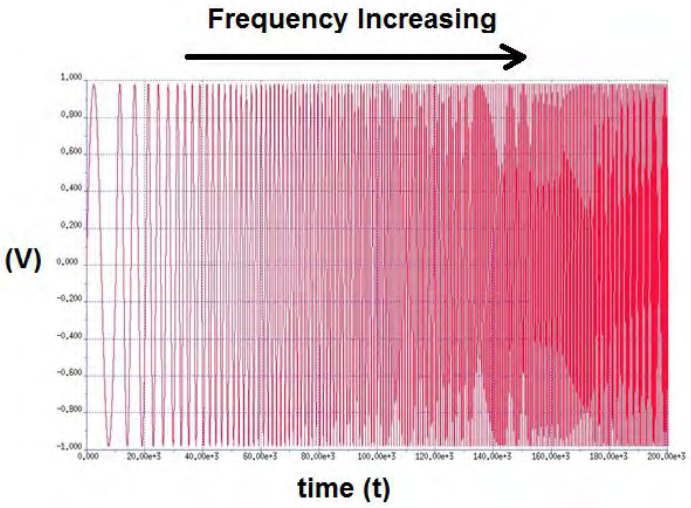



**Figure 34 - Time Domain of a Discrete Stepped Chirp (Crest factor: 1.4)**

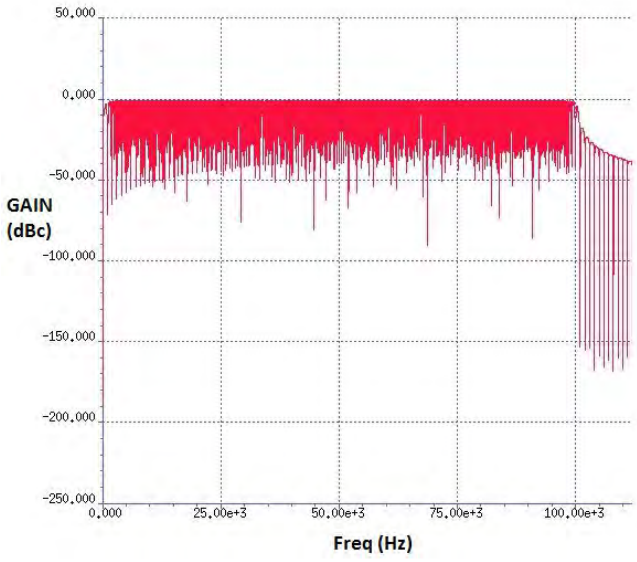



**Figure 35 - Frequency domain of a discrete stepped chirp**

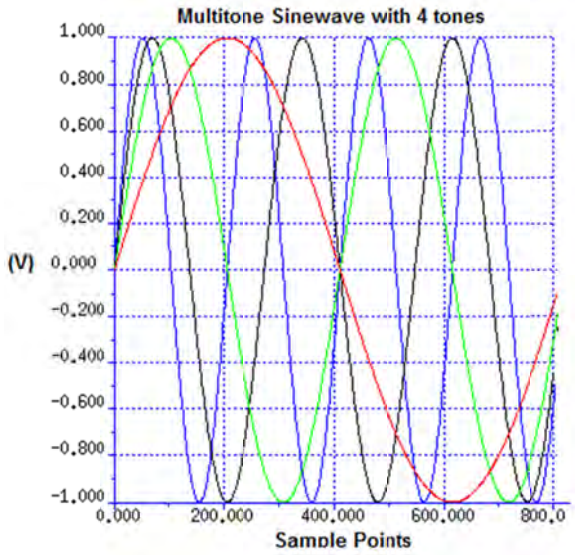



**Figure 36 - A Multitone Waveform - component**

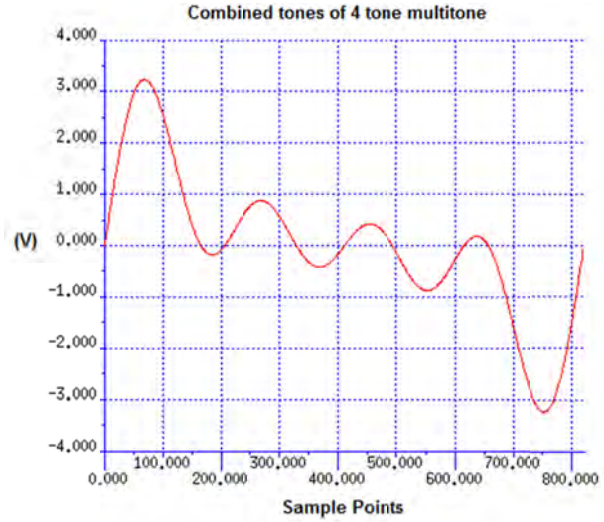



**Figure 37 - A Multitone Waveform – composite**

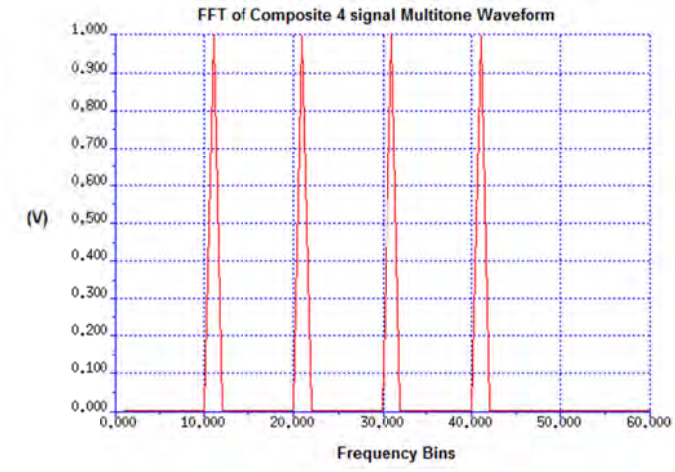



**Figure 38 – FFT of Composite Multitone Waveform**

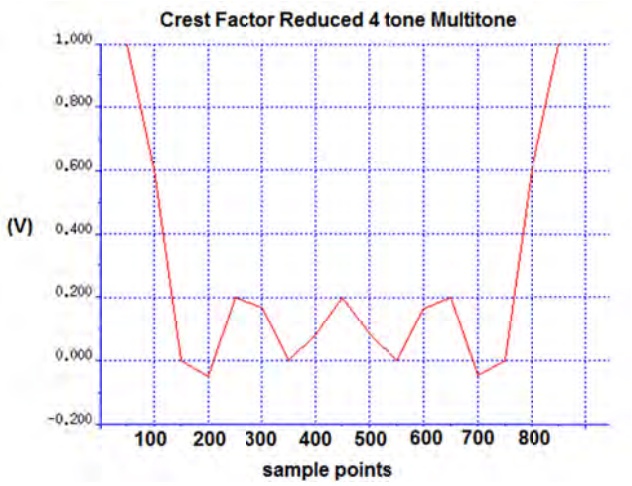



**Figure 39 – A Crest Factor Reduced Multitone Waveform using Schroder, Newman method**

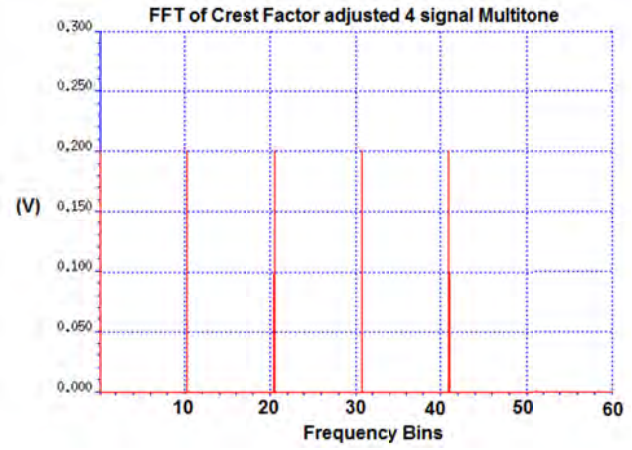



**Figure 40 – FFT of Crest Factor Reduced Multitone Waveform using Schroder, Newman**

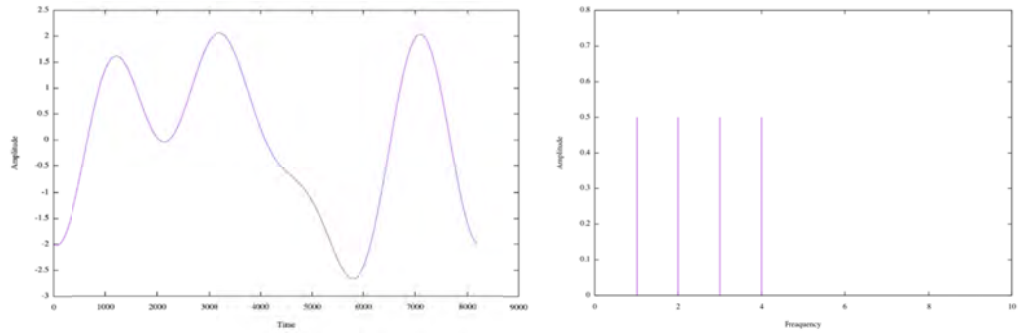


**Figure 41 – Crest Factor Reduced 4 tone Multitone Waveform using H. Kitayoshi algorithm (Crest Factor: 1.89), left time domain, right power spectrum**

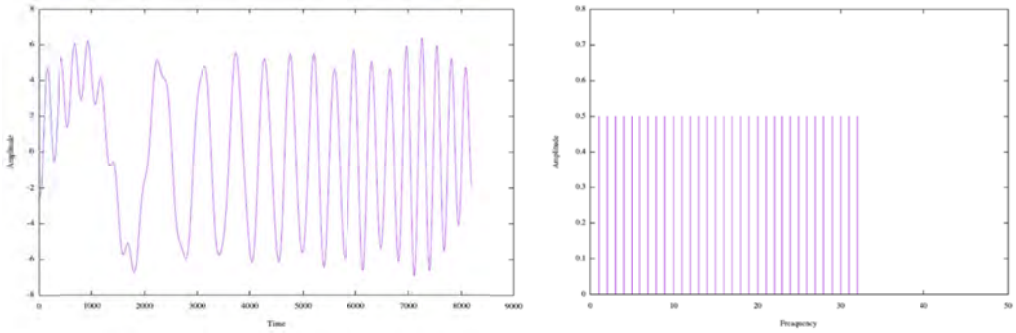


**Figure 42 – Crest Factor Reduced 32 tone Multitone Waveform using H. Kitayoshi algorithm (Crest Factor : 1.73), left time domain, right power spectrum**

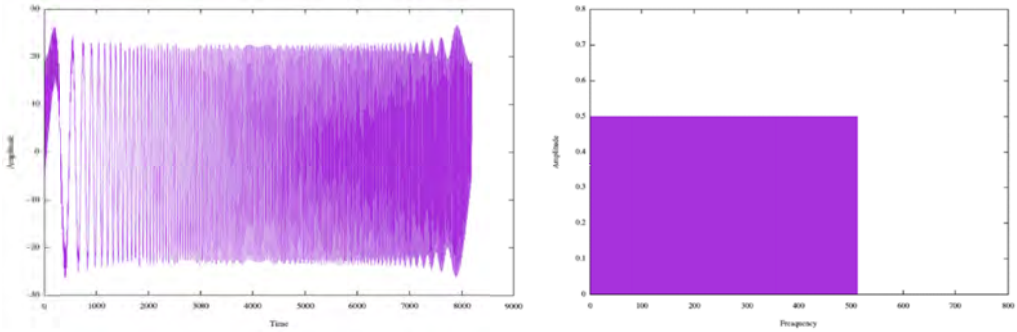


**Figure 43 – Crest Factor Reduced 512 tone Multitone Waveform using H. Kitayoshi algorithm (Crest Factor: 1.66)**

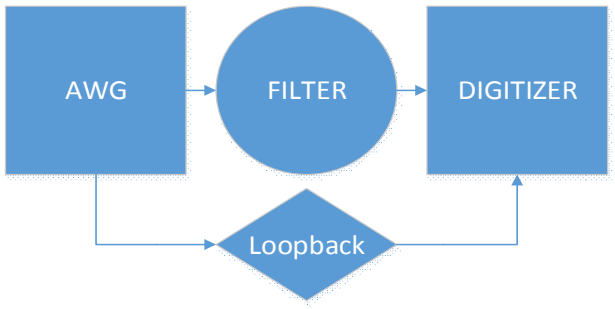



**Figure 44 - Test setup of chirp test of filters**

The author's contribution to the state of the art here is in two parts,

- The use of a loopback switch allows the calibration of the AWG whilst using the chirp signal that allows the removal of the starting phase of the AWG. This loopback path will also remove any distortions originating in the AWG and signal path to the DUT, as this distortion is both common to the device input and loopback measurement.
- By creating a normalised stepped chirp signal, where the signal starts at 1 kHz, the signal can be easily reused within a test program without having to recreate multiple signals that cover similar bandwidth or the signal can be

used at higher sampling frequencies to cover different frequency bands. Multiple different signals can be made for the perfect signal, however; this approach can save AWG memory when limited. This can also be go some way to make IP reuse within a company to help drive faster time to market.

## 3.1.2 Other Chirp Generation methods

### 3.1.2.1 Chirp Generation Using Spectral Warping

Using digital impulses to test the frequency response has many issues, such as the required power the impulse requires to generate enough power in the frequency tones [21,22] for the technique to be valid which can cause damage to the device. However, there are some literatures, [21 - 23], that discuss the use of transforming an impulse response by the means of spectral warping to generate a chirp. Unfortunately, by using this method, the chirp that is produced has frequency components that are not a multiple of a specific sampling frequency and hence any signal generated in this way results in some form of non-coherence at certain frequencies. The authors discuss using windowing techniques to help in coherence, however by using windowing techniques there is loss of signal amplitude. In addition, the signal generated in this way does not produce equal amplitudes in the frequency domain and hence if used to test transfer function of a filter will result in an error if the assumption that the spectrum is equal. Due to the above reasons, this method of generating the chirp described is not suitable for highly accurate filter test metrics.

### 3.1.2.2 Chirp Generation using a Time Stretched Pulse method

Reference [24] describes a method for generating decreasing frequency signal using a inverse FFT for testing of the frequency response of speakers and microphones. This method ensures the power spectrum of a decreasing frequency waveform is flat, however, the issue associated with the spectral warping technique, [21 - 23], will still be an issue here as the signal frequencies are stretched and will result in some form of non-coherence at certain frequencies. Also, [24], has no analysis of the phase performance of the signal as this signal was designed only to test the magnitude response speakers and microphones in the last century. As such, the new methods of applying chirps to filters, as described in this chapter, circumvent the work described in [24].

### 3.1.3 Comparison to previous works

Reference [13,14] discuss using chirps to test audio devices that require a large signal to noise ratio and hence state that using multitones are not sufficient to test these devices efficiently enough. However, there are some similarities to these works to the work undertaken within this study. Reference [13] shows a similar system that is put in place by the author to allow the calibration of the starting phase of the AWG used within the users ATE system. The authors of [14] state that this is needed to do a calibration of the DAC used within their system but no specifics are given as to the actual calibration undertaken. In this aspect, this work is an advancement on this prior work as details of how to make the measurement using this technique are specified.

One of the main reasons to undertake the research into chirp for test of filters was to decrease the test time of these devices under test at ams AG. No such study was undertaken in [13], this work clearly shows the benefits of using a stepped chirp technique in production compared to using standard method of testing individual frequencies sequentially. In this case, one stepped chirp waveform was loaded into memory and used for multiple different filters. This was possible by simply changing the sampling frequency of the AWG and scaling the Frequency Resolution. Repeatability data of the solution was presented, which is not present in [12], which demonstrates the robustness of scaling technique and as such is a advancement on the prior art. The drawback back of using a stepped chirp compared to either a discrete chirp or multione in terms of memory usage is that a stepped chirp occupy a multiple of the number of steps in the chirp compared to either a multitone or a discrete linear chirp and as a result will take the same multiple longer to execute.

This work shows real world examples of using both stepped and discrete linear chirps in production with repeatability data and comparison data to laboratory results. No such comparison to laboratory data for a specific device is shown and as such, this work is an advancement on [13,14]. We have also shown that the stepped chirp can be easily stretched in time by simply scaling the Sampling Frequency of the AWG whereas the discrete linear chirp can save waveform memory at the expense of more complicated waveform generation. Both solutions are available as IP to test engineers at ams AG and as such is state of the art.

Reference [14] shows that the measurement technique used could measure group delay to an accuracy in the millisecond region. The authors work shows that the method

described using a discrete linear chirp and a loopback calibration can produce a stable, repeatable, and reliable measurement in the microsecond region that is probably the greatest advancement on this prior work achieved by the authors work.

### 3.1.4 Chapter Organisation

This chapter is organised in the following way: Chapter 3.2 will discuss how a phase calibration circuit was developed such that the phase dependencies of measurement instruments can be removed such that only the phase change of a signal through a DUT will be seen. Chapter 3.3 will go through the creation of discrete stepped and linear chirps in DSP and discuss in detail how the set of filter characterisation tests are done. This chapter will discuss results from actual silicon and compare them to lab results, and then continue to analyse test stability and reproducibility. Chapter 3.4 will discuss the effect on test time reduction and savings. Conclusions will then be drawn.

## 3.2 Phase Calibration Circuit

The starting phase of an AWG and the measuring instrument (digitizer) can be removed by doing a simple loopback, Figure 45, on the test board where the excitation signal is measured directly by the digitizer, the signal is then switched to the DUT without stopping the AWG and the signal is captured via the DUT i.e. there are two switches. In addition, if there is any amplitude variation present in the signal due to loadboard parasitics, using the Phase Calibration Circuitry, Figure 45, can help to calibrate out the amplitude variation as only the change through the filter will be observed.

It should be noted that by switching in a mechanical relay, that there would be a propagation time through the relay, however, as this is mechanical and not a semiconductor component, the propagation delay is negligible compared to the chirp signal properties**,** therefore no adverse impacts would be observed on the measurement due to this switch. Also, there is the same relay in the path to the device, therefore the same delay is present in the measurement of the filter and hence, even if there is a difference between the propagation delay of the two relays the combined effect will be extremely negligible. This is because the effect is cancelled during the subtraction of the phase response of the filter, and the loopback of the test system.

The control of switch is very simple, as the UTP of the chirp is known and as the AWG is triggered within the test program, it is a simple case of waiting the UTP time plus a small amount before activating the switch. By observing Figure 57, in the middle section labelled “relay switch”, a small part of the filter waveform can be seen before a glitch and then the same loopback characteristic can be seen. This glitch is into the 2nd cycle of the chirp with this extra duration added to the wait time that includes the UTP time as the contributing factor.

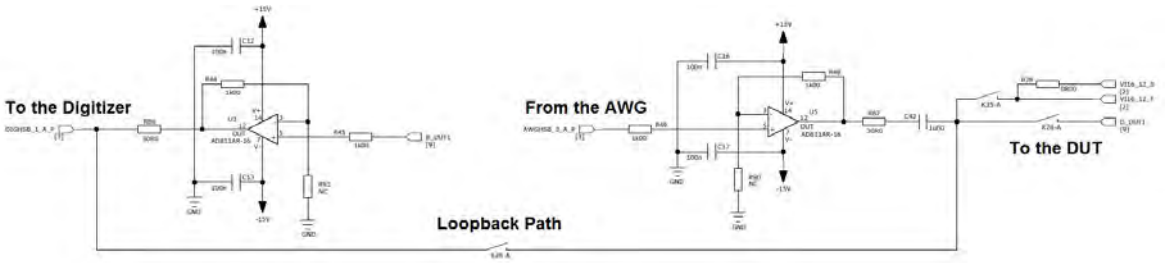



**Figure 45 - Phase calibration circuit**

## 3.2.1 Calibration Circuit Specifications

As most things in industry, many things are taken for granted and are known that they will work just from intuition. The buffer used for the AWG is a video bandwidth amplifier from Analog Devices with very low distortion up to 10MHz and is used for testing RF devices with high signal bandwidth requirements, and as such is ideal for this. The relays used are standard off the shelf components that are used in every day ATE loadboards. Therefore, the circuit could be used to group delay up to 100ns without any additional distortion, however, as the circuit uses a difference methodology, any distortion added by this buffer would be removed during the subtraction process as the buffer is used to decouple the 50R impedance of the AWG from the device.

The choice of using two Single Pole Single Throw (SPST) relay is a simple one but in this instance not critical, compared to Single Pole Double Throw (SPDT) and SPST has a much faster switching time by a factor of greater than 10x. Using a SPDT relay would save space however. As the conducting medium of a relay is metal, the propagation through it is effectively the speed of light and as such, no specification would be listed in the relay specification, as it would be meaningless.

### 3.2.2 Phase Calibration Procedure

To remove the initial phase and capture a clean signal through the DUT, one should consider sourcing three cycles of the excitation waveform (shown in Figure 57) as follows:

1. Capture signal through the DUT; this will eventually be subtracted from 3) so that group delay of the DUT can be calculated.
2. Switch the relay or switch mechanism that allows the signal to be passed to the DUT; this capture will be distorted from the switching of the relay, it is only required to keep the waveform running without stopping the AWG, and hence will not be used for DSP. This intermediate capture is essential to be able to keep the AWG running but at the same time as switching the loopback relay.
3. Capture loopback signal with DUT disconnected; this is the capture of the source signal so that the initial "all-pass" phase delay can be removed

## 3.3 Implementation of Stepped and Discrete Chirp Proposed Techniques

To build a stepped chirp signal in DSP, several parameters need to be understood:

1. What is the required accuracy of the measurement for the corner frequencies of the bandpass filter?
2. Which features of the test instrument support the measurement – memory depth, maximum sampling frequency, waveform stitching?
3. What is the required phase resolution so that a good phase unwrap can be achieved?

With the syntax available in the latest tools for developing test code, such as MATLAB, it is quite straightforward to build a chirp waveform in DSP. Syntax usually provides an option to build a sinusoidal waveform using built-in-functions (BIF) based on the number of samples, sampling frequency, array size, phase delay, and Sinx/x correction. By putting this code into a loop and linearly increasing the bin number, a set of waveforms will be generated with a linear increase in the tone frequency ($F_t$). As long as coherence is guaranteed by ensuring that the coherency theorem, Reference [16], is met, i.e. the signal starts and stops at the same value, then these waveforms can be stitched together to make a stepped chirp signal.

Equation (10) is the coherence condition or sampling theorem, Reference [16], and by using this, the Fourier frequency can be derived as equation (11). For this example, this is the frequency resolution of the waveform, $F_f$ . Hence, each frequency step in a stepped chirp will increase by $F_f$. The UTP of each signal, defined as equation (12), is simply the inverse of the Fourier Frequency, $F_f$. For a stepped chirp, the total UTP will be the number steps in the chirp multiplied by the UTP of one segment, Figure 46. The phase resolution, $\theta_r$, of the waveform can be calculated using equation (13):

$$F_s/F_t = N/M \quad \textbf{(10)}$$

$$F_f = F_s/N \quad \text{or} \quad F_f = F_t/M \quad \textbf{(11)}$$

$$UTP = 1 / F_f \quad \textbf{(12)}$$

$$\theta_r = 2\pi/N \quad \textbf{(13)}$$

where M is the number of cycles, N is the number of samples, $F_s$ is the sampling frequency of the Arbitrary Waveform Generator (AWG) and $F_t$ in is the discrete frequency of the chirp at any one time. If there is no BIF for generating the individual segments, this can be done easily by using the sampling theorem, equation (10), Reference [16]. This requires an array in which M is increased in a linear fashion, Figure 46 & equation (14). To be able to generate a sinewave from this an inverse FFT needs be performed. An FFT works by producing the real and imaginary numbers of a waveform and placing them in an array in the order of the increasing bin number, M, through the array, Reference [16]. By converting these Cartesian numbers to polar format will produce the phase angle of a sinewave at the given frequency associated with the bin number, M. If one element of the two elements associated with the bin number, M, is set to 1 and an inverse FFT is done, Reference [16], then a sinewave is produced i.e. the inverse of doing an FFT. This form of generating a sinewave is much more convenient than using a software “for loop” to generate a sinewave in the time domain.

### 3.3.1 Stepped Chirp

One of the aims of this engineering exercise was to generate a test scheme that the IP could be re-used across multiple projects. The strategy to provide for re-use was to base the scheme on a normalized waveform that goes from 1 kHz to 100 kHz at 1 kHz step. To normalise something is to scale something to one, such as taking the maximum value in a waveform and dividing the whole waveform by the maximum value, such that the

maximum value is now one and the rest of the waveform is scaled accordingly. By normalising the frequency to 1 kHz the frequency resolution can easily be scaled up by up sampling the waveform – that is, by increasing the sampling frequency, $F_s$ of the AWG. This feature is used whenever the frequency range is outside of the 100 kHz range. However, attention needs to be made when making a group delay measurement when using this feature as equation (13) could be violated in that the phase resolution becomes too large to allow an adequate phase change such that the phase unwrap algorithm can no longer be performed accurately.

$$x_\beta = sin(2\pi f_f \beta t)\ 0 < t < 1,\ \beta = 1,2..100 \quad Assuming\ f_f\ is\ an\ integer \qquad \textbf{(14)}$$

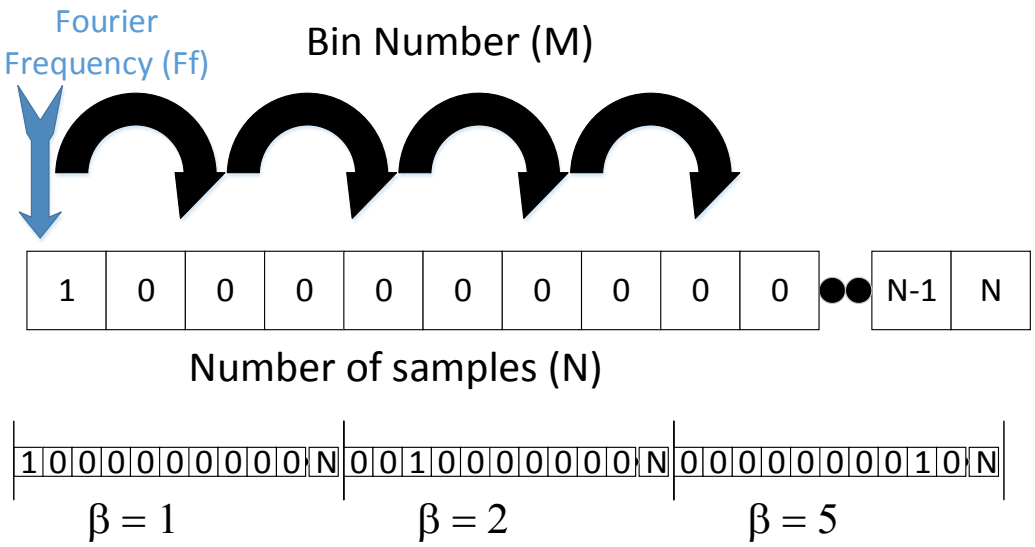



**Figure 46 – Illustration Stepped Chirp step creation – Equation (14)**

Equation (14) details mathematically Figure 46 diagrammatically, how a stepped chirp waveform would be created using a sequence. As described above, the chirp is designed to be stepped by the Fourier Frequency, $F_f$, equation 10, with the starting frequency being $F_f$, so $F_f$ must form the frequency part of the sinusoidal definition. The sequence then creates a series of sinusoids that increases $F_f$ by an integer, $\beta$. The sequence will end with the defined maximum frequencies required, in this case where $\beta$ = 100. $F_f$, in this example is 1 kHz. With a start and stop frequency of 1 kHz and 100 kHz respectively. As a full cycle is applied, t = 1, then each segment,$\beta$ starts and stops at a value of 0. This ensures the segments are coherent with the stepped chirp.

### 3.3.2 Discrete Linear Chirp

To save memory a discrete version of the linear chirp can be produced. By rearranging equation (10) the bin number, M, can be calculated for the start, $f_0$, and stop, $f_1$, frequencies defined in equation (8) as $m_0$ and $m_1$. The frequency resolution, $F_f$, is now simply dictated by the choosing the appropriate array length, N, with the used sample frequency, $F_s$, as defined in equation (11). Using equation (15), a discrete version of equation (8) can be made. As the change in signal frequency is made by incrementing by the frequency resolution, $F_f$, the waveform remains coherent with the measurement instrument using the same sampling conditions. Ensuring that there is a small change between frequency changes will ensure phase jump between frequency changes will not be large and allow a phase unwrap to be done a captured signal using this stimulus.

$$d_n = \cos(\theta_r * (m_0 + k * n) * n) \; n = 1,2..N-1 \qquad \textbf{(15)}$$

where k $= \frac{1+(m_1-m_0)}{2N}$

### 3.3.3 High Fidelity, Discrete Linear Chirp Creation

A high speed AWG in the LTX MX tester has a sampling frequency of 250 MHz, therefore it would be possible to generate a very high fidelity discrete linear chirp signal from 25 kHz to 12.5 MHz such that it exhibits very high phase resolution. This is needed to be able to perform a successfully phase unwrap of the measured signal, [29, 30]. Referring to Chapter 3.3 and using Equations (10-13, 15), a discrete version of equation (8) can be made to modulate the RF Source.

Reference [31] discusses the issues pertaining to phase reversal using Double Sideband-Suppressed Carrier (DSB-SC). Care needs to be taken when generating the chirp, such that when both the mixing signal and the chirp signal pass through zero, phase reversal does not occur as this would cause the phase unwrap process, described above, too fail.

Depending on the difference between the start and stop frequencies within the chirp, ΔF, Equation (16), and the time taken transition between these two frequencies, N*UTP, will result in the chirp waveform exhibiting sidelobes after the waveform start and just before the waveform end. The ΔF *UTP is often referred to as the time-bandwidth product (tbp), Equation (17), of the chirp. Equation (17) can also describe

the roll-off of the sidelobes. As the sidelobes fall away at 3 dB over a tbp interval, the linear chirp flat gain (lcfg) and the roll off can be described as Equation (18) & Equation (19), Figure 48 & Figure 49. Therefore, it is possible to calculate, at any frequency, what the sidelobe power would be and any undesired influence this may cause to a system using this scheme. The longer it takes to transition from the start to the stop frequency results in a larger tbp, resulting in larger sidelobe bandwidth and hence shallower roll off. The resulting flat bandwidth of the chirp will be the bandwidth minus the starting frequency, $2*F_{start}$, Equation (18). $2*F_{start}$ is subtracted to remove the large amplitude change over the 1st tbp as shown in Figure 49.

$$\textit{linear chirp bandwidth (lcbw)} = \Delta F\ (3dB) \quad \textbf{(16)}$$

$$tbp = \Delta F*UTP \quad \textbf{(17)}$$

$$\textit{linear chirp flat gain (lcfg)} = \Delta F - 2*F_{start} \quad \textbf{(18)}$$

$$\textit{sidelobe attentuation} = 3\ dB/tbp \quad \textbf{(19)}$$

where $\Delta F = F_{stop} - F_{start}$, chirp -3 dB bandwidth

As the key requirement of a chirp generated for Group Delay would be to have a small phase change between frequency steps, will result in a small frequency change between points, Figure 97. Referring to Equation (9), it can be seen that the phase change of a linear chirp is a function of sin ($t^2$). In Figure 47, the graph of the sin ($t^2$) function can be observed. It can be seen that the phase changes aggressively at the beginning of the waveform. This is due to t being small and the sine of small number being also small, as t increases as a function of a square, and the sine function increasing as the number also increases means that the starting phase of the signal changes rapidly. This can be observed also in the captured phase signal of the chirp in the measurement later in Figure 97.

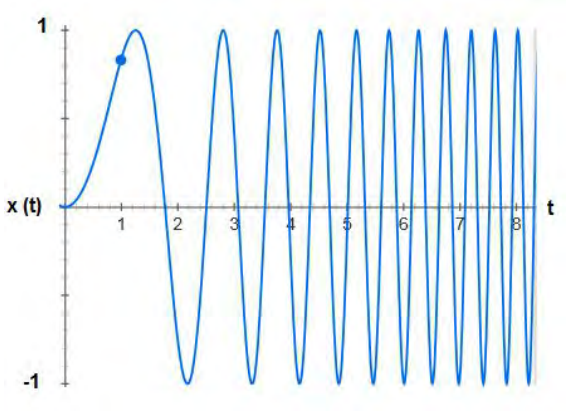



**Figure 47 –Phase characteristic of a linear chirp**

To enable as small a phase change between sample points leads to the requirement of a large number of points, N. Referring to Figure 48, it can be seen that there is a sharp fall off of the sidelobes, at the stop frequency. Using the setup of the chirp creation for the 12.5 MHz frequency sweep, a time-bandwidth product of 25000 can be calculated, as there are 200000 points with a time per sample of 10 ns. By observing the generated chirp in dBc, it is possible to look at the 3 dB bandwidth points. The linear chirp will have a ripple of 3 dB from the start and stop frequencies of the chirp. Referring to Figure 48 & Figure 49, it can be seen that the sidelobe occupies approximately 25 kHz of spectrum bandwidth, 25 kHz at the start and stop frequencies, -3 dB fall off, of the generated 12.5 MHz chirp signal, which is 0.4 % bandwidth. This means that the response is flat from 25 kHz to 12.475 MHz, therefore of the 12.5 MHz of the generated chirp signal, the maximally flat portion is from 25 kHz to 12.475 MHz which is a bandwidth of 12.45 MHz, Equation (18). This is confirmed by using Equations (16-19), with a tbp of 25000 would result in roll off point starting at 12.475 MHz, finishing 3 dB down at the stop frequency of 12.5 MHz. However, if a smaller tbp is used will result in the sidelobe bandwidth occupying less bandwidth and hence, a sidelobe would occupy more bandwidth with a larger the tbp.

The corresponding time domain representation for a discrete linear chirp can be observed in Figure 32. The frequency spectrum of a discrete linear chirp can be observed in Figure 48.

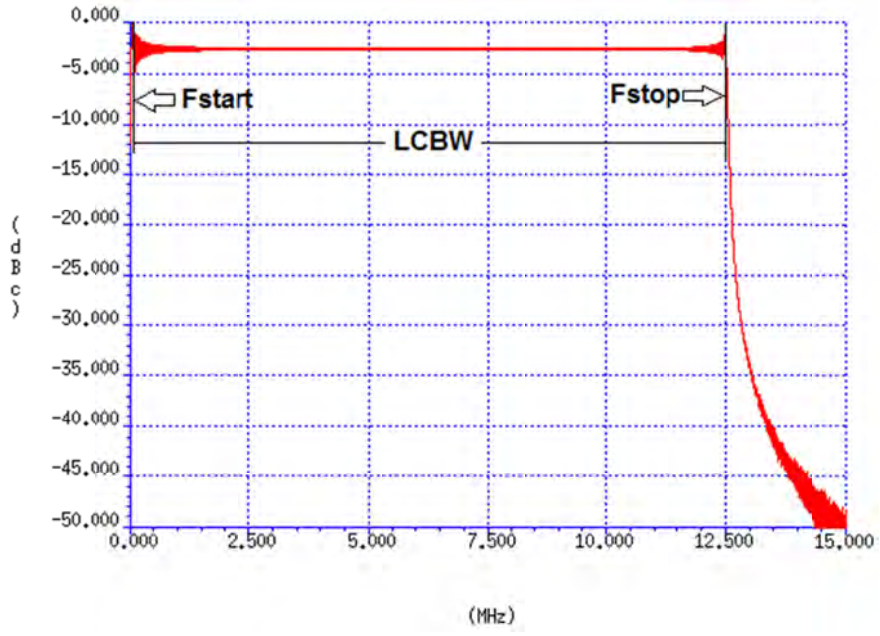



**Figure 48 –Frequency Domain representation of a discrete linear chirp**

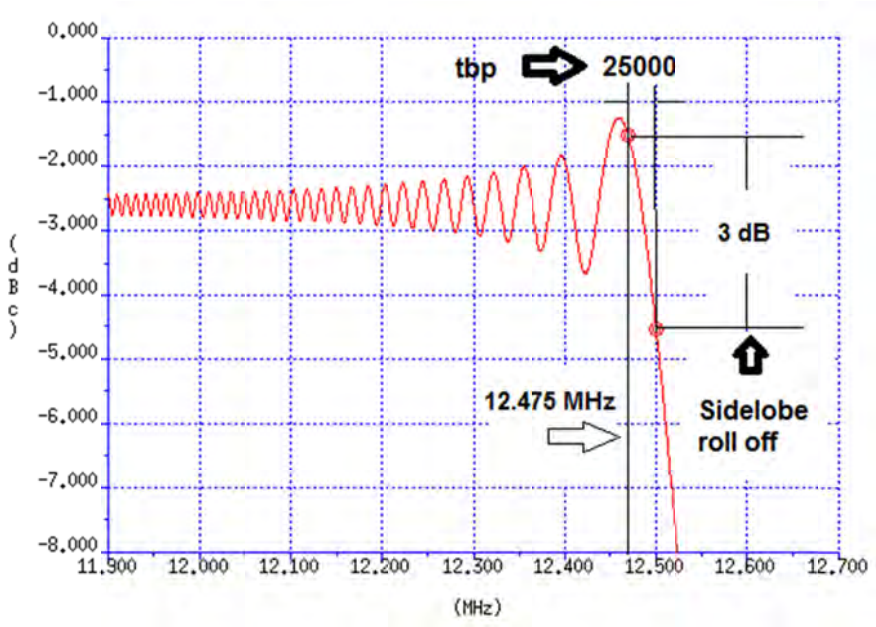



**Figure 49 –Zoom in on the chirp upper sidelobe**

### 3.3.4 Amplitude Measurement Implementation

Using a bandpass filter as an example for this technique, the analysis of the device showed that the worst case filter measurement was for the narrowest frequency range bandpass filter. This filter had a centre frequency of 40 kHz and a bandwidth of around 80 kHz. It was required to measure the low and high 3 dB points to an accuracy of better than 5 kHz. Based on the worst case, a chirp was built that had 100 discrete frequencies with a frequency resolution of 1 kHz. The stepped chirp signal, Figure 34 & Figure 35, was used with the chirp frequency components starting at 1 kHz and finishing at 100 kHz. This signal was used as there is no ripple associated with the stepped chirp as there is with the discrete chirp and as such would present the same amplitude across all frequencies to the filter, i.e. no crest factor. With these frequency components, an amplitude capture of the 40 kHz bandpass filter should be attained with no observable ripple in the pass band.

### 3.3.5 Amplitude Measurement Calculation

Figure 52 shows a one-shot capture of the time domain characteristic curve of a bandpass filter using the chirp created in Figure 34 as the stimulus. The bandpass filter characteristic can now clearly be observed. It is therefore conceivable that mathematics can be developed in DSP to calculate the 3 dB points of the filters cut off frequencies and the bandwidth. After performing an FFT on the captured array and performing a

running average every 6 points on the waveform to smooth out unwanted noise, Figure 53, the exact 3dB point of the filter can be calculated using DSP with the sequence in Figure 51.

Using a feature of equation (11), we can use the same waveform in the memory of the AWG to test higher frequencies. By simply multiplying the sampling frequency, $F_s$, of the AWG by a factor of N, it is possible to test a filter response that has a frequency response with multiple of N higher. Obviously, the multiplying factor has to be within the specified range of the maximum sampling frequency of the AWG. Figure 54 demonstrates this feature. However, it now must be considered that the frequency resolution, $F_f$, is also a multiple of N higher and may not be suitable for group delay measurements due to the phase resolution, $\theta_r$, being too large.

3.3.5.1 Amplitude measurement DSP algorithm

```
ary_len = dimsize(capture_float,1)
capture_move = capture_float

freq_domain = mag_fft (capture_move)
freq_domain[1] = 0.0
running_average( freq_domain ,  freq_domain, 6 )
mag_to_dbc ( freq_domain ,  freq_domain )
freq_domain = freq_domain + attenuation
limit = 1.e-17
vp_clip(limit,  freq_domain, 1,  clip_ary, 1, ary_len/2+1)
xtrm_ary = xtrm(clip_ary)
lowCfreq = xtrm_ary[2]
vp_reverse(clip_ary,1,ary_len/2+1)
xtrm_ary = xtrm(clip_ary)
highCfreq = float(ary_len)/2. + 1.0 - xtrm_ary[2]
filt_band = highCfreq - lowCfreq  -- bandwidth of filter

out_ary[1] = lowCfreq
out_ary[2] = highCfreq
out_ary[3] = filt_band
```

**Figure 50 – Amplitude Measurement DSP code**

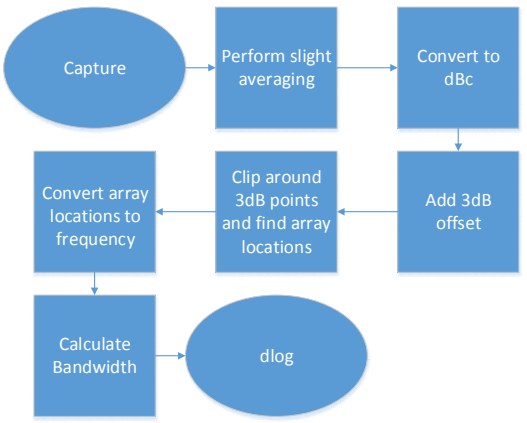



**Figure 51 - DSP algorithm flow for 3 dB point calculation of bandpass filter**

Once an FFT has been done on the captured array, DSP can done on the data to calculate our required parameters, Figure 51 & Figure 50. To find the exact 3 dB point of the filter the following steps in DSP were executed,

1. Perform slight averaging to smooth the array to ensure no sharp changes in the array, make a running average of every 6 points.
2. Converted the captured array from volts to log dbc.
3. Add 3dB to the whole array, -3 dB point is now at 0 dB.
4. Execute clip array around a very small number.
5. Xtrm analysis on the clipped array, this finds the array location of the low pass 3 dB points
6. Reverse array, this puts the high pass part of the filter at the beginning of the array
7. Xtrm analysis on the reversed clipped array, this finds the location of the high pass 3 dB point but it must be remembered that this value should be subtracted from the total number of points.
8. Multiply the array locations by the Fourier Frequency and then the results are the ±3 dB points in Hz, Figure 53.
9. Subtract the 3 dB point values and this result gives the Bandwidth of the filter.

To optimise the execution time of an FFT, in this case we only use a magnitude FFT so that the phase components are ignored and only the magnitude components are returned, this was done because the Group Delay is not of interest in the NFC filter test list. To optimise the test time further, all FFT's are executed in the background of the test program. This means that the calculation is taking place whilst another measurement is being done in parallel. This effectively means that there is no test time associated with calculation of an FFT as they are done in parallel to measurements. This is also sometimes known as interleaving.

3.3.5.2 Amplitude measurement setup

**Table 1 – AWG Test Setup for Stepped Linear Chirp**

| ***AWG Setting*** | ***Value*** |
|---|---|
| Sampling Frequency ($F_s$) | 10.24 MHz |
| Number of Points (N) per segment | 10240 |
| Fourier Frequency ($F_f$) | 1 kHz |
| Number of segments ($\beta$) | 100 |
| Total Waveform size | 1024000 |
| Start Freq | 1 kHz |
| Stop Frequency | 100 kHz |

For this example, the sampling frequency, $F_s$, was set at 10.24 MHz, number of samples per waveform segment, N, was set to 10240. The start and stop frequencies were chosen to be 1 kHz and 100 kHz respectively, which is 100 kHz bandwidth. With a frequency resolution, $F_f$, of 1 kHz, would result in 100 segments. Hence, the resulting chirp waveform size being 1024000 points, see Table 1. As this was an NFC device, with a very low input amplitude requirement with the need to have a very accurate measurement of the corner frequencies of the BPF's justified using a stepped chirped signal as the frequency spectrum of a stepped chirp is flat with no crest factor, As can be seen in Figure 35. However, this resulted in a test time of over 100 ms but with the benefit that the amplitude response of the filter could be tested with high fidelity as will be demonstrated with the measurement results later in this chapter.

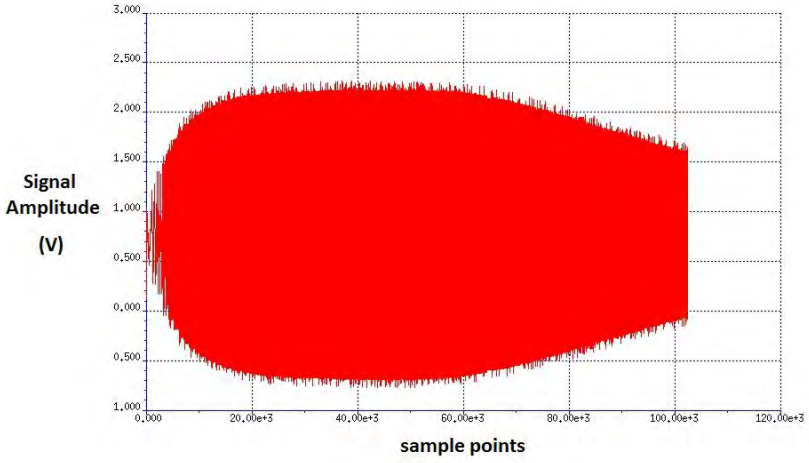



**Figure 52 - Time domain capture through bandpass filter**

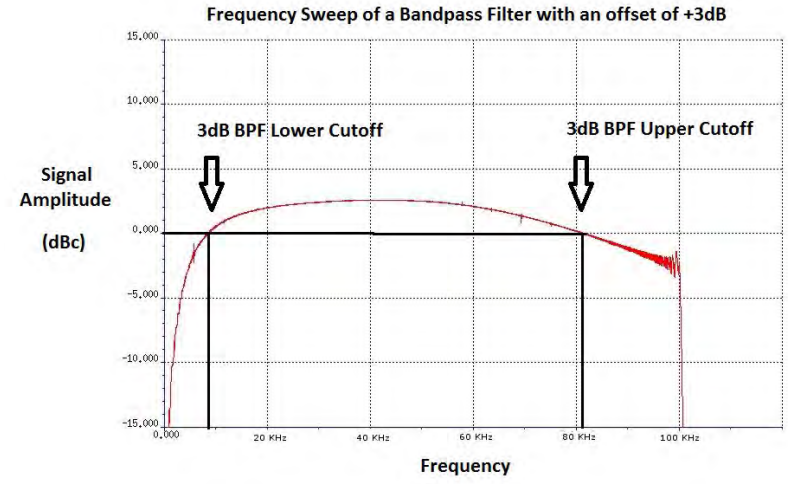



**Figure 53 - Frequency domain capture of bandpass 40 kHz filter**

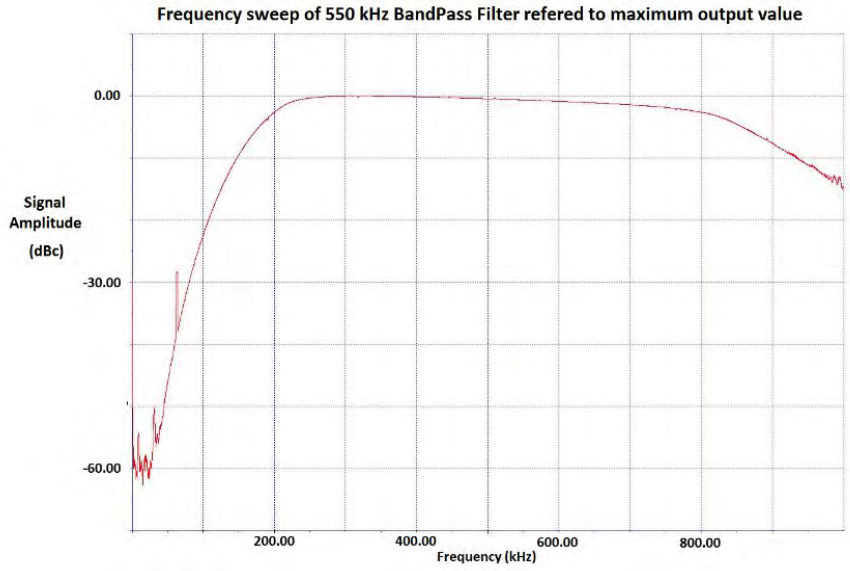



**Figure 54 - Frequency domain capture of 500 kHz bandpass filter**

### 3.3.6 Amplitude Measurement Results & Correlation

A specification is needed to be able to judge if the stepped chip technique was adequate to test filters in production. For these following examples, an accuracy goal of less than 1% is required to be fulfilled to prove that the test technique is production worthy.

Referring to Table 2, it can be seen that there is a very strong correlation between the laboratory test and production test setups. With only a 300 Hz error for both the upper and low 3 dB points, the stepped chirp solution is more than adequate for a production test solution due to its flat response in the frequency domain i.e. no crest factor. Referring to Table 3, it can be seen that there is a larger error; however this is simply because the sampling frequency is increased and hence the resolution of the waveform decreased, i.e. Fourier Frequency increased. As the cut off frequencies of filter, referred to in Table 3, are an order of magnitude higher than the filter referred to Table 2, a larger error is tolerable. In this example, the error is approximately 2 kHz that is more than acceptable for a production test program as absolute accuracy is not required. In the 40 kHz bandpass example, the upper cut-off frequency measurement has a very stable result of 84 kHz with a standard deviation of 35 Hz (Figure 55), resulting in a very stable laboratory correlated test result. It can be observed in Figure 56 that the lower cut off frequency measurement stability of the 550 kHz filter has a similar standard deviation of the 40 kHz filter although the frequency resolution is 10 times lower.

For all measurements taken, the absolute error is less than 1 % of the cut off frequencies of all the filters measured and hence is acceptable for production test. This is because a stepped chirp has no intermodulation distortion products, no crest factor issues, and no ripple on the signal.

**Table 2 - 40 kHz Band Pass Filter 3dB Point Results**

| Test | 40 kHz Bandpass Filter | |
|---|---|---|
| | ***Chirp (ATE)*** | ***Lab*** |
| Lower 3dB point (kHz) | 7.72 | 7.48 |
| Upper 3dB point (kHz) | 83.98 | 84.32 |

**Table 3 - 550 kHz Band Pass Filter 3dB Point Results**

| Test | 550 kHz Bandpass Filter | |
|---|---|---|
| | ***Chirp (ATE)*** | ***Lab*** |
| Lower 3dB point (kHz) | 196.6 | 198.3 |
| Upper 3dB point (kHz) | 822.0 | 819.3 |

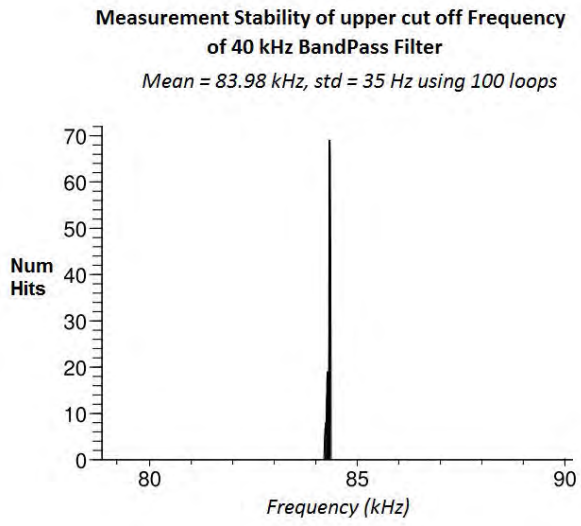



**Figure 55 - Measurement stability histogram of the 40 kHz BPF filter upper cut off frequency**

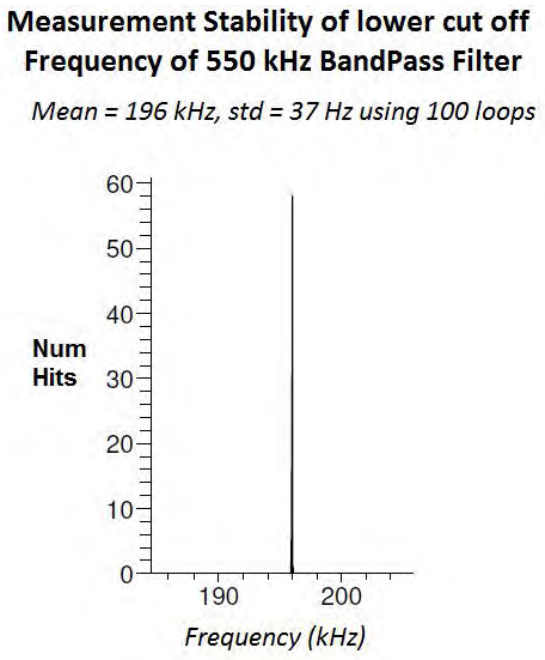



**Figure 56 - Measurement stability histogram of the 550 kHz BPF lower cut off frequency**

## 3.3.7 Group Delay Measurement Implementation

Equation (20), group delay, $T_g$, is a useful measure of time distortion, and is calculated by differentiating, with respect to frequency, the phase response of the device under test (DUT); the group delay is a measure of the slope of the phase response at any given frequency. Variations in the group delay can cause signal distortion, just as deviations from linear phase cause distortion.

$$T_g = -d\varphi / d\varpi \quad \textbf{(20)}$$

where $d\varphi$ is the change in phase and $d\varpi$ is the frequency aperture or frequency change. In this case, this frequency aperture is the frequency step size of the chirp that is for this example the Fourier frequency of the signal, $F_f$, (defined in equation (11)) or the frequency resolution.

The starting, "all-pass", phase of an ATE AWG is unknown and if this delay is not removed from the calculation, the group delay will indicate the propagation delay through the system plus the frequency response dependent phase derivative. For this example, a discrete chirp was used.

Group delay is measured in production by measuring the phase change between two frequencies, one that is the reference and the frequency of interest, if multiple

frequencies are required then the phase change can be measured at multiple points and compared to the reference such that the group delay can be calculated between these points. The loopback circuitry is still required to take out the influence of the starting phase of the AWG as well as the relay switch. Therefore, to do a full sweep using the chirp does not add much to the overhead, but much is gained in that the full data set is available in one sweep and test time is saved by doing one shot capture.

### 3.3.8 Group Delay Measurement Calculation

To be able to initiate the group delay calculation, two captures from the continuous acquisition need to be extracted that contain the filter and loopback information (Figure 57 & Figure 58). This will facilitate DSP on the array to extract the phase information for each frequency of the chirp.

Now for the phase unwrap to work accurately, it is critical that the size of the integration step or phase resolution, $\theta_{r,}$ is high resolution, i.e., short steps between phase changes. Therefore, by using the Fourier frequency of the waveform as the frequency step, all requirements of the DSP are inherently fulfilled.

$$V = X + jY \quad \textbf{(21)}$$

The phase information is extracted by performing an FFT on the data arrays giving the real and imaginary components, Equation (21), for each signal frequency in the array. These signal components or complex number representation of a vector, Figure 59, needs to be extracted for each individual frequency. Once the real and imaginary components of the signal are extracted, the phase can be calculated for each individual frequency by doing an Arctan(Y/X). This will produce an array of results that visually shows a phase changing over frequency; however, this is useless for analytical purposes as the phase jumps, $\theta_w$, every rotation of π (Figure 62).

The next step is to phase unwrap the overall phase response to produce real values,$\theta_{uw}$ for both the filter (Figure 63), and the source signal (Figure 64). This is achieved by simply detecting if the sample-to-sample difference in the array is greater than or less than $\pi$, if so, the phase jump value in the unwrapped array has $\pm 2\pi$ added to it until the next phase jump, k, is detected, then the process is repeated.

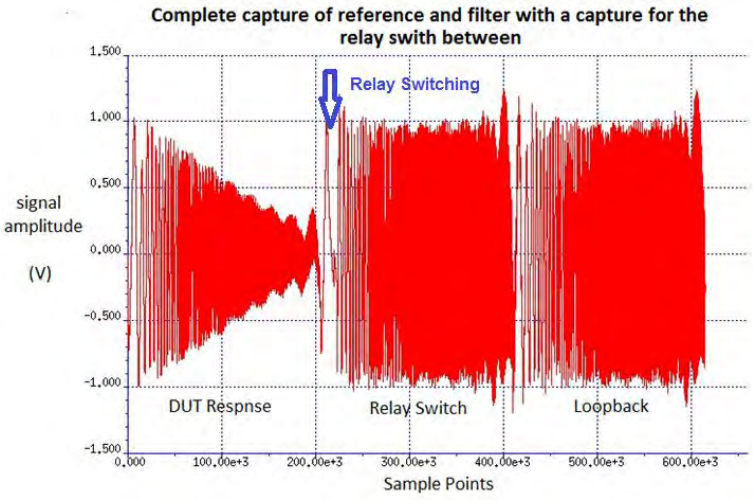



**Figure 57 - Capture of continuous waveform**

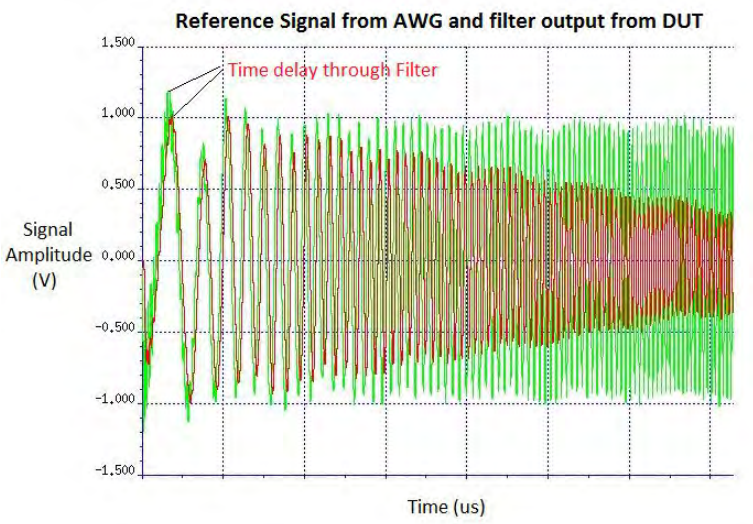



**Figure 58 - Capture of continuous waveform**

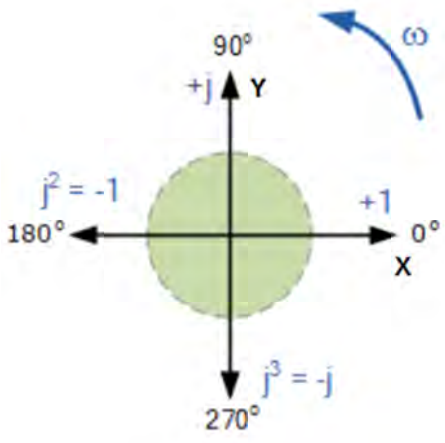



**Figure 59 – Vector diagram represented as complex numbers**

$$\theta_{uw}(f) = \begin{cases} \theta_w(f) + 2\pi k & \Delta\theta_w > \pi\,\Delta f \\ \theta_w(f) - 2\pi k & \Delta\theta_w < -\pi\,\Delta f \\ \theta_w(f) & otherwise \end{cases} \quad \textbf{(22)}$$

There are many different algorithms to phase unwrap an array of phases with phase jumps, [15]. However, there are generally only certain algorithms available between different software applications. For this example a Schafer's phase unwrap based algorithm, [16], was used, equation (22). This algorithm limits the discontinuities of a vector of consecutive phase angles by monitoring the change in angle along the array. This method depends very critically on the size of the integration step $\theta_r$.

3.3.8.1 Group Delay measurement DSP algorithm

Once an FFT has been done on the loopback or filter array, DSP can done on the data to calculate the required time related parameters,

1. Pick out the Real and Imaginary components from the FFT.
2. Calculate the phase, Arctan, of each frequency component of FFT, the frequency step is the Fourier Frequency $F_f$, Figure 62.
3. Phase unwrap the result of the phase calculation to ensure a continuous phase change, equation 21, Figure 63 & Figure 64.
4. Differentiate the array of phase values by dividing by the linear ϖ gradient of the chirp to give the all pass time of the signal through the filter and in loopback. Figure 65 shows both the reference and filter on the same plot.
5. Subtract the filter all pass time response from the filter all pass time response to get the group delay through the filter, Figure 66.

```
i_values      = 0.0
q_values      = 0.0

in_size       = dimsize(in_fft,1)

size          = num_freq -- in_size/2-1
mag_values[1] = in_fft [1]

vp_pick (in_fft[3:], 1, 2, i_values, 1, 1, size)
vp_pick (in_fft[4:], 1, 2, q_values, 1, 1, size)

psize = size + 1
vp_atn2 (i_values, 1, q_values, 1, phase_values, 1,  size)      -- phase of each component

vp_phase_unwrap_rad (pi_local, phase_values[1:size],1, i_values[1:size], 1,  size)

vp_line(zero, max_val, q_values, 1, num_freq)                   -- array of 2 * PI * frequencies
vp_div(i_values, 1, q_values, 1, time_values, 1, num_freq)     -- progressive time delay

time_values = abs(time_values)
```

**Figure 60 – Time Measurement DSP code**

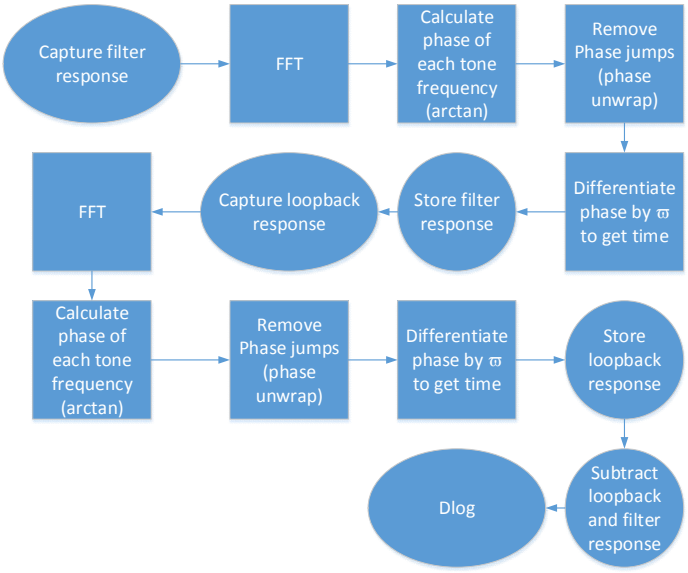



**Figure 61 – Group Delay Measurement flow chart**

3.3.8.2 Group Delay measurement setup

Going back to initial setup of the chirp, it was decided to produce a high-resolution frequency steps size, $F_f$, so that the characteristic curve of the filter could be measured. Equation (11) says that the larger the number of points, N, the lower the Fourier frequency, $F_f$, hence the finer the frequency resolution.

Observing the results from the phase unwrap of the reference in Figure 63 & Figure 64, and filter phase plots, one can see that the phase is changing non-linearly over time as equation (8) shows.

It has been previously suggested in, [17], that the phase needs to be differentiated twice to remove the effect of the square on the phase to get the group delay. This is in fact not correct for the two arrays are seeing the same constant phase change over time, hence performing a differentiation of both reference and filter arrays as shown in Figure 65, then making difference measurement (Figure 66) will produce the actual group delay through the DUT, [20]. As the AWG and digitizer have been running continuously and by using the loopback from AWG to the digitizer any group delay effect associated with the source and measurement instruments has subsequently been cancelled out.

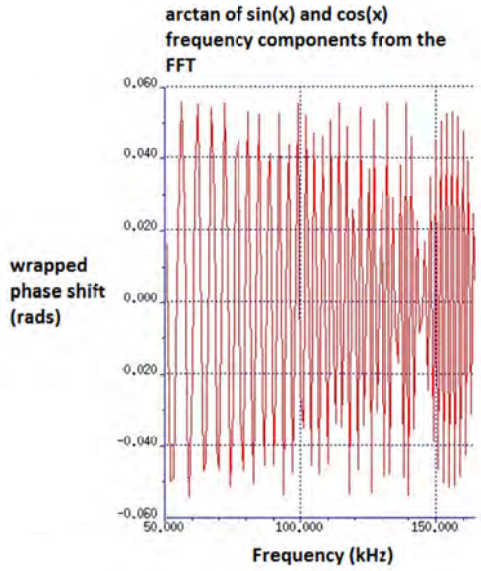



**Figure 62 - Phase change of filter without phase jumps removed**

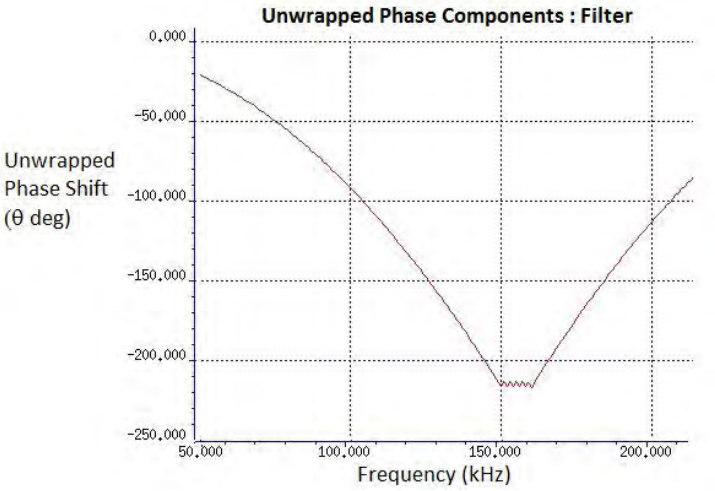



**Figure 63 - Phase change of filter**

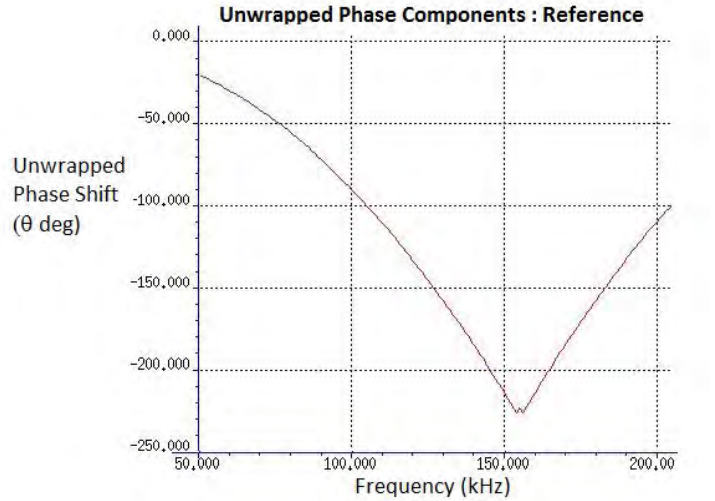



**Figure 64 - Phase change of reference**

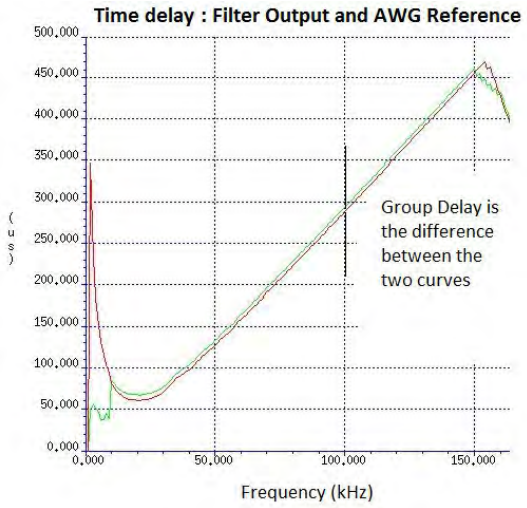



**Figure 65 - Time change of both output of filter and reference**

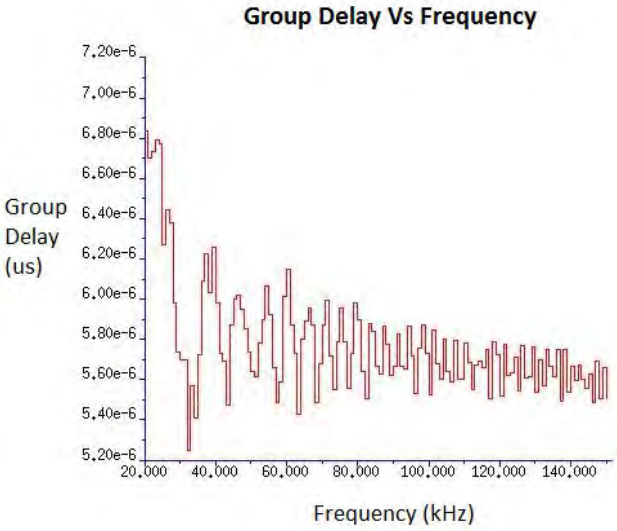



**Figure 66 - Group Delay of the Filter – Band 1**

Differentiating the two phase arrays, equation (20), results in the time array with the "all-pass" effect still present (Figure 65). It can be seen that the two curves match with a small shift, this small shift being the group delay. Hence, the group delay can be calculated by simply subtracting the filter output and reference waveforms, equation (23), Figure 66.

$$\frac{\delta\theta ref - \delta\theta filter}{\delta\omega} = group\ delay \qquad \textbf{(23)}$$

The differentiation was performed by simply dividing the phase response by$\varpi$. This is possible due to the spectral density of the chirp in the frequency domain thus allowing an approximate differentiation by a simple subtraction. In DSP, this was implemented by drawing a line from the start to end frequency and multiplying by 2π. The phase

array was then divided by this gradient, ϖ. See Figure 60 & Figure 61 for Group Delay measurement code and flow diagram.

Using a sampling frequency, $F_s$, of 102.4 MHz and a required Frequency Resolution, $F_f$, of 500 Hz results in the needed Number of Points, N of 204800 for the discrete linear chirp creation, See Table 4.

**Table 4 – AWG Test Setup for Discrete Linear Chirp**

| *AWG Setting* | *Value* |
|---|---|
| Sampling Frequency ($F_s$) | 102.4 MHz |
| Number of Points (N) | 204800 |
| Fourier Frequency ($F_f$) | 500 Hz |

### 3.3.9 Group Delay Measurement Results & Correlation

Figure 66 shows a typical group delay curve. However, this requires comparison to the expected results from the design simulation environment, Cadence Spectre, and the measured results of the characterisation from the lab. For this DUT, two filters were measured using the chirp group delay technique; one can see by comparing the results in Table 5 and Table 6 that the correlation is good to a few hundred nanosecond error between the ATE solutions using the chirp technique and the conventional method, Figure 30, used in the lab. This is where only a reference and point of interest phase are measured and the difference calculated. Considering the variables involved in measuring group delay and the influences of different hardware and parasitics, this is an excellent correlation.

For a production worthy solution, the test result has to be repeatable; depending on how repeatable the result is in comparison with the limits of the test dictates the yield the test solution will produce. Hence, there would be no point in having a faster test solution if the yield were to be lower due to poor repeatability and increased guardbands.

Figure 67 shows the repeatability of the chirped group delay test. One device was run 100 times to get a sense of the repeatability. It can be seen that the stability of the measurement is 0.1 % of the 5.4 us group delay that is more than adequate for a production measurement. This is a result of having very good phase resolution due to the small signal frequency spacing within the chirp. Having such repeatable result enables a very production worthy solution.

Figure 68 shows the reproducibility between two different test contact sites on the ATE test board. It can be clearly seen that the result is independent of the test site used; hence, the solution is hardware independent. This comes in part due to the feedback loop using the loopback relay. Any difference in the waveform on the input would be reflected on the output leaving only the result of the device itself. What can be seen is that the standard deviation or stability of the measurement on the second site has increased by around three times to 20 ns that is still only 0.3 % of the measured group delay, however for this example this is insignificant.

**Table 5 – Band 1 Group Delay Results**

| Freq (kHz) | Group Delay (us) | | |
|---|---|---|---|
| | *Simulation* | *Chirp (ATE)* | *Lab* |
| 50 | 5.4 | 5.7 | 5.6 |
| 100 | 4 | 5.6 | 5,5 |
| 120 | 3.2 | 5.5 | 5.4 |

**Table 6 – Band 2 Group Delay Results**

| Freq (kHz) | Group Delay (us) | | |
|---|---|---|---|
| | *Simulation* | *Chirp (ATE)* | *Lab* |
| 40 | 7.0 | 6.7 | 7.57 |
| 70 | 6.5 | 5.5 | 6.7 |

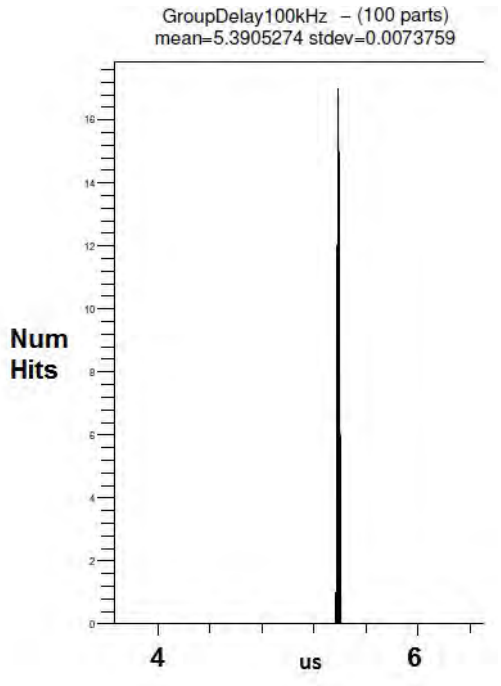



**Figure 67 - Repeatability of the group delay chirp solution with ATE.**

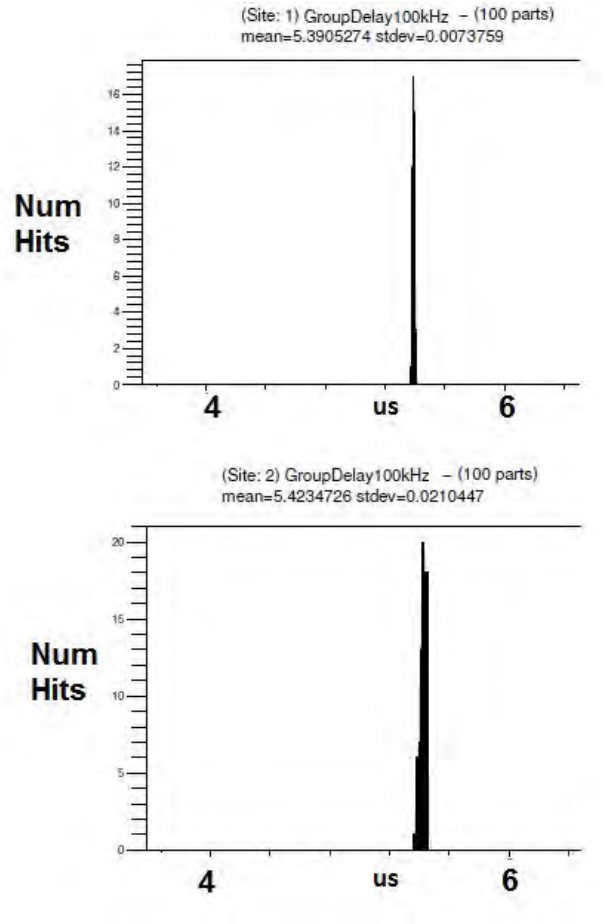



**Figure 68 - Reproducibility of the group delay chirp solution with ATE**

## 3.4 Test Time Saving

The original test program of the bandpass filter code used discrete sequentially applied frequencies with the AWG and Digitiser start and stopped 5 times. The test time for one filter test was 250 ms. With the start, stop and wait times for the instrument being fixed then it is easy to appreciate that a one shot capture will be quicker and that there is more information available from the test. Using a chirp with one start and stop the test time is dependent on the UTP time of the source that for this example is 100 ms. Hence this is example is 2x more efficient using the chirp compared to the original. However if it would be possible to reliably use a multione with 100 tones then the test time would approximately be 100x faster as all the tones would be applied instantaneously. However, this is not feasible due to the low voltage input of the NFC device that would result in the individual signal levels of the multitone being below a level that the AWG could produce accurately. However as a stepped chirp was used, a very accurate input signal voltage was present on the input of the device, for each

frequency change, that resulted in extremely accurate and repeatable measurement of the device filters.

However, if a comparison were made to a multitone; again, in theory, the test time of the multitone would be 150x faster as all 150 tones used in the chirp would be instantaneously supplied in the multitone. However, this is also not feasible in reality due to the low voltage input requirements of the device that would result in the individual signal levels of the multitone being below a level that the AWG could produce accurately. In addition, to make the group delay measurement requires 3x the test time of just the amplitude measurement as the starting phase of the AWG needs to be calibrated as shown in Figure 30.

## 3.5 Conclusion

This chapter has presented that by using either a stepped or discrete linear chirp signal with a small phase change between the frequency steps of the chirp, a good correlation can be achieved for filter characterisation purposes. If a stepped chirp signal is designed in such a manner that it can be normalised, the same waveform can also be used for higher frequencies with lower resolution based on the manipulation of the sampling frequency of the AWG, making this approach extremely flexible and suitable for characterisation and testing application to a wide range of analogue/RF devices.

Using the discrete linear chirp technique with loopback results in the most efficient and cost effective way to characterise a filters phase response that is available to the ATE test engineer of today, in the authors opinion it cannot be bettered.

In the future, RF components will be going into autonomous vehicles and thus the need to test these group delay specifications in RF devices will become much more necessary, rather than using alternative test methods [2-4]. For today, using a chirp for group delay measurements gives reliable, accurate and reproducible results that are not only test time efficient but also give full characterisation data that is priceless to today's semiconductor integrated circuit designer.

# 4 MEASURING GROUP DELAY OF FREQUENCY DOWNCONVERTER DEVICES USING A CHIRPED RF MODULATED SIGNAL

## 4.1 Introduction

In today's market place, group delay specifications of Radio Frequency (RF) receivers are generally not tested in production due to the prohibitive nature that these tests exhibit. In addition, as these specifications are usually guaranteed by design, the tests of such parameters are generally not needed. In general, catastrophic defects in a LPF can be detected by simply testing the amplitude response; however, latent defects that cause a small change in the amplitude response can cause a larger effect in the group delay and hence need to be caught. Also, this goes a long way to improving the quality of devices in the field as over time a latent defect will become a catastrophic defect causing the entire system to fail. Therefore, in the future, as automobiles incorporate more RF receivers and as many of these devices will be going into autonomous vehicles, and as described in 2.8, all datasheet parameters of an automotive device have to be tested in production; therefore the need to test these group delay specifications in RF devices will become mandatory.

Comparing techniques for measuring group delay using conventional Vector Network Analysers (VNA's), such as the Aeroflex 6480 Microwave System Analyser provide phase or group delay data relative to a golden device [25,26]. Other companies such as

Keysight technologies have developed their own calibration standards for downconverters using three broadband standards, a power meter as a magnitude standard, a comb generator as a phase standard, and an S-parameter calibration kit [27, 28]. Other techniques such as using a transmitter with the same characteristics as the receiver under test have been used so that the same frequency range can be used with a VNA to measure the group delay directly [25,29]. There have been efforts made to measure the group delay of a system without having the system phase locked by using a comb generator [28], however this is not feasible at ATE as no such calibrated standard is available on an Automated Test Equipment (ATE) test system. There has also been a drive to use a two-tone signal where measuring the phase change between the two tones at both baseband and RF will indicate the group delay through the device [29]. The problem with using [29] is the problem of how to measure the phase of the RF input signal. At the end, the user is restricted to down converting the input by means of the testers RF measure, which results in the inability to calibrate a part of the signal path due to the weaknesses of the standard RF ATE tester infrastructure (Figure 70). In addition, using the two-tone option of an ATE would be very slow to sweep across the band. This is because only two tones are applied at the same time and thus to build a complete picture of a device then the two tones would need to be swept across the band that take much time due to the settling time of the generators. However, if the idea in [29] is expanded upon it, then a scheme for measuring and characterizing group delay in production could be developed. The fundamental idea in [29] is that the phase of the first frequency is not important to measure but the phase delta between two frequencies. By modulating an RF source with a two-tone signal would give controllability over the phase would only require one generator and an Arbitrary Waveform Generator (AWG), where the Digital to Analog Converter (DAC) resides, to do the same as described in [29] but with less cost and more flexibility. However, a two-tone signal between a large bandwidth could have a large error. The technique in [29] can be further improved if a high-resolution chirp is used to modulate an RF generator. To use such standards and techniques [27-29] is dependent on the architecture of the test equipment available to an ATE test engineer and for the majority of the ATE testers available today; one possible solution would be to use a golden sample correlation method to work around the weaknesses of the standard RF ATE tester infrastructure.

One major benefit of an ATE test system is that all instruments are phase locked to the tester 10 MHz reference clock, Figure 71. The main items to note in Figure 71 are

the SMATE, which are RF Generators, SMA100A, which is used as the Local Oscillator (LO) for down-converting RF signals. The Digitizer (DIG), which is used to capture analogue Baseband (BB) signals or digitize a down-converted RF signal and the Arbitrary Waveform Generator (AWG), which is used to supply analogue signals to a Device Under Test (DUT) or to modulate an RF Source. There are two RF bricks has 4 port modules each with 2 ports, labelled A + B, per port module. As there are 4 RF generators, this allows the possibility to source 4 RF signals simultaneously.

Another feature of the RF Source is that it is possible to apply a modulation waveform to the IQ modulation inputs of the generator, Figure 92. For a limited band sweep, such as 25 MHz,

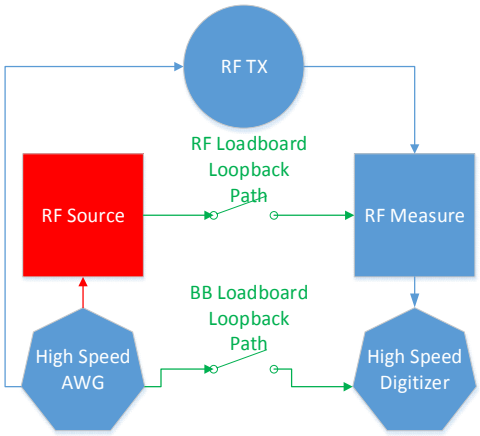



**Figure 69 – RF tester setup showing the weakness with calibrating the full RF path whilst testing RF transmitters.**

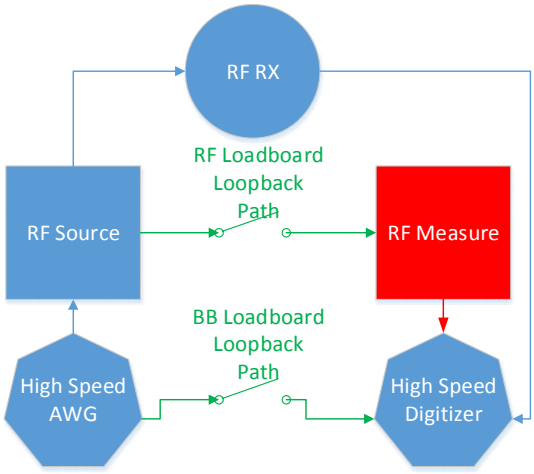



**Figure 70 – RF tester setup showing the weakness with calibrating the full RF path whilst testing RF receivers.**

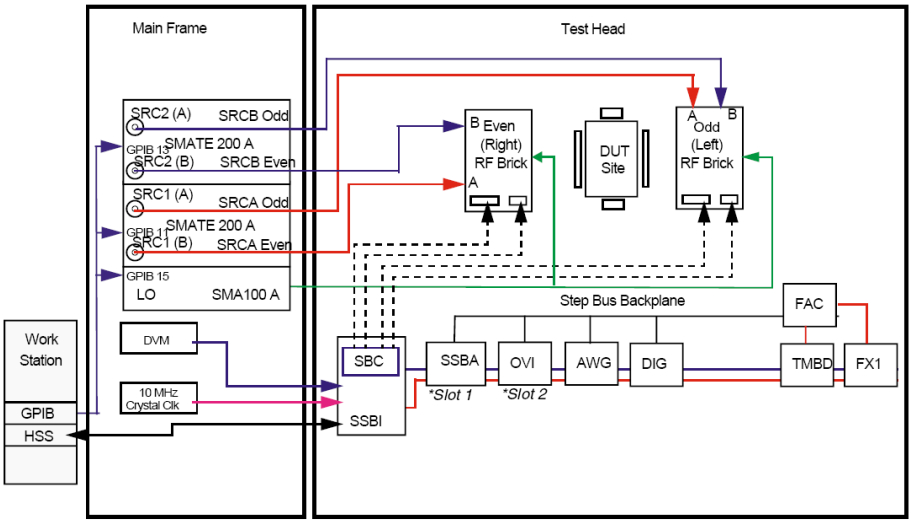



**Figure 71 –Standard RF Tester Overview**

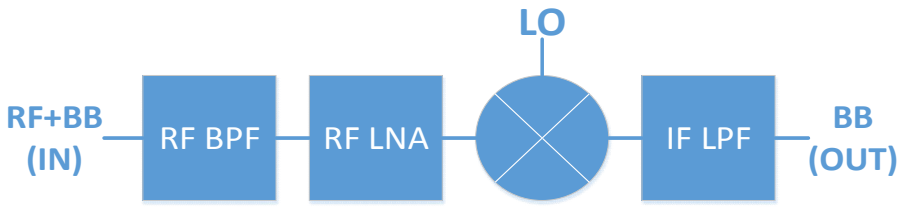



**Figure 72 –RF RX Block Diagram**

it would be possible to chirp [17,20,29,30] the RF source frequency hence sweeping the frequency of the generator in a fast and efficient manner compared to using a two-tone setup [28].

By utilizing the benefits of the test system and using a proven technique from the laboratory, it would be possible to produce a reliable and equivalent test on ATE. This would result in high confidence of the end quality that is the starting requirement of the shipped devices to end customers in the automotive industry for use in autonomous vehicles where a slight error in a communication could result in a fatality.

One of the features of a chirp is its distortion properties. The distortion seen at the carrier frequency, $\varpi 1$, is too far out of band to interfere with the signal integrity, therefore, distortion properties of a RF amplifier can be ignored.

In this chapter, we will discuss how to implement a Group Delay test on a Very High Frequency (VHF) or Ultra High Frequency (UHF) RF Receiver, Figure 72, using todays RF ATE tester installed base. Figure 72 shows a simplified block diagram of a RF Receiver. The device consists of a RF bandpass filter (BPF), RF Low Noise Amplifier

(LNA), Mixer and a BB, Low Pass Filter (LPF). The main two direct contributors that could cause a group delay failure of the system described would be due to the LPF and BPF. As these components primary function is to alter the phase at specific frequencies, if there is a defect that changes this characteristic, it could be possible that a group delay failure could go undetected by testing only the magnitude response. However, we will also show, through simulation, that if the non-linearity components of the mixer are not tested that this could have a tremendous impact on the measured group delay due to distortion of the signal.

This chapter is organized in the following way, Chapter 4.2 will describe the basics of mixer theory, and how a chirp as described in [29, 30] can be used to modulate an RF generator to be able to sweep a RF downconverters input. Chapter 3.3.3 will discuss how to generate a discrete linear chirp; Chapter 4.3 will describe the effects of a chirp on a non-linear system, Chapter 4.4 will describe the effect of a defect on a filters performance and how the chirp technique can detect the fault. Chapter 4.5 will describe the architecture of a standard RF tester, and the issues pertaining to making an RF group delay test of an RF downconverter. Chapter 4.6 will detail some measurements using the RF tester in loopback to verify the technique and then conclusions will be drawn.

## 4.2 RF Frequency Sweep

A side effect of the mixing process can used as an advantage and be used to save some AWG bandwidth. By using the chirp waveform, as described in [29, 30], for every step of discrete linear chirp waveform we would observe the frequency of the RF Source Generator increase in steps of the chosen Fourier Frequency, Equation (11), up to and including the maximum frequency in the chirp waveform. A side effect of the mixing process is the creation of an image of the modulation signal at the negative side of the RF signal, f1 - f2, as well as in the positive, f1 + f2, Equation (24), Figure 74.

$$cos\,\omega_1\,cos\,\omega_2 = \frac{cos[\omega_1+\omega_2]+cos[\omega_1-\omega_2]}{2} \quad \textbf{(24)}$$

where cos(ω1t) and cos(ω2t) are the time-domain representations of f1 and f2. ω = 2πf

The ½ factors simply show that the input amplitude is divided between the two output terms. By placing the RF frequency of the generator in the centre of the measurement band, it is possible to sweep either side of the RF carrier frequency. Therefore, if we

would like to sweep to the maximum possible frequency using this technique using the standard tester configuration, we would need to create a chirp that can be measured by the maximum Digitizer rate.

$$F_s \text{ (Digitizer)} > 2F_t \text{ (AWG) or } F_s \text{ (Digitizer)} = 4F_t \text{ (AWG)} \quad \textbf{(25)}$$

The Sampling theorem dictates that to oversample a signal correctly, we need at least 2 points per cycle, Equation (25), however, in reality; slightly more is needed to avoid the Nyquist frequency itself. A better rule is 4 points per code. This means that we are limited to measuring to a maximum frequency of 25 MHz if the maximum sampling rate of the digitizer is 100 MHz. As we will get a doubling of the signal bandwidth due to the mixing effect, we can generate just a 0 to 12.5 MHz chirped signal.

For the modulated chirp, Equation (15) needs to be substituted for ω2 in Equation (23) which results in Equation (26), Figure 73.

$$s(\omega 1t) = d(t)\cos(\omega 1t) \quad \textbf{(26)}$$

Where $\omega 1$is the carrier frequency and d$(t)$ is the DA conversion output of $d_n$, Equation (15)

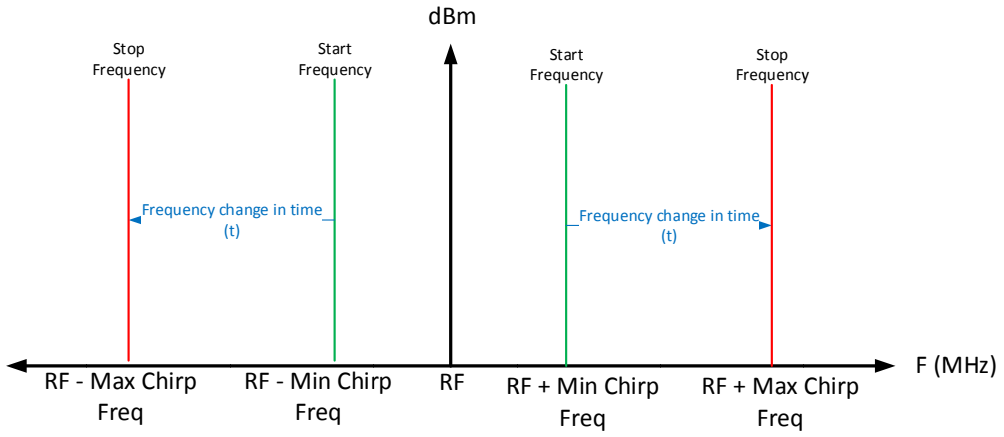



**Figure 73 –Frequency Spectrum of a chirped RF source without quadrature signal on the IQ inputs.**

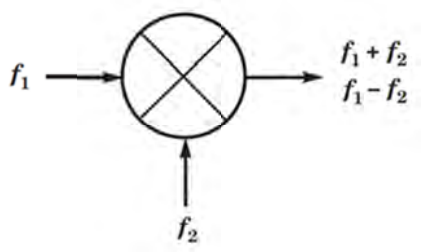



**Figure 74 –Frequency Mixer schematic diagram**

## 4.3 Understanding the Effects of a Chirp on A RX Model with Distortion

Suppose that the PA in the RX, Figure 72, has the 3rd, 5th, 7th, 9th, and 11th order distortion.

$$Y_{pa}(n) = a_{1\_pa}D_{in}(n) + a_{3\_pa}D_{in}(n)^3 + a_{5\_pa}D_{in}(n)^5 + a_{7\_pa}D_{in}(n)^7 + a_{9\_pa}D_{in}(n)^9 + a_{11\_pa}D_{in}(n)^{11} \quad \textbf{(27)}$$

where $a_{1\_pa}$ = 10000,

$a_{3\_pa}$ = -500 x 10-6,

$a_{5\_pa}$ = -15.0 x 10-6,

$a_{7\_pa}$ = -18.0 x 10-6,

$a_{9\_pa}$ = -22.0 x 10-6,

$a_{11\_pa}$ = -4.00 x 10-6.

and the mixer has also the 3rd, 5th, 7th, 9th, and 11th order distortion.

$$Y_{mix}(n) = a_{1\_mix}D_{in}(n) + a_{3\_mix}D_{in}(n)^3 + a_{5\_mix}D_{in}(n)^5 + a_{7\_mix}D_{in}(n)^7 + a_{9\_mix}D_{in}(n)^9 + a_{11\_mix}D_{in}(n)^{11} \quad \textbf{(28)}$$

where $a_{1\_mix}$ = 0.5,

$a_{3\_mix}$ = -500.0 x 10-9,

$a_{5\_mix}$ = -50.0 x 10-9,

$a_{7\_mix}$ = -5.00 x 10-9,

$a_{9\_mix}$ = -0.50 x 10-9,

$a_{11\_mix}$ = -0.05 x 10-9.

and is stimulated with a chirp that has start and stop frequencies of 0.1 Hz and 1 kHz respectively, Figure 75. If this chirp were then modulated on to a carrier, $f1$, of 25 kHz, Equation (26), with low power, so that it can be amplified by the RX amplifier, would result in Figure 76. What can be seen here is the lower and upper sidebands created due to the mixing process and result in effectively doubling the bandwidth of the chirp, Equation (26). Now when this signal is amplified by 40 dB, with the distortion properties, Figure 77, as defined in Equation (27), we can no visible distortion to the

chirp relative to the non-distorted modulated signal. This is due to the fact the distortion at the carrier frequency, $\varpi 1$, is too far out of band to interfere with the signal integrity, therefore, distortion properties of a RF amplifier can be ignored. Now, the model would down-convert this modulated signal by multiplying the modulated signal by the carrier frequency + maximum frequency of the modulated signal. A mixer has also distortion properties; we have modelled this with Equation (28). Figure 78 shows a distorted signal, as defined in Equation (28), which is down-converted, plotted against a non-distorted signal. What we can see is there is a great deal of outband distortion and that the signal in the inband has been corrupted. This is due to the mixer folding in frequency components into the inband as part of the odd intermodulation distortion components, i.e. $3^{rd}$, $5^{th}$, $7^{th}$, $9^{th}$ and $11^{th}$. This is different to the amplifier case as there is no mixing component so there is no intermodulation distortion. To ensure this the proposed technique is sound, the proposed technique has to be immune to such distortion, as this is inherent to all silicon devices in different forms. This signal would then be filtered using a 20 order FIR LPF. Comparing the down-converted signal, Figure 79, with the low passed signal, it can be seen that the filter has removed a significant amount of the outband noise due to the distortion caused by the Mixer. However, we need to see how the inband distortion affects the group delay of the system at a later stage, Figure 82 & Figure 83. If we now compare what the original modulation would look like with no distortion compared to the distorted version, Figure 80. We see that although the chirp clearly has some distortion but the overall shape of the chirp is maintained. Therefore, using such a waveform with a device with distortion properties would still be permissible.

Equation (20), group delay, Tg, is a useful measure of time distortion, and is calculated by differentiating, with respect to frequency, the phase response of the device under test (DUT); the group delay is a measure of the slope of the phase response at any given frequency. Variations in the group delay cause signal distortion, just as deviations from linear phase cause distortion. In this case, this frequency aperture is the frequency step size of the chirp that is for this example the Fourier frequency of the signal used for the digitisation of the measured signal using the digitiser, $F_f$, (defined in Equation (11)) or the frequency resolution. For measured results of a system, the group delay of the signal needs to be subtracted from the measured group delay that results in the group delay of the system, as defined in Equation (23).

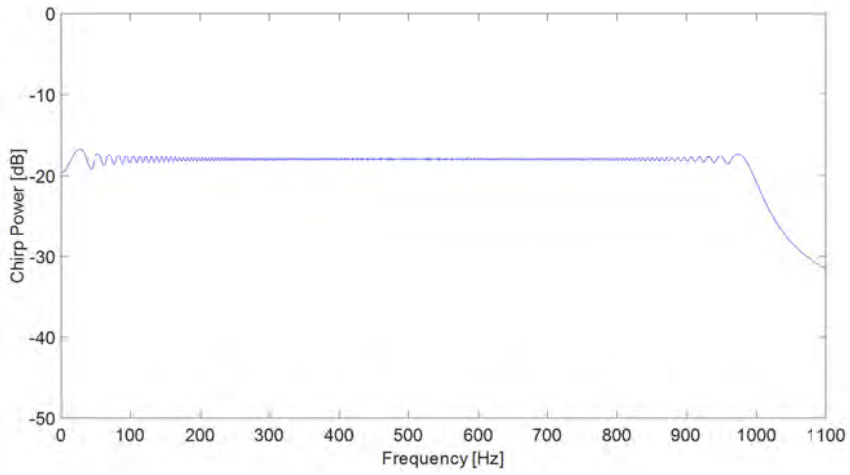



**Figure 75 –Discrete linear chirp used in RX model simulation**

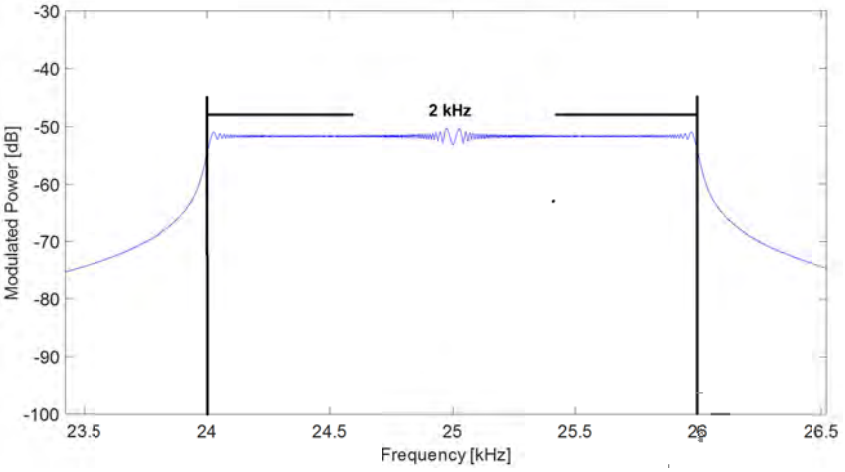



**Figure 76 –Simulated RF chirp modulation signal**

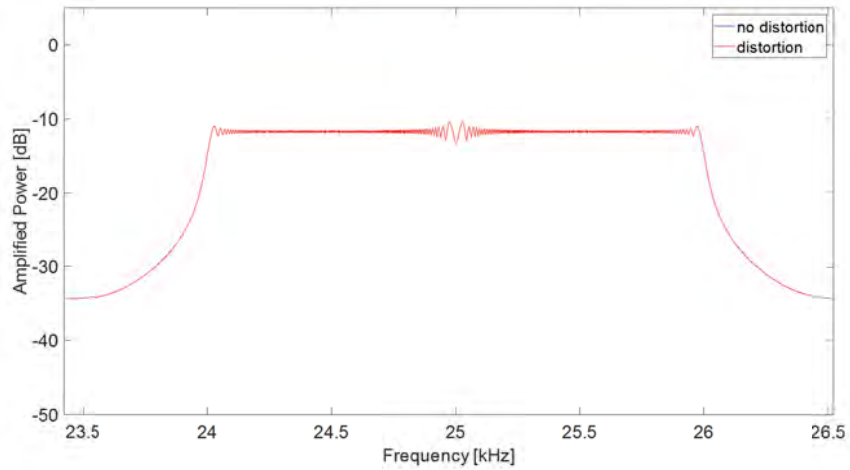



**Figure 77 –Simulated Amplified RF chirp modulation signal, with distortion, compared to non-distorted signal**

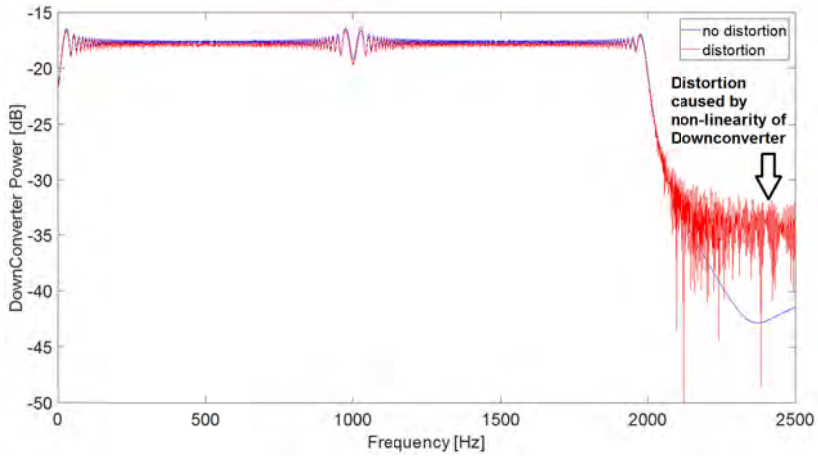



**Figure 78 –Distorted Downconverted Chirp Signal Vs non Distorted Downconverted signal**

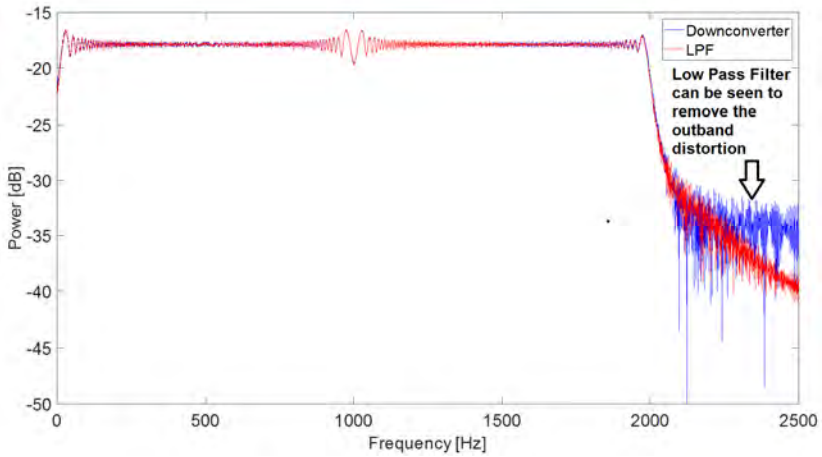



**Figure 79 –Distorted Downconverted Chirp Signal Vs Low Passed Distorted Downconverted signal**

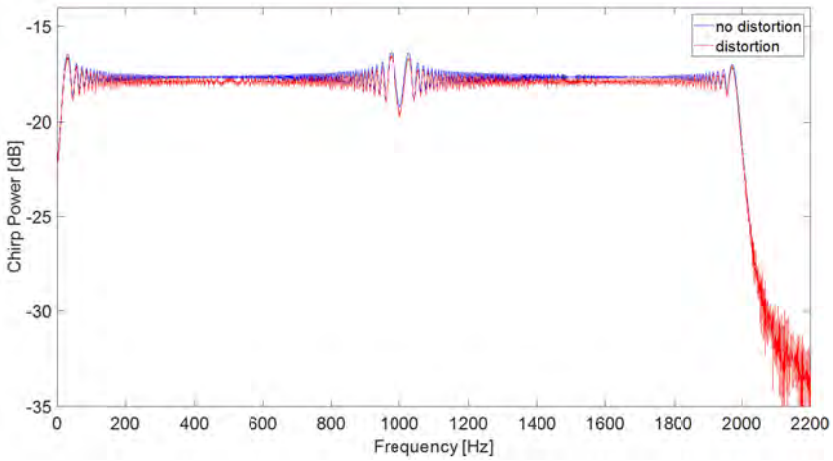



**Figure 80 –Distorted Downconverted Chirp Signal Vs Non Distorted Downconverted signal**

By calculating the group delay of both the distorted and non-distorted waveforms, Figure 81, we can observe if the distortion would have had an impact on the chirp signal such that it would degrade the chirp signal sufficiently, that this technique could not be used. This would be done, as one would calculate the group delay through the device, by simply subtracting the group delay of the signal from the group delay of the output, Figure 82. Figure 81 shows that the curves of both the distorted and non-distorted linear chirp signal are in fact linear; however, with the presence of quite a large amount of distortion, remains linear. The phase of the signal can be seen to reverse at the downconvertered carrier frequency, $\varpi 1$, or, the max frequency of the linear chirp, $F_{stop}$, where the modulation signal is doubled. To be able to calculate the group delay, the ideal signal with no distortion, needs to be subtracted from the measured, therefore for the calculation, the modulation signal needs to be mathematically upconverted and downconverted in the same way as would be done in the actual device but without distortion components. Figure 82 shows the subtraction of these two waveforms, with the non-linearity's defined in Equations (27, 28) included in the measured signal. As the subtraction of an ideal signal with no distortion is from a signal with distortion then ultimately the distortion is not cancelled and is still included in the Group Delay result.

What can be observed is that the group delay is constant over the band of the chirp except for the start and stop of the chirp; this is due to the large phase changes that take place during the first and last tbp periods of the chirp, Figure 47, as discussed in 3.3.3.

Alternatively stated as from when the chirp rises and falls from the -3 dB points of the chirp. Therefore, for any group delay test using a chirp, the bandwidth where the chirps gain is flat, lcfg, should only be used, Equation (18). Thus the Group Delay measurement Start and Stop frequencies points are defined as Equations (29, 30) respectively. Using this technique, two sides of a bandpass filters can be tested at the same time within the time to test half the bandwidth.

$$F_{start\_GD\_Meas} = 2*F_{start} \quad \textbf{(29)}$$

$$F_{stop\_GD\_Meas} = F_{stop} - F_{start} \quad \textbf{(30)}$$

If $11^{th}$ order non-linearity component of the mixer, $a_{11\_mix}$, is increased by factor of 5, results in complete destruction of the spectral purity of the chirp signal, Figure 83, making it impossible to test a Group Delay metric with such a device unless some post processing is carried out to the signal to remove random effects. Figure 83 shows effect

of smoothing the measured waveform slightly that enables the measurement of the Group Delay.

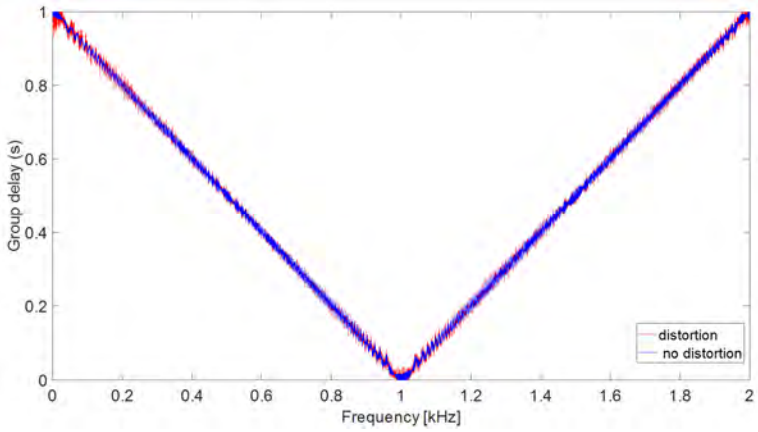



**Figure 81 –Simulated distorted chirp signal compared to non-distorted version**

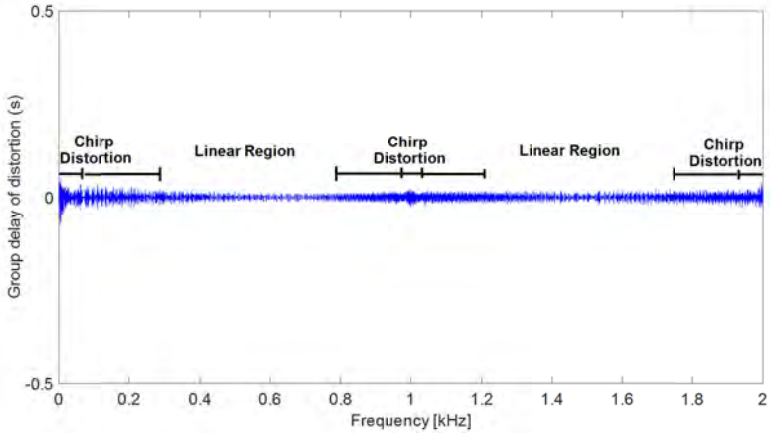



**Figure 82 –Group Delay through the modelled RX under the influence of distortion**

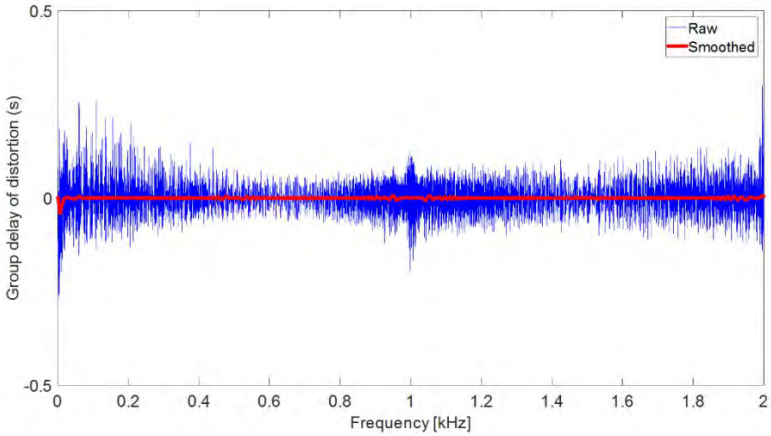



**Figure 83 –Simulated increase of the 11th order product of the downconverter distorted chirp signal compared to smoothed version.**

## 4.4 Measuring the Effects of a Chirp on a RX Model with Filter Defects

All of the previous analysis was considering that the system was working correctly and hence it was not possible to observe the bandpass filter, Figure 72. To be able to analyse the effect of the bandpass filter, the cut off frequencies need to be brought in. A 20-order FIR bandpass filter was simulated with start and stop cut off frequencies designed to be at 75% of $\varpi 1 - \varpi 2$ and $\varpi 1 + \varpi 2$ of the RF chirp bandwidth, Figure 84. This will allow observing the filters phase characteristic using the chirp sweep as the chirp sweeps bandwidth is larger than that of the filters setting. The non-linearity effect has been resolved as slight averaging was applied to the measured signal that removed the random effects of the non-linearity. Observing Figure 85, one can see the expected group delay of the BPF. The peaks of the phase change are clearly seen and should be easily seen in any algorithm produced to measure such a metric. It can be seen that the Group Delay Peaks at 283.6 samples at 24.375 kHz and 25.724 kHz. Calculating these frequencies in the down-converted spectrum leads to 0.375 kHz and 1.724 kHz. Observing Figure 86 shows that the measured group delay was 283.4 samples at 0.376 kHz and 1.726 kHz. This is an error Frequency of 1 Hz and 2 Hz and Group Delay error of 0.2 samples. As the error is so negligible, these errors are attributed to the resolution of the measurement used in the simulation. If more points and more computation time were used in this simulation i.e. the same as used in the test program, this error would be reduced and should not be seen in the measurements on a real ATE tester.

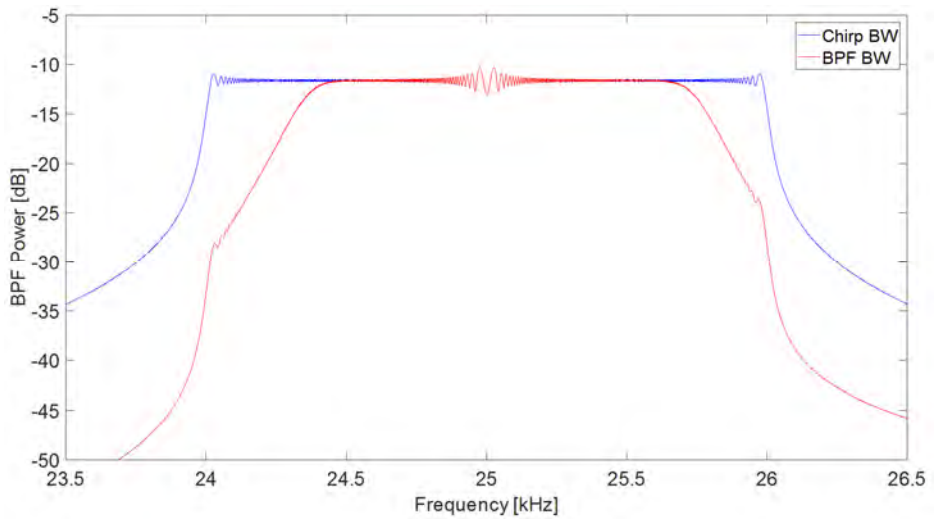



**Figure 84 –Simulated chirp signal with bandwidth greater than bandpass filter**

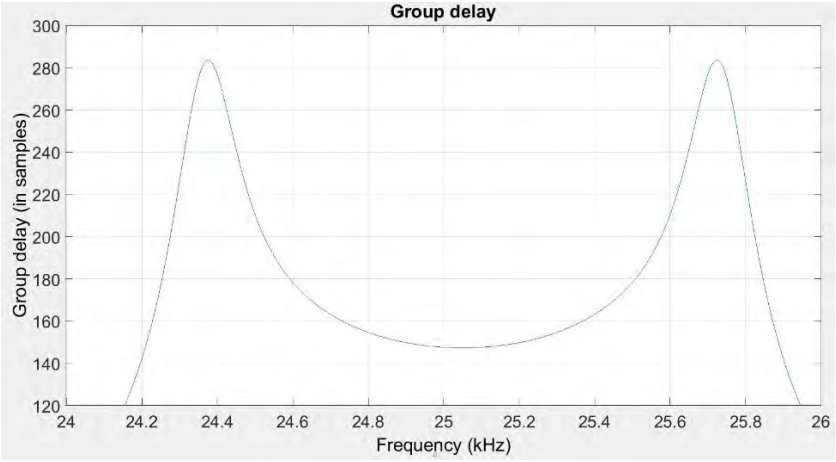



**Figure 85 –Group Delay plot of 20 order Bandpass Filter**

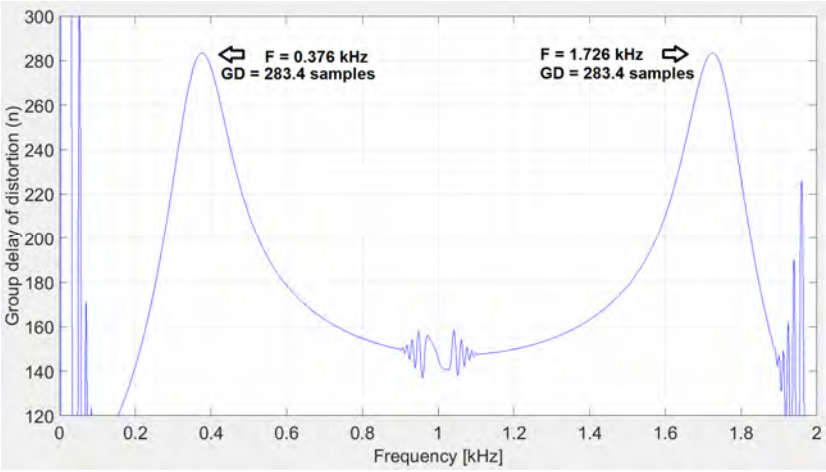



**Figure 86 –Group Delay plot of 20 order Bandpass Filter measured at downconverted frequency using Chirp Sweep**

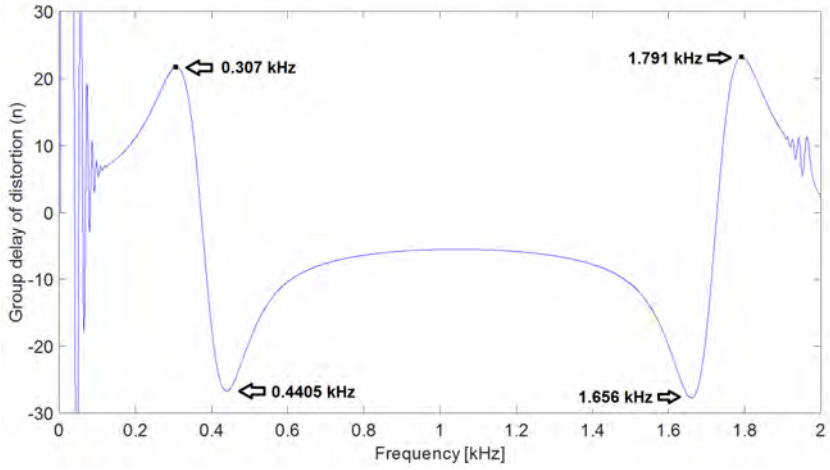



**Figure 87 –Group Delay difference plot of two, 20 order Bandpass Filter with corner frequency differences of 1%**

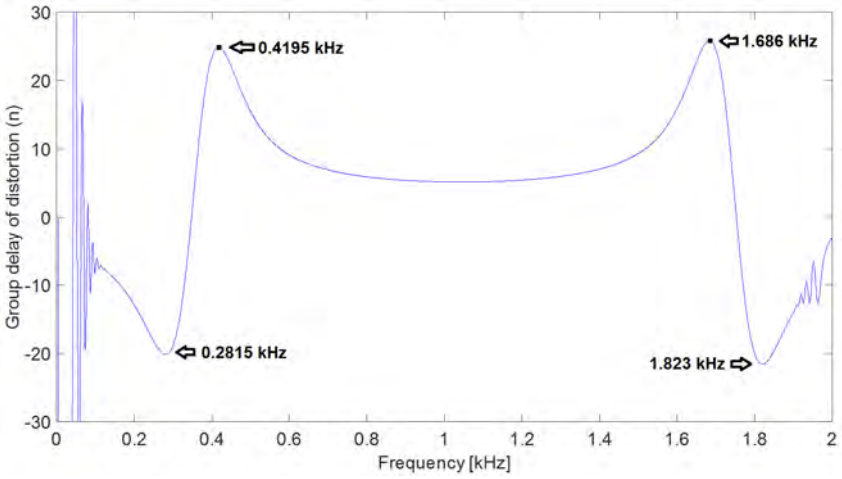



**Figure 88 –Group Delay difference plot of two, 20 order Bandpass Filter with corner frequency differences of 1%**

It is easy to measure the profile of a good device; the challenge is to measure the effect of a defect within a system or in this case the BPF. If the simulation model is changed slightly so that the filter would represent a real world application where the bandwidth would be outside of the expected signal bandwidth, but then this is compared to a filter that has an unexpected defect that inadvertently changes its filter pole. This can then be used to see how effective the algorithm is at detecting slight group delay changes that could inadvertently affect the device performance.

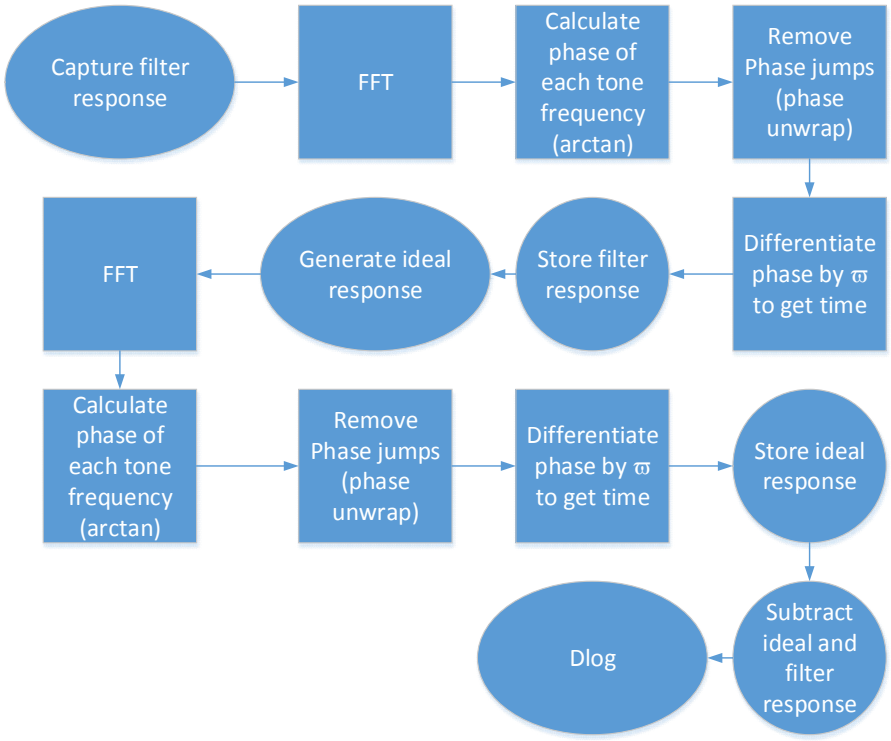



**Figure 89 – RF Group Delay Measurement flow chart**

By reducing the filter corners by 1 % to Min and Max frequencies of 24.384 kHz and 25.714 kHz respectively results a down converted signal at Min and Max frequencies of 0.384 kHz and 1.714 kHz respectively. Using the chirp measurement technique resulted in Min and Max frequencies of 0.400 kHz and 1.700 kHz this time an error of around 14 Hz. Although there is a negligible error, the change in the filters corner frequencies was easily detected. If the ideal and the altered filters are compared, then this error should be clearly seen. Observing Figure 87, the difference of the ideal and altered filter are shown.

It can be seen that the algorithm, Figure 89, is showing that measured filter compared to the ideal is inside of the expected Range, as the group delay switches from positive to negative on the corner frequencies of the ideal filter, where it intersects with the corner frequencies of the altered filter within the bandwidth of the ideal. Figure 88 shows the altered filters corner frequencies lying outside of the ideal filter corner frequencies by 1 %. The difference response is close to the mirror image, which would be expected, as now group delay transitions earlier than expected, then is intersected by the ideal curve, hence causing an initial difference group delay due to the unexpected change compared to the ideal.

What this shows is that it is possible to measure, using a chirp, the filters response, and if available, plot it against the expected filter response, however this is seldom available to a test engineer. In addition, in practice, the above exercise is not feasible due to process variation of the silicon foundry and hence the component values that make up a filters design will change by more than 1 % due to the silicon foundry process operating windows, also known as corners. However, this information can be used to set precision limits accordingly to spice models of the filter used within the design to ensure that a defect, that cause the filter to behave outside of the process variation, fails and does not enter the field.

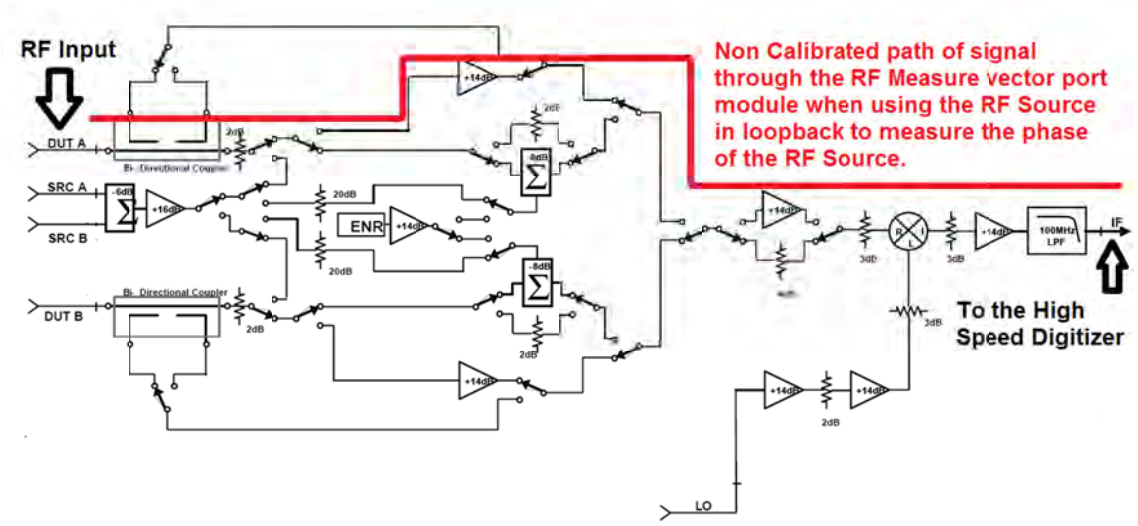



**Figure 90 –Diagram of RF Measure Port Module**

## 4.5 RF Measure Path

A standard RF ATE tester is configured in the following way, in the tester is a RF Vector Port Module; this can allow either a RF port to be configured as a Source or Measure Port. If the port is configured as a Source Port, the signal from the RF Generator is switched through the port module to the device. However, if the Port is configured as a Measure Port the signal is sent through significant amplification and attenuation steps to allow the signal to be conditioned for down conversion by the mixer,

Figure 90. For a RF group delay test, a possible way to measure would compare a down converted copy of the RF input signal and compare it to the output of the down converter under test, [29, 30]. However, if the clock of the DUT and the RF Measure are not phase locked, would result in the frequency of the DUT being shifted, however, this problem is addressed in Chapter 2.3.13. In addition, any non-linearity of the Measure module would need to be compensated for; however, as long as the measure module was used with the correct headroom then these contributions should be negligible. In the measure module, there is also a non-switchable 100 MHz LPF, any phase change of the signal due to this filter would need to be compensated for also, however, as the signal that would be measured would have a bandwidth of 25 MHz, any influence of the filter would be extremely negligible. Due to these issues, another solution should be found; however, the RF measure module can be used as replica of a standard RF RX and thus,

could be easily used to debug any measurement technique or algorithm developed for testing a RF Downconverter.

## 4.6 Measurement Setup

To verify this technique described in chapter 4.4, the tester can be configured such that an RF RX is replicated by using the measure module to measure a chirp modulated RF source.

An RF measure port is designed to be able to cover a large range of input power, 120 dB, and this is done by using different amplifier and attenuator setting within the port module, the aim is to have approximately 1 V coming from the port module so the digitizer can sample this signal correctly. Unfortunately, these amplifiers cause distortion and hence if a good measure of group delay performance is required as a reference from the port module, then the range that uses the least amount of active components should be selected. In addition, to ensure that the highest linear region of the ADC in the digitizer is utilized, only half of the ADC's range should be used. For this example, an input level of the RF source needs to be set such that it delivers enough power, resulting in 0.5 V at the digitizer input. Using the tester in such a manner should allow to produce a baseline performance of the technique such that a minimum error can be established.

By applying the same signal to both IQ inputs, i.e. not in quadrature, or just to the I channel, with nothing on Q we will achieve a standard frequency double sideband AM modulated signal, Figure 73. By using AM modulation produces twice the bandwidth of the generated chirp and hence gives the possibility of making greater bandwidth measurements with test equipment that has lower ranges hence potentially extending the range of lower cost testers. Hence, to be able to cover more twice the bandwidth than that of the AWG, by keeping the IQ waveforms, Figure 91, in phase i.e. not 90 degrees apart but 0 degrees phase difference, allows to produce a DSB AM modulated signal and hence when demodulated, produces twice the bandwidth of the generated chirp signal. To be able to calculate the group delay of the RF subsystem, the tester needs to be modelled accurately in math, Equation (26) & Figure 93, and hence produce the ideal response to compare against the measured signal. This can be expressed mathematically as Equation (26) and is shown in Figure 93, $s(\omega 1t)$ is a discrete modulated chirp, as would be seen at the output of an RF generator when modulated by Equation (15) using

an AWG. Using the RF measure to down converter and measure the amplitude response shows the double modulation bandwidth.

Figure 93 may look like a multitone, but it is a generated discrete linear chirp as described in generated Equation (26). The reason it looks like a multitone is due to the setup of spectrum analyser to capture the signal. This was the best setup that could be made to capture the signal. If Figure 93 & Figure 96 are compared, it can be seen that the same amplitude envelope is present and hence Figure 93 is an AM modulated discrete chirp.

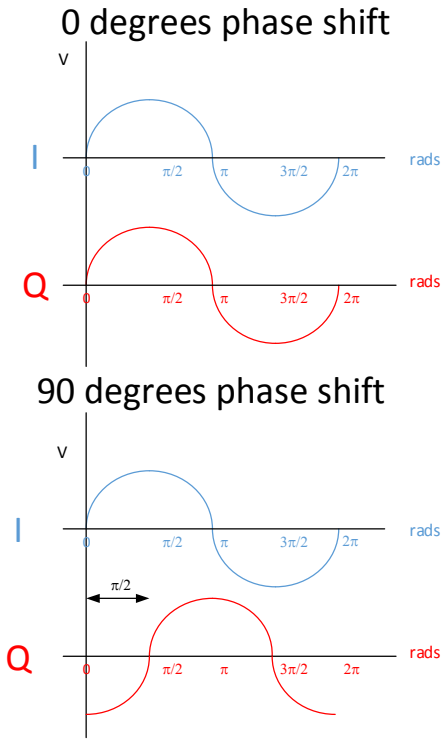



**Figure 91 –Diagram of IQ waveforms with 0 and 90 degrees (π/2) phase shift**

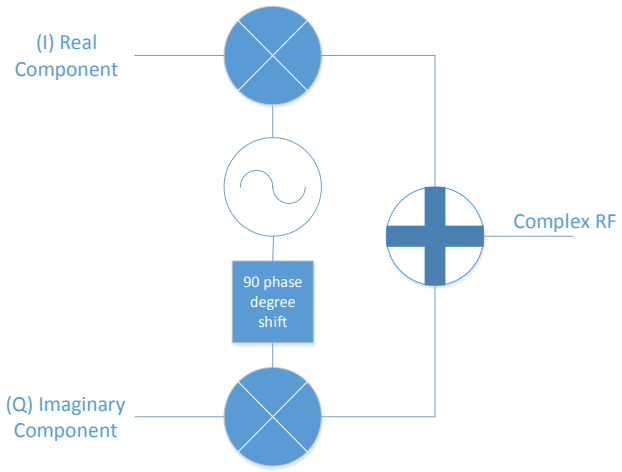



**Figure 92 –RF Generator with IQ modulation**

In mixed-signal/analogue circuit test, it is very common to trigger an AWG and digitizer such that they start at the same time, [29, 30], unfortunately in an RF system this is not possible as the digitizer is part of the RF subsystem and thus another method for capturing the data is needed. If the setup of the digitizer is made such that twice as many points, N*2, Figure 94, are captured compared to the chirp signal but with the sampling rate, Fs, then two cycles of the chirp should be present. By locating the position where the chirp stops would allow to extract the signal, Figure 95, for post processing and thus solving the triggering issue. To be able to mix down the RF modulated signal, Figure 96 , the LO of the RF Measure,

Figure 90, needs to be set at the Carrier Frequency of the modulated signal, $\varpi 1 \pm$ Chirpmax_freq. In this case the LO frequency was set to the Carrier Frequency - Chirpmax_freq.

As the chirp signal is a user-defined signal, with its own bandwidth definitions. It is required to setup the RF measure in such a way that will allow capturing this signal but also allowing for easy manipulation of the signal in Digital Signal Processing (DSP) once it is digitized.

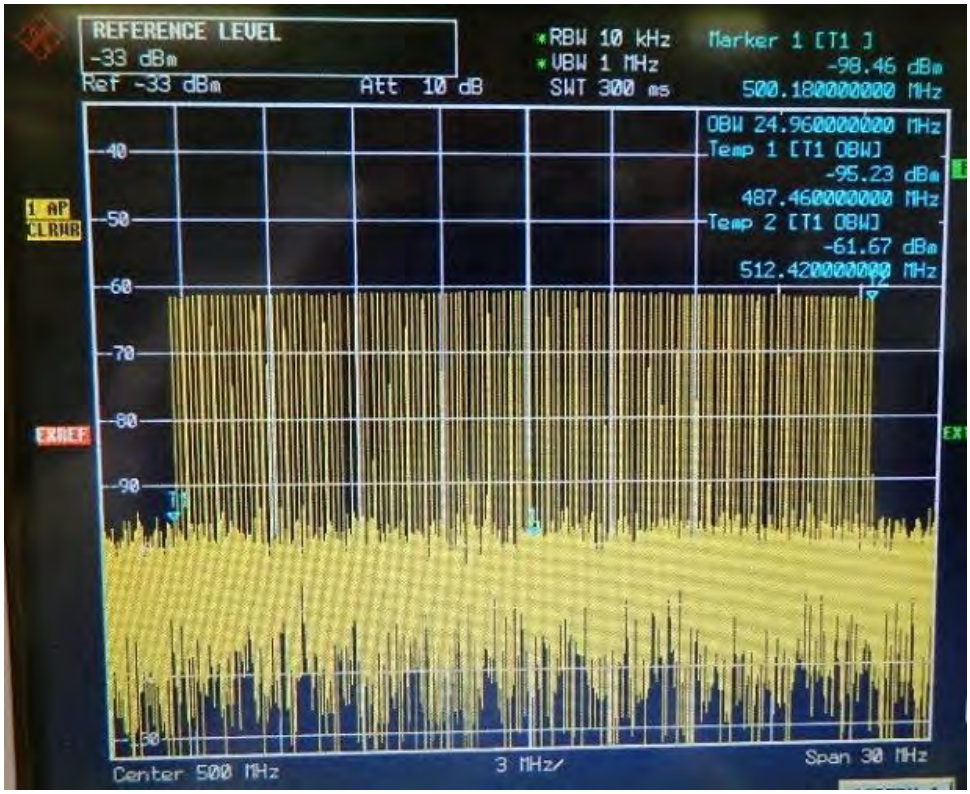



**Figure 93 –DSB-SC chirp modulation waveform at 500 MHz carrier frequency**

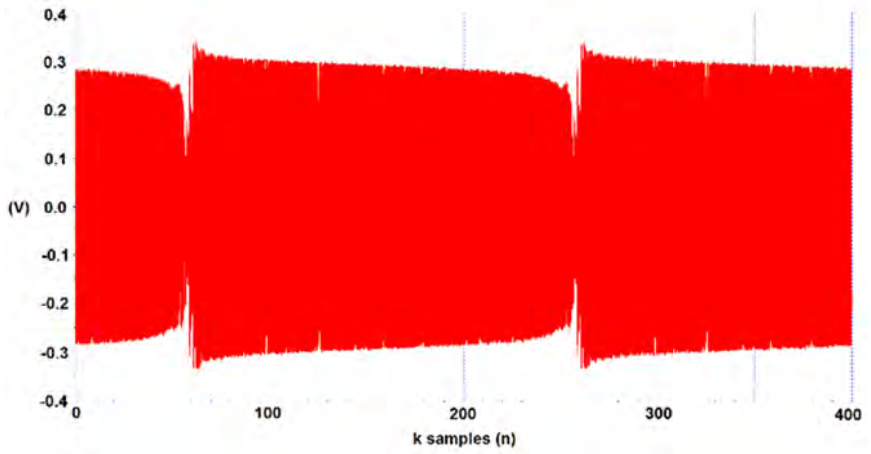



**Figure 94 –Raw capture of chirp signal from RF Measure Port**

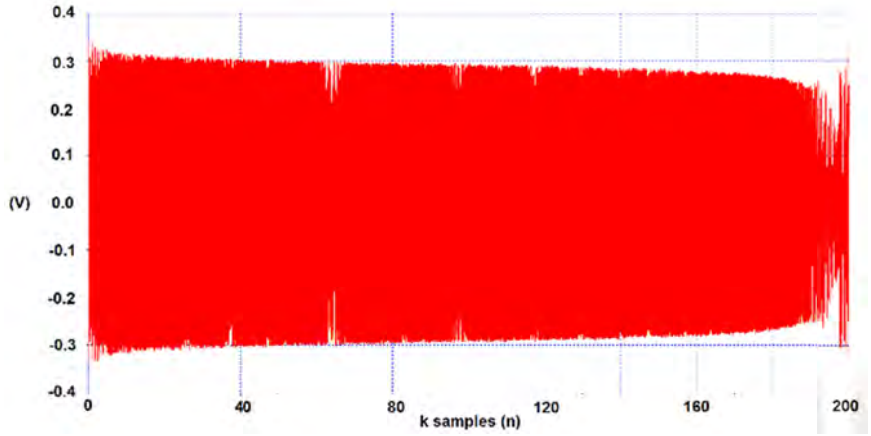



**Figure 95 –Extracted data from Raw Capture**

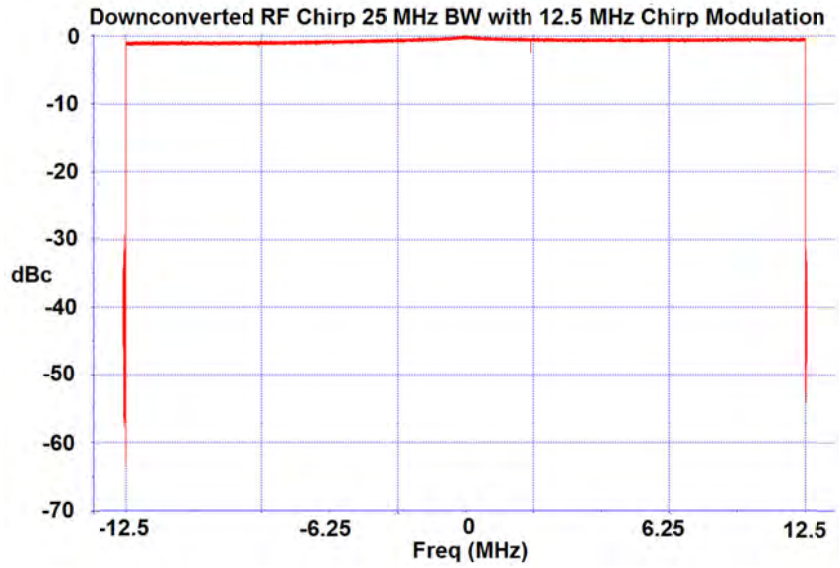



**Figure 96 –Demodulated DSB chirp modulation waveform centred at the max frequency of the chirp**

The RF measure of a tester will usually have a user-defined range that allows setting the Intermediate Frequency (IF), which the RF will mix the frequency down to, the sampling frequency and number of points, the digitizer will be programmed to use for the capture of the IF, Table 7. By following Equation (25) and setting the IF to Fs/4, it is possible to verify Equation (26) by multiplying the digitized signal by a co-sinusoid that has a harmonic frequency of the IF. This will extract the co-sinusoidal component of the chirp from the captured signal. If there is no sinusoidal component, by multiplying with a sinusoid that has a harmonic frequency of the IF should result in zero.

**Table 7 – RF Chirp ATE settings**

| | ***AWG Setting*** | ***Digitizer Setting*** | ***RF Setting*** |
|---|---|---|---|
| *Number of Points (N)* | *200,000* | *400,000* | - |
| *Sampling Frequency (Fs)* | *100 MHz* | *100 MHz* | - |
| *Intermediate Frequency (IF)* | - | - | *25 MHz* |

Now that the captured time domain data is extracted, a Fourier Transform can be performed such that the real and imaginary components of each frequency can be analysed. Plotting the magnitude of each component will show the frequency response of the captured signal, Figure 96. By performing an Arctan of the real and imaginary components of each frequency will produce the phase change across the frequency sweep. Arctan produces phases between -180 and 180 degrees, thus when the phase crosses these boundaries will result in phase jumps. This can be resolved by performing a phase unwrap, [29, 30], Figure 97. However, due to chirp phase reversing at $Chirp_{max\_freq}/2$, results in phase distortion. This has the effect that a small part of the chirp bandwidth being unusable and requires to be nulled so that the real results can be displayed correctly. Figure 98 shows the zoomed in part of the group delay capture, at the end of the frequency sweep, where the affect would be the most apparent compared to the reference. This shows the group delay of the RF measure becoming increasingly worse after around 10 MHz. This makes sense as the useable bandwidth of the Xcerra RF subsystem is around 20 MHz, i.e., ±10 MHz, thus proving that this technique is viable. Subtracting the measured group delay of the tester from the group delay of the reference signal, Figure 99, gives the group delay of the Xcerra Fusion MX RF subsystem. As no filter is in the path, the only explanation for the roll off across the

band shown in Figure 99, for what is observed, is the group delay of the mixer used to down-convert the signal.

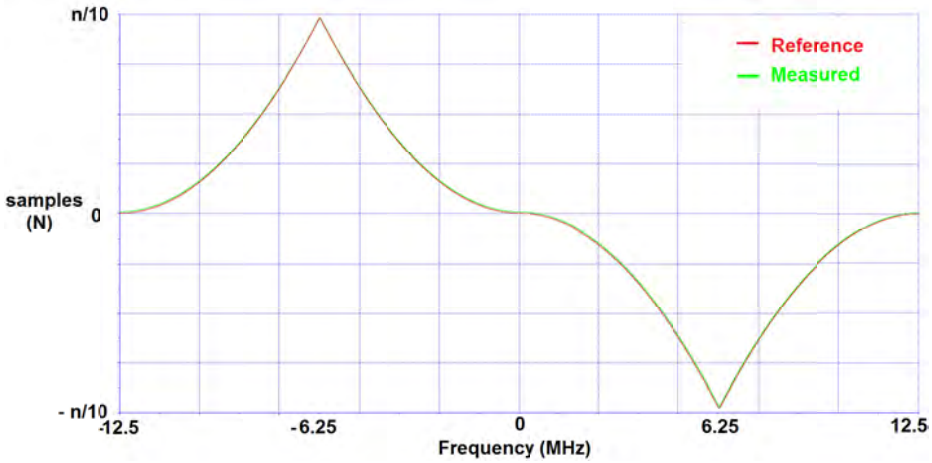



**Figure 97 –Phase response of Captured RF Chirp vs Expected RF Chirp**

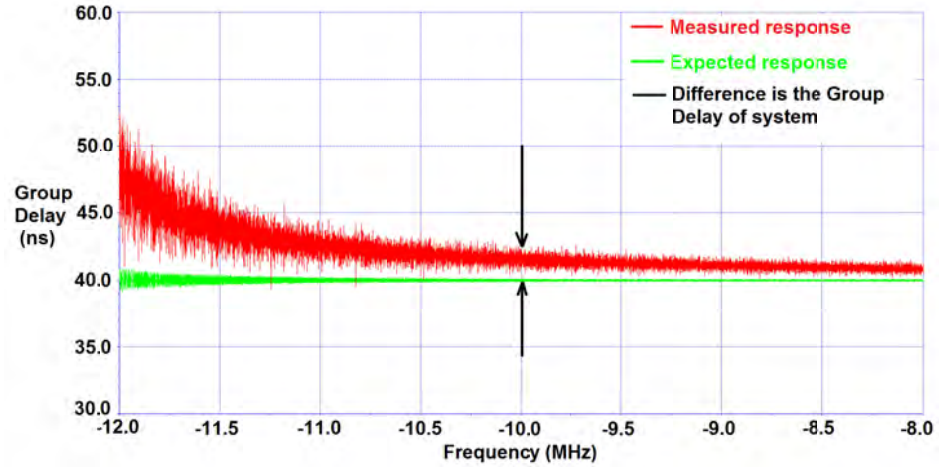



**Figure 98 –Zoom in of the measured group delay profile of the measurement vs expected profile using a chirp signal**

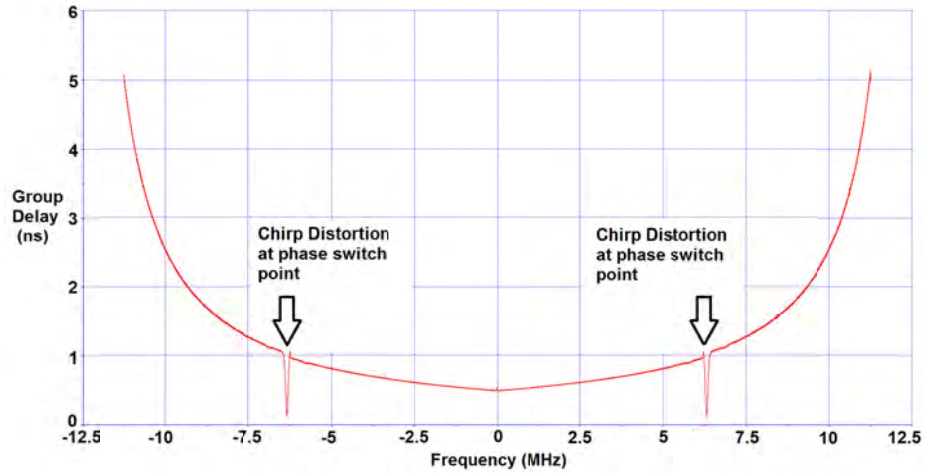



**Figure 99 –Group delay of Xcerra RF measure port.**

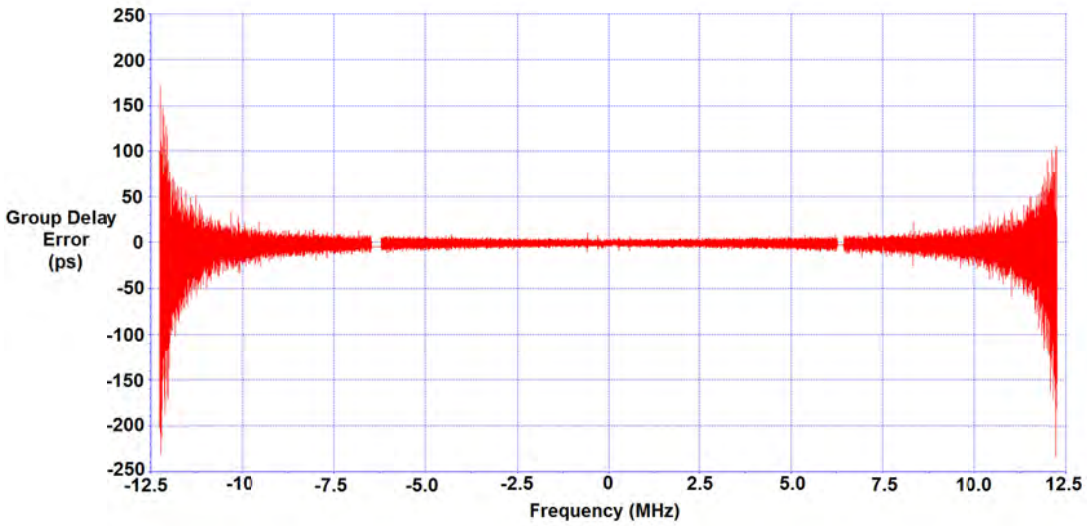



**Figure 100 –Measurement error of group delay of IF path of the RF measure port.**

If a true and accurate representation of a filter is required, then this characteristic needs to be calibrated out. This is easily achieved by subtracting the group delay of any filter from the RF subsystem with no filter in the path. Before this can be done, the measurement accuracy of the system needs to be understood. This can be done by measuring the IF path with no filter twice and subtracting the two responses, theoretically resulting in the measurement error. Figure 100 shows the measurement error, it can be seen that the measurement error is in the picosecond region thus 10x more than the required measurement accuracy for the filter specifications targeted by this technique. To test any RF RX, the above needs to be considered at all times. If an RF BPF is required to be tested, then some way of switching this filter out of the RF path needs to be available, as part of the DFT, or the influence of the down-convert mixer will always be present in the measured result, the same consideration is also needed for the LPF post mixer.

If a filter in the IF path is now switched in, the characteristic of this filter should be clearly identifiable. Observing Figure 101, one can see a typical group delay characteristic of a LPF filter. In this case, the specification is stated to be 10 MHz Bandwidth, it can be clearly seen that this filters 3 dB point is at 9.75 MHz but the group delay starts to degrade at 7.5 MHz, which can be seen to be the knee of the filter, where attenuation starts to occur.

Observing Figure 102, a typical group delay characteristic of a bandpass filter can be observed. In this case, the specification is stated to be 1 MHz±300 kHz Bandwidth, however, it can be clearly seen that this filter has only about 500 kHz of 3 dB

bandwidth. In addition, the group delay changes by 20 ns between the -3 dB points and hence this filter could not be used where a flat group delay is a requirement.

Observing Figure 103, the amplitude response and group delay of the Mini Circuits BPF-B48+ bandpass filter are shown. Making a comparison of the measured result, to the expected curves from the datasheet, [32], it can be observed that using the chirp technique results in a very good correlation to the expected result. The expected curve should have a starting group delay of 167 ns at 46 MHz, have a minima at of 127 ns at 48.5 MHz and then return to 155 ns, at 50.5 MHz. The measured curve fits this exactly except for a 5 ns error from the expected at the maximum frequency of 50.5 MHz. This error is due to the phase distortion that occurs at $Chirp_{max\_freq}/2$.

Using these two examples covers all the components shown in Figure 72 of a RF RX.

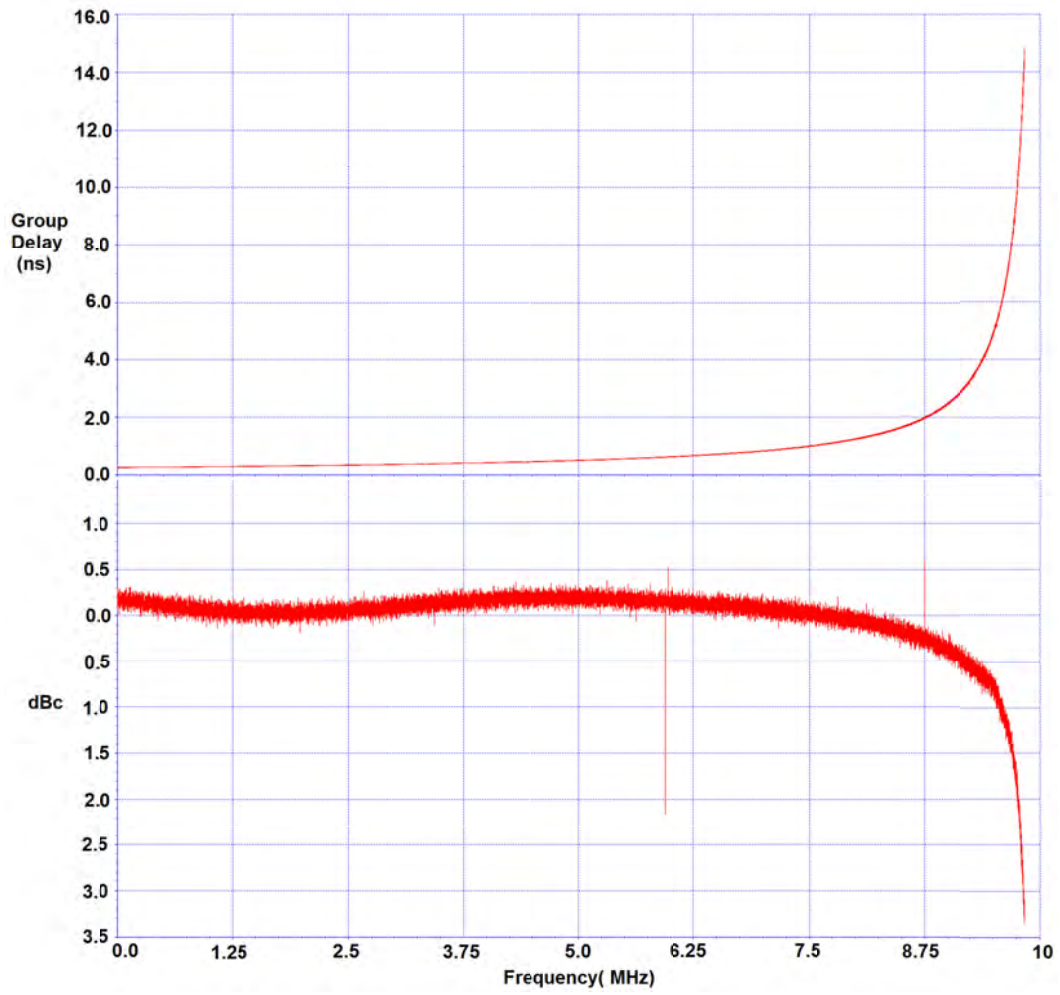



**Figure 101 –Group Delay of 10MHz LPF in the IF path of the Xcerra RF Measure port**

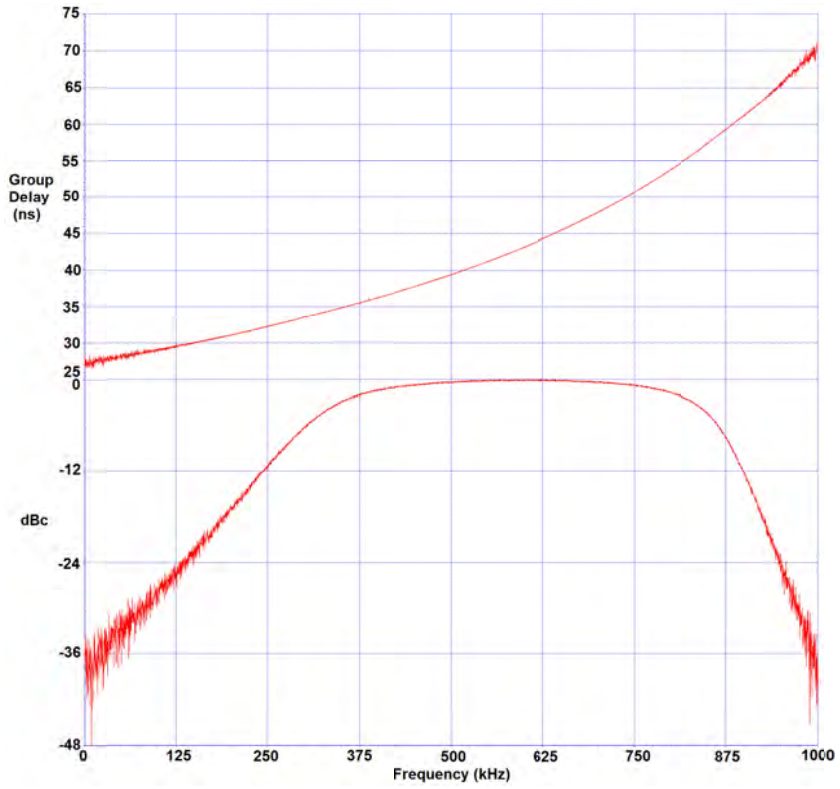



**Figure 102 –Group Delay of 1 MHz BPF in the IF path of the Xcerra RF Measure port**

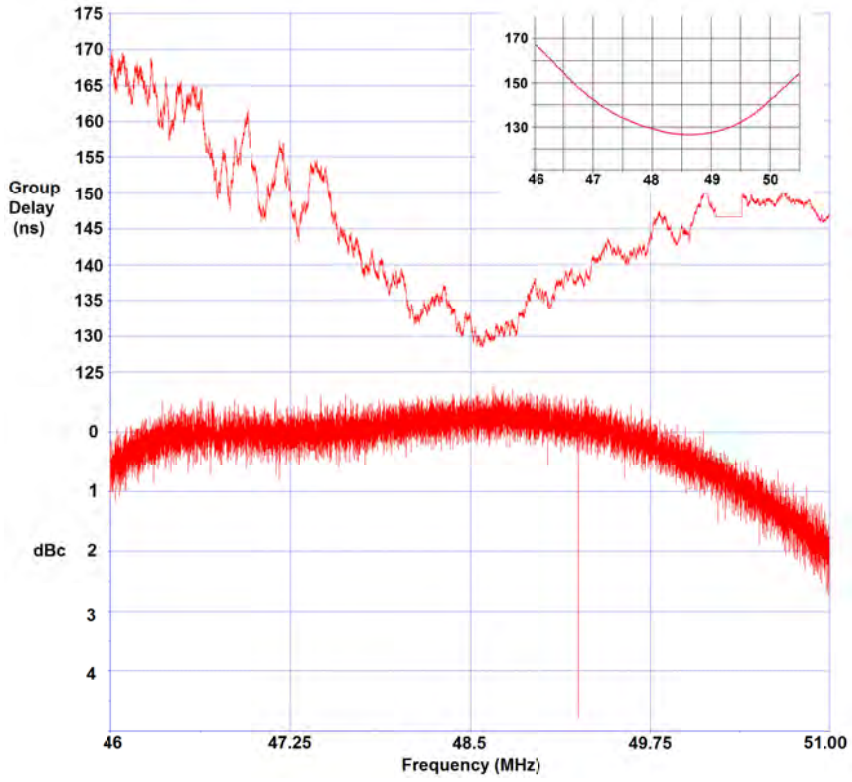



**Figure 103 –Group Delay of 2 MHz BPF in the RF path of the Xcerra RF Measure port.**

## 4.7 Conclusion

By using an RF tester, such that DSB modulation is produced, the user can potentially extend the range of their production equipment if the system AWG used for modulation has a limited bandwidth, thus allowing older test equipment that has potentially lesser specification compared to newer equipment to test today's devices. Using the technique specified, results in a very accurate and fast method of characterizing an RF RX devices group delay metric even if a device has a high level of intermodulation distortion as this distortion can be simply averaged to expose the response of a device under test from the distortion. By subtracting the measured response from the ideal, a real comparison can be made. Whereas today, a few frequencies of a filter would be compared against a threshold and potential defects allowed into the field. As such, the quality level of a tested RF RX, with any RF or IF, filter should be significantly increased.

# 5 USING A DISTORTION SHAPING TECHNIQUE TO EQUALIZE ADC THD PERFORMANCE BETWEEN ATES

## 5.1 Introduction

Many techniques are available to automated test equipment (ATE) test engineers to combat the drawbacks of a test system that they have to use for testing of a device whose performance is better than the test system. Either from under-sampling of an arbitrary waveform generator (AWG) for the frequency range extension or noise filtering to increase the effective number of bits (ENOB), but until recently there has been no mainstream way to reduce the harmonic content of an AWG by way of digital signal processing (DSP) techniques. However, there comes a time when there simply is no choice but to invest in new test equipment and hence this is directly responsible for driving up the test costs of the next generation devices when the industry is under constant pressure to continually reduce test costs at a time when it is harder to guarantee device specifications.

This paper presents that the techniques described in [33-35] improve the reproducibility of 12-bit ADC THD testing results across ATEs test sites, and also significantly increase the test accuracy of 16-bit ADC THD within an actual ATE environment. The previous work in [33-35] describes the phase switching algorithms with theory and simulation as well as limited experiment results at the laboratory level. However, this paper verifies their effectiveness quantitatively at the extended experiments at the mass production level using ATE systems and shows that the phase switching techniques can realize low-cost and high-quality testing of ADCs in industry. The experimental verification of simultaneous multiple harmonic cancellation algorithms is also presented. (Note that in [34] only HD3 cancellation experiment was described.)

Testing THD in production of any analogue/mixed-signal device is generally restricted to the base performance of the instrument testing the device under test (DUT); in this case, it would be an AWG. If the performance of the AWG is worse than the DUT, the performance of the AWG will be measured rather than that of the DUT. Alternatively, if the device performance is marginally better than the AWG, the problem of yield loss will occur due to cutting into distribution of the performance of the AWG. This would lead to the need for guard-bands that would result in yield loss due to failing devices that were close to the edge of performance but were not necessarily defective. In either case, the performance of the device is not being reflected accurately in the measured results and the true measure of the device performance is not being accurately portrayed.

What is required in the short term is a way to make low cost tester resources perform better than their actual specifications. This paper shows the development of the techniques, described in [33-35] into a production style or ATE environment. Several other methods for low distortion signal generation have been proposed [36-44]. The aim is to provide a mechanism to make the digital-to-analogue converter (DAC) THD performance of an AWG in the ATE look better than it actually is; this results in the ability to test higher performance ADCs without the need for higher end equipment (such as high performance analogue filters [45]). The techniques in [33-35] cancel the harmonics caused by the DAC non-linearity inside the AWG only with a waveform memory data change without the DAC non-linearity identification or hardware modification. In other words, a digitally pre-compensated technique is used. A positive side effect of doing this work is that by artificially improving the performance of the

AWGs, the result is to equalize the performance over the installed tester base of the same manufacture resulting in the need for less stringent guard-banding requirements to guarantee device performance over the companies' tester base.

There is always the argument that specification based testing is not required as a defect can be found, [39], by other means that would cause a THD failure. However, automotive requirements dictate that specifications with a minimum and maximum value have to be tested as specified in the datasheet. Therefore, although structural test would help to eliminate such tests in the future, until automotive requirements change, new cost effective ways will continually need to be developed to test such specifications.

## 5.2 Theory

### 5.2.1 Phase Switching Technique

By implementing a signal that switches between two defined phases [33-35], the harmonic content of the generated signal of the AWG can be moved to a sub-frequency of the Nyquist rate. This results in effectively shifting the harmonic content of the signal to a higher frequency range. As the Nyquist rate is usually much higher than the generated fundamental frequency results in the ability to filter off the harmonic content with standard filtering of built in filters of an AWG. By the use of up-sampling and on-board filters of the AWG, the task of the phase switching technique becomes a DSP exercise, hence making the implementation a simple task. By analysing the harmonic contents of an AWG, a decision can be made as to what harmonics need to be suppressed for a given application. One can decide to suppress only one or multiple harmonics at once using the formulae given in [33-35], which require only a simple waveform memory content change. The only restriction is that as more harmonics are suppressed results in a proportional decrease of the available signal bandwidth from the AWG. As an example, if the user requires suppressing just the $2^{nd}$ harmonic, a 2-phase signal interleaving method can be used with $X_0$ in case of even n, and $X_1$ in case of odd n as the AWG input signal. This will result in the AWG output signal Y as defined in Equations (31-33) using the same notation developed in [36]. If we assume that the DAC in the AWG has the second-order distortion then we can develop the following

$$X_0(n) = A\sin\left(2\pi f_{in} nT_s + \frac{\pi}{4}\right) \quad \textbf{(31)}$$

$$X_1(n) = A\sin\left(2\pi f_{in} nT_s - \frac{\pi}{4}\right) \quad \textbf{(32)}$$

$$\begin{aligned} Y(nT_s) &= a_1 D_{in}(n) + a_2 D_{in}(n)^2 \\ &= a_2 A^2 + \frac{\sqrt{2}}{2} a_1 A \sin(2\pi f_{in} nT_s) \\ &+ \frac{\sqrt{2}}{2} a_1 A^2 \cos\left\{2\pi\left(\frac{f_s}{2} - f_{in}\right) nT_s\right\} \\ &+ \frac{1}{2} a_2 A^2 \sin\left\{2\pi\left(\frac{f_s}{2} - 2f_{in}\right) nT_s\right\} \end{aligned} \quad \textbf{(33)}$$

By analysing Equation (32), it can be seen that the 2nd harmonic is now a function of the sampling frequency, $f_s$, and hence can be filtered off in the same way that the Nyquist frequency would be filtered off in a production test program using the standard filters of an AWG. Also, notice how the 2nd harmonic is suppressed by phase switching by $\frac{\pi}{2}$, i.e. by $\pm\frac{\pi}{4}$ or $\pm\frac{\pi}{2HDx}$.

Developing this further, we can use the same reasoning for suppressing the 3rd harmonic. By using the same 2-phase signal interleaving method using $X_2$ in case of even n, and $X_3$ in case of odd n as the AWG input signal, we can build the AWG output signal Y as defined in Equations (34-36) using the same notation introduced in [34]. Again, if we assume that the DAC in the AWG has the third-order distortion and using the above reasoning, if we phase switch by $\pm\frac{\pi}{6}$, then we should see the 3rd harmonic suppressed.

$$X_2(n) = A\sin\left(2\pi f_{in} nT_s + \frac{\pi}{6}\right) \quad \textbf{(34)}$$

$$X_3(n) = A\sin\left(2\pi f_{in} nT_s - \frac{\pi}{6}\right) \quad \textbf{(35)}$$

$$\begin{aligned} Y(nT_s) &= a_1 D_{in}(n) + a_3 D_{in}(n)^3 \\ &= \frac{\sqrt{3}}{2}\left(a_1 A + \frac{3}{4} a_3 A^3\right) \sin(2\pi f_{in} nT_s) \\ &+ \frac{1}{2}\left(a_1 A + \frac{3}{4} a_3 A^3\right) \cos\left\{2\pi\left(\frac{f_s}{2} - f_{in}\right) nT_s\right\} \\ &- \frac{1}{4} a_3 A^3 \cos\left\{2\pi\left(\frac{f_s}{2} - 3f_{in}\right) nT_s\right\} \end{aligned} \quad \textbf{(36)}$$

Again, we see that the harmonic has been shifted to a sub-harmonic of the Nyquist frequency and is easily filtered off using the standard filters if the Nyquist rate is much

greater than the fundamental that is generally the case. Using the Equations (31-35), we can combine them to produce a unified equation for all cases Equation (37, 38) by substituting the harmonic, $HD_x$, of choice for suppression.

$$X_4(n) = A\sin\left(2\pi f_{in} nT_s + \frac{\pi}{2HDx}\right) \quad \textbf{(37)}$$

$$X_5(n) = A\sin\left(2\pi f_{in} nT_s - \frac{\pi}{2HDx}\right) \quad \textbf{(38)}$$

If we now look at suppressing two harmonics simultaneously and address both the 2nd and 3rd harmonics, we can use a 4-phase signal interleaving method with $X_6$, $X_7$, $X_8$, and $X_9$ as the AWG input signal, and build the AWG output signal Y as defined in Equations (39-44) using the same notation as was introduced in [34]. If we assume that the DAC in the AWG has the second and third-order distortions, then we can develop the following:

$$X_5(n) = A\sin\left(2\pi f_{in} nT_s - \frac{5}{12}\pi\right) \quad \textbf{(39)}$$

$$X_6(n) = A\sin\left(2\pi f_{in} nT_s - \frac{1}{12}\pi\right) \quad \textbf{(40)}$$

$$X_7(n) = A\sin\left(2\pi f_{in} nT_s + \frac{1}{12}\pi\right) \quad \textbf{(41)}$$

$$X_8(n) = A\sin\left(2\pi f_{in} nT_s + \frac{5}{12}\pi\right) \quad \textbf{(42)}$$

$$Y(nT_s) = a_1 D_{in}(n) + a_2 D_{in}(n)^2 + a_3 D_{in}(n)^3$$

$$= \begin{cases} a_1 X_4(n) + a_2 X_4(n)^2 + a_3 X_4(n)^3 & (n = 4k) \\ a_1 X_5(n) + a_2 X_5(n)^2 + a_3 X_5(n)^3 & (n = 4k+1) \\ a_1 X_6(n) + a_2 X_6(n)^2 + a_3 X_6(n)^3 & (n = 4k+2) \\ a_1 X_7(n) + a_2 X_7(n)^2 + a_3 X_7(n)^3 & (n = 4k+3) \end{cases} \quad \textbf{(43)}$$

$$\mathrm{Y}(nT_s) = \frac{1}{2}cA^2 + \frac{\sqrt{6}}{4}\left(aA + \frac{3}{4}bA^3\right)\sin(2\pi f_{in} nT_s)$$
$$-\frac{\sqrt{6}+\sqrt{2}}{8}\left(aA + \frac{3}{4}bA^3\right) \times \left[\sin\left(2\pi\left(\frac{1}{4}f_s + f_{in}\right)nT_s\right) + \cos\left(2\pi\left(\frac{1}{4}f_s + f_{in}\right)nT_s\right)\right]$$
$$-\frac{\sqrt{6}-\sqrt{2}}{8}\left(aA + \frac{3}{4}bA^3\right) \times \left[\sin\left(2\pi\left(\frac{1}{4}f_s - f_{in}\right)nT_s\right) + \cos\left(2\pi\left(\frac{1}{4}f_s - f_{in}\right)nT_s\right)\right]$$
$$-\frac{1+\sqrt{3}}{8}cA^2 \times \left[\sin\left(2\pi\left(\frac{1}{4}f_s + 2f_{in}\right)nT_s\right) - \cos\left(2\pi\left(\frac{1}{4}f_s + 2f_{in}\right)nT_s\right)\right]$$
$$+\frac{1-\sqrt{3}}{8}cA^2 \times \left[\sin\left(2\pi\left(\frac{1}{4}f_s - 2f_{in}\right)nT_s\right) - \cos\left(2\pi\left(\frac{1}{4}f_s - 2f_{in}\right)nT_s\right)\right]$$

$$+\frac{\sqrt{2}}{16}bA^3 \times \left[\sin\left(2\pi\left(\frac{1}{4}f_s + 3f_{in}\right)nT_s\right) + \cos\left(2\pi\left(\frac{1}{4}f_s + 3f_{in}\right)nT_s\right)\right]$$
$$-\frac{\sqrt{2}}{16}bA^3 \times \left[\sin\left(2\pi\left(\frac{1}{4}f_s - 3f_{in}\right)nT_s\right) + \cos\left(2\pi\left(\frac{1}{4}f_s - 3f_{in}\right)nT_s\right)\right]$$
$$-\frac{\sqrt{2}}{4}\left(aA + \frac{3}{4}bA^3\right)\cos\left(2\pi\left(\frac{1}{2}f_s - f_{in}\right)nT_s\right)$$
$$-\frac{\sqrt{2}}{8}bA^3\cos\left(2\pi\left(\frac{1}{2}f_s - 3f_{in}\right)nT_s\right) \quad \textbf{(44)}$$

Again, we see that the 2nd and 3rd harmonics have been shifted to a sub-harmonic of the Nyquist frequency and are easily filtered off using the standard filters if the Nyquist rate is much greater than the fundamental that is generally the case. Now by unifying these equations so that they can be used for any situation i.e. $HD_{x1}$ and $HD_{x2}$ are the two different harmonics of choice, then we can rewrite Equations (39-42) as Equation (45–48).

$$\mathbf{X_{10}(n) = A\sin(2\pi f_{in} nT_s - X)} \quad \textbf{(45)}$$

$$X_{11}(n) = A\sin(2\pi f_{in} nT_s - Y) \quad \textbf{(46)}$$

$$X_{12}(n) = A\sin(2\pi f_{in} nT_s + Y) \quad \textbf{(47)}$$

$$X_{13}(n) = A\sin(2\pi f_{in} nT_s + X) \quad \textbf{(48)}$$

Where $X = \frac{\pi}{2HDx1} + \frac{\pi}{2HDx2}$

And $Y = \frac{\pi}{2HDx1} - \frac{\pi}{2HDx2}$

### 5.2.2 Simulation

Figure 104 shows graphically the simulation result for HD3 suppression using the 2-signal interleaved approach defined in Chapter 5.2.1 using Equations (34, 35). Figure 105 shows graphically the simulation result for the simultaneous HD2 & HD3 suppression using the 4-signal interleaved approach defined Chapter 5.2.1 using Equations (45-48). Also, refer to [34].

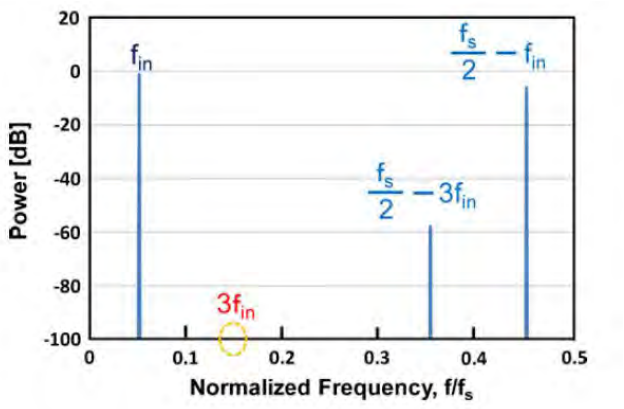



**Figure 104 - Spectrum of 2-signal interleaved 3rd harmonics suppressed signal**

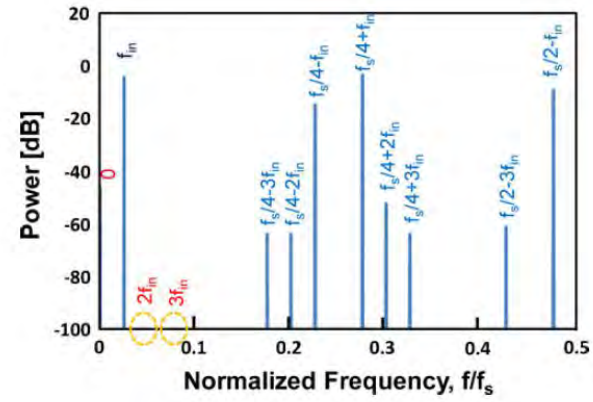



**Figure 105 - Spectrum of 4-signal interleaved 2nd & 3rd harmonic suppressed signal**

## 5.3 Measurement of the System

Before an exercise is undertaken to improve the measurement of a mixed signal block using the approach described, a thorough analysis of the influence of the prescribed method on the harmonic content of the AWG should be carried out (in this case the LTX-Credence 16-bit AWGHSB). This was possible using the loopback calibration path of the loadboard, as shown in Figure 106. This path enables the user to loopback the AWG into the digitizer (in this case the LTX-Credence 14-bit DIGHSB), and measure the spectrum directly. The following plots show the influence of the harmonic suppression on the harmonic level compared to the reference using only a conventional sinus that we generated using the AWG and measured using the digitizer. It should be noted that although the digitizer resolution is only 14 bits compared to the 16 bits of the AWG that as long as the digitizer can detect the change in level then the actual

resolution is of no consequence. As the expected change in the harmonic level is in the range of dB's, then the resolution and range are of no concern, what would be of concern would be the noise level. If the harmonics were below the noise level of the digitizer measurement range, then it would be pointless to try to measure the harmonics to begin with and hence a different solution would be required. It can be seen from the below plots that the harmonics are generally at least 10 dB above the noise floor so measuring the influence of the new algorithm in this manner should be feasible.

Figure 107 shows that by suppressing just the 2nd harmonic using a 2-phase signal interleaving method with $X_0$, and $X_1$ to build the signal Y as defined in Equations. (31-33). We can see from Figure 107 that around 6 dB improvement can be achieved to the 2nd harmonic coming from the AWG using the phase switching technique.

Figure 108 shows the measured spectrum improvement by suppressing just the 3rd harmonic using a 2-phase signal interleaving method with $X_2$ and $X_3$ to build the signal Y as defined in Equations (34-36). It can be seen in Figure 108 that around a 16 dB improvement can be achieved to the 3rd harmonic coming from the AWG using the phase switching technique.

Table 5 contains all the improvements that were achieved using the suppression technique obtained by using Equations. (37,38) for the different harmonics, 2nd thru 5th, that would be important to a semiconductor test engineer. In this example the sampling frequency was set to 250 MHz, therefore the sub-harmonic frequencies were shifted out to around $\frac{f_s}{2}$ which is approximately 125 MHz. As we have a 50 MHz low pass filter (LPF) available, it is very easy to filter off these Nyquist sub-harmonics.

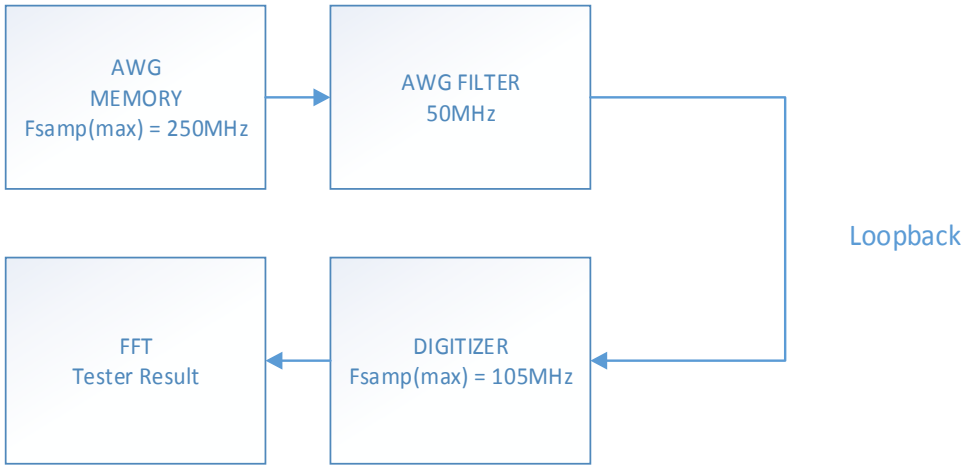



**Figure 106 - AWG to digitizer loopback**

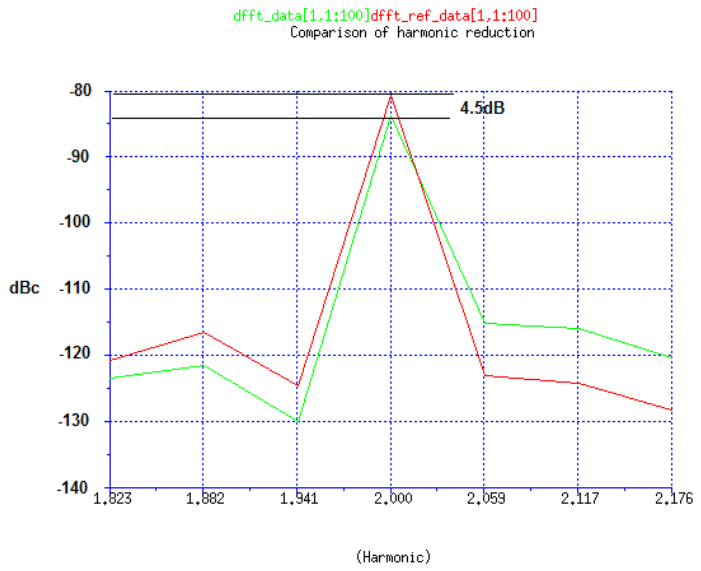



**Figure 107 - Measured spectrum level improvement of the reference vs suppressed 2nd harmonic for AWG.**

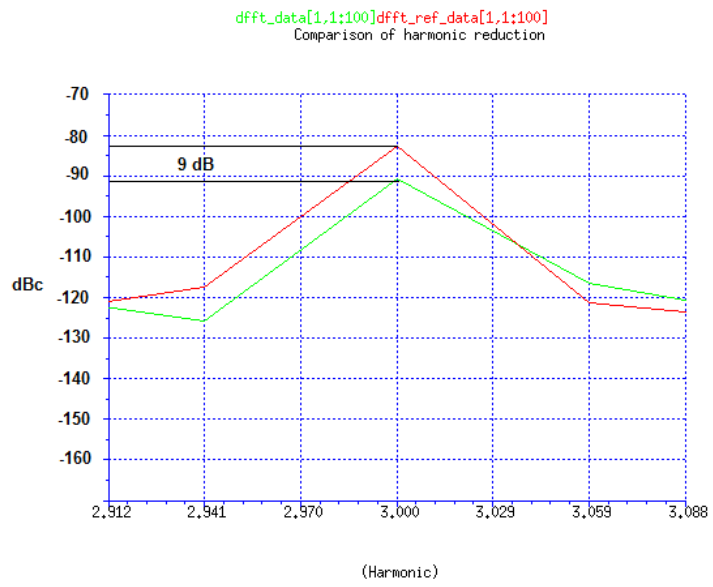



**Figure 108 - Measured spectrum level improvement of the reference vs. suppressed 3rd harmonic for AWG.**

In addition, the results in Table 5 show a great improvement of individual harmonics but it would be much more beneficial if multiple harmonics could be suppressed at once, due to there being bandwidth available, it would be possible to use a 4-signal interleave approach. Using a 4-signal interleaved approach and using the same sampling frequency would push the Nyquist sub-harmonics out to approximately $\frac{f_s}{4}$; this would result in the Nyquist sub-harmonic being around 75MHz, again easily filtered off by the AWG's 50 MHz LPF. However, this is the limit to where we can go; if we were to go to

an 8-signal interleaved methodology, an on-board filter would be needed at the output of the AWG. This is due to the 50 MHz LPF being too high to filter off the Nyquist sub-harmonics generated by the 8-signal interleaved approach and hence for a generic setup we are limited to a 4-signal interleaved approach. As such, we do not make any experiments using anything above the 4-signal interleaved algorithm.

Figure 109 and Figure 110 show the measured harmonics improvement by suppressing the 3rd and 5th harmonics simultaneously using a 4-phase interleaved signal composed of $X_4$, $X_5$, $X_6$, $X_7$ as defined in Equations (39-43) to build the signal Equation (44). We see that suppression in both harmonics can be achieved with similar results to the single harmonic suppression Y.

As can be observed in Table 5, an improvement across the band was possible therefore; one would expect that an improvement should be observed in the dynamic performance of an ADC if the performance of the DUT is better than that of the AWG.

**Table 8 –Measured harmonic performance improvement an AWG**

| *Harmonic* | *Improvement* |
|---|---|
| 2nd | 4.5 dB |
| 3rd | 9 dB |
| 4th | 9 dB |
| 5th | 12.5 dB |

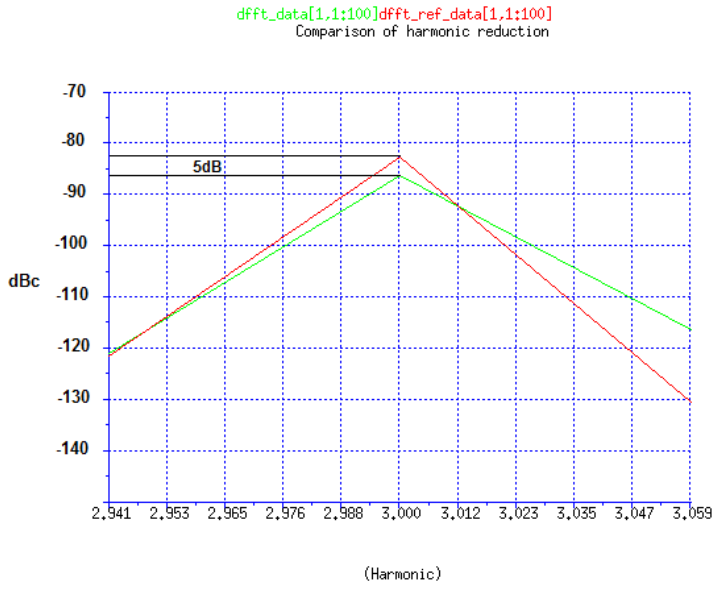



**Figure 109 - Comparison of 3rd harmonic when both 3rd and 5th harmonics are suppressed for AWG.**

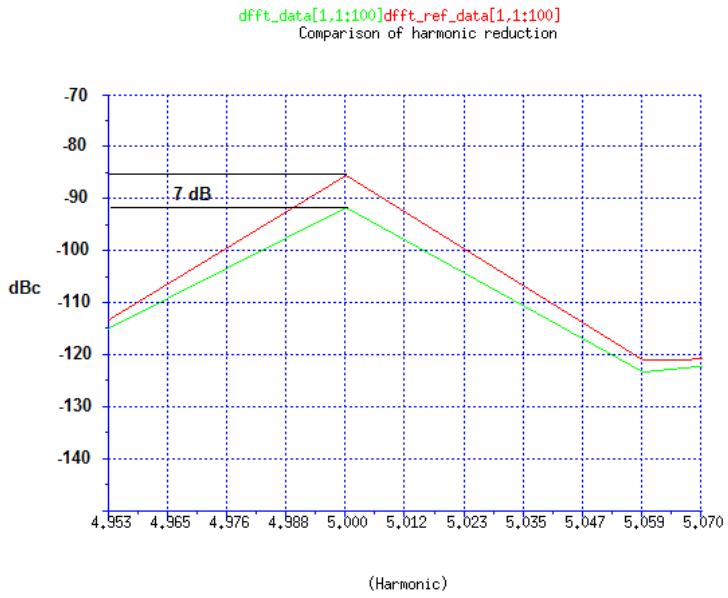



**Figure 110 - Comparison of 5th harmonic when both 3rd and 5th harmonics are suppressed for AWG.**

## 5.4 16-Bit ADC Example

No hardware was designed for these experiments, rather we used what was available to use from a production hardware standpoint with available devices for that test solution. For this 16-bit ADC test case, the ADC was tested using a 16-Bit AWG. The point of this exercise was to see the effect of the phase switching had on the THD performance of the AWG and the resulting effect on the performance of an ADC. As the THD performance of the AWG was much more inferior to the ADC performance that was chosen, it is of no consequence that usually an ADC is tested with an AWG with usually 2 more bits of resolution that the ADC under test.

Using a 16-bit ADC as a benchmark for this technique and using the standard technique for measuring THD, a value of -80 dB as shown in Figure 112, was measured with a tone frequency of 10 kHz and a sampling frequency of 259.2 kHz; however, we expected a much lower value from the simulated results. Next, we implemented the phase switching technique. However, it has to be remembered that the generated Nyquist sub-harmonic frequency components need to be filtered off as described in [33-35], so just applying the technique without an adequate filter would make the results look better than they should. Therefore, to implement this correctly, we needed to up sample the generated signal by a factor of 600 so that we could use a sampling frequency of 155 MHz. This would allow us to use a 2-signal interleaved approach as

$\frac{f_s}{2}$ would be approximately 77.5 MHz and would allow the use of the AWG's 50 MHz LPF, as shown in Figure 111. However, due the sampling restriction a 4-signal interleave approach was not possible and we were restricted to only suppressing a single harmonic. After careful verification it was seen that the 2nd harmonic was dominant (Figure 112), and a 2-signal interleave approach was adequate for this example. By suppressing the 2nd harmonic at this sampling scheme, an improved THD of -89.5 dB, as shown in Figure 113, was measured. To ensure that the observed results were repeatable over several tester platforms, this experiment was rerun on different testers of the same type. Referring to Table 6, one can see that similar results were observed for all systems used.

For the sake of completeness, it was desired to see the influence of the phase switching on the ADC itself and what influence it would have on the THD result. It would be expected that the result would improve if the filter were disabled due to the phase switching suppressing the harmonics of the ADC as well as the AWG. By disabling the filter with the initial sampling scheme of 259.2 kHz where the filter would have no influence, the THD was improved by 5 dB. In other words, if no filter is set, the results can be made to look artificially better than they should be. In this example, the sampling frequency of the ADC is the same as the AWG; if the sampling frequencies are different, the same rule cannot be applied. In other words, we observed experimentally that, when the sampling frequencies of the AWG and the ADC under test were equal, an analogue LPF filter followed by the AWG was required to measure ADC THD accurately with the phase switching technique; otherwise, the obtained ADC THD data was better than the actual, and this was shown in [34].

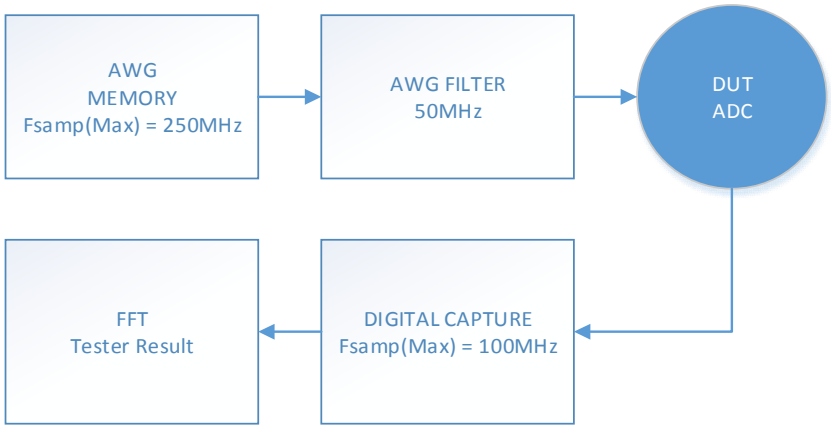



**Figure 111 - Tester setup with ADC**

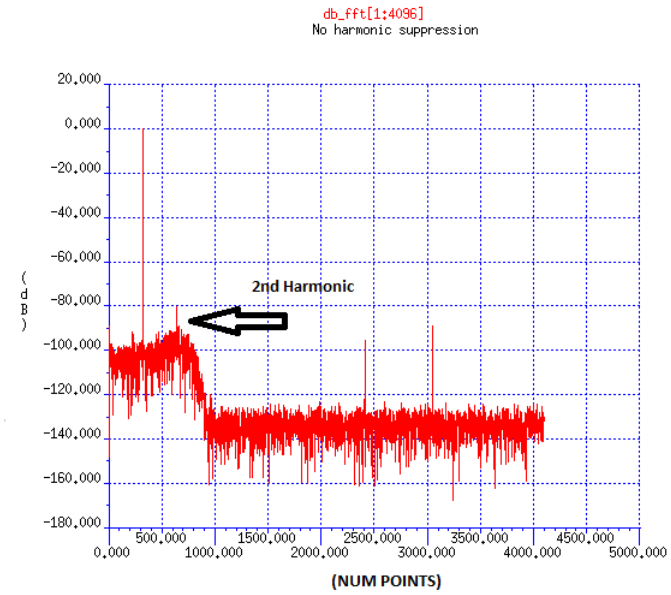



**Figure 112 - Spectrum of 16-bit ADC without suppression**

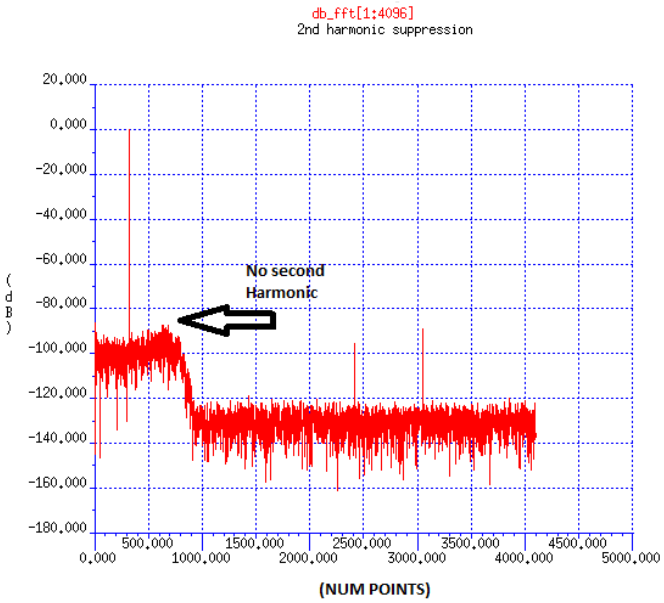



**Figure 113 - Spectrum of 16-bit ADC with 2nd harmonic suppressed**

**Table 9 – Tester to tester for 16-bit ADC**

| | THD vs Tester | |
|---|---|---|
| | *ref* | *suppression* |
| Tester 1 | -80 dB | -89.5 dB |
| Tester 2 | -84 dB | -88.0 dB |

## 5.5 12-Bit ADC Example

For this technique to be adopted, the solution needs to be universally applicable for all settings of an AWG. In the next application we use apply the technique to 12-bit ADC. Here we demonstrate the effect of suppressing each individual harmonic as well as a combination of harmonics. We show the effect that was seen across test sites of a load-board as well as tester-to-tester variation.

### 5.5.1 THD Measurement

Ideally, the harmonic performance of a 12-bit ADC should not be as good as a 16-bit ADC and hence the improvement in results should be much less. However, this is dependent on the architecture of the ADC and the resulting effective number of bits (ENOB) that would result from the signal-to-noise and distortion (SINAD) metric. Here we will show some important observations that can be made about the 0.35um CMOS silicon technology used based on the following results. The issue here was not the value of THD that was measured; this value was in the expected region of measurement.

However, there was not much margin between the influence of the AWG and the measured value of the ADC. By using the phase switching technique, it can be seen that using a combination of suppressing the $2^{nd}$ and $4^{th}$ harmonics leads to the best result as referring to Table 7. Referring to Table 8, it can be observed that the result correlates between testers also. However using the standard approach, the results show differences between the testers, and this suggests that the results using the standard approach are not accurate due to the influence of the harmonics. This was also true with the 16-bit ADC, referring to Table 6. It can be seen that there is large difference in result without suppression, but with suppression, we see the same result across tester platforms. This is great news for both correlation and gauge repeatability and reproducibility (GRR) exercises.

### 5.5.2 No Filter

For the 12-bit ADC the sampling frequency of the ADC is asynchronous to the tester and runs at its own speed. Therefore, in this example the sampling scheme is no longer the same for the ADC and AWG. It would not follow that by removing the filter in this instance would make the observed result better as the effect on the harmonic of the ADC would be different due to the AWG and ADC having different sampling rates. For this exercise, removing the filter actually made the result of the ADC worse by 2 dB.

### 5.5.3 Site to Site Correlation

Unfortunately, the 16-bit ADC example was only available in a single site test solution. However, the 12-bit ADC example was available as a dual site solution. This gave the possibility to check the test result improvement not only from a tester-to-tester point of view with the same model of instrument but also to check the 2nd channel of the AWG. For this exercise, we made a repeatability analysis to show the stability of the measurement between sites. Using the 2nd and 4th harmonics suppression, one device was looped a 100 times in site 1 as shown in Figure 114, and then site 2, as shown in Figure 115. It can be observed from Figure 114 and Figure 115 that the distribution of the result correlates very well and that the mean value measured is a good match, as shown in Table 9. The standard deviation result allows a perfect method for a test solution in that we have gauge repeatability, reproducibility, and most of all it is possible to correlate across sites and to the bench.

**Table 10 – THD versus suppressed harmonic tester 1**

| *suppression* | *THD* |
|---|---|
| ref | -89.7 dB |
| 2nd | -89.8 dB |
| 3rd | -89.3 dB |
| 4th | -89.5 dB |
| 5th | -89.2 dB |
| 2nd, 4th | -91 dB |
| 3rd, 5th | -89.7 dB |

**Table 11 –Tester to tester THD measurement for 12-bit ADC using 2nd & 4th harmonic reduction**

| | **THD vs Tester** | |
|---|---|---|
| | ***ref*** | ***suppression*** |
| Tester 1 | -89.7 dB | -91.0 dB |
| Tester 2 | -87.7 dB | -91.0 dB |

### 5.5.4 Standard Deviation Improvement

For the 12-bit ADC, there would not be much point to apply the harmonic suppression technique if the absolute limit was not so critical; however, this is only half of the problem. As already discussed, it can be observed that there is already quite a good standard deviation of the measured distribution. However, there has been no comparison to the initial standard deviation with the unsuppressed harmonics. In

addition, no comparison has been drawn with how the standard deviation of the THD measurement changes with different harmonic suppressions. To start it would be interesting to compare how the standard deviation of the measured result varies with the results observed in Table 7. Using the same testers and suppressing the 2nd and 4th harmonics, it can be also observed that the repeatability of the measurement also improves by 0.1 dB (see Table 9). This in itself is not a large improvement. However if we sacrifice some gain in harmonic suppression for a better repeatability of the result, we can improve the overall solution. Table 10 shows the harmonic reduction using only the 4th harmonic suppressed. It can be seen that the improvement of the THD is reduced by about 1.5 dB, resulting in no improvement at all of tester 1. However, Table 11 shows a great improvement to the repeatability of measurement of 0.5 dB. This in conjunction with the fact that the THD has equalized across the tester platforms results in a superior test program compared to the non-harmonically suppressed alternative.

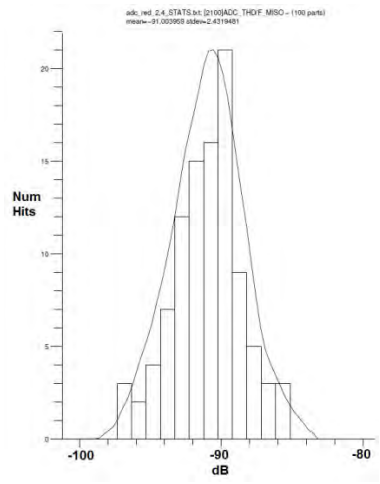



**Figure 114 - Repeatability of THD on site 1 of the 12-bit ADC example**

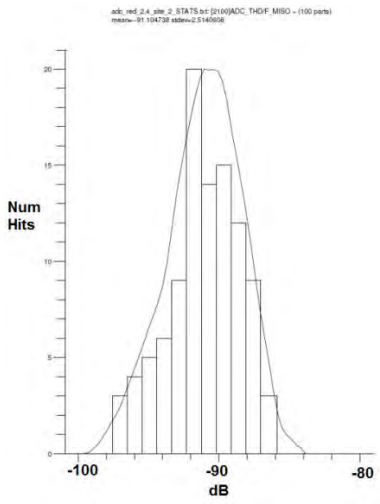



**Figure 115 - Repeatability of THD on site 2 of the 12-bit ADC example**

**Table 12 – Tester to tester STD measurement for 12-bit ADC using 2nd & 4th harmonic reduction**

| | THD(std) vs Tester | |
|---|---|---|
| | *ref* | *std* |
| Tester 1 | 2.5 dB | 2.4 dB |
| Tester 2 | 2.5 dB | 2.4 dB |

**Table 13 –Tester to tester THD measurement for 12-bit ADC using 4th harmonic reduction**

| | THD vs Tester | |
|---|---|---|
| | *ref* | *THD* |
| Tester 1 | -89.7 dB | -89.5 dB |
| Tester 2 | -87.7 dB | -89.3 dB |

**Table 14 –Tester to tester STD measurement for 12-bit ADC using 4th harmonic reduction**

| | THD(std) vs Tester | |
|---|---|---|
| | *ref* | *std* |
| Tester 1 | 2.5 dB | 2.0 dB |
| Tester 2 | 2.5 dB | 2.1 dB |

## 5.6 0.35um CMOS Technology for ADC

If one pays attention to the actual results of both the 12-bit and 16-bit successive approximation register (SAR) ADC harmonic performance results described in Chapter 5.4 & Chapter 5.5, both graphically and numerically. It can be seen that the actual measured result for both technologies is around -90 dBc, which is clearly independent of the number of bits of the converter. Although the focus of this paper is not the design performance of an ADC, what one can summarize is that the ENOB for both converters would be the same. Looking closer at the 12-bit design, it can be seen that the architecture utilizes 16-bits of resolution but throws away 4 bits in the conversion process, whereas the 16 bits converter uses all of the available bits. With a value of -90 dBc one can estimate that the ENOB should be around 14.7 bits. As can be seen from the two different designs this value is tied to the silicon process and the architecture. These findings were confirmed by simulation. Although not directly relevant to the topic, it is of interest for future debugging of such problems.

## 5.7 Drawbacks

As with any technique, there are always drawbacks. In this case, the harmonic performance of the AWG varies across the Xcerra MX test platform. From tester to tester, comparing the equivalent harmonic performance between AWG's can be up to 10 dB different and as such would be unreasonable to expect the same performance increase for each of these testers. What was found was that if the harmonic performance were already at the best that could be expected from the AWG, then by adding phase switching to the DSP would make the performance worse rather than improving it. What this means is that prior to loading the DSP waveforms into the AWG's memory, the AWG needs to be measured. This is done using the loopback shown in Figure 106 prior to testing. A decision can then be made in the run time code as to load either the standard sinewave or a phase switched version depending on the measured performance of the AWG.

## 5.8 Conclusion

It has been shown in this paper that by using a simple DSP technique in combination with appropriate analogue filters that the performance of an AWG can be extended to test higher performance analogue circuitry. The author has developed and industrialised an otherwise academic research exercise and has shown that by using this technique the performance of the AWGs within an Xcerra MX tester can be equalized. Issues that have plagued test engineers for years such as GRR between testers and between load-board test sites can now be minimized with the industrialised approach. As the difference in results between testers are now much smaller allows for smaller guard-bands, resulting in better yield and confidence in the delivered quality of product. This was demonstrated with 12-bit and 16-bit ADC testing with ATE systems.

# 6 A DISTORTION SHAPING TECHNIQUE TO EQUALIZE INTERMODULATION DISTORTION PERFORMANCE OF INTERPOLATING ARBITRARY WAVEFORM GENERATORS IN AUTOMATED TEST EQUIPMENT

## 6.1 Introduction

In many cases, the specifications of ATE instruments are much worse than the actual performance that is achieved within a test system. This is because less stringent specifications can be written into the tester instrumentation specifications. This results in the ability to supply boards that do not perform as efficiently as possible without the risk of having a board returned due to not meeting specifications that are more stringent; therefore, the yield is improved and the rework rate is reduced. These boards need to operate across many working areas so that they can be flexible to the user; therefore, the instrument design is a compromise amongst cost, performance, and flexibility. This

results in certain operating regions of the instrument actually being better than its specification. However, the ATE providers will not warrant these criteria because the warranty puts pressure to guarantee the specification with additional calibrations and checkers that are sometimes simply not possible in a production environment without much more additional work.

In today's world, there are many instances that the guaranteed ATE performance is no longer adequate to test new devices coming to the market without the need to invest in newer and more expensive test equipment. This in turn is driving up the test cost of certain analogue/mixed-signal semiconductor devices. As the performance of modern analogue/mixed-signal devices gets ever increasingly better, the pressure to keep reducing the test cost continues unabated; therefore, this results in the continuing drive to guarantee more complex device specifications with the existing tester installed base.

This paper expands on previous works in phase switching techniques [33-35, 37-38, 40-45] and shows how IMD products can be improved using two-tone phase switching techniques [34,35,37]. However, these existing literatures on the phase switched two-tone IMD testing only describe how to add the two-tone signals together, but they do not address the issue of AWGs using special architecture converters such as interpolating DACs [46-48]; an example is the DAC1617D1G0HN from Integrated Device Technology. This paper extends our previous works and derives suppression equations that demonstrate how the IMD tones of the non-linear DAC can be suppressed mathematically. In addition, this paper evaluates the effectiveness of the developed phase switching algorithms in the two-tone input case for mixed-signal circuit testing in an ATE environment quantitatively and shows how these algorithms can be used effectively to equalize the performance of the AWGs across the installed tester base. There has been some work done using hardware solutions to produce predistortion to suppress IMD products [37-38] but this is very specific to each individual setup, frequency dependent and not feasible for existing instruments installed in the field. This paper further develops the work done in [49], by further development of the mathematics pertaining to interpolating DACs and producing a characterization of the effect of the phase on the IMD performance of the AWG.

The restriction of testing IMD of any analogue/mixed-signal semiconductor device is limited to the base performance of the instrument for testing the Device Under Test (DUT); in this case, that is an AWG. If the AWG performance is worse than the DUT,

then the performance of the AWG will be measured rather than that of the DUT. If a dual-site test program is deployed and one AWG is better than another within the system, it is feasible that for one test site, the true ADC performance will be measured, whilst on the second site the performance of the AWG is reflected through the ADC; hence the true IMD of the DUT is not measured correctly. This would lead to site-to-site mismatch and yield issues if the second site was marginally good. However as long as the IMD tones of the AWG are well below the specified limit of the AWG, a true test of the ADC is possible although with a poor Gauge Repeatability and Reproducibility (GRR) between sites. We will show how it is possible to equalize the IMD tones of a set of AWGs, such that the IMD tones are below the level required to guarantee the performance of an ADC. In addition, there is the case of dual or I-Q ADCs that need balanced spectrum so that they operate at their most efficient, and having two AWGs (that have different spectrum) could potentially cause issues for these types of DUT; hence, a way of equalizing the performance of the AWGs would be beneficial.

The benefit of this work is such that the performance of an AWG can be significantly improved by using an alternative algorithm for generating the two-tone signal with the use of on-board filters that all AWG's have included in their design architecture. This results in equalizing the Intermodulation Distortion (IMD) performance of poor performing AWG's that are marginally worse than other AWG's installed within the tester fleet. This in turn results in the need for less stringent guard banding due to tester instrument performance differences. As such, the confidence in measuring the same IMD value across a company's tester fleet around the globe is increased and the effort in producing Gauge Repeatability and Reproducibility (GRR) is significantly reduced. Suppression of the IMD products of an AWG also increases the observed Effective Number of Bits (ENOB) of the instrument; hence, this extends the lifetime of the tester as the need to go to the next generation tester is delayed, and thus reduces the cost of purchasing new test instruments or even more critically a completely new test platform.

This chapter is organized in the following way. Chapter 6.2 derives the mathematics of the suppression techniques used in the described solution. It is shown graphically and mathematically how the suspected behaviour of an interpolating DAC reacts to the derived mathematics and how this could improve the performance. Chapter 6.3 shows how the system reacts to the suppressed waveforms and how this goes someway to prove the developed theory in Chapter 6.2. This solution is then demonstrated across

multiple AWGs within one test system, and it is shown how an improvement of the intermodulation distortion can be made using the developed solution. Chapter 6.4 describes a characterization procedure carried out on the AWG where the phase and frequency offset of the two tone suppressed signal were swept and the resulting improvement/degradation plotted on a 3D chart. Chapter 6.5 shows how the perceived performance of an ADC is improved by applying the developed techniques to the AWG that is testing the ADC and then conclusions are drawn.

This chapter further develops the work done in Chapter 5, [34 ,40], by further development of the mathematics pertaining to interpolating DACs and producing a characterization of the effect of the phase on the IMD performance of the AWG.

## 6.2 Theory

### 6.2.1 Two-tone Phase Switching Algorithm

Using the techniques in the current literature [40], an algorithm for the 3rd order intermodulation distortion suppression can be derived using a two signal interleaved algorithm. Mathematically it is possible to generate two-tone with the 3rd order tones fully suppressing them. Suppose that the DAC in the AWG has the 3rd, 5th, 7th, 9th , and 11th order distortion.

$$Y(nT_s) = a_1 D_{in}(n) + a_3 D_{in}(n)^3 + a_5 D_{in}(n)^5 + a_7 D_{in}(n)^7 + a_9 D_{in}(n)^9 + a_{11} D_{in}(n)^{11} . \quad \textbf{(49)}$$

where a1 = 1.0,

a3 = -5.0 x $10^{-6}$,

a5 = -1.5 x $10^{-6}$,

a7 = -1.8 x $10^{-6}$,

a9 = -2.2 x $10^{-6}$,

a11 = -0.4 x $10^{-6}$.

We use the DAC input in the AWG as follows:

$$D_{in}(n) = \begin{cases} X_1(n) \text{ in case } n\text{: even} \\ X_2(n) \text{ in case } n\text{: odd} \end{cases} \quad \textbf{(50)}$$

Then we have the AWG output as follows:

$$Y(nT_s)\begin{cases} a_1X_1(n) + a_3X_1(n)^3 + a_5X_1(n)^5 + a_7X_1(n)^7 + a_9X_1(n)^9 + a_{11}X_1(n)^{11} \\ \text{in case } n\text{: even} \\ a_1X_2(n) + a_3X_2(n)^3 + a_5X_2(n)^5 + a_7X_2(n)^7 + a_9X_2(n)^9 + a_{11}X_2(n)^{11} \\ \text{in case } n\text{: odd} \end{cases} \quad \textbf{(51)}$$

Consider using the following, $X_1$ (n), $X_2$ (n), where *IMD*=3, A & B are the amplitude values of each tone:

$$X_1(n) = A\sin\left(2\pi f_1 nT_s + \frac{\pi}{2\,IMD}\right) + B\sin\left(2\pi f_2 nT_s - \frac{\pi}{2\,IMD}\right). \quad \textbf{(52)}$$

$$X_2(n) = A\sin\left(2\pi f_1 nT_s - \frac{\pi}{2\,IMD}\right) + B\sin\left(2\pi f_2 nT_s + \frac{\pi}{2\,IMD}\right) \quad \textbf{(53)}$$

It can be seen graphically, from simulation in Figure 116, that 2f1-f2, 2f2-f2 are aggressively driven into the noise, and 3f1-2f2 and 3f2-2f1 are slightly reduced. However, the tone frequencies are also reduced by the same amount so this is not a real reduction in the IMD3 tones as the difference from the tone frequencies themselves remains. This is important and needs to be considered for all phase-switching implementations; otherwise, the change in spectral components could be simply due to reducing the input power of the signal to a device.

Now using $X_1(n)$, $X_2(n)$, Equations (52 – 53) where IMD=5, A & B are the amplitude values of each tone, it can be seen graphically, from simulation in Figure 117, that 3f1-2f2, 3f2-2f1 are aggressively driven into the noise and 2f1-f2 , 2f2-f1, 4f1-3f2 and 4f2-3f1 are slightly reduced. However, in this case, the difference between the tone power and the spectral components is much larger and can be considered for use.

Now using $X_1(n)$, $X_2(n)$, Equations (52 - 53) where IMD=4, A & B are the amplitude values of each tone, it can be seen graphically, from simulation in Figure 118, that 2f1-f2 , 2f2-f1 and 3f1-2f2, 3f2-2f1 are slightly reduced and would be perfect solution where IMD3 and IMD5 are needed to be reduced but not aggressively driven into the noise. Using the techniques in the current literature [40], an algorithm for the 4$^{th}$ order intermodulation distortion suppression can be derived. It can be seen that 2f1-f2, 2f2-f1 and 3f1-2f2, 3f2-2f1, tones are reduced, Equation (54), by the amount seen in Figure 118, Table 15, and hence mathematically it is possible to generate a two-tone signal with the 3$^{rd}$ and 5$^{th}$ order tones reduced by significant amount.

Where $\alpha = 2\pi f_1 nT_s$, $\beta = 2\pi f_2 nT_s$, $\varphi_1 = \frac{\pi}{8}$, and $D_{in}(n) = A\sin(\alpha) + B\sin(\beta)$

$$Y(nT_s) = cos(\varphi_1)\cdot Aa_1 sin(\alpha) + cos(\varphi_1)\cdot Ba_1 sin(\beta) + cos(3\varphi_1)\cdot\frac{3}{4}A^2Ba_3 sin(2\alpha-\beta) + cos(3\varphi_1)\cdot\frac{3}{4}AB^2a_3 sin(\alpha-2\beta) + cos(5\varphi_1)\cdot\frac{5}{8}A^3B^2a_5 sin(3\alpha-2\beta) + cos(5\varphi_1)\cdot\frac{5}{8}A^2B^3a_5 sin(2\alpha-3\beta) \quad \textbf{(54)}$$

**Table 15 –Derived Intermodulation Product Reduction**

| IMD products | Attenuation | |
|---|---|---|
| | ***Formula*** | ***Value (dB)*** |
| IMD3 | $cos\,3(\varphi_a)$ | -8.34 |
| IMD5 | $cos\,5(\varphi_a)$ | -8.34 |

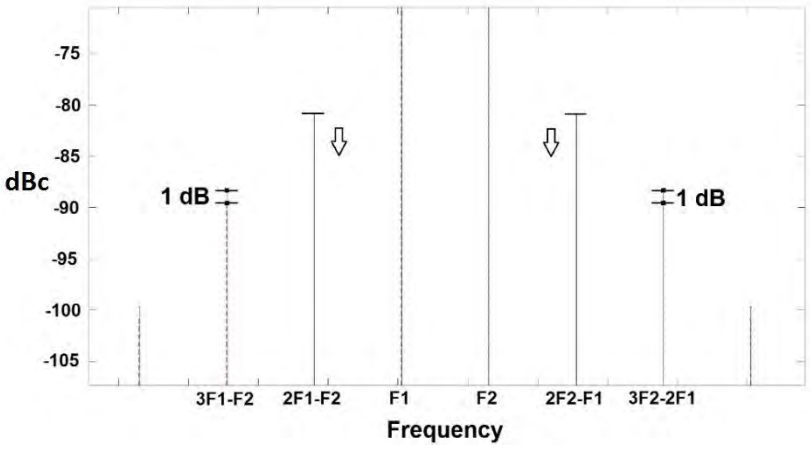



**Figure 116 - Simulation of conventional two-tone generation vs 2-signal interleaved IMD3 suppression technique**

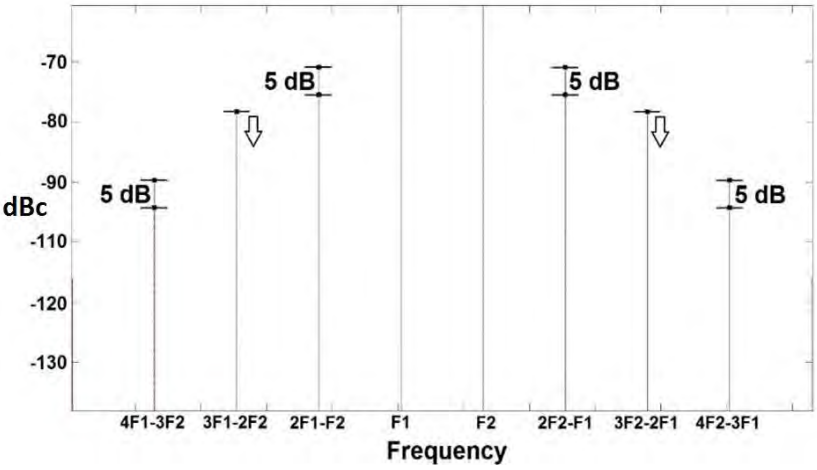



**Figure 117 - Simulation of conventional two-tone generation vs 2-signal interleaved IMD5 suppression technique**

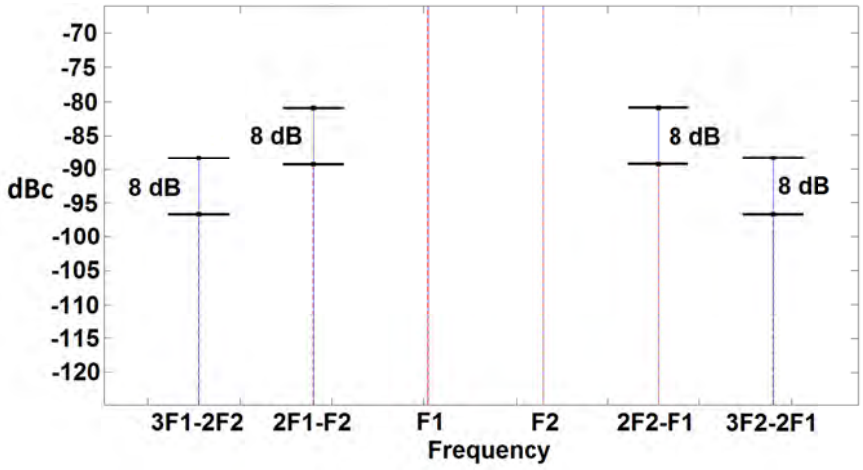



**Figure 118 – Simulation of conventional two-tone generation vs 2-signal interleaved IMD4 suppression technique**

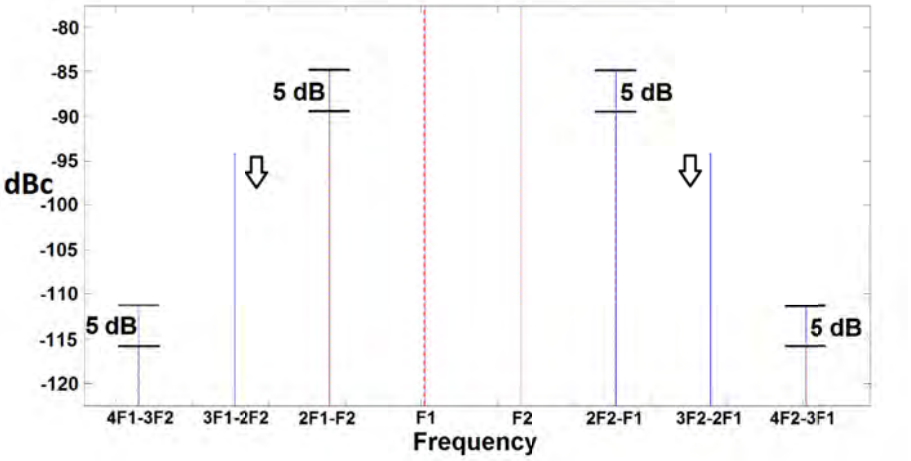



**Figure 119 - Simulation of conventional two-tone generation vs 2-signal interleaved IMD3 suppression technique with only 7th order non-linearity.**

### 6.2.2 2-Signal Interleaved effect on higher order distortion products

Now consider a device that does not exhibit 3rd or 5th order distortion but for the sake of argument, has only 7th order distortion, as shown in Equation (55). Due to spectral components that 7th order distortion occupies, distortion components will be seen in the same locations, $\sin(2\pi(F1 - 2F2))$, $\sin(2\pi(2F1 - F2))$, $\sin(2\pi(2F1 - 3F2))$, $\sin(2\pi(3F1 - 2F2))$, $\sin(2\pi(3F1 - 4F2))$, $\sin(2\pi(4F1 - 3F2))$, as shown in Equation (56), IMD components highlighted in bold, as would be seen if there were 3rd and 5th order distortions in Figure 119. However, the phase switching algorithm does not discriminate with how the frequency content was generated, in that if there is spectral content in the location targeted by the two-tone switching algorithm, then it will

be suppressed. Figure 119 shows the IMD5 being suppressed with the spectral content only being generated from 7th order distortion. Figure 119 shows graphically, through simulation, the spectral content of a two-tone signal, with only 7th order non linearity. The 5th order distortion suppressed phase switched version is overlaid. The suppression of the 3rd and 7th tones are approximately the same as the values shown in Figure 117, which has distortion components defined in Equation (49). However, IMD5 is fully suppressed even though the spectral content is derived from a 7th order distortion. From a design perspective, once very good 3rd and 5th order distortions are achieved, it is then extremely important to ensure there is an even higher level of 7th and 9th order performance; otherwise the 3rd and 5th order terms generated by the higher order distortion terms will swamp the carefully crafted self-generated 3rd and 5th order distortion products. It can be seen from the IMD terms highlighted in bold that 3rd order terms increase by a factor of $a_7 \frac{\mathbf{735}}{\mathbf{64}}$ i.e. and the 5th order by $a_7 \frac{\mathbf{245}}{\mathbf{64}}$ i.e. the 3rd and 5th order components attributed to the 7th order distortion increase by a factor of 10 and 4 respectively. Therefore, it is easy to imagine a device with very small 3rd order distortion could see its IMD3 spectral component dominated by contributions from 7th order distortion.

$$Y(nT_s) = a_1 D_{in}(n) + a_7 D_{in}(n)^7 \quad \textbf{(55)}$$

Where $\alpha = 2\pi f_1 nT_s$, $\beta = 2\pi f_2 nT_s$ , $D_{in}(n) = A\sin(\alpha) + B\sin(\beta)$

$$\begin{aligned}
&Y(nT_s) = a_1[A\,sin(\alpha) + B\,sin(\beta)] \\
&+a_7[\,\mathbf{\frac{735}{64}sin(\alpha - 2\beta)} + \mathbf{\frac{735}{64}sin(2\alpha - \beta)} - \mathbf{\frac{245}{64}sin(2\alpha - 3\beta)} \\
&+\mathbf{\frac{245}{64}sin(3\alpha - 2\beta)} - \mathbf{\frac{35}{64}sin(3\alpha - 4\beta)} + \mathbf{\frac{35}{64}sin(4\alpha - 3\beta)} - \frac{7}{64}sin\,(\alpha - 6\beta) \\
&+\frac{147}{64}sin(\alpha - 4\beta) - \frac{147}{64}sin(4\alpha - \beta) + \frac{7}{64}sin(6\alpha - \beta) - \frac{735}{64}sin(2\alpha + \beta) \\
&+\frac{147}{64}sin(4\alpha + \beta) - \frac{7}{64}sin(6\alpha + \beta) - \frac{735}{64}sin(\alpha + 2\beta) + \frac{245}{64}sin(3\alpha + 2\beta) \\
&-\frac{21}{64}sin(5\alpha + 2\beta) + \frac{245}{64}sin(2\alpha + 3\beta) - \frac{35}{64}sin(4\alpha + 3\beta) + \frac{147}{64}sin(\alpha + 4\beta) \\
&-\frac{35}{64}sin(3\alpha + 4\beta) - \frac{21}{64}sin(2\alpha + 5\beta) - \frac{7}{64}sin(\alpha + 6\beta) + \frac{1225}{64}sin(\alpha) \\
&+\frac{1225}{64}sin(\beta) - \frac{441}{64}sin(3\alpha) - \frac{441}{64}sin(3\beta) + \frac{49}{64}sin(5\alpha) + \frac{49}{64}sin(5\beta) - \frac{1}{64}sin(7\alpha) \\
&-\frac{1}{64}sin(7\beta)]
\end{aligned} \quad \textbf{(56)}$$

### 6.2.3 4-Signal Interleaved Two-tone IMD Suppression

Producing a 4-signal interleaved that combines Figure 116 and Figure 117, has been written in the previous literature [40], and nothing would be new reproducing this. However, there has been no investigation in using 4-signal interleaved signals to reduce the harmonic content by a fraction rather than fully suppressing the distortion components. We will discuss the need for such a technique in next sections but for now, we will discuss the theory behind such a proposal. Again, assuming a DAC in the AWG has the 3rd, 5th, 7th, 9th, and 11th order distortions with the coefficients defined previously and using a 4-signal interleaved approach with the 4th and 5th intermodulation distortion products suppressed, we can substitute Equations (59 - 62) into Equation (58). By simulating Equation (58) with Equations (59 - 62), Figure 120 reveals that the IMD3 components have been reduced by around 13 dB, the IMD5 components have been fully suppressed and the IMD7 by around 5 dB. Using the techniques in the current literature [40], an algorithm for the 4th and 5th order intermodulation distortion suppression can be derived. It can be seen that 2f1-f2, 2f2-f1, 3f1-2f2, 3f2-2f1, 4f2-3f1 and 4f1-3f2 tones are reduced, Equation (63), by the amount seen in Figure 120, Table 16, and hence mathematically it is possible to generate a two-tone signal with the 5th order tones fully suppressed along with the 3rd and 7th order tones reduced by significant amount.

$$Y(nT_s) = a_1 D_{in}(n) + a_3 D_{in}(n)^3 + a_5 D_{in}(n)^5 + a_7 D_{in}(n)^7 + a_9 D_{in}(n)^9 + a_{11} D_{in}(n)^{11}. \quad \textbf{(57)}$$

$$= \begin{cases} a_1 X_3(n) + a_3 X_3(n)^3 + a_5 X_3(n)^5 \\ + a_7 X_3(n)^7 + a_9 X_3(n)^9 + a_{11} X_3(n)^{11} \ (n = 4k) \\ a_1 X_4(n) + a_3 X_4(n)^3 + a_5 X_4(n)^5 \\ + a_7 X_4(n)^7 + a_9 X_4(n)^9 + a_{11} X_4(n)^{11} \ (n = 4k+1) \\ a_1 X_5(n) + a_3 X_5(n)^3 + a_5 X_5(n)^5 \\ + a_7 X_5(n)^7 + a_9 X_5(n)^9 + a_{11} X_5(n)^{11} \ (n = 4k+2) \\ a_1 X_6(n) + a_3 X_6(n)^3 + a_5 X_6(n)^5 \\ + a_7 X_6(n)^7 + a_9 X_6(n)^9 + a_{11} X_6(n)^{11} \ (n = 4k+3) \end{cases} \quad \textbf{(58)}$$

$$X_5(n) = A\sin\left(2\pi f_1 nT_s - \frac{9}{40}\pi\right) + B\sin\left(2\pi f_2 nT_s - \frac{9}{40}\pi\right). \quad \textbf{(59)}$$

$$X_6(n) = A\sin\left(2\pi f_1 nT_s - \frac{1}{40}\pi\right) + B\sin\left(2\pi f_2 nT_s - \frac{1}{40}\pi\right). \quad \textbf{(60)}$$

$$X_7(n) = A\sin\left(2\pi f_1 nT_s + \frac{1}{40}\pi\right) + B\sin\left(2\pi f_2 nT_s + \frac{1}{40}\pi\right). \quad \textbf{(61)}$$

$$X_8(n) = A\sin\left(2\pi f_1 nT_s + \frac{9}{40}\pi\right) + B\sin\left(2\pi f_2 nT_s + \frac{9}{40}\pi\right). \quad \textbf{(62)}$$

Where $\alpha = 2\pi f_1 nT_s$, $\beta = 2\pi f_2 nT_s$, $\varphi_a = \frac{\pi}{8}$, $\varphi_b = \frac{\pi}{10}$

and $D_{in}(n) = A\sin(\alpha) + B\sin(\beta)$

$$Y(nT_s) = \frac{1}{2}(cos(\varphi_a + \varphi_b) + cos(\varphi_a - \varphi_b)) \cdot Aa_1 sin(\alpha) + \frac{1}{2}(cos(\varphi_a + \varphi_b) + cos(\varphi_a - \varphi_b)) \cdot Ba_1 sin(\beta) + \frac{1}{2}(cos\,3(\varphi_a + \varphi_b) + cos\,3(\varphi_a - \varphi_b)) \cdot \frac{3}{4}A^2Ba_3 sin(2\alpha - \beta) + \frac{1}{2}(cos\,3(\varphi_a + \varphi_b) + cos\,3(\varphi_a - \varphi_b)) \cdot \frac{3}{4}AB^2a_3 sin(\alpha - 2\beta) + \frac{1}{2}(cos\,5(\varphi_a + \varphi_b) + cos\,5(\varphi_a - \varphi_b)) \cdot \frac{5}{8}A^3B^2a_5 sin(3\alpha - 2\beta) + \frac{1}{2}(cos\,5(\varphi_a + \varphi_b) + cos\,5(\varphi_a - \varphi_b)) \cdot \frac{5}{8}A^2B^3a_5 sin(2\alpha - 3\beta) + \frac{1}{2}(cos\,7(\varphi_a + \varphi_b) + cos\,7(\varphi_a - \varphi_b)) \cdot \frac{35}{64}A^4B^3a_7 sin(4\alpha - 3\beta) + \frac{1}{2}(cos\,7(\varphi_a + \varphi_b) + cos\,7(\varphi_a - \varphi_b)) \cdot \frac{35}{64}A^3B^4a_7 sin(3\alpha - 4\beta) \quad \textbf{(63)}$$

**Table 16 –Derived Intermodulation Product Reduction**

| IMD products | Attenuation | |
|---|---|---|
| | ***Formula*** | ***Value (dB)*** |
| IMD3 | $\frac{1}{2}(cos\,3(\varphi_a + \varphi_b) + cos\,3(\varphi_a - \varphi_b))$ | -12.95 |
| IMD5 | $0(cos\,5(\varphi_a + \varphi_b) + cos\,5(\varphi_a - \varphi_b))$ | ∞ |
| IMD7 | $\frac{1}{2}(cos\,7(\varphi_a + \varphi_b) + cos\,7(\varphi_a - \varphi_b))$ | -5.3 |

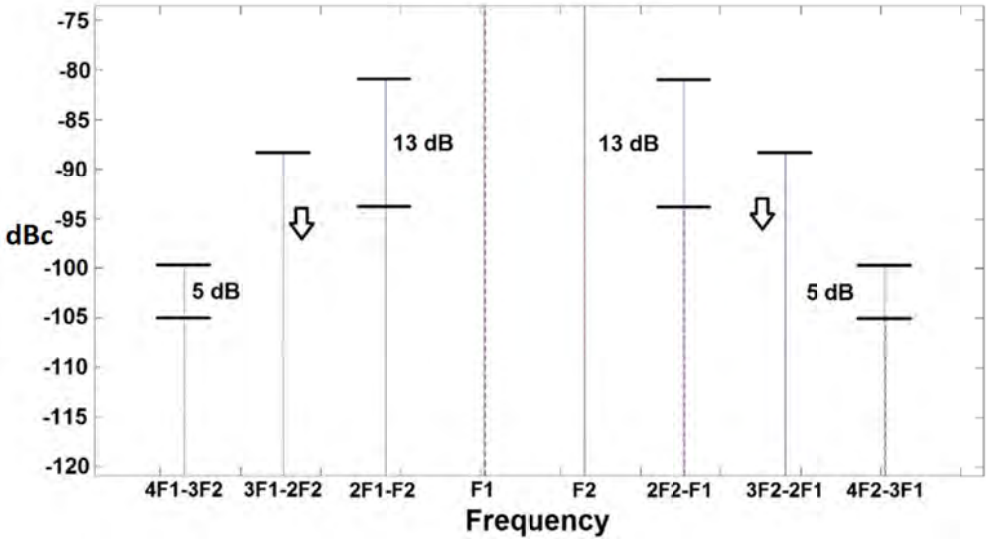



**Figure 120 - Simulation of conventional two tone generation vs 4-signal interleaved IMD4, 5 suppression technique**

### 6.2.4 Filter Considerations

Based on the current literature [40], we need to consider what filter needs to be used to remove the spectral components that have been generated at the Nyquist frequency. The required filter for the two-signal and four-signal interleaved waveforms are different. This is due to phase switching effect moving the spectral energy from the harmonic to sub Nyquist frequencies. Obviously, the more spectral energy that is moved the larger the occupied bandwidth used at the sub Nyquist frequencies. Hence, the cut off frequency of a low pass filter used to filter these sub Nyquist frequencies will become lower. The cut off frequency of the filter is actually dictated by both the sampling frequency, Fs, used and the signal frequencies used in the two-tone signal. However, if the sampling frequency is much greater than the tone frequencies, then the cut off filter requirement is dictated by the sampling frequency.

Considering the distortion components defined in Equation (49), the maximum distortion component is the $11^{th}$ order term; this means that there will be a frequency component at 11(f1 + f2) + (6f2-5f1) and thus, with the phase switching, a component will be present at Fs/2 -11(f1 + f2) - (6f2-5f1), Figure 121. As these components are a part of the suppression, then it is necessary to filter these components off to be able to use the signal as a real sinewave, otherwise the harmonic components of the device under test will also be suppressed, leading to the incorrect measurement of intermodulation distortion frequency components of a device. This can be written as Equation (64).

To be able to filter the sub Nyquist frequencies efficiently, there needs to be a separation between them and the generated standard distortion components. Therefore, the maximum distortion component needs to be considered when calculating the sampling frequency, $F_s$. This can be written as eq (65).

If we use the same reasoning for a 4-signal interleaved waveform, the occupied bandwidth in Figure 122 is now double that of the 2-signal interleaved waveform and hence now the image is now reflected to $F_s/4$. Due to this, it can be seen that the cut off frequency of the filter now needs to be 2x lower.

$$LPF_{Cutoff} = \frac{F_S}{2^{NTS}} - Dist_{MAX}(f_1 + f_2) - F_{IMD}\lfloor Dist_{MAX}\rfloor \quad \textbf{(64)}$$

$$F_s = 2^{NTS} \text{x} Dist_{MAX} \quad \textbf{(65)}$$

Where NTS stands for Number of Tones Suppressed and $Dist_{MAX}$ denotes the maximum distortion component present in system. $F_S$ is the sampling frequency of the AWG and $f_1, f_2$ are the two tone frequencies. $F_{IMD}[Dist_{MAX}]$ is the frequency of the maximum IMD component.

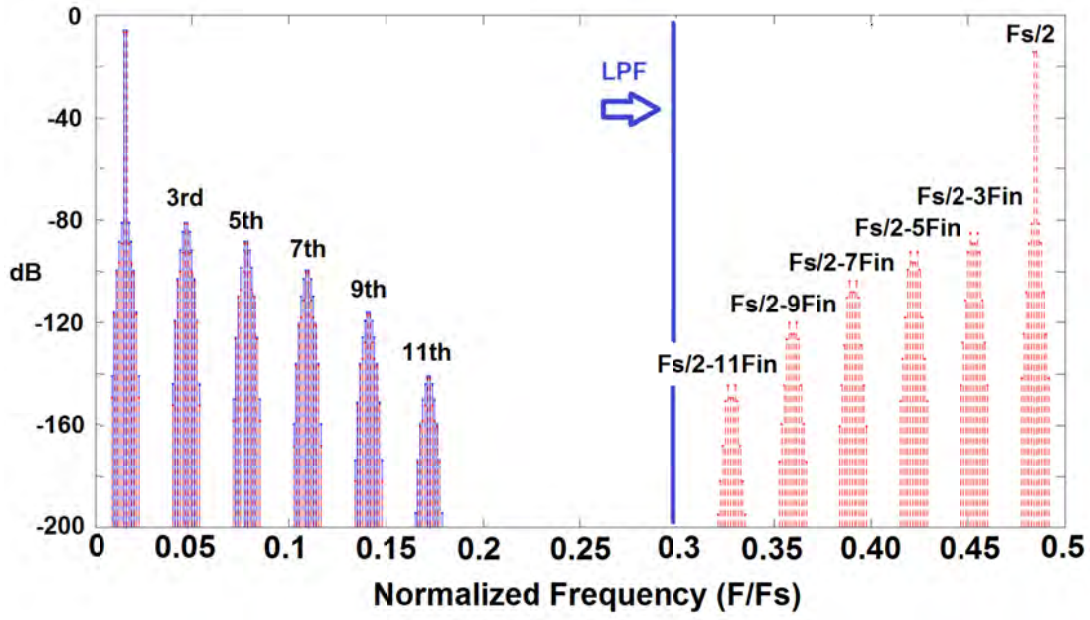



**Figure 121 - Distortion and sub Nyquist components of 2-signal phase switching**

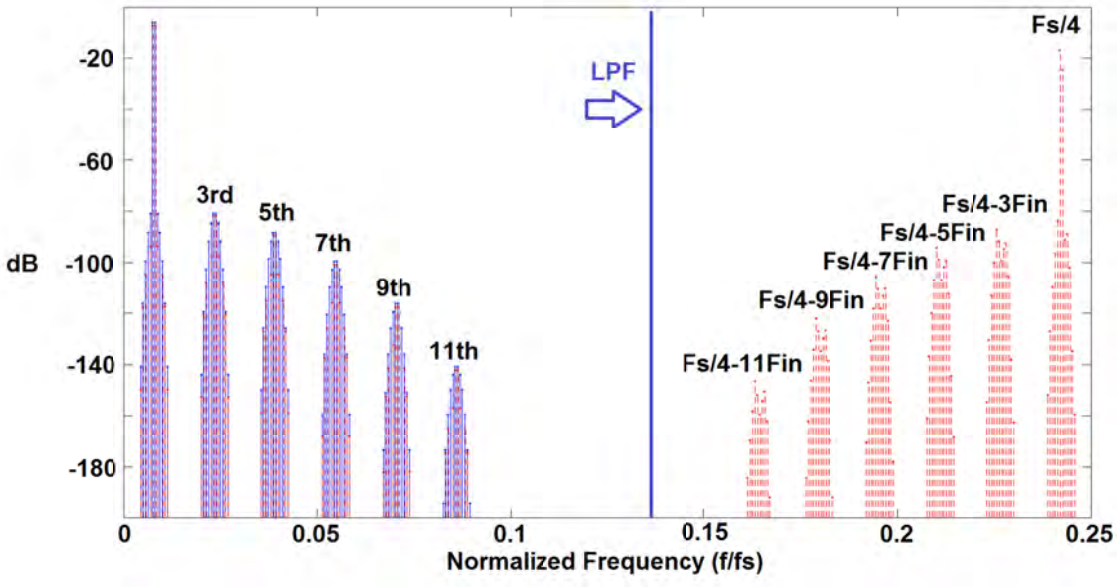



**Figure 122 - Distortion and sub Nyquist components of 4-signal phase switching**

### 6.2.5 Interpolation DAC Architecture

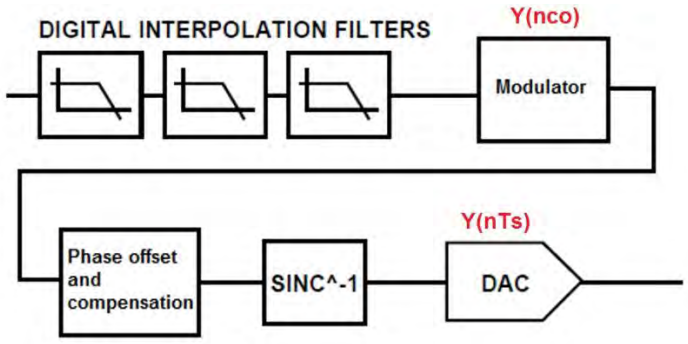



**Figure 123 - Architecture of an interpolating DAC**

An interpolating DAC, shown in Figure 123, usually incorporates an up sampler before the actual DA conversion [44-46]. One possibility of this functionality is to pad the incoming signal with the corresponding number of zeros with regard to the up sampling setting used, as shown in Figure 123. The amount of padding used corresponding to the chosen up sampling rate, i.e., if an up sampling rate of 2 is selected, then for every sample a zero is inserted. The up sampling process produces aliases by a factor of the up sampling, as shown in Figure 124. Figure 125 shows a segment of the time domain waveform of the phase switched signal with $4^{th}$ and $5^{th}$ IMD suppression that has been up sampled by a factor of 4. An interpolation filter is used to smooth the added zeros before the DA conversion, which is usually a finite impulse response (FIR) filter. When a signal is up sampled, in the frequency domain the sinc frequency response is moved out by a factor of the up sampling. As a copy of this signal produced every reciprocal of the up sampling of Nyquist, there would be no potential conflicts with the sub Nyquist frequencies of the phase switching process. This is because the sub Nyquist frequencies are also shifted down by a factor of the up sampling rate. Hence, the up sampled version of the 4-signal interleaved phase switched signal in Figure 121 would be shifted down by a factor of the up sampling rate. Therefore, the FIR filter would need to filter off up sampling but leave the sub Nyquist frequencies of the phase switching untouched; otherwise, a non-phase switched sinewave would be applied to the digital-to-analogue converter. If it is ensured that, the interpolation is a factor lower than the phase switching process, it further guarantees that the interpolation process will not interfere with the phase switching algorithm. This is because the phase switching causes and images at $\frac{F_S}{2^{NTS}}$, whereas the upsampling causes images at $\frac{1}{Intp_{fact}}$. Therefore, by making the NTS equal to the upsampling rate

will force the images of the upsampling outside of the spectrum of the sub Nyquist frequencies of the phase, hence allowing filtering of the up sampling. Using an example, this could be implemented using a 2x up sample in Figure 124, with a 4-signal interleaved phase switched signal in Figure 126 and an FIR interpolation by a factor of 4 in Figure 127. After FIR filtering, the output waveform can be seen to have maintained the up sampled frequency. Here the modulation rate has been removed leaving the waveform up sampled, but the zeros are smoothed to give an interpolated and continuous waveform, as shown in Figure 128. However, it is possible to disable the interpolation process within these DAC's and hence it is possible to use the phase switching technique to reduce harmonic content if the above equations are not possible to fulfil or cause degradation to the harmonic performance of the standard DAC. The down sample process is achieved by increasing the update rate of the DAC while keeping the data rate a lower frequency, as shown in Figure 129. However, comparing Figure 129 and Figure 116, and, Figure 120 and Figure 130, it can be seen that the interpolation process has been detrimental to the intermodulation products.

Figure 120 shows an interpolation process of a standard sinewave versus a phase switched version with the 3$^{rd}$ order tones suppressed. However, what can be seen is that the tones always done to the highest distortion product have been reduced by a similar amount, approximately 10 dB. Although Figure 130 shows a 4$^{th}$ and 5$^{th}$ order suppression compared to a standard sinewave, no observable difference in the total suppression of the tones is seen compared to Figure 129, the total suppression being approximately 60 dB across all intermodulation products. As the main reason for using interpolation is to reduce noise, the trade-off between using interpolation or not is dictated by the requirement for lower noise or lower distortion. Using interpolation or not, it has been shown that phase switching will help to reduce the intermodulation distortion products generated by the non-linearity of a DAC, but there is no need to use more than a 2-signal interleaved approach, as it delivers the same performance for interpolated signals as would a 4-signal interleaved approach. It was observed that this was independent of the up sampling factor used and the resulting reductions were approximately equal for all up sampling rates simulated.

The signal is then modulated using a Numerically Controlled Oscillator (NCO) to use the best frequency region of the DAC bandwidth. An inverse sinc filter is then applied to filter off all frequencies above Nyquist.

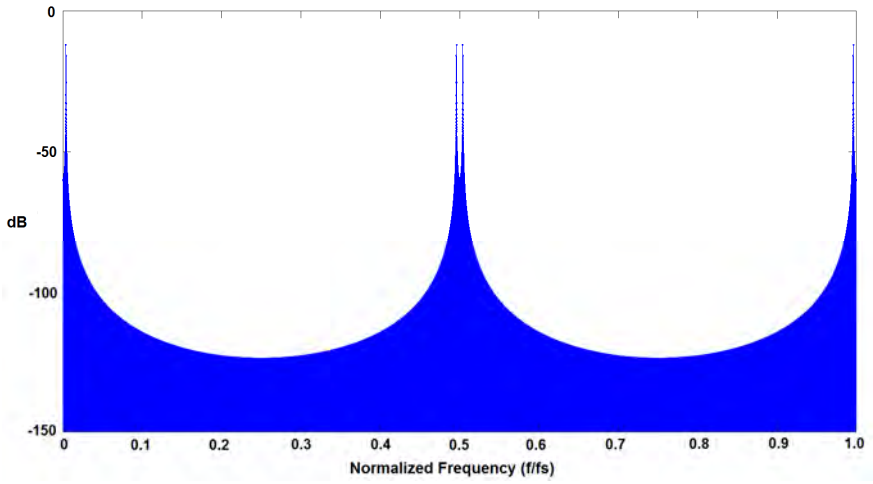



**Figure 124 - Frequency domain of a two tone signal up sampled by 2.**

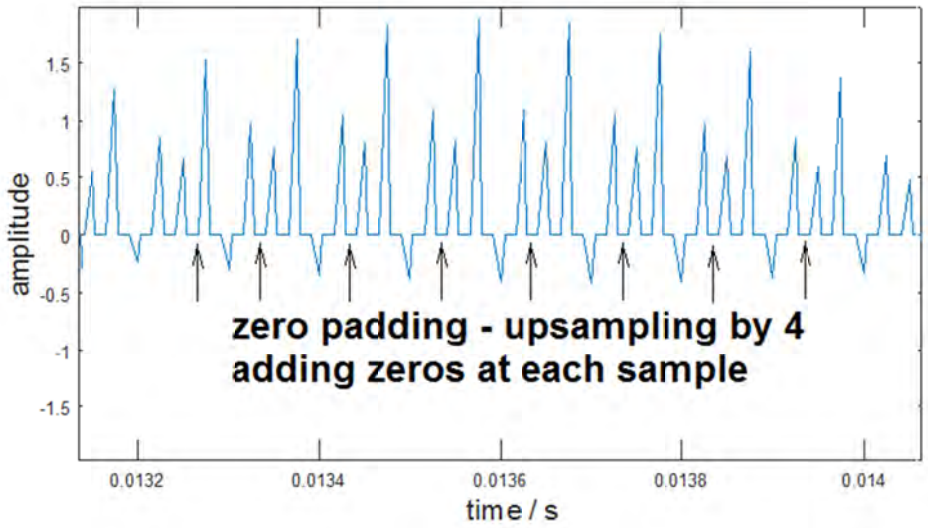



**Figure 125 - 4x Upsampled version of 4-signal phase switching**

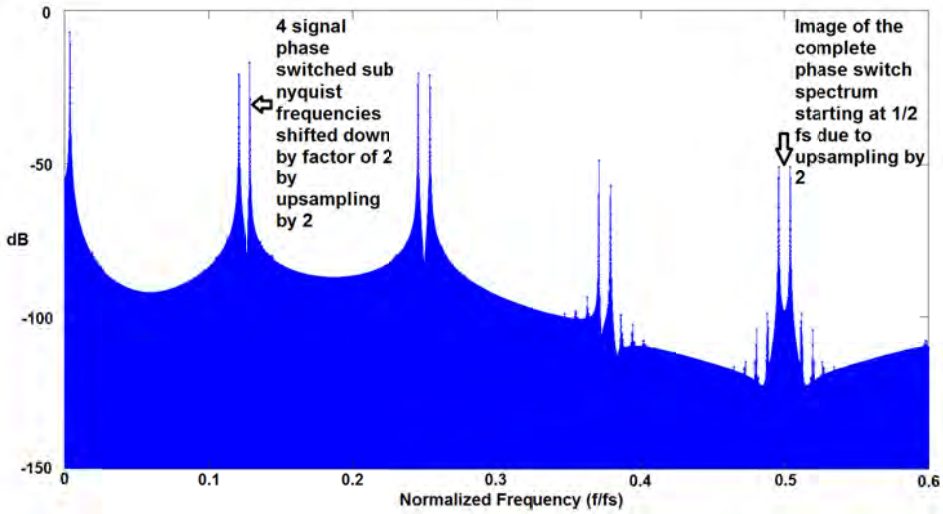



**Figure 126 - 4-signal phase switched, 2 two tone signals, up sampled x2**

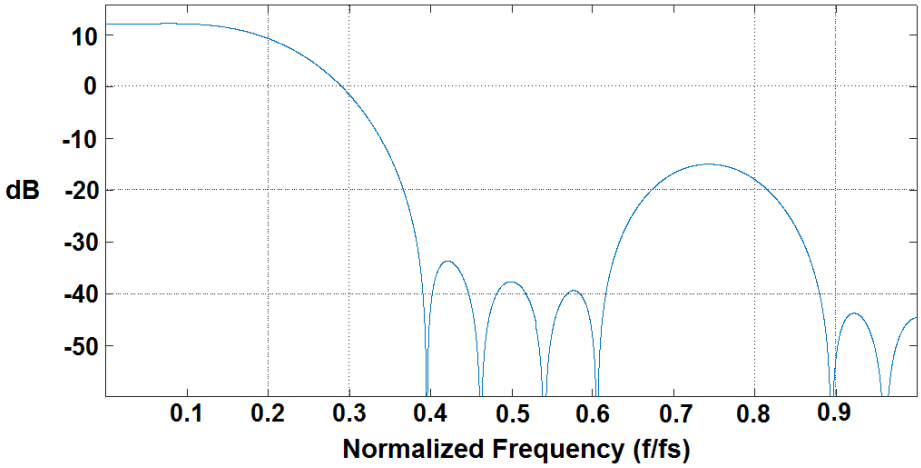



**Figure 127 - FIR with 4x interpolation, used to filter 2x up sampling**

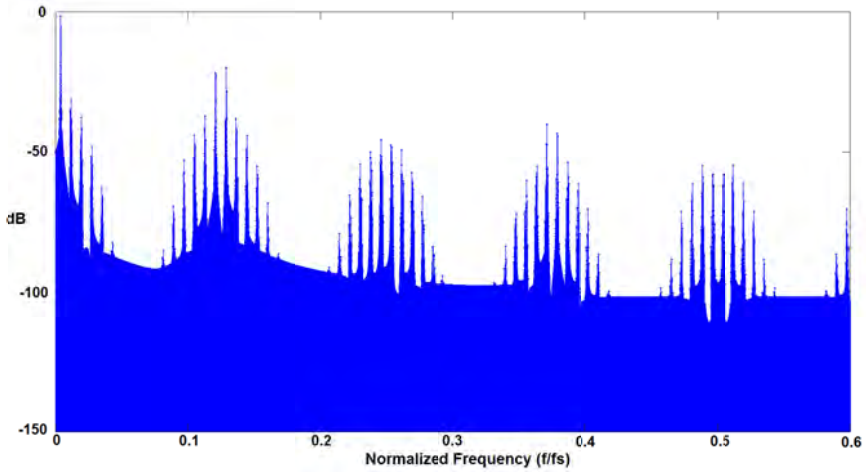



**Figure 128 - 4-signal phase switched signal with distortion, up sampled x2 and FIR filter interpolated by 4, 2x up sampling filtered, inserted zeros smoothed**

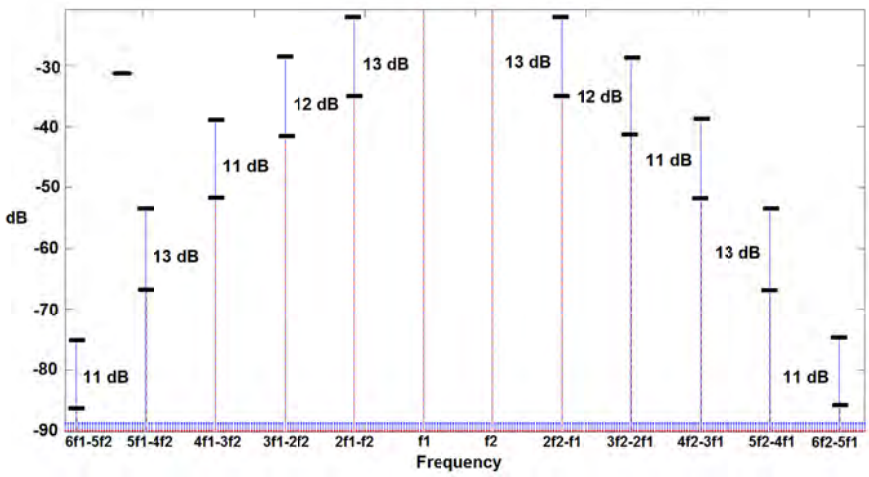



**Figure 129 - Interpolation process effect on both standard and 3rd suppressed phase switched sinewave**

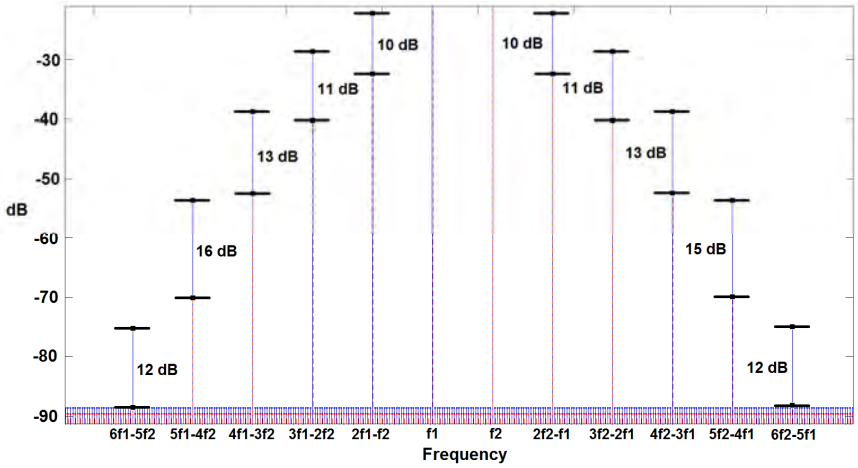



**Figure 130 - Interpolation process effect on both standard and 4th, 5th suppressed phase switched sinewave**

## 6.3 Measurement System

### 6.3.1 Intermodulation Distortion Spectrum Measurements

Before an exercise is undertaken to improve the measurement of a mixed signal block using the above described approach, one should check directly the influence of the prescribed method on the harmonic contents of the AWG directly (in this case the LTX-Credence 16-bit AWGHSB). To do this, the loopback calibration path of the load-board was used (Figure 106). This enables the user to loopback the AWG into the digitizer (in this case the LTX-Credence 14-bit DIGHSB), and measure the spectrum directly. The following plots show the influence of the harmonic suppression on the harmonic level compared to the reference using only a conventional sinus.

Figure 131 shows the effect of directly suppressing the IMD3 of the interpolating DAC using the 3rd order suppression technique. It can be seen that the spectral components in the IMD3 frequency space were degraded by around 5 dB by suppressing the 3rd order tones. In addition, Figure 132 shows the effect of directly suppressing the IMD5 of the interpolating DAC using the 5th order suppression technique. It can be seen that the spectral components in the IMD5 frequency space were degraded by around 5 dB by suppressing the 5th order tones. In both cases, by aggressively suppressing the intermodulation products resulted in an actual degradation of the performance rather than an improvement. However, what can be seen in both Figure 131 & Figure 132 is that there is an improvement in the other spectral components with regard to the suppressed. In Figure 131, the IMD5 component has

improved and in Figure 132, the IMD3 has improved, observing Figure 116 and Figure 117, whilst suppressing certain components also resulted in subtly reducing other distortion components. Although the effect of suppressing the IMD tone in question is not behaving as expected, the effect on the other IMD tones seems to be behaving as expected. This leads to the idea that maybe an interpolating DAC in Figure 123 cannot improve beyond a certain point and a less aggressive suppression is required. Figure 133 shows the effect of not suppressing either IMD3 or IMD5 fully but reducing slightly using IMD4 suppression, as shown in Figure 118. The result is an improvement in the intermodulation performance by an amount similar to the simulation results of around 5 dB.

Figure 131 shows an unexpected imbalance of the IMD5 tones. Generally, these tones should be balanced, as shown in Figure 116 - Figure 120 i.e., the component at 3f2-2f1 should be the same amplitude as 3f1-2f2. Reference [50] details that by having a phase changing circuit results in the intermodulation distortion tones becoming unbalanced. This would be the equivalent of the phase offset compensation circuit, within the interpolating DAC (Figure 123), that is causing the same effect.

If a 4-signal interleaved approach was used to reduce two intermodulation components at the same, it could be possible to completely clean up the bandwidth consisting of the IMD3, IMD5 and IMD7 frequency components if a less aggressive approach was taken to the suppression. Figure 134 shows the effect of suppressing the $4^{th}$ and $5^{th}$ order intermodulation distortion products and also shows approximately a 10 dB improvement in both IMD3 and IMD5. It can also be see that there is around 15 dB improvement in IMD7, which does not fully agree with the simulations shown in Figure 117 and Figure 118, as they only predict around a 5 dB improvement. However, the simulation shown in Figure 120 predicts a 25 dB improvement in IMD5.

For IMD5 we believe that the physical limit of the digitizer performance has been reached, therefore if any improvement beyond this performance limit has been achieved, it will not be observed. For the IMD3, as stated previously, we believe that if a too aggressive suppression is used, the result is a degradation rather than improvement in the IMD3 product and hence by using the $4^{th}$ and $5^{th}$ IMD suppression algorithm allows the user to not overly stress the interpolating DAC. As for the IMD7, the same response is seen in both Figure 132 and Figure 134 as well as the simulation showing the same expected 5 dB improvement in Figure 117 and Figure 120. Although not

expected, the improvement is seen repeatably, using different schemes and hence is systematic, the improvement compared to the simulation is suspected to be due to the effect of phase compensation circuitry.

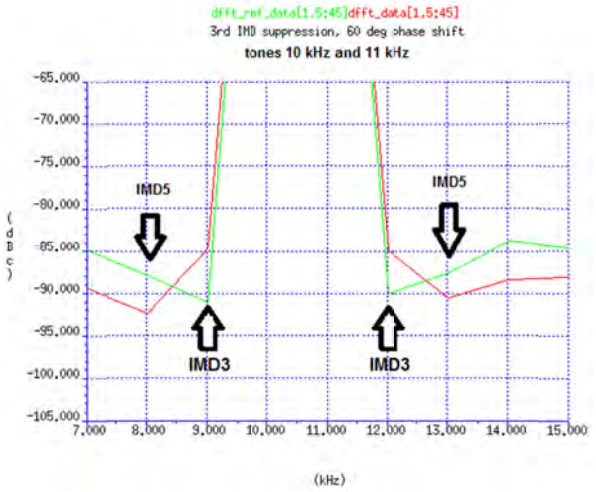



**Figure 131 – IMD3 degradation caused by suppressing 3rd order tones**

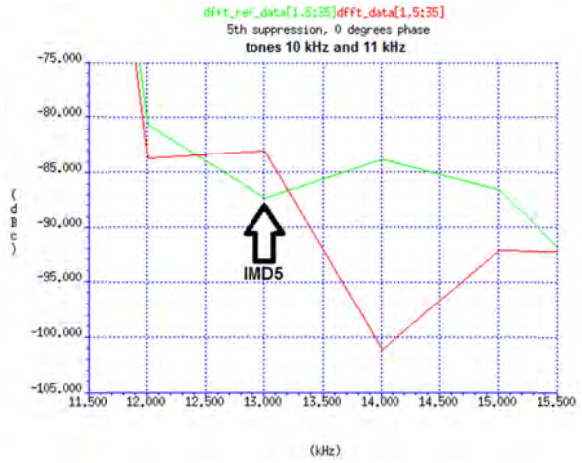



**Figure 132 - IMD5 degradation caused by suppressing 5th order tones**

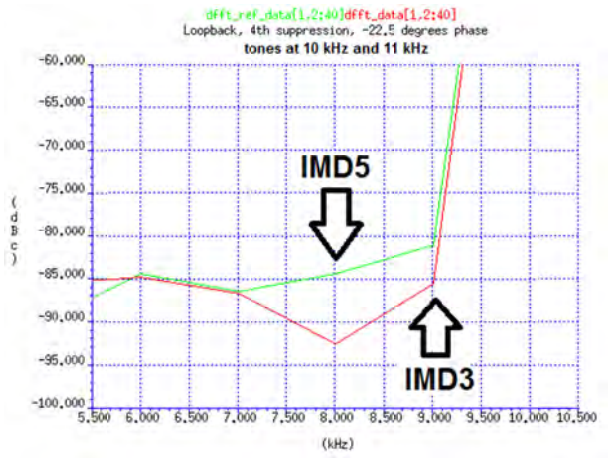



**Figure 133 - IMD4 suppression reducing 3rd and 5th order tones**

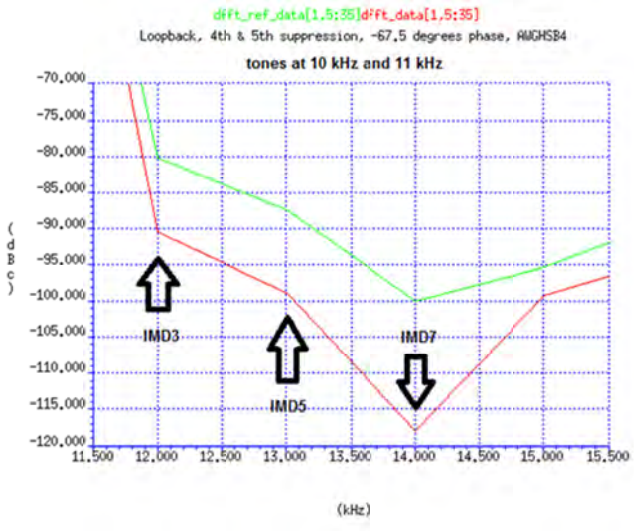



**Figure 134 - IMD3, 5, 7, reduction using 4th & 5th order 4-signal interleaved suppression technique**

## 6.3.2 Starting Phase analysis on Suppression

If equations (60-63) are modified slightly such that a starting phase can be added to the overall waveform results in equations (66– 69). Then we have the following:

$$\boldsymbol{X_5(n) = A\,sin\left(2\pi f_1 nT_s - \frac{9}{40}\pi + \theta\right) + B\,sin\left(2\pi f_2 nT_s - \frac{9}{40}\pi + \theta\right)} \quad \textbf{(66)}$$

$$X_6(n) = A\,sin\left(2\pi f_1 nT_s - \frac{1}{40}\pi + \theta\right) + B\,sin\left(2\pi f_2 nT_s - \frac{1}{40}\pi + \theta\right). \quad \textbf{(67)}$$

$$X_7(n) = A\,sin\left(2\pi f_1 nT_s + \frac{1}{40}\pi + \theta\right) + B\,sin\left(2\pi f_2 nT_s + \frac{1}{40}\pi + \theta\right). \quad \textbf{(68)}$$

$$X_8(n) = A\,sin\left(2\pi f_1 nT_s + \frac{9}{40}\pi + \theta\right) + B\,sin\left(2\pi f_2 nT_s + \frac{9}{40}\pi + \theta\right). \quad \textbf{(69)}$$

By observing the polar plot in Figure 135, Figure 136 and Figure 137, the effect of adding the starting phase in equations (66 - 69) can be seen on the IMD3,5,7 on an IMD4 suppressed signal. By changing the starting phase of the IMD4 suppressed signal results in the measured suppression oscillating with a repetition every 11.25 degrees. It is speculated that the starting phase somehow influences the phase offset and compensation circuit of the interpolating DAC. Reference [50] draws some similar

comparisons in that changing the phase of the two-tone RF signal improves the IMD performance by use of a phase shifter that is the equivalent of the phase compensation circuit within the interpolating DAC (Figure 123). Investigations into the crest factor of the signal showed that there was no influence and that the crest factor of the waveform could be ruled out as a possible contributing factor.

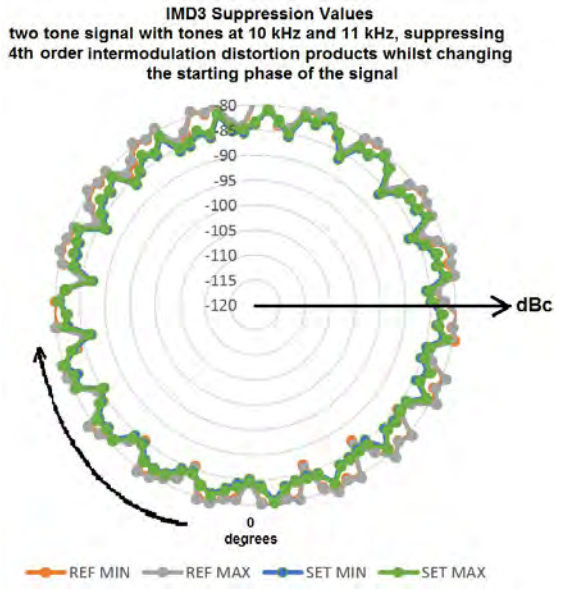



**Figure 135 - The effect of adding a starting phase to an 4th order suppression on the IMD3 tones**

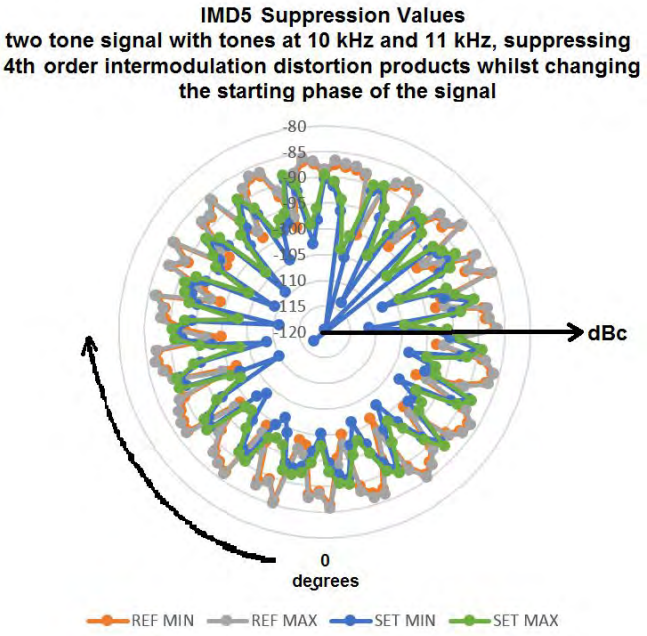



**Figure 136 - Effect of adding a starting phase to a 4th order suppression on the IMD5 tones**

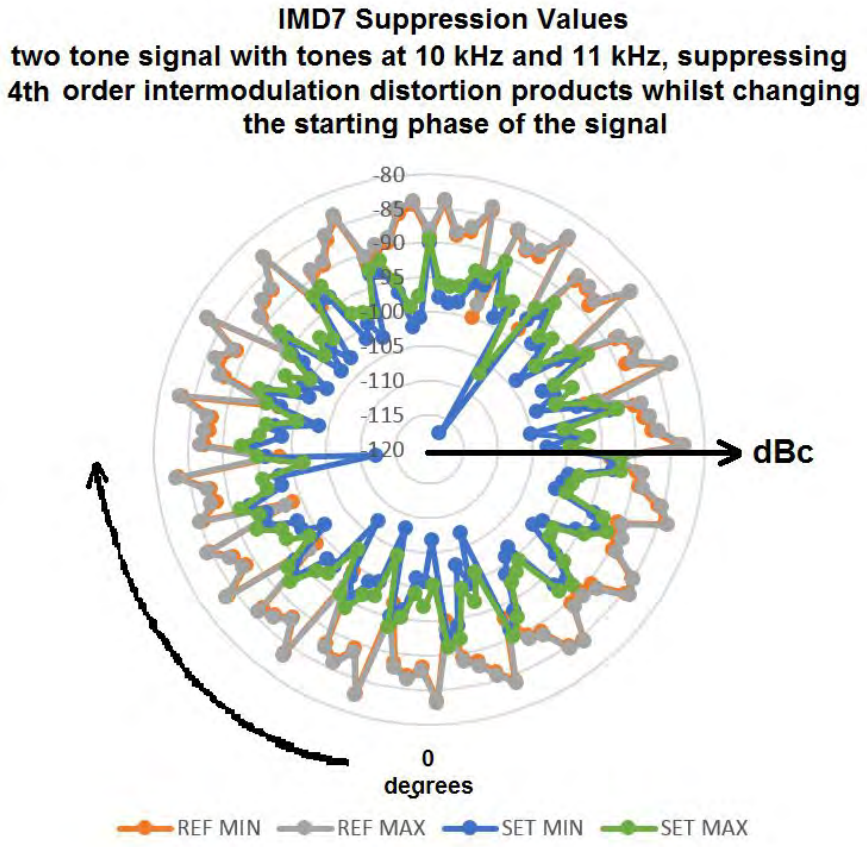



**Figure 137 - Effect of adding a starting phase to a 4th order suppression on the IMD7 tones**

## 6.3.3 Reproducibility Between AWGs

Figure 138 shows the effect of suppressing the 4th and 5th order intermodulation tones with a starting phase of the waveform of 37.5 degrees (refer to equations (66-69)). Obviously, it is inconceivable that the suppression would result in the same effect on the intermodulation tones for every AWG. However, what Figure 138 shows is that the intermodulation distortion is significantly reduced on both AWGs. Using a pair of AWGs on multiple testers and using the same approach as above produced a guaranteed improvement by 6 dB of IMD tones from IMD3 to IMD7. Hence, this allows the possibility to test higher performance ADCs or analogue/mixed-signal devices using this technique with one-shot suppression of all tones. In addition, if the test time is not so critical, the above-described techniques can be used to suppress individual tones to obtain the best possible performance for each individual intermodulation distortion component. Using a three-pass approach to suppress the individual tones resulted in saving of approximately 10 dB across the band.

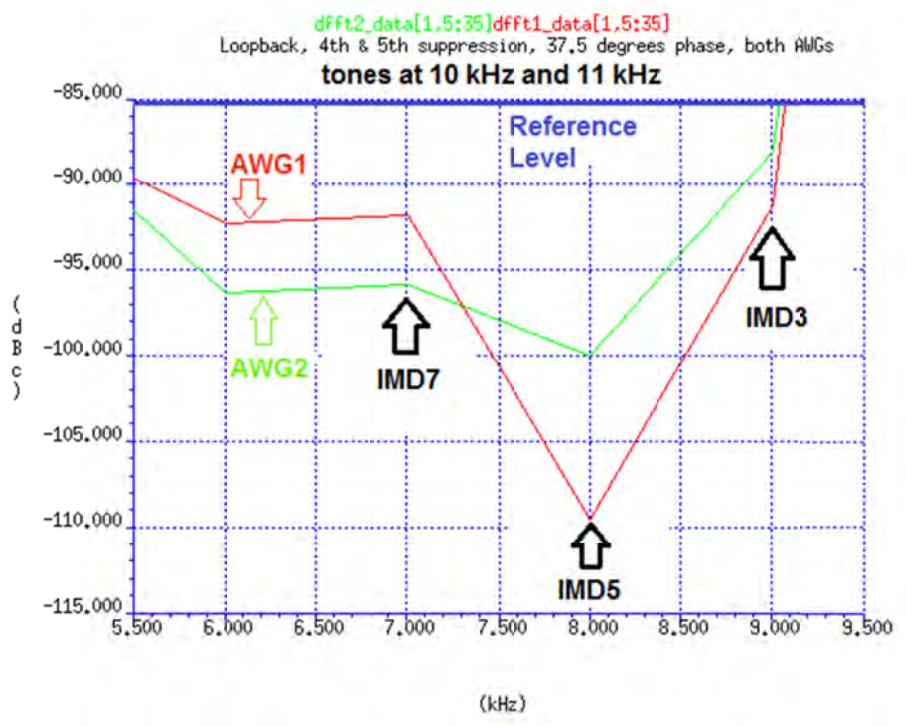



**Figure 138 - Comparison of 4-signal interleaving suppression technique on two AWG's within the same test system**

## 6.4 Frequency/Phase Characterisation of new procedure

To understand the robustness of this technique a sweep of the phase and frequency is required to characterise the proposed routine.

Figure 139, Figure 140, Figure 141 shows the effect of suppressing the $4^{th}$ and $5^{th}$ order tones that had on the IMD3, IMD5 and IMD7. What this shows is an extension of the behaviour observed whilst experimenting with just the starting phase of suppressed waveform. It can be seen that there is a repetition effect occurring and that some starting phases cause an extreme suppression of the intermodulation distortion components used in combination with the distortion shaping phase switching algorithm. However, this ripple effect can be observed across frequency as well as phase that leads to some frequency combinations that are worse than others. From what can be observed from Figure 139, Figure 140, Figure 141, it seems that if the lower tone of the two-tone signal is an odd number causes a worse IMD performance than if the first tone is an even number.

In Figure 139, Figure 140, Figure 141, the y-axis shows the relative calculation between the reference IMD value, measured with the standard sinewave and the IMD measured with the phase switched signal. The peaks and troughs of the chart shows

where there is an improvement compared to where the measurement becomes worse. The negative values show and improvement in the IMD compared to the reference whereas the positive values show degradation in the IMD compared to the reference.

The oscillatory nature of the two-tone intermodulation distortion can also be observed without any phase switching (Figure 142), and thus is independent of the phase switching algorithm, however, when the phase switching is used, the peaks and troughs of this characteristic are extenuated and driven lower.

Referring to Figure 123, it can be seen a typical architecture of an Interpolating DAC, although from an AWG perspective, there is no direct access to interrogate what is happening within the AWG itself, the only explanation, for the observed behaviour, can be attributed to the phase compensation circuit, as described in 6.3.1.

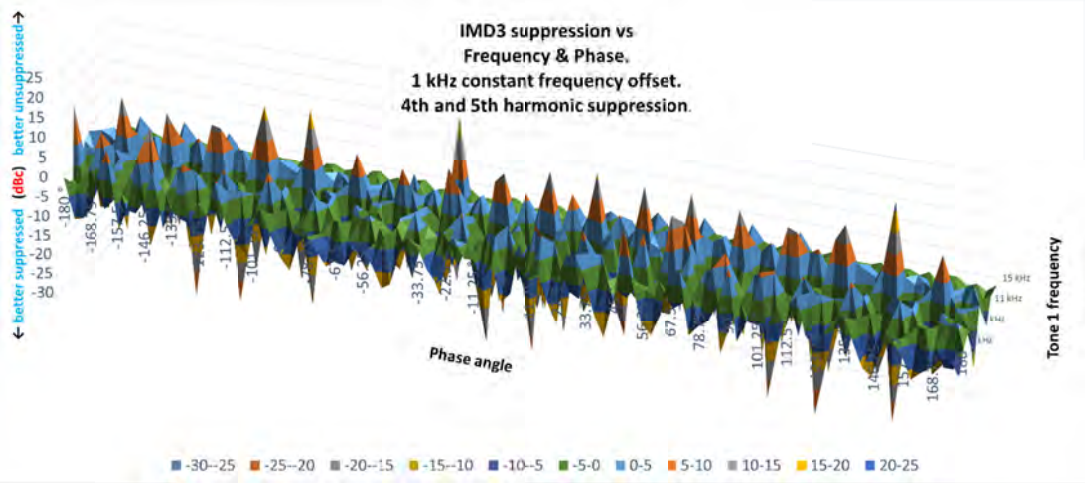



**Figure 139 – IMD3 of suppression vs phase and frequency**

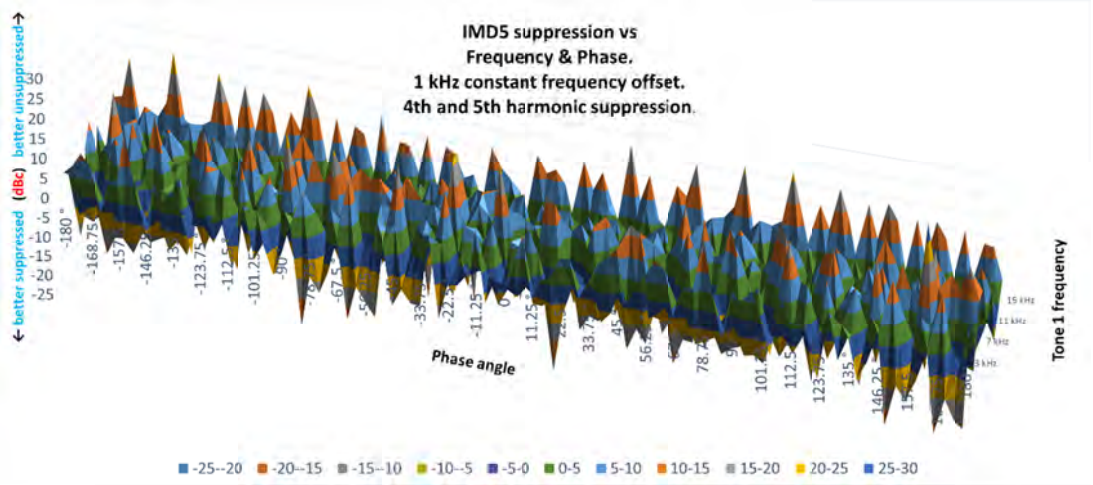



**Figure 140 – IMD5 suppression vs phase and frequency**

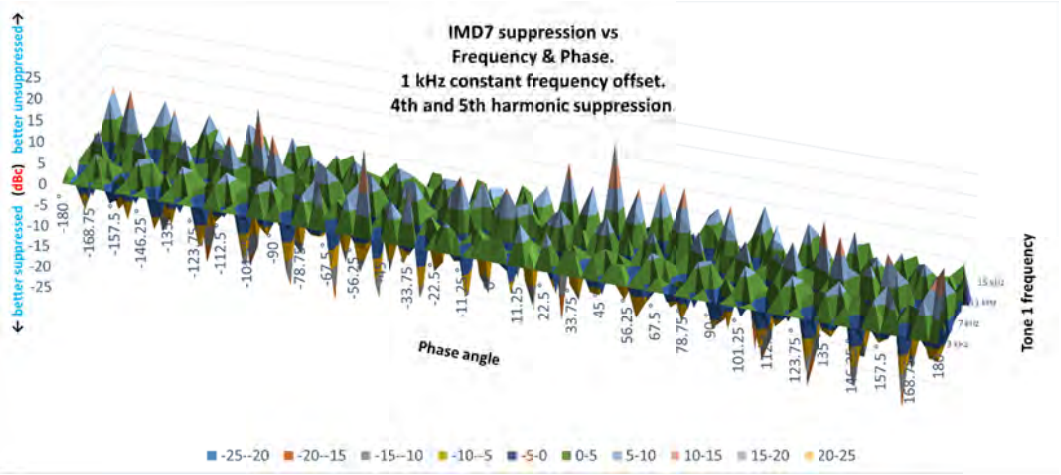



**Figure 141 – IMD7 suppression vs phase and frequency**

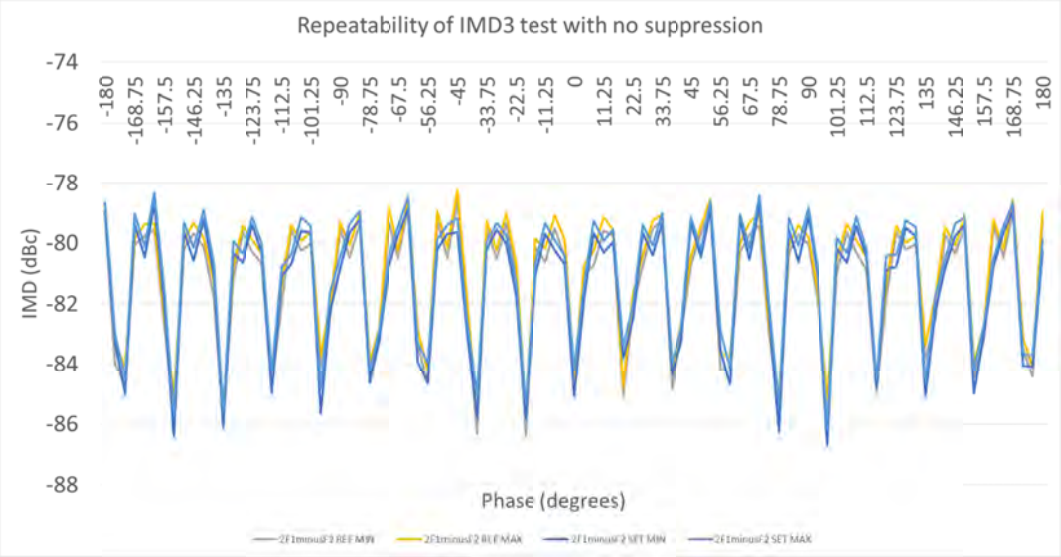



**Figure 142 – Repeatability measurement of IMD3 using standard sinewave with added starting phase.**

It can be seen by observing Figure 143 that by suppressing IMD3, and sweeping the starting phase of the waveform, the degradation in the IMD3 performance is constant and not specific to one phase. However, the oscillatory nature of the signal seen without any suppression is still present whilst using a suppressed waveform and hence the phase relationship is not related to the suppression algorithm but general to the interpolating DAC as shown in Figure 123. Figure 144, Figure 145 and Figure 146 show a constant improvement in IMD3, IMD5 and IMD7 using a phase switched 4-signal interleaved IMD4 and IMD5 suppressed waveform as shown in Figure 120. Figure 144, Figure 145 and Figure 146, all exhibits the oscillatory nature seen in the previous figures that show

it is independent of the phase switching methodology and more related to the interpolating DAC architecture shown in Figure 123.

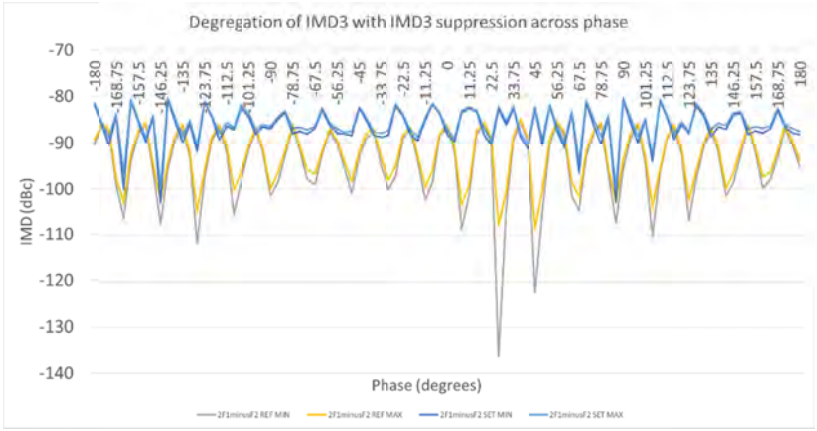



**Figure 143 –Degradation of IMD3 performance across phase of standard sinewave vs IMD3 suppressed version**

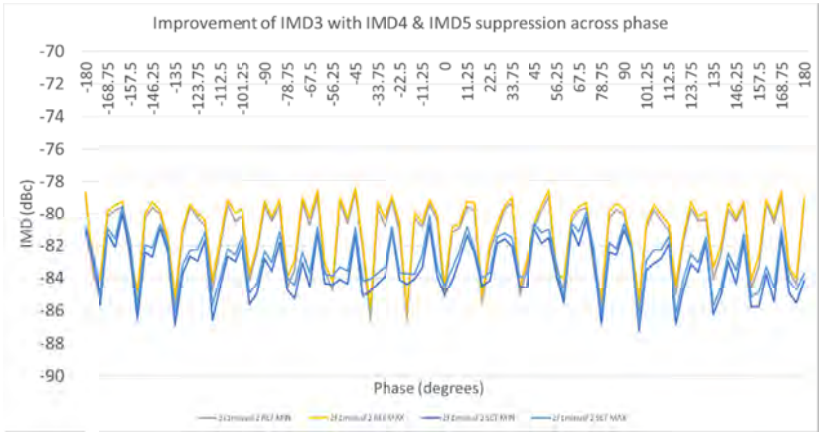



**Figure 144 –Improvement of IMD3 performance across phase of standard sinewave vs waveform with IMD4 and IMD5 suppressed**

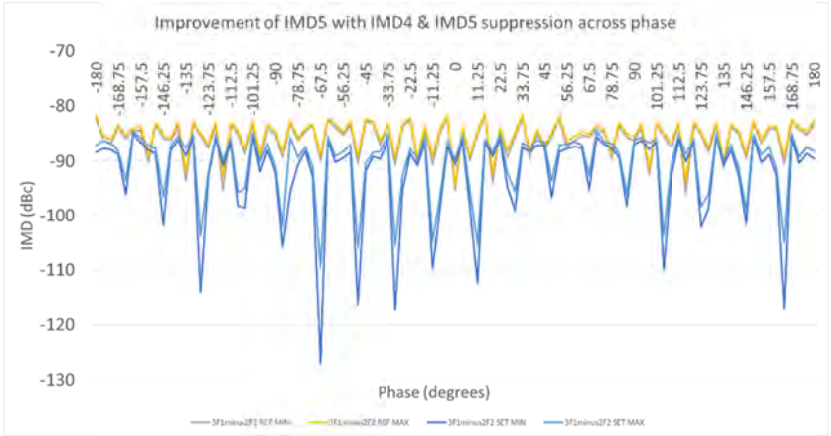



**Figure 145 –Improvement of IMD5 performance across phase of standard sinewave vs waveform with IMD4 and IMD5 suppressed**

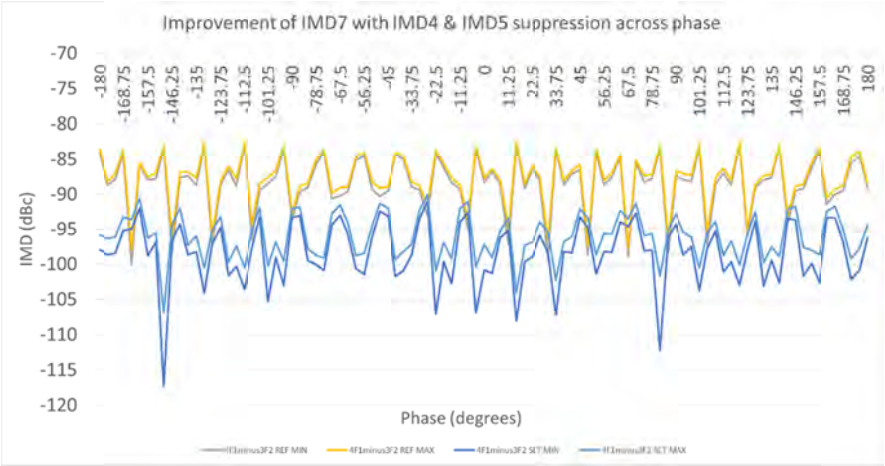



**Figure 146 –Improvement of IMD7 performance across phase of standard sinewave vs waveform with IMD4 and IMD5 suppressed**

## 6.5 16-Bit ADC Example

Using a 16-bit ADC as a benchmark for this method and using the standard technique for measuring IMD3 using a standard two-tone signal, an average value of -80 dBc was measured with the test setup shown in Figure 111. Figure 147 shows the IMD components, with tone frequencies of 10 kHz and 11 kHz and a sampling frequency of 259.2 kHz; however, the measured value of the IMD3 components was -80 dBc but the expected value was much lower. Next, the proposed phase switching technique was tried; it has to be remembered that the generated frequency components need to be filtered off as described in [33-35], so just applying the technique without an adequate filter would make the results better than they should be. See test setup in Figure 111.

As the minimum filter of the AWG used was much higher than the sampling frequency, an oversampling approach was adopted so that the generated sampling frequency components can be filtered off. By using a sampling frequency of 155 MHz, i.e. oversampling the signal 600 times, a 50 MHz filter could be used to filter off the sampling frequency components, (Figure 111). By using a 4-signal, interleaved approach with the $4^{th}$ and $5^{th}$ intermodulation products suppressed would require a 38.5 MHz LPF filter, Equation (64), whilst using a sampling frequency of 155 MHz. The AWG used in this example only has a 50 MHz LPF available and hence would be too high; however, it was shown in Figure 118 and Figure 133 that the $3^{rd}$ and $5^{th}$ intermodulation distortion products could be reduced by simply suppressing the $4^{th}$ order intermodulation products with a 2-signal interleaved approach. For this 16-bit example, reducing the $3^{rd}$ and $5^{th}$ order components should be ample as the noise floor

of the 16-bit ADC is too high to see anything above the 5th order components. Using Equation (64) with a 2-signal interleaved approach with a sampling frequency of 155 MHz yields that a filter of less than 77.5 MHz is needed and hence the internal 50 MHz of the AWG should be sufficient.

By suppressing the IMD4 tones, as described in 6.2.1 with this sampling scheme, an improvement in the IMD3 components by 3 dB can be observed as shown in Figure 148. In addition, the IMD5 tones of the ADC were driven down into the noise, ~10dB. This results in the true IMD performance of the ADC being observed as shown in Figure 148. The results of the 3rd order IMD products are in the range of the expected value that was measured in the laboratory with benchtop equipment. This proves that a more accurate IMD3 and IMD5 for an ADC can be measured. Unfortunately, the IMD7 and IMD9 tones of the ADC were much too low in comparison to the noise floor of the 16-bit ADC to be able to measure any influence of any phase switching algorithm.

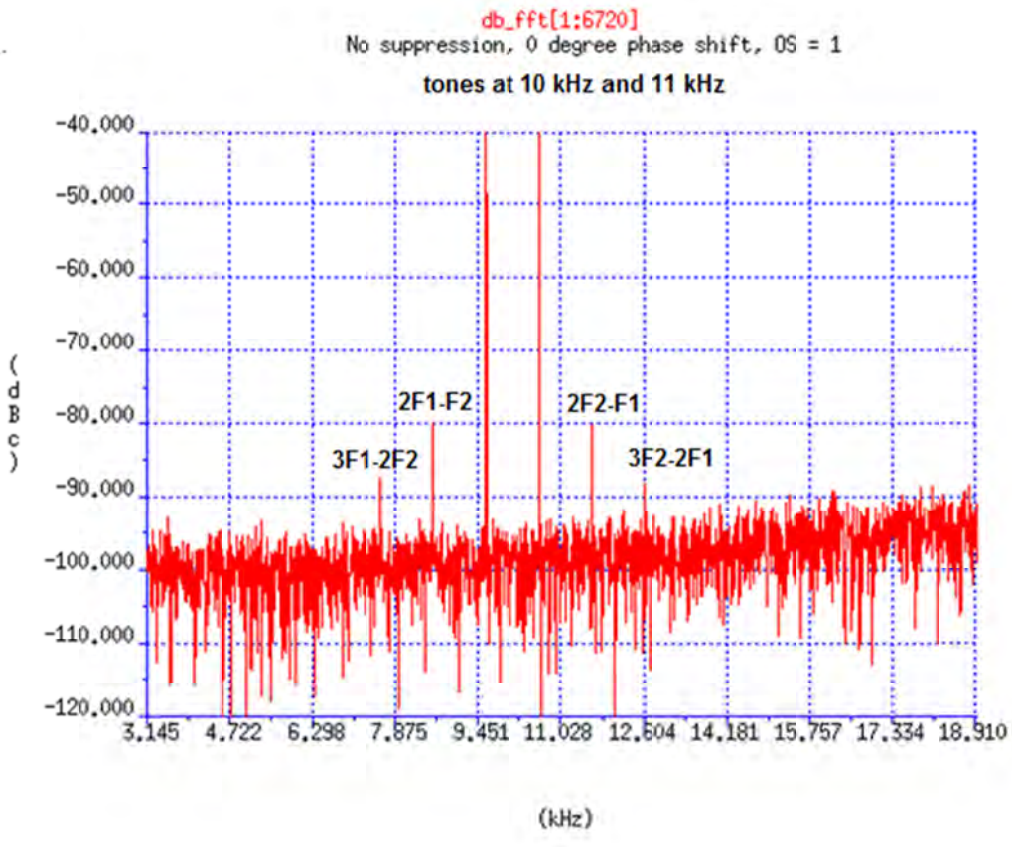



**Figure 147 – Intermodulation distortion products seen by ADC caused by poor performance of AWG. This is not the real performance of the ADC.**

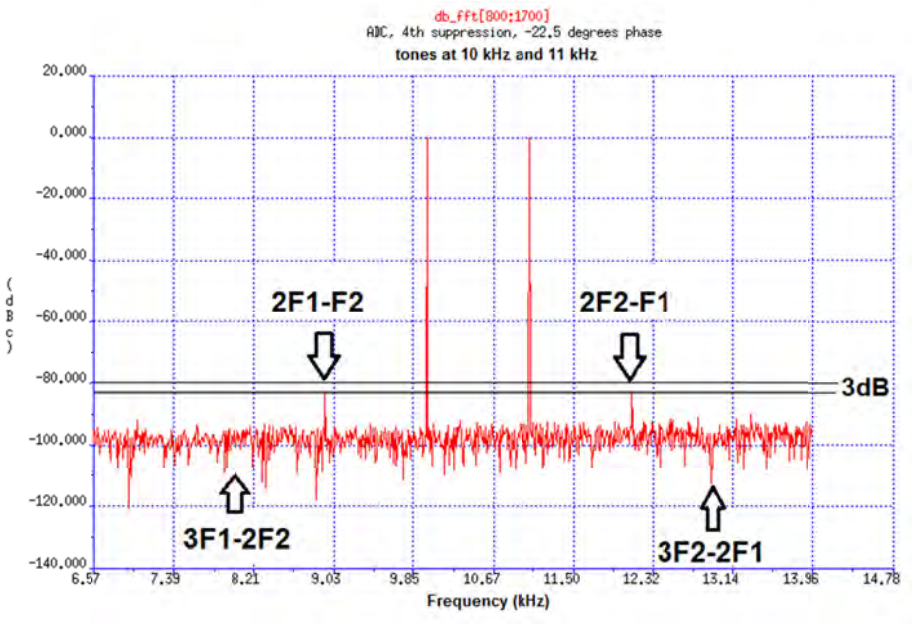



**Figure 148 – Real Intermodulation distortion performance of ADC due to improving the performance of the intermodulation distortion of the AWG.**

## 6.6 Conclusion

This chapter has presented that the phase switching technique for an existing AWG with an interpolating DAC can improve the IMD testing performance of analogue/mixed-signal DUTs without additional hardware and cost. We would like to say in general that by understanding the architecture of a test instrument, it is possible to extend its performance by applying some commonly used techniques. In this case, it was possible to extend the intermodulation distortion specification of the installed base of the AWGs in the LTX-Credence testers at ams AG by using simple phase switching techniques, but applying them in a unique way. This allows a test solution to be deployed that utilizes the current test equipment but with an extended range that would not have been possible without purchasing new test equipment.

By installing these algorithms into the current ageing tester installed base around the globe, would possibly allow extending the range and life of such testers, and hence recycling the tester base in some way. Testers such at the Teradyne Catalyst were purchased in huge quantities. The described solution could help extend the lifetime of such testers; this could reduce the need to procure newer test systems and hence reduce the investment budget of an Integrated Device Manufacturer (IDM) or a contract provider of semiconductor test operations.

# 7 AUTOMOTIVE GRADE SCREENING PROCEDURE FOR ANTI-FUSE GATE-OXIDE NON-VOLATILE MEMORIES

## 7.1 Introduction

An efficient and effective way is needed to screen devices at one temperature to enable cost effective testing so that multiple test stages are not required to capture devices that would fail at the specified temperature extremes. However, this should not be done at the cost of significantly increased failure rate. Many techniques have been developed over the years that show different ways of capturing defects that are seen either at hot or cold testing. These defects have to be detected. If the safe launch stage of automotive device testing fails, this leads to an expensive and time-consuming backend flow, which also leads to very long cycle times, as all devices will need to have multiple test stages over the temperature extremes. What this means is that where the standard method of testing at one test stage with temperature guard bands (see Figure 149), would not be possible and the device would have to be 100% tested at all the device temperature specifications. This can happen when the device has a bimodal effect where it flips from good too bad at temperature extremes. This results in a nonlinear degradation such as would be seen with a standard Gaussian distribution which prevents the use of guard bands and requires screening at the temperature extremes.

Figure 149 shows how a devices performance may change with temperature. The device parameters of an automotive device are specified at the extremes of the temperature range, and all devices must adhere to the values stated in the datasheet otherwise they will be returned from the customer otherwise known as Returned Material Advice (RMA). As discussed in Chapter 2.8.6, for an automotive device, for the first 100,000 devices, all production devices have to be tested at the temperature extremes therefore devices failing at these temperatures will be automatically captured. However, for high volume production, this is not possible and the test flow needs to be reduced to the typical specified temperature only that has the possibility of screening the corner temperature failures also. This can be done by a means of guard banding the test result by the shift seen at temperature extremes as shown in Figure 149. For the means of demonstration, Figure 149 shows a device that has marginal performance at temperature extremes and thus the Gaussian distribution is cut. To be able to guarantee the devices that pass the typical temperature only test, the guard band as shown has not only to cut away parts the parts that would fail the temperature extremes but also a few close to the edge of this limit so no marginal parts can get through the testing procedure. This is why the limit is cutting the Gaussian distribution slightly more extreme at the typical than would be at the maximum and minimum temperatures.

An issue was found with a product running in production that contained a one-time programmable (OTP) memory block that used oxide breakdown technology. It was found that the device population experienced a certain defect parts per million (dppm) level when reading the memory contents at low temperature (Figure 150). As this was an automotive device shipping in the millions of units per month, the experienced defects in the field were unacceptable and a cost effective solution was required.

The complexity of the test of an OTP in production entirely depends on the amount of test access that is installed in the OTP itself. Hence the test coverage is linearly proportional to the design for test (DFT) that the designer implements. However, more DFT and test access come at the cost of bigger die, which thus results in a higher cost of silicon. Therefore, many intellectual property providers provide the smallest die possible to be attractive to greatest market share.

In this example, the OTP only had four test modes where the leakage current of the blown oxide could be checked against a reference level. The design weakness here is that the oxide may be weakly blown, as the leakage current decreases over temperature.

The fuse can then go from being seen as a programmed bit to an un-programmed bit, hence, there is no effective way to guard band this effect, as would be the case with a Gaussian distribution as shown in Figure 149. The result is that the weakly blown fuse now looks more like a resistive bridge between the fuse elements rather than being open or short. This is a huge issue because if a poorly blown device reaches the field then the contents of the OTP could become corrupted in the event of extreme cold, hence causing an issue for an automotive suppliers and users.

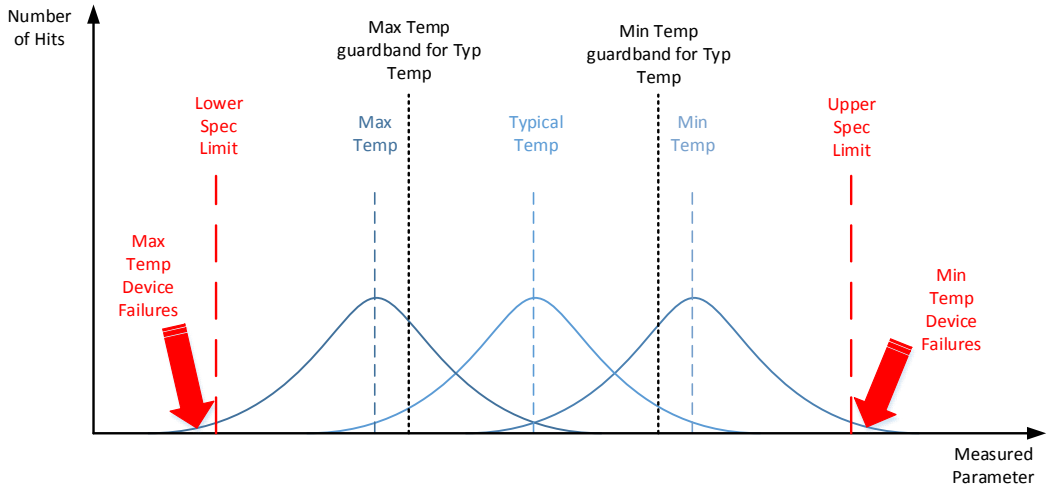



**Figure 149 –Graphical representation of how device performance metrics shift with temperature and how guardbands are applied at the standard test temperature to guard against these failures entering the field**

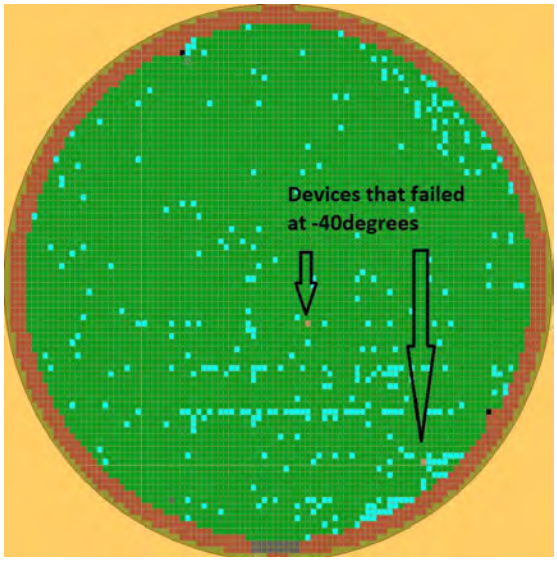



**Figure 150 -Wafermap of initial wafer probe OTP screen at -40 deg @ 2.2 V**

This chapter shows how by applying a guard band to VDD and hence making a very low voltage test (VLV) of the OTP contents post burn places the OTP block into a mode

such that a badly blown fuse is now observable at the temperature of the standard test stage. Therefore enabling a single test step without the need to screen at the minimum test temperature. It is then shown how the resistive bridges were the result of a layout weakness of the cell. This weakness in layout was not seen when the cell was first automotive qualified. However when the cell was transferred to a second foundry location, the alignment accuracy of the installed equipment was not as controlled as at original foundry, hence causing the design weakness to become observable.

It is shown how an increase in the die to active area improves the manufacturability of the cell, and how, due to this weakness, the actual VLV screen is more accurate at capturing failures than the cold temperature screen.

This test approach could be easily adopted for all OTP's that have an anti-fuse architecture so that an automotive grade screening is achievable; this is where we claim this is a state of the art test approach for anti-fuse OTP's.

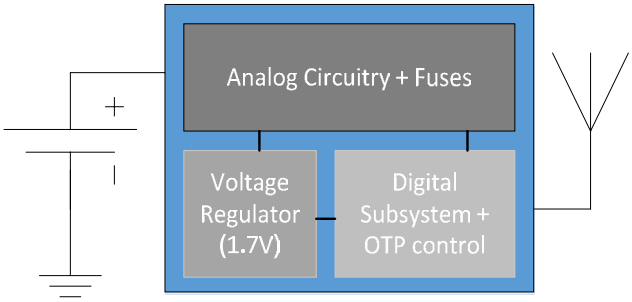



**Figure 151 –Block Diagram of the Device**

Several papers have previously been published on these topics. The references [51-54] show a relationship between VLV testing and cold temperature failures, and the reference [55] shows an anti-fuse breakdown OTP similar to the one used in this example, while the reference [56] discusses the actual reliability of an anti-fuse gate-oxide breakdown technology. The reference [57, 59] goes further and shows how oxide failures can be discovered with VLV testing while the reference [58] discusses the improvement of the design of a non-volatile memory (NVM) with respect to the die to active area region. Reference [59] is a published paper by the author on this chapter.

This chapter is organized in the following way. Chapter 7.2 demonstrates the correlation of the low voltage screen to the cold temperature test step. Chapter 7.3 goes into details of the root cause of the problem. Chapter 7.4 discusses how the results from the cold test stage were not very accurate and how additional neighbouring die to failing

die locations required manual removal in order to have higher confidence in the shipped die prior to the VLV test introduction. This is further developed to show how the failure signature has a certain gradient over the wafer and due to this how the manual inking strategy could be relaxed to enable better yield but still high confidence in the shipped die. Chapter 7.5 describes how the change to the layout influenced the yield and how the introduced VLV test resulted in a very high yield with an extremely low dppm level. Conclusions are then drawn.

## 7.2 Background

### 7.2.1 Chip Architecture

The device is battery operated RF communications device for the automotive market. The device is designed such that as the battery discharges over lifetime that the device can continue to operate. Based on the chip design and the low leakage features, the device can be used for a guaranteed period of 10 years without the need to change the battery. After 10 years, the user would simply throw away the unit.

The device is designed on a 0.18 um low power RF process that uses the 1.8 V transistors. The digital is powered by an on chip voltage regulator that regulates the incoming battery voltage. Although the device is specified to work correctly at 2 V in the application, the device can operate, for its basic functions, as long as the voltage regulator is operating correctly. From experimentation, this can be as low as 1.75 V. See Figure 151. By driving the supply below the specified voltage, gives the possibility of stressing the OTP fuses such that badly blown oxides can be detected without effecting the normal operation of the digital control circuitry that drives it.

### 7.2.2 Development of Very Low Voltage OTP screen

During the safe launch of this automotive device, we experienced a large lot-to-lot variation in the failure rate of the OTP read-back seen at cold temperature testing at wafer-sort. This is indicated with the dies marked in grey in Figure 152. Because failures were being seen at the verification stage of the production flow (see appendix 5), questions were being raised with regard to the root cause, and why the failures were not being caught at the production stage where the OTP fuses were being blown. This led to some confidence issues in the final solution. However, as the product was in safe launch and there was a cold temperature test step the risk of shipping bad die was

extremely low. Even so, a new approach was required to screen these failures at the first probe stage, which was run at the standard test temperature of 35 degrees for wafer-sort.

As the test modes designed into this product were not sufficient to capture the failures that were being experienced at the cold test step, another stimulus was needed to force these defects to be observable at the standard test temperature. The initial thought was to make a guard-band on the power supply voltage to try to force the observability of these digital failures. However after some research, it was discovered that very low voltage testing was used to find resistive bridge faults in low geometry technologies [51-54] and that the same could be applicable to this problem.

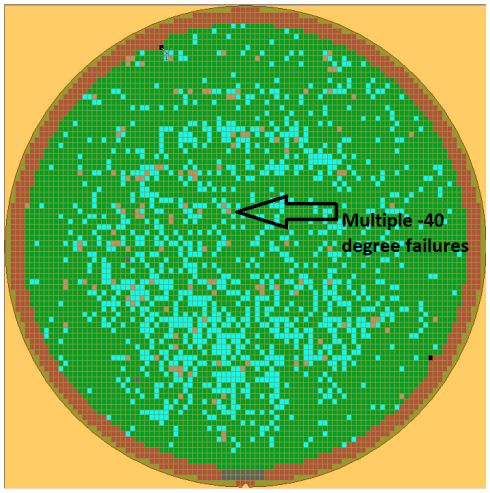



**Figure 152 - Wafer-map of multiple -40 deg failures @ 2.2 V**

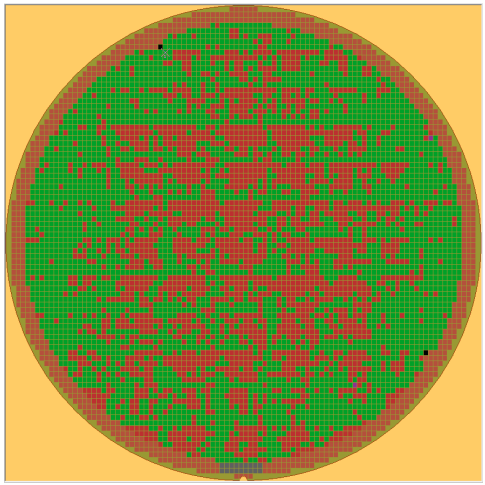


**Figure 153 - Wafermap showing OTP failures at 35 deg @ 2 V – step field**

The first test step setup for this device was as follows: with VDD @ 2.2 V, a pattern runs to select the hardest reference level to check the memory contents against then the data is read and checked against what was programmed. At the standard test

temperature, there was no issue seen. However, when the verification step was made at -40 degrees using the easiest reference setting of the OTP block, the read back of the memory did not match what was programmed at the standard test temperature of 35 degrees. The power supply level for the verification stage was the same as at the standard test temperature. Therefore, the only issue could be within the block itself, hence eliminating any real questions regarding the actual quality of the test program. Experimenting with a very low supply voltage of 2 V for this technology gave rise to a good correlation of failures at cold. However, another issue was identified at the same time with respect to the step field (Figure 153).

## 7.3 Route Cause

The OTP cell in question works with oxide breakdown technology (anti-fuse), i.e., when a fuse is programmed (Figure 154b) [55], the oxide is blown through causing a short or a higher leakage of at least 100x compared to the non-blown fuses (Figure 154a). In the case of badly blown fuses or poor oxide breakdown, the short looks more like a resistive bridge resulting in only approximately 10x higher leakage compared to a non-blown fuse (Figure 154c). Reference [56] discusses and demonstrates a large variation of the current seen by the oxide breakdown at low temperatures, which also leads the reader to conclude that the resistance of the breakdown has a large variation.

For Figure 157 - Figure 165, the following colour codes for dies shown on these wafer are as follows,

| | | |
|---|---|---|
| Yellow | - | Manually inked die by an engineer |
| Red | - | Device failure, standard failure, **not VLV OTP failure** |
| Purple | - | Continuity failure, missing diode from ground or to VDD. |
| Light Blue | - | Masked device failure from 35 degrees test stage, so as to avoid retesting the dies at -40 degrees |
| Grey | - | Optical Inspection failures after 35 degrees test stage, |
| Black | - | Reference die, used for alignment orientation of wafer during prober setup |
| Orange | - | Special bin to identify die that would fail OTP test at -40 degrees either by testing at -40 degrees or with VLV test |

At room temperature, with typical supply voltage, the difference is still detectable and a programmed fuse will be seen as a blown fuse. However, once the temperature is lowered to -40 degrees the leakage current of the blown fuse decreases. Hence, if an oxide is not totally broken down, [57], it can be seen as a non-blown fuse, resulting in bits being seen to flip from programmed to not programmed (Figure 155). By reducing the voltage to a lower value, the leakage current is decreased and the possibility of detecting a badly blown fuse is significantly increased. Therefore the question is by how much would the supply voltage need to be reduced at 35 degrees to ensure that all quality issues, which resulted from badly blown fuses, would be detected without affecting the normal operating conditions of the device? Fortunately or not, depending on the point of view, a layout rule versus process capability problem resulted in the above-described issue being much worse.

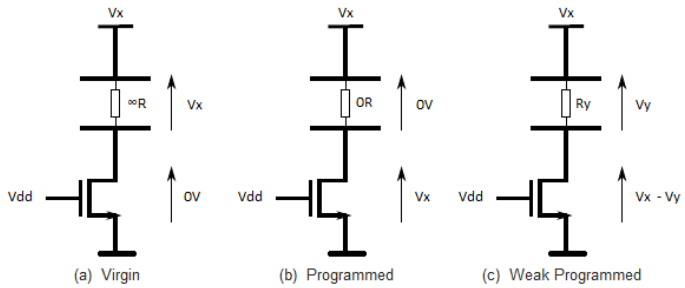



**Figure 154 - Schematics of different conditions of gate-oxide breakdown anti-fuse**

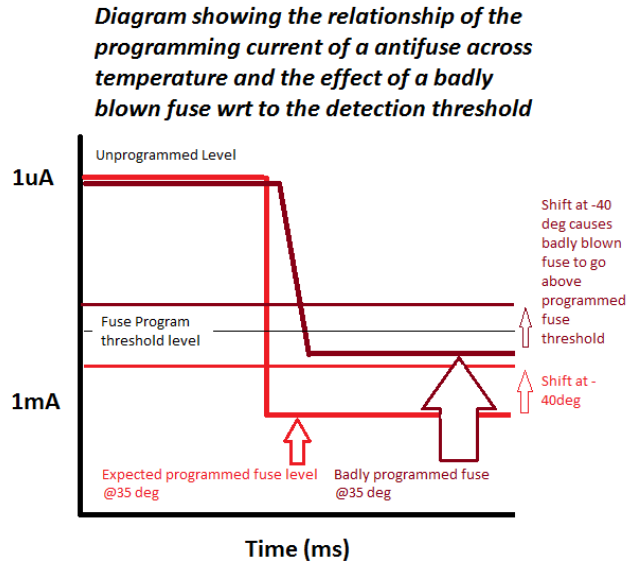



**Figure 155 - Diagram showing the programmed current shifts wrt temperature**

This resulted in some wafers exhibiting a large step field signature that caused many wafers to have yields of as low as 5 %, which therefore required an even greater confidence in the quality of the shipped dice of high yielding wafers. The step field issue was attributed to the accuracy alignment of the lithography step, where in the original foundry location the stepper had a better accuracy than in the second foundry. Hence, the issue was not experienced when the original OTP block was qualified and released. It was found that the NWELL did not have a large enough process window to allow for marginal mask alignment issues in the fabrication process and this led to the step field issues seen in production. This was causing issues with the oxide breakdown [58], hence causing resistive bridges, rather than shorts for programmed fuses. By increasing the NWELL to ACTIVE AREA spacing, there is more tolerance for mask misalignment [58], thereby eliminating the issue.

The solution to the problem was simple and just required a layout change (Figure 156). However, as there were many lots in the production line a way to screen the failures at 35 degrees was needed so that the result at -40 degrees could be monitored. This was used to increase the level of confidence in the devices that were being shipped to the field from wafers exhibiting the step field.


**Figure 156 - Old NWELL to new NWELL process window**

## 7.4 Step Field and Nearest Neighbourhood Die

In order to enable shipping of production wafers in the line with sub-optimal processing, the nearest neighbour dice were electronically marked for removal in the production flow. This seriously reduced the number of good dice on a bad wafer lot (Figure 157). However, the suspect defect dice were picked out for further correlation analysis to be done later. After testing the additionally marked dies, it was found that there was a high probability that the nearest neighbour dice would have a quality risk, and so for this production step to be removed, it would need to be proved that these nearest neighbours would be caught by other means. Having launched all of the counter

measures into production to ensure a secure and defect free supply line, an enhanced correlation exercise was undertaken.

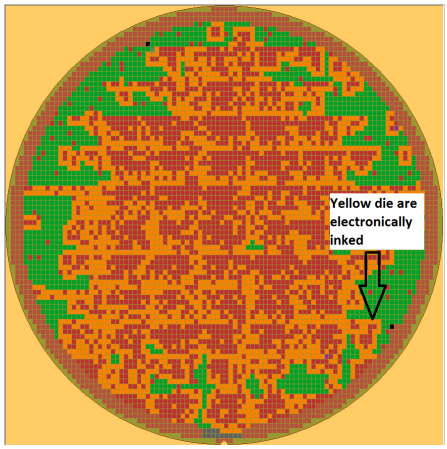



**Figure 157 - Electronic inking of suspect neighbour die of step field wafer**

The goal of this exercise was to move as many as possible of the failures at the verification step at cold temperature to the first production stage at the standard test temperature, so that if desired later the verification step could be removed. What was required was a wafer that had good yield, but also had a failure at cold. This was needed in order to be able to find a good correlation at 35 degrees and hence demonstrate the impact on the yield of the wafer because of capturing this failing die. Figure 158 shows such a wafer with only one die found at the verification step as an OTP memory failure. Having isolated this wafer from the production lot, a correlation exercise was carried out on the appropriate die. The power supply voltage on this die was ramped from 1.725 V to 2 V. It was thus shown that the failure could be captured with a VLV of 1.8 V (Figure 159). For completeness, Figure 160 shows where the device flips and becomes a good device. However, it can be seen that the yield is also reduced with a decreased voltage.

The major issues here are the wafers with the step field signature and the need to ship wafers with a high confidence. Using a step field issue wafer, one can see a large difference between fails at cold (Figure 161) and with the 1.8 V screen at 35 degrees (Figure 162). What Figure 161 shows is the complete central area, approximately 1500 dice, of a wafer failing the cold temperature screen. By running the VLV test at 35 degrees, Figure 162, we see a correlation to the failure at -40 degrees. However, it is clear that the VLV screen is more aggressive at finding the problem die than the actual cold screen itself, an increase of 500 dice in and around the step field that would

ordinarily be inked. The question now is whether the VLV test can replace the manual inking step. If one compares the electronic inking of the nearest neighbour dice in Figure 157 and the fact that these parts were confirmed to be a quality threat then one can appreciate that the very low supply voltage screen can detect these parts. It can also be seen that the problem area of a step field wafer is centrally located and will exhibit VLV test failures, and as such, manual inking can be relaxed. It can be seen that the effect of the step field goes from a north to south direction, and hence the dice that present a quality risk are located on the south side of the step field. Due to this and the fact that the new VLV screen is able to capture the previously uncaught nearest neighbour dice, a new electronic inking strategy can be put in place to further guarantee quality and increase the yield. A new inking strategy was put in place to select the south side of a step field issue (Figure 163), hence increasing yield without affecting the reliability and quality of the delivered product.

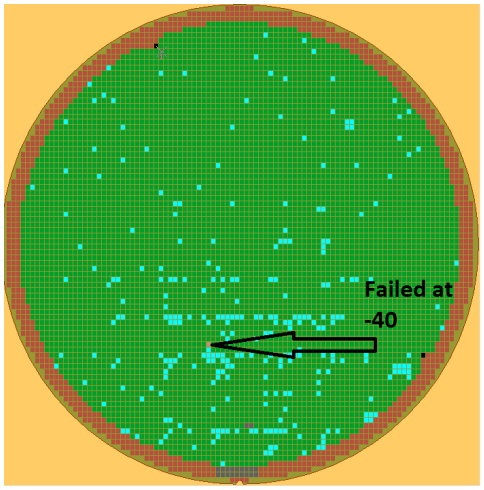



**Figure 158 - -40 failure @ 2.2 V for correlation against voltage at 35 deg**

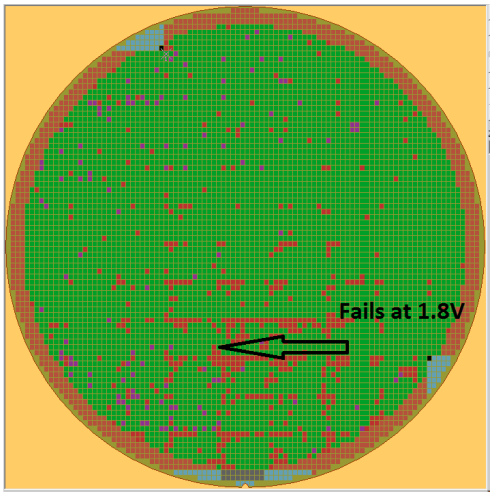



**Figure 159 - -40 failure failing at 35 deg @ 1.8 V**

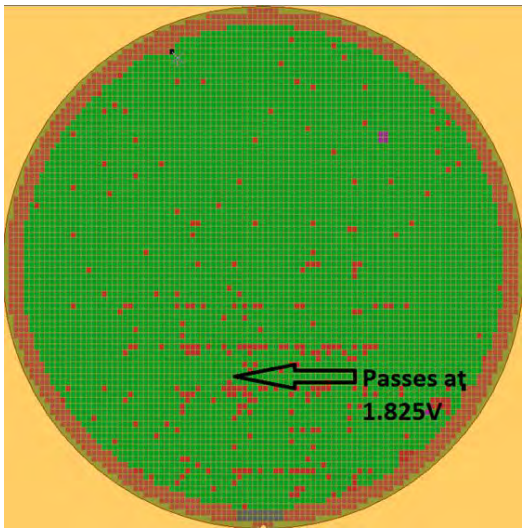



**Figure 160 - -40 deg failure passing at 35 deg @ 1.825 V**

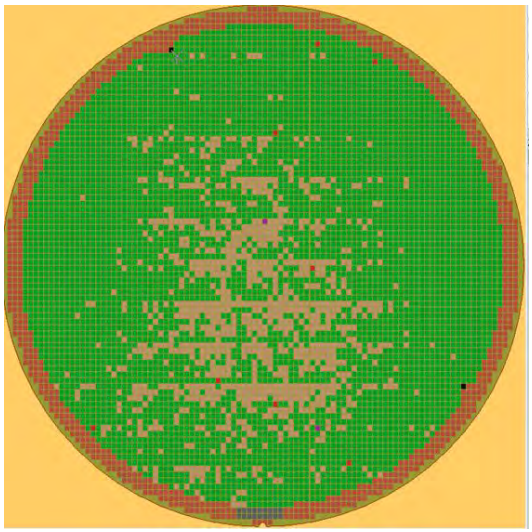


**Figure 161 - Wafermap of a step field wafer fully tested at -40 deg @ 2.2 V**

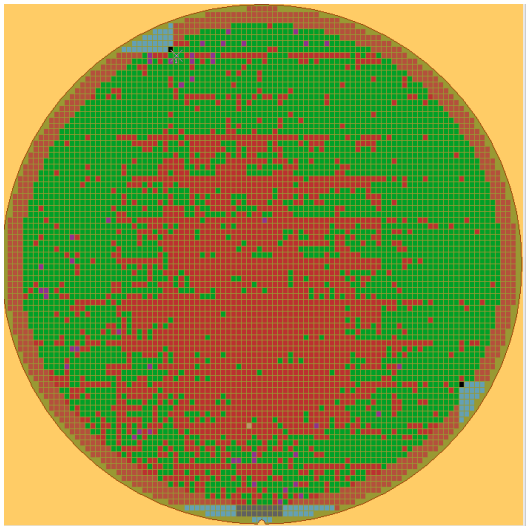


**Figure 162 - Wafermap of a step field wafer tested at 35 deg @ 1.8 V**

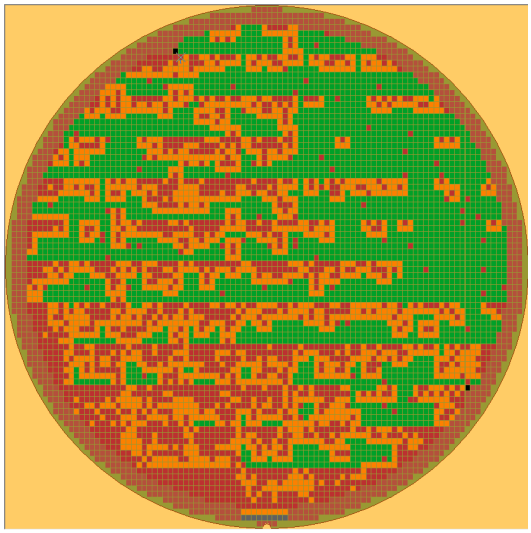


**Figure 163 - New nearest neighbor inking strategy**

## 7.5 Results of the Layout Modification

Once wafers were received with the modified layout, the wafers were subjected to the standard flow that was developed to identify the issues at cold. Figure 164 shows the results from the cold screen of a wafer with the modified process. It can be seen that all evidence of a step field has been eliminated and that the fails are randomly distributed as expected.

Although the results from the cold screen are very good, it would be ideal if the cold screen could be skipped completely, hence making a significant reduction in production costs. Examining Figure 165 shows that all the cold failures are covered by the VLV screen. These dice have been identified by a blue circle for ease of identification.

The other failures that are shown in Figure 165 are associated with standard production tests such as IDD, IDDq, SCAN and other memory tests. Having completed all the standard GRR and release to production checks this test procedure was released into production. To date over 500 lots of material have been tested and shipped to customers, approximately 50 million units, with this solution with a stable yield of approximately 96 %, Figure 166, and no customer returns thus showing that the solution is robust, production worthy and most importantly, technically sound. It can be concluded from this that the cold screen can now be considered for complete removal, as the very low supply voltage screen captures all the cold failures at 35 degrees. However, this can only be done after multiple wafer lots have been checked for low dppm fallout at the cold screen.

Observing Figure 164, a light blue failure can be seen in the bottom right hand corner of the wafer; this is an OTP failure at cold that passed at the 2 V at 35 degrees but failed with the 1.8 V VLV screening. This die can be seen as an VLV failure as it is colour coded orange in 35 degrees screen as can be seen in Figure 165 circled in blue. With the new design, we are really looking for dppm rates of failures and thus this test does not affect yield but increases the quality level. The yield is unaffected as there are approx. 4000 die per wafer therefore if there is one failure on a wafer this equates to 0.025 % of the total dies of a wafer. In reality, only 1 die per wafer lot could be expected which is approx. 0.001 % of the total lot size as there are 25 wafers per lot i.e. 100K dies. Either 1 die per wafer or 1 die per wafer lot, the loss is negligible but with a large quality increase.

It can be concluded from this that the cold screen can now be considered for complete removal, as the very low supply voltage screen captures all the cold failures at 35 degrees. However, this can only be done after multiple wafer lots have been tested with low dppm during the cold screen.

Now a solid and reliable test program is in production and the test solution could be left to run allowing statistics to be collected on the tested wafers for analysis at a later stage. After monitoring production in batches of 30 lots (which is the equivalent of 4.6 million devices), it was found by the use of a stack map (Figure 167) that 15 devices failed at the cold temperature, randomly distributed across the 750 wafers. This means that the failures are not systematic and in all likelihood are false failures. Without the cold screen, the actual quality level would be 15 failures in 4.6 million devices, which would equate to 3 ppm. This is without the manual inking step as the VLV screen removed the need for inking of the improved design.

Due to the fact that cold temperature probers are expensive, and usually as there are only a few available on a production test floor. By removing the cold test stage, the benefits to the company are not only a reduced overall test time but also the delivery time of a product is significantly reduced. As there would be no longer any need to wait for test equipment, that is fully loaded to become free, appendix 6, therefore resulting in shorter lead times by as much as up to 2 days.

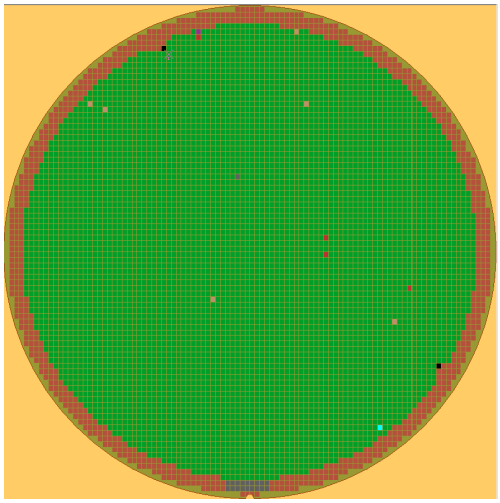


**Figure 164 - -40 deg probe with new nwell designed OTP cell @ 2.2 V**

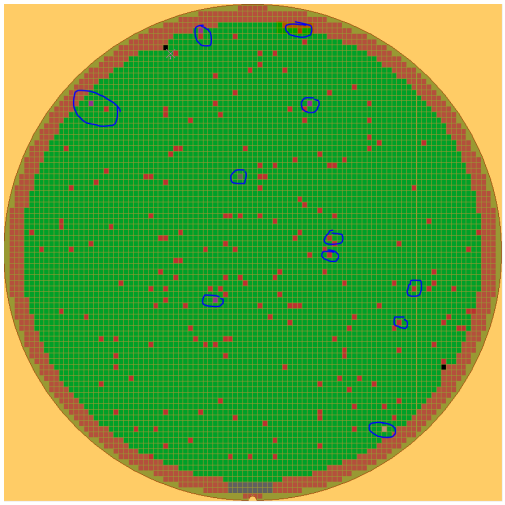


**Figure 165 - 35 deg very low voltage screen @ 1.8 V**

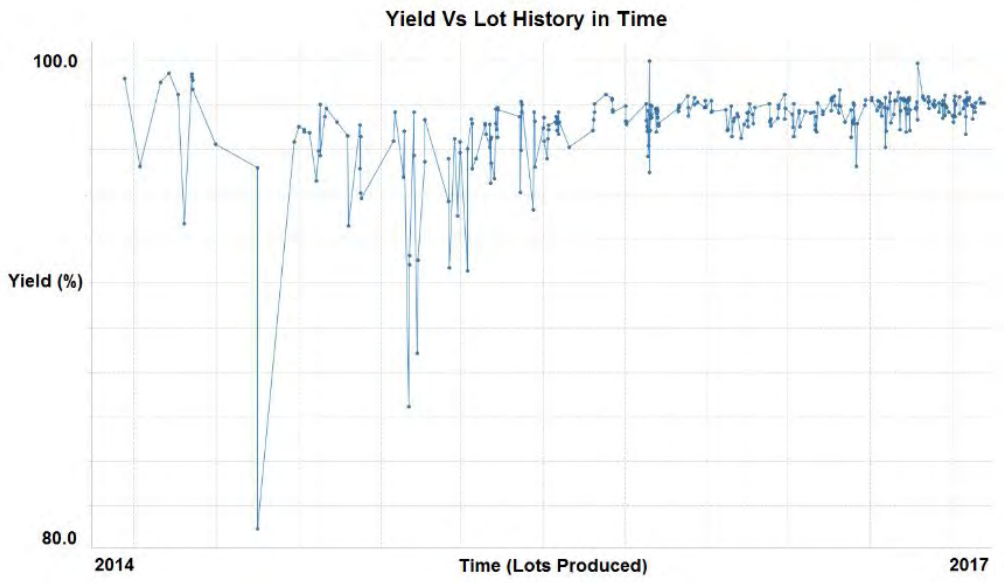



**Figure 166 - Plot over time showing the improvement in the product yield**

## 7.6 Conclusion

This chapter has shown that by using a very low supply voltage screen, it was possible to produce a test solution that was capable of screening OTP memory failures that could only be normally observed at cold temperature at the standard production temperature. The effect was to make a parametric guard band of a digital parameter that was sensitive to cold temperature. This allowed screening of material that had a design issue such that devices could still be sent to the field with 100 % confidence in the delivered die.

Through a thorough analysis of the issues with the design and an understanding of the problem caused by the design issue a reliable and cost effective technique could be developed to remove such problem devices resulting either from process issues or from randomly distributed defects. In combination with an intelligent nearest neighbour die removal strategy, one can have a high degree of confidence in shipping high quality and reliable dice even when suspect wafers have been produced.

By introducing a Very Low Voltage temperature-screening procedure of the OTP, a solution was produced that would find fuses that did not experience the full burning cycle such that they would be effected by the higher leakage seen at lower temperatures. The result being that a burnt fuse was indistinguishable from an unburnt fuse. This resulted in the contents of the OTP being corrupted but with the very low voltage test could now be caught at the standard test temperature without the need to make a $2^{nd}$ screening at minus temperature.

The remaining problems are to verify if the actual 3 dppm failure is real or whether these devices were false failures. There is a limit on how far the VDD can be guard banded due to the voltage regulation limit of the reference voltage of the OTP. Therefore, there is a limit to how far the dppm could be reduced; this would need to be ascertained. For now, there is a good compromise between yield and potential failures in the field.

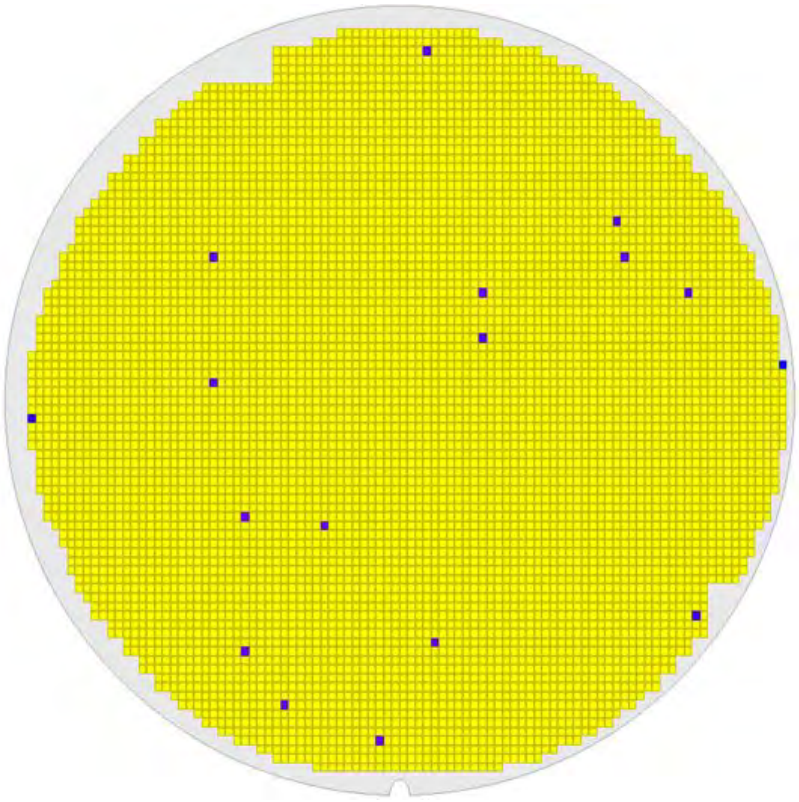


**Figure 167 - Stacked map of failing die over 30 lots – 15 devices in 4.6M – 3ppm**

# 8 FAST BIT SCREENING OF AUTOMOTIVE GRADE EEPROMS - CONTINUOUS IMPROVEMENT EXERCISE

## 8.1 Introduction

EEPROMs are able to store data even when the supply is turned off. Each EEPROM bit cell is turned either on or off by changing the threshold voltage, which itself depends on the charge on the floating gate of the bit cell. Two mechanisms are widely used to charge and discharge the floating gate (FG): one is hot carrier charge injection and the other is Fowler-Nordheim tunnelling. In both cases, a high voltage is needed to charge and discharge the floating gate. In this sense, EEPROM bit cells are special high voltage devices that require high quality oxides to ensure the correct functionality of the memory. It has been shown in the past that EEPROMs using Fowler-Nordheim tunnelling have advantages over those using hot carrier technologies with reference to the maximum number of endurance cycles in combination with data retention. Hence using the tunnelling technology to charge and discharge the EEPROM bit cells is the best option for those smaller memories in the kilobyte range used in sensor applications.

The EEPROM technology shown in this paper is based on a p-channel electrically erasable programmable read-only memory cell that is implemented in a 0.35 µm CMOS technology.

The EEPROM array has a cross-point structure and the floating gate is charged via the Fowler-Nordheim tunnelling effect [59-62]. A fully differential memory approach is a good choice for high reliability EEPROMs. In a fully differential memory, one data bit consists of two EEPROM bit cells. A data bit is always written such that one bit cell is programmed and the other bit cell is erased. This results in a good readout margin.

A major challenge in establishing a stable EEPROM process is how to realise a compact cost efficient technology without losing quality. To achieve this, a highly sophisticated production test screening procedure must be used. The EEPROM test flow has to screen out process weaknesses, [63], therefore cannot be reduced to standard digital memory testing [64]. The test flow also has to ensure that the endurance and data retention specifications, [65, 66, 74, 75] quoted in the data sheet can be met. Most of the test procedures are focused on the oxide integrity that is the major process parameter for endurance and data retention. Another important point is the quality of the contacts used in the memory plane that can generate serious problems in the field if they are not properly screened [67]. This is in contrast to EEPROM screening procedures at other IDM's around the globe where the voltages used for programming and erasing the bit cells can only be generated using the internal charge pump of the device hence not guaranteeing each device sees the exact same voltage. In addition, there are no test modes available that allows the current of each bit cell to be measured; hence being able to guarantee the charge that bit cell is capable of storing. By being able to manipulate the actual voltage the bit cell sees allowed the author to be able to make leading edge characterisations of the bit cells that in turn allowed to fine tune to a process of detecting fab lots that could be a risk to ams AG automotive customers. These EEPROM test features, both in DFT and test IP, have been positioned ams AG as a leader in the EEPROM reliability field. As such, many patents have been and are currently in the process of being filed on the topic.

For standard EEPROM technologies, it is very common that an EEPROM could have a lifetime of 1 million write cycles at room temperature with guaranteed retention of 10 years. However, when the specification changes to 150 degrees, for under the hood automotive applications, this changes somewhat. To guarantee 50 thousand write cycles

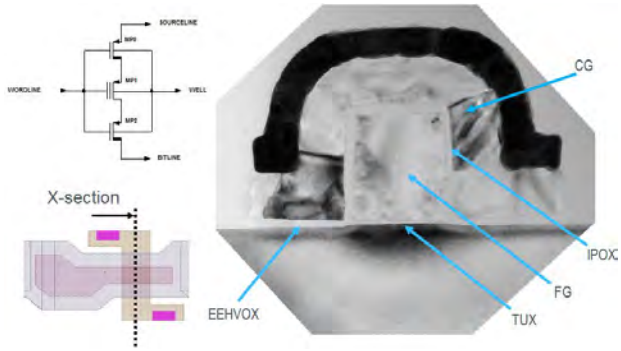



**Figure 168 - Electrical equivalent circuit, layout, and TEM image of the EEPROM bit cell.**

is leading edge and depending on the technology used, sometimes impossible to achieve. It is only due to the combination of the bit cell architecture, the DFT and test procedures in place at ams AG that allows them to specify such a hard specification.

For automotive applications, a fully differential memory approach is the preferred solution. For this chapter we will concentrate solely on this architecture.

This chapter is organized in the following way; Chapter 8.2 will discuss the physics of the bit cell. Chapter 8.3 will discuss the design for test (DFT) that enabled accurate testing of the bit cell. Chapter 8.4 will discuss the effect that a test time reduction exercise had in capturing a maverick lot and how it detected an issue that charge accumulation had masked within the bit cell during test. Chapter 8.5 will go into details of how a process shift affected the test program and how yield was improved by considering these shifts. Chapter 8.6 and Chapter 8.7 will discuss defects that affect and EEPROM and how tets coverage is measured. Conclusion will then be drawn.

## 8.2 Physics of a Bit Cell

The oxide screening procedures depend on the technology used. A transmission electron microscope (TEM) image of the actual bit cell technology shows the different oxides needed to create the floating gate (Figure 168). The tunnel oxide (TUX) and the interpoly oxide (IPOX) surrounding the floating gate to isolate it from the rest of the bit cell. This isolation guarantees that electrons can only move into or move out of the floating gate by applying a high voltage of approximately 13.5 V. The movement of charge should always take place on the interface TUX to NWELL. The IPOX only isolates the floating gate, but no charge moves over this barrier.

The high voltage oxide (EEHVOX) is needed to build the two select transistors that are integral to the bit cell; this is a feature of the technology. These two select transistors ensure the immunity of the bit cell to disturbances.

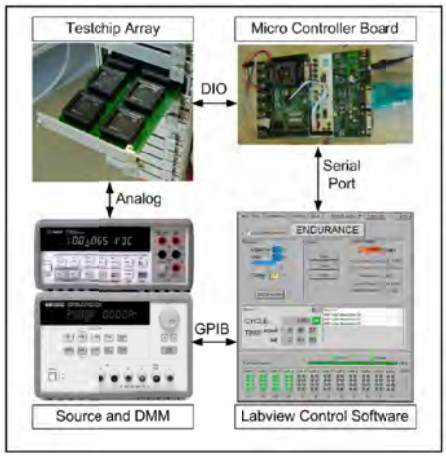



**Figure 169 - Laboratory test equipment for EEPROM bit cell fast bit testing**

It is essential that the test routine should be able to screen for possible defects in the oxides of the bit cell and high voltage transistors. The most critical oxide is the tunnel oxide, and in many high reliability EEPROM technologies, a special oxide is grown to achieve the required quality. A test concept described in [63, 68] determines a way to screen for defects in the TUX and IPOX which are the most important oxides of the EEPROM bit cell. In principle, the test routine checks for bit cells that can be fast erased or fast programmed. It can be shown that such fast bits will be unable to operate for the specified number of endurance cycles defined in the data sheet. Therefore, such EEPROM cells must be screened out. Depending on the technology used, a fast erasable bit cell may not be a fast programming bit cell; therefore, the test routine must check for both polarities of the bit cell. As stated in [68] this fast bit screening should take place at the beginning of the EEPROM testing.

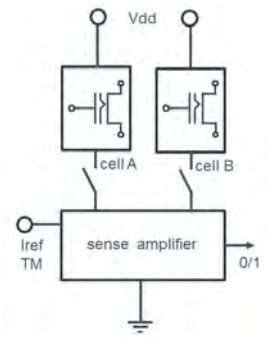



**Figure 170 - Schematic diagram of the differential bit cell concept and the sense amplifier.**

Due to the architecture of the bit cell, it is possible to repeatedly program or erase the bit cell with an increasing voltage as one would do with a with a standard semiconductor characterisation test program. For ease of programming and the test time, this approach has its advantages. However, the charge accumulation will result in a voltage shift between the apparent voltages that the bit cell is programmed or erased with. This in turn can cause further correlation issues when dealing with suspicious production lots that could have a quality or reliability issue. Due to the charge accumulation issue, it is therefore a strict requirement from a quality and reliability standpoint that for every programing operation of the bit cell an erase operation at the max voltage is done to ensure that any charge accumulation is cleared. The same criterion is required for an erase operation in that the bit cell has to be fully programmed at the maximum voltage to ensure that the charge accumulation is cleared. This guarantees that a true correlation to lab equipment can be done without the need for unreliable correlation factors (Figure 169) [69].

## 8.3 Design for Test

There are three important criteria for the quality and reliability of an EEPROM and its contents over lifetime.

i. Maximum write cycles – the number of write operations the EEPROM can withstand until the EEPROM can no longer be written
ii. Write operation program and erase voltages
iii. Charge contained within each bit cell after a write operation.

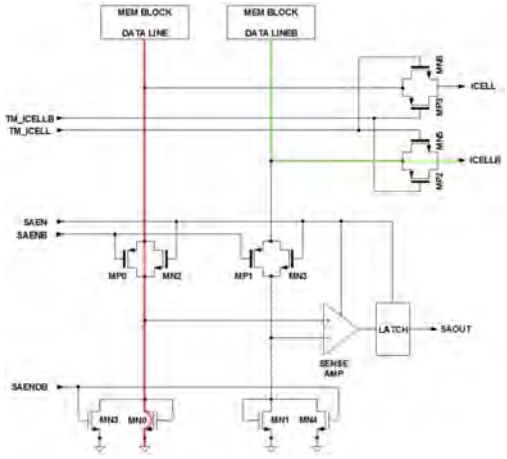



**Figure 171 - Fully differential sense amplifier schematic. Red line: current path in sense amplifier mode for DATA LINE. Green line: current path in external cell current measurement mode for DATA LINEB.**

To enable reliable testing of the bit cells, a test mode is added to the EEPROM block so that test and verification can be done easily. This is implemented by having a pin called MEAS that can be directly connected to the on chip high voltage generator, in such a way that the voltage needed to program and erase the bit cells can be driven from the outside world. Once the bit cell is programmed or erased, the bit cell current can be measured on the same pin, MEAS, connected to the node IREF, (Figure 170), and if the charge is insufficient for guaranteed lifetime the device can be failed easily.

With all of the implemented DFT methods, every chip will see the same voltage for program and erase operations. This would not be the case if the internal charge pump of the device were used for program and erase because the internal charge pump output voltage varies between wafers and over lots. Hence, due to the implemented test mode, the test quality of this block is improved.

If the bit cell current is greater than the reference, a digital high signal is sent. If the bit cell current is smaller than the reference, a digital low signal is sent and a digital pattern can be used to verify all bits are programmed or erased correctly. Using this technique, a quick digital check that all bit cells are programmed or erased correctly is available. Furthermore, a margin of the ICELL current is available so that the digital test has a guard band ensuring greater product quality.

In the bit cell current test mode, the sense amplifier is turned off. Both transmission gates (TG) MP2 / MN5 and MP3 / MN6 are closed (i.e. conducting) and the bit cell currents of the selected fully differential bit cells connected to DATA LINE and DATA LINEB can be measured on the pins ICELL and ICELLB against ground. As an example, the current path for DATA LINEB is shown by the green line in Figure 171, [67]. During these measurements, the TGs MP0 / MN2 and MP1 / MN3 are open. In the standard sense amplifier mode, the TGs MP2 / MN5 and MP3 / MN6 are open and the TGs MP0 / MN2 and MP1 / MN3 are closed. Both transistors (MN0 and MN1) are in diode configuration. Depending on whether the selected floating gate of the bit cell connected to DATA LINE is programmed or erased, some current or no current flows through the diode MN0. The applicable current path is marked red in Figure 171. The same is true for the diode MN1 connected to DATA LINEB. If the EEPROM bit cell connected to DATA LINE is programmed, a current flows through diode MN0 which leads to a voltage drop on MN0 depending on the level of the current and the dimension of the diode MN0. This voltage drop on MN0 reduces the drain-source voltage of the

selected EEPROM bit cell connected to DATA LINE. The reduction of the drain-source voltage leads to a reduced drain-source current of approximately 20 % in the chosen technology and explains the value used for the external reference current, which will be discussed in Chapter 8.5.

The second problem is the drain-source current degradation of the EEPROM bit cells caused by the two NMOS diodes (MN0, MN1) in the sense amplifier configuration. The process variations of the two NMOS diodes are not tracked with the bit cell current test procedure because the sense amplifier is turned off in that test mode (Figure 171).

## 8.4 Release to Production

### 8.4.1 Fast Bit Testing

A fast bit is detected when the bit cell programming or erasing voltage is very low or the resulting bit cell charge current is below a certain threshold. Although these bits work correctly when the EEPROM is initially programmed and erased, a problem can occur when this is done multiple times; due to the low programming voltage, the bit cell could wear out quickly and not perform to specification. This is why fast bit testing is done to ensure that these weak bits are found and the affected device effectively screened out. These parts would fail any specification-based endurance loop testing, and hence would not meet specification to the maximum number of write cycles [63-68, 72, 74]. During the test program and device qualification, marginal devices were subjugated to the specified maximum write cycles plus a 10K addition guard band stress test to guarantee the maximum write cycles, after which they were baked for 4000 hours to ensure that the devices met the data retention specification. Limits for the fast bit tests were chosen based on this qualification and verification process.

Figure 172 shows the distribution across a wafer of the required voltage to erase all the EEPROM bit cells. It can be clearly seen that there are some hard failures. However, it also can be seen that there are some devices, which have the potential to be marginal. Those devices require special attention to be certain no parts will be shipped field that could fail the endurance specification in the field.

For the initial release to production of the EEPROM, the bit cells were effectively characterized by incrementing the voltage, using a ramp, red line shown in Figure 174, according to the voltage profile (Figure 173), to find the actual minimum program and

erase voltages of all bit cells, for all devices, on all wafers produced. In addition, for test time reasons, the voltage was linearly increased in 100 mV steps, using a ramp, red line shown in Figure 174, without clearing the bit cell programming, which resulted in an accumulated charge within the bit cell. This resulted in a 1 V difference between the lab and the Automated Test Equipment (ATE).

The test procedure described above was released to production after correlation to the lab equipment, Figure 169. The test described captured fast bit devices that would have failed the endurance loop specification of 50,000 write cycles at 150 degrees for 10 years lifetime. This test became the benchmark for all ams AG EEPROM designs for single ended and pseudo differential bit cells.

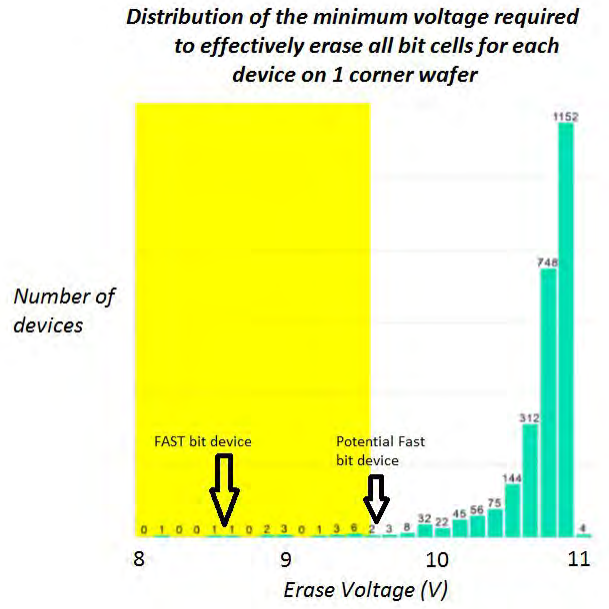



**Figure 172 - Minimum voltage to erase all bit cells for each device on one wafer**

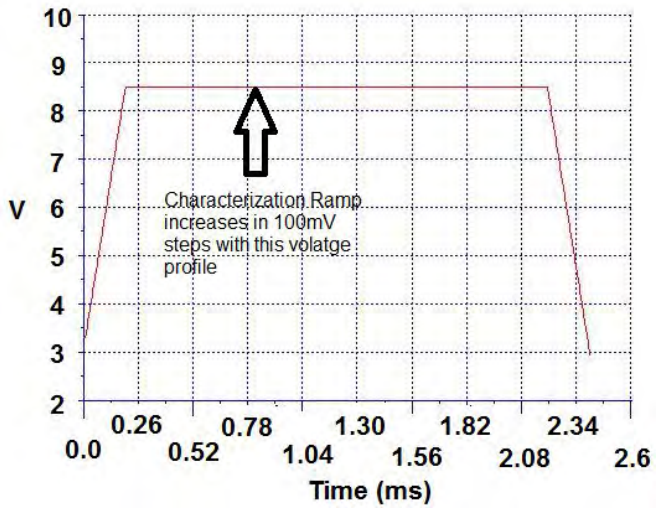



**Figure 173 - Programming profile of EEPROM bit cell**

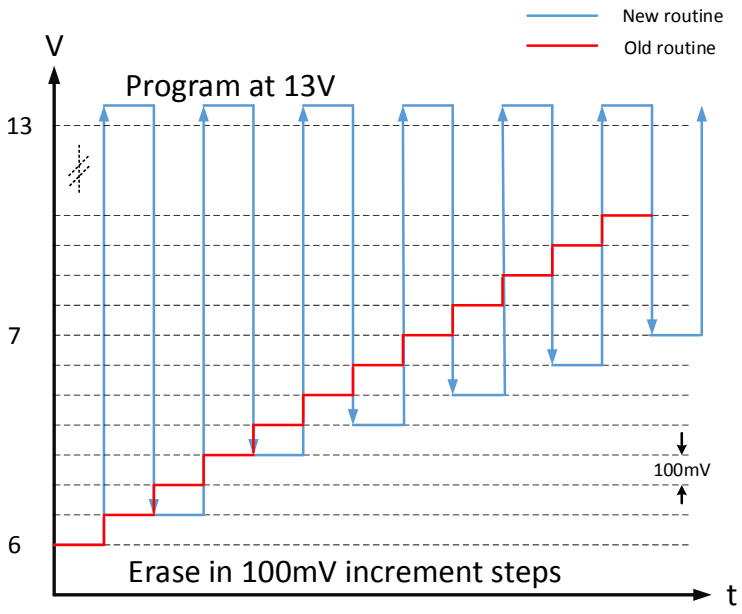



**Figure 174– Full Programming profile of EEPROM bit cell**

## 8.4.2 Ramp Up

The EEPROM production test program Intellectual Property (IP) was designed as a 16-site solution (Figure 175), to exploit the maximum pin count tester configuration. This ensured that the most test cost effective solution possible was produced at the time.

The fast bit test guarantees the lifetime of the bit cell, however it cannot guarantee the reliability in the field. An additional test step is required to guarantee the integrity of the bit cell so that latent defects are captured and do not enter the field. This was done by additional bake and verification steps.

During the first production test stage, the bit cell currents were directly measured using the available test modes (Figure 171). The values of the bit cell currents were then stored in a database for use at a later stage. A digital pattern was also stored in the whole EEPROM memory as part of this procedure so that this data could be read-back as part of a second stage verification.

During the characterisation of the EEPROM, where wafers were baked for 4000 hours to guarantee the 10-year lifetime specification, it was found that 24 hours were only required to trigger latent defects so that they became detectable. To detect any failures triggered by the bake step, the bit cell currents were measured as part of the second test step. The measured bit cell currents were then checked against the bit cell

currents measured during the first test stage and delta measurements performed. Any deviation of a bit cell current greater than 1 uA would result in the device being failed.

To detect any failures triggered by the bake step, the bit cell currents were measured as part of the second test step. The measured bit cell currents were then checked against the bit cell currents measured during the first test stage and delta measurements performed. Any deviation of a bit cell current greater than 1 uA would result in the device being failed.

With this characterisation style test program in place, several maverick lots from the foundry were successfully captured, which resulted in over 10 million EEPROMs (average 300 billion bit cells) being tested without one RMA coming back from a customer.

### 8.4.3 Charge Accumulation Characterisation

A mini characterisation exercise was undertaken to develop an optimized test program, based on a one-shot test approach to reduce test costs. This was due to the large test time of the ramp-up characterisation test program that was in production. As there was full access to all the characterisation data, it was possible to make a very good estimation of the required parameters for the one-shot tests at the beginning of the cost optimization exercise.

**Figure 175 – 16-site EEPROM wafer sort probe head**

As described in Chapter 8.4, the original test program consisted of a ramped voltage, using a ramp, red line shown in Figure 174, which found the minimum erase and program voltage by use of a test mode within the EEPROM. Unfortunately, due to charge accumulation within the bit cell, the correlation to a one shot measurement

would not be completely accurate and some form of endurance loop characterisation would need to be carried out.

The first task of the test time reduction was to find a minimum voltage, where the device would program and erase reliably, but at the same time guarantee the performance of the device. Due to the charge accumulation, this could only be estimated due to the 1 V difference experienced during this initial release. However, by producing a ramp that cleared the charge accumulation and characterised where the device would fully program and erase, blue line shown in Figure 174, an accurate measurement could be made for the one shot measurement and show the correlation to the original implemented ramp, red line shown in Figure 174.

Based on historical data and experimenting with a one shot measurement of the fast bit test, an optimal setting was found for each test, these setting correlated to the lab test setup (Figure 169) and the original test program. A corner wafer was then run, and the wafer map and production data were saved. Running the same wafer with the old characterisation style test program, one can see where a die flips from fail to pass. Any device that flips from a fail to a pass is cause for concern and needs to be verified by endurance loop characterisation. Figure 176 shows a binflip report of the old to the new test program results. The non-green dice either are showing where a device has flipped from a pass to fail or fall to pass. The binflip map indicated that 54 devices that required special attention; however, it was shown that 49 out of the 54 were marginal parts and only five required detailed analysis. The remaining five devices were subjugated to 60,000 endurance write cycles (the specified number of write cycles plus a 10,000 guard band), at 150 degrees. The wafer was then baked and the five parts checked to see if the data held within the device was not corrupted and the bit cell currents had not shifted as described in Chapter 8.4.

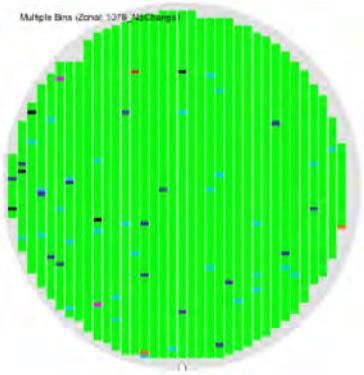


**Figure 176 - Binflip report with fail to pass flips of EEPROM corner wafer**

All devices passed the endurance loop verification and the test program released into production with a 50 % decrease in the first stage test time. The test time drop of 50 % was due to being able to remove two characterization tests that were running in the production test program that contained 80 to test two parameters and replaced them with two one-shot measurements making the whole fast bit testing much more accurate and efficient. Also due to the charge accumulation issue seen with the initial test program, it can be stated that a truer correlation to the laboratory test setup was achieved, hence not only giving a cost reduction but an increased confidence in the quality of the shipped die.

### 8.4.4 Production Monitoring and Maverick Lots

After over 50 wafer lots were tested, a question was raised to the validity of the new test program. This came about due to a maverick lot seen from the foundry. The new test program was showing a step field issue (Figure 177) with a lot from the foundry. However, when the original program was used for verification no step field was obvious (Figure 178). This caused some internal issues as the old program was saying that the wafer lot could be shipped whilst the new was showing a reliability issue. In the commercial world, this is a huge issue as any wafer not shipped affects the profit and loss of a company. Hence, there is a pressure to ship as much material as possible.

The original test program had the effect of the charge accumulation that caused a shift in the actually voltage the bit cell was erased at, and in this case was causing an issue in the processing of wafer to be seen as a good quality processing.

The original test program was modified such that, for every step taken in the linear search (Figure 173), of the erase voltage applied to the bit cell, that there was a full program cycle after each step. This was to avoid the issue of charge accumulation within the bit cell. This resulted in a close correlation to the wafer map seen in Figure 177. This proved that both the maverick wafers were correctly captured and that the new cost optimized test program was trust worthy. Further analysis of this lot was needed.

After testing 52 suspect dice with all 3 test program, it was observed that the old program had no correlation with regard to the bad material with either the one shot measurement or the modified ramp. Therefore, there were risks in capturing such a maverick lot shown with the old test solution. As the modified ramp had correlated to the new shot fast bit, no further investigation was warranted. This demonstrates that the

in-depth correlation that was carried out during the release phase of the optimized test program was correct. However, an issue was only seen with the erase voltage of the bit cell. One would expect that with such a wafer signature that the programming voltage would also be affected but the test did not show any obvious sign of this.

After running all three-test programs for the program fast bit test, 11 suspect devices were identified to have further analysis. It was observed that there was not a 100 % correlation between the one shot and the new modified ramp and no correlation to the original ramp test. The problem here was that the voltage used in the one shot test to program the bit cell was marginally too low and needed to be decreased by 100 mV to achieve 100 % correlation and achieve the required test coverage.

However, it has to be noted that due to the original characterisation and the fact that marginal parts were aggressively guard banded such that a marginal part would fail that the chances of reliability risk device being shipped into the field with the issue seen in Figure 177 was extremely low.

## 8.5 Yield Optimisation

After several years of constant production, the material review board (MRB) reported a maverick lot in the silicon foundry, which had a high yield loss due to oxide integrity tests (Figure 179). The blue dots in the wafer map indicate continuity fails and the yellow dots show leakage failure, while the red dots refer to EEPROM fails. The light green areas of the wafer map indicate specially stressed parts; these parts are inked out before the wafer is shipped into the field.

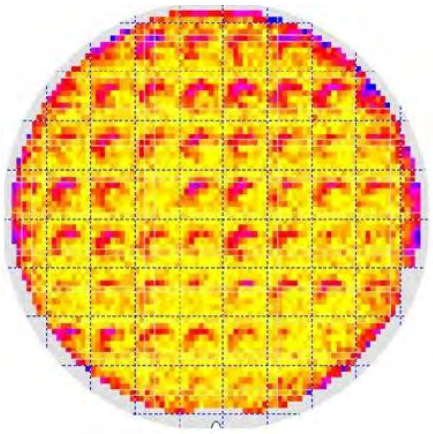


**Figure 177 - Wafer exhibiting a step field reported by cost optimized test program**

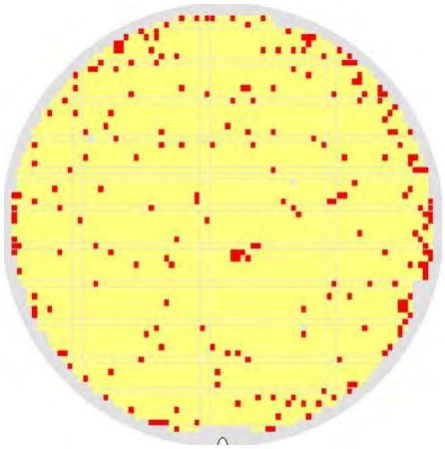


**Figure 178 - Wafer map of step field wafer with the original test program**

### 8.5.1 Analysing the Oxide Reliability Test Results

A stable running EEPROM technology normally suffers only from standard defects or maverick events in the silicon foundry. Therefore, a wafer map showing fast bits should show them randomly distributed across the wafer (Figure 180). This further emphasises the failures shown in the wafer map found by the material review board (Figure 179). In order to gain an understanding of the root cause of the failures, all relevant wafer acceptance test (WAT) parameters were screened. Clearly, all WAT parameters must be within the specified limits; otherwise, the abnormality would have been found before the lot was moved to the production test stage. A TUX QBD (charge to breakdown) SLM (scribe line monitor) measurement is part of the WAT parameters. A comparison of the actual run with previous lots displayed no abnormality (Figure 181).

The only parameter that showed an anomalous distribution over the lot and was close to the specified limit, was the drain-source current of the EEPROM bit cell (Figure 182). A small 4x4 EEPROM bit cell array is part of the SLM structures. All terminals of the bit cell (source, drain, gate, and bulk) are accessible in the SLM. The measured drain-source current is typically in the range of 37.5 uA. The mean value of the cell current of the maverick lot was around -30 uA, a difference of 20 %.

Two test procedures in the production test flow are related to the drain-source currents of the EEPROM bit cells. Prior to the required bake step, to check the data retention performance of the EEPROM, a special data retention pattern is written into the memory. At address ‘0’, a checkerboard is written, the rest of the EEPROM array is set to logic ‘0’. At address ‘0’, a checkerboard is written, and the rest of the EEPROM

array is set to logic ‘0’. In the technology used, a neutral bit cell having no charge on the floating gate can still provide a drain source current. Therefore, a logic ‘0’ in the EEPROM memory map can flip to a logic ‘1’ by losing its charge. However, a logic ‘1’ will remain a logic’1’. To check for drain source current changes, special test modes are implemented in the EEPROM control logic. With these test modes, each bit cell in the memory can be addressed and the bit cell current referenced to ground can be measured with the production tester on a dedicated analogue test pad [63].

The measured drain source currents of the bit cells are slightly smaller, compared to the WAT measurements. The reason for the difference is the HV select transistors in the y-path of the memory. A comparison of the drain source current value wafer map in Figure 183, with respect to the wafer map in Figure 179, shows a good correlation.

The bit cell current distribution of a standard manufactured lot shows an average value around -24.8 uA, Figure 184. A comparison of the bit cell current distribution with the maverick lot shows a shift of the average value around 20 %, Figure 185, which resulted in confirmation of the WAT results. It was then necessary to determine whether the fast bits screened by the test procedure on the maverick lot were real weak bits or if there was another effect that was masked by the fast bit test. The maximum defined number of endurance cycles for this specific memory is 50k cycles at 150 deg.

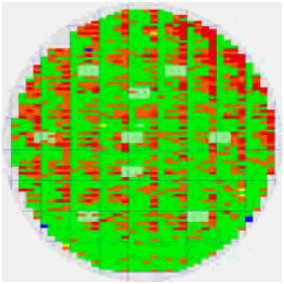


**Figure 179 - Wafer map of an EEPROM showing abnormal fail pattern**

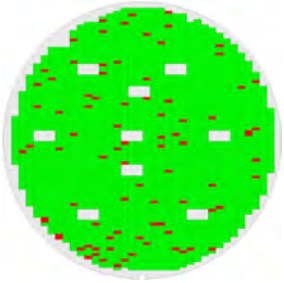


**Figure 180 - Wafer map of an EEPROM showing random distributed fast EEPROM bit cells (red dots).**

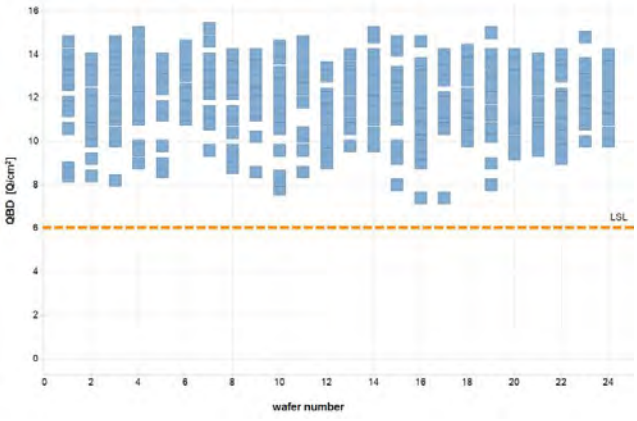



**Figure 181 - TUX QBD WAT measurement results of maverick lot. (LSL lower specification limit).**

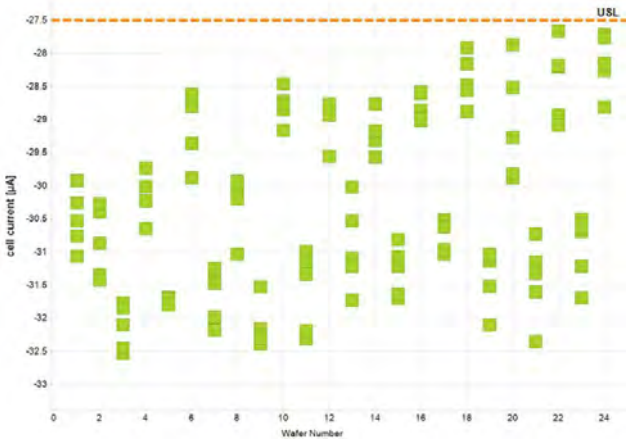



**Figure 182 - Drain source current distribution of the EEPROM bit cells across the lot.**

Data retention is specified to be 10 years at 150 deg. On five randomly chosen touchdowns of the wafer including fast bit fails, [70], a reliability stress test was applied. The devices (16 per touchdown) were cycled up to 60k endurance cycles (10k margin to the spec limit) at 150 deg. After the cycle test, a data retention bake of 440 h at 250 deg was performed and the data has been checked after the bake. All devices passed the reliability stress test successfully. The results of the test confirmed the QBD measurements. Therefore, it was proved that the fast bit routine masked an effect that the actual EEPROM screening test program interpreted as a fast bit failure.

## 8.5.2 Root Cause Analysis

A closer look to the fast bit procedure shows the reason of the masking effect. To judge if an EEPROM bit cell is a fast programming or fast erasing bit cell, an external reference current is used. To find fast erase EEPROM bit cells, the memory array is

programmed with a standard program pulse to ensure that all bit cells are properly programmed. The next step is a soft erase pulse with a defined voltage (Figure 173). The voltage level of the erase pulse depends on the technology and on the defined reliability limits. To find fast erasing bits, the bit cell current of the soft erased EEPROM bit cells is compared against an external reference current. The comparison of the bit cell currents is done with the standard sense amplifier used for normal read operations inside the memory (Figure 171). The sense amplifier must be modified such that it can be used to compare the selected bit cell with the external reference current in a special test mode [67]. The value used for the external reference current is defined by the bit cell current of a properly programmed EEPROM bit cell. An area efficient EEPROM technology always uses the smallest possible geometries available in the process node. The lower limit of the bit cell current is derived from the maximum allowable read access time plus the degradation of the bit cell current generated by the endurance. To enable efficient testing a fixed value of the reference current was used in the actual test program. This value was based on a 3-sigma process variation of the bit cell current plus the endurance degradation. The ideal value of the reference current is half of a correctly programmed bit cell current, assuming that an erased bit cell current provides no current at all. The definition of the reference current being equal to the half of a correct bit cell current results in a symmetrical fast program and fast erase test. Additionally a correctly programmed bit cell current has the same offset to the reference current as a correctly erased bit cell current. The value defined in the production test program has been chosen to be -10.4 uA. Comparing this value with the cell current distribution of a typical programmed bit cell current (Figure 184) shows that the value differs by 2 µA from the average value divided by 2 (12.4 uA). The reason for this difference is determined by the different configurations used for measuring the bit cell current and for using the bit cell in sense amplifier configuration (Figure 171) explained in Chapter 8.3.

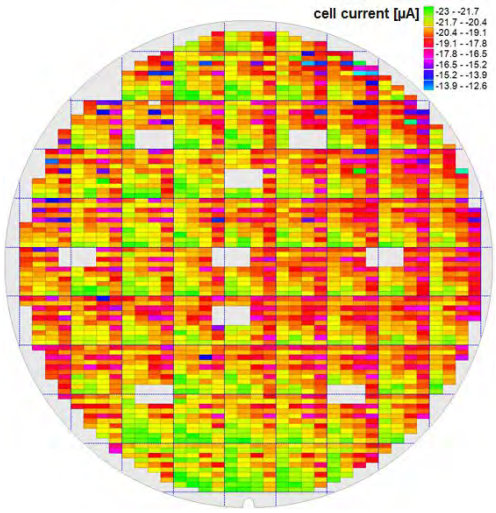



**Figure 183 - Drain source current wafer map of the EEPROM bit cells.**

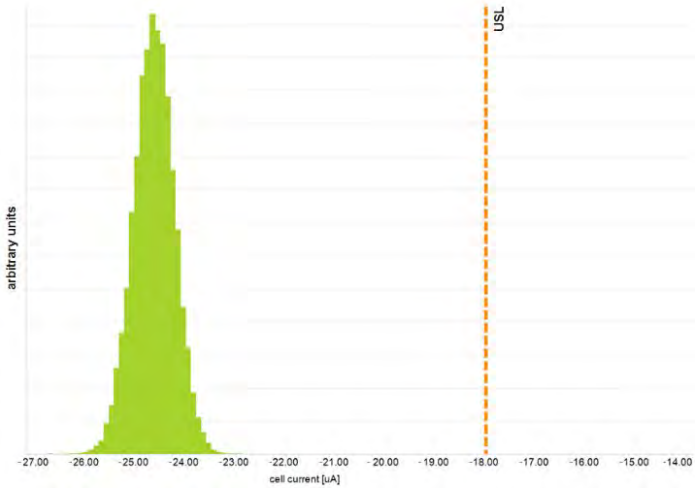



**Figure 184 - Bit cell current distribution (programmed cell) of a typical lot.**

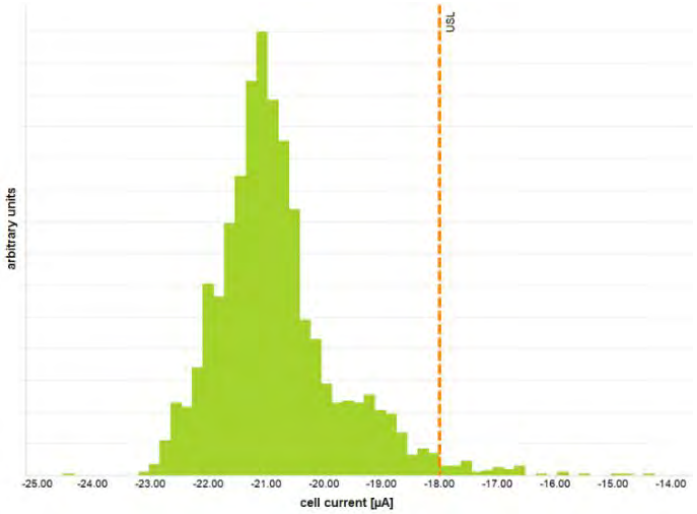



**Figure 185 - Bit cell current distribution (programmed cell) of the maverick lot**

### 8.5.3 Improved Screening Procedure

To improve the actual fast bit test procedure, it was necessary to find a way for the external reference current to track the process variation. The bit cell current of a properly programmed bit cell is a good candidate to calculate the reference current needed for the fast bit test. To overcome the variation of the bit cell current in the memory plane (Figure 185), ideally the average value of all measured bit cell currents should be extracted. This would entail measuring a significant number of different bit cell currents to amass enough statistics to calculate the mid value. As the bit cell current measurement is an analogue current measurement against ground and takes roughly 1ms per measurement this would result in a dramatic increase in test time. A second problem is the drain-source current degradation of the EEPROM bit cells caused by the two NMOS diodes (MN0, MN1) in sense amplifier configuration. The process variation of the two NMOS diodes are not tracked with the bit cell current test procedure because the sense amplifier is turned off in that test mode, Figure 171.

To determine the correct value for the external reference current, a test mode could be used, where the sense amplifier could be turned on and the bit cell current flowing from a programmed bit cell into the NMOS load diode (MN0, MN1) could be measured, depending on which bit cell of the differential EEPROM bit cell pair was selected. A special test mode using an external reference current can solve this problem. Assume that during a sensing operation, one bit cell of the differential EEPROM bit cell pair is disconnected and the load diode of this branch is connected to a current source controlled by the production tester. If the current is set to a smaller value than that of a properly programmed bit cell current on the other branch, the sense amplifier switches its output to logic ‘1’. If the external current is higher than the bit cell current of the connected bit cell, the sense amplifier switches to logic ‘0’. If the exact current of the connected EEPROM bit cell is not known, a sweep of the external current from a defined low value until the sense amplifier switches from logic ‘1’ to logic ‘0’ is needed. The sense amplifier changes its state when the external current is slightly higher than the bit cell current of the connected programmed EEPROM bit cell. Hence, the internal bit cell current can be identified. The procedure also works for the other bit cells that are disconnected. With a second test mode, the bit cell currents can be switched and the bit cell current of the second bit cell can be measured. The advantage of this test procedure is that it is based on a digital decision. This enables fast testing because the whole address range of the memory can be read at maximum defined speed

by using the external current. A more detailed description is given in [68, 74]. The modification needed for the sense amplifier to enable this test mode is described in detailed in [71].

The test described above, [68, 74, 76], can be used to calculate the required reference current for the fast bit test. Therefore, this test must be executed before the fast bit test by changing the position of the contact screen test in the production test program. The contact screening test is done after a proper program pulse and before the soft erase test is executed for the fast bit screening. The reference current is calculated from the mid value of the measurements made during the sweep described above. The increase in test time is minimal because the calculation of the average value is fast. As the contact screening test correlates with the bit cell current test the wafer maps should show similar results, Figure 183 and Figure 186. The last requirement of the investigation and yield improvement exercise was to prove that the outgoing quality level of the new procedure is equal to or greater than that of the old procedure. By running the improved test program on the wafer showing fast bits (Figure 180), the new test routine showed a perfect correlation with the test time optimized routine. The new fast bit test procedure was able to find all fast bits, so there is no risk of a reliability escape rate (Figure 187). Running the new test program on the maverick lot increased the yield by around 15 %, (Figure 188), to the around expected yield based on the defect density of the ams 0.35 um EEPROM process. The rest of the failing parts are due to there being simply not enough bit cell current to fulfil the read access specification of the EEPROM at the end of the life cycle.

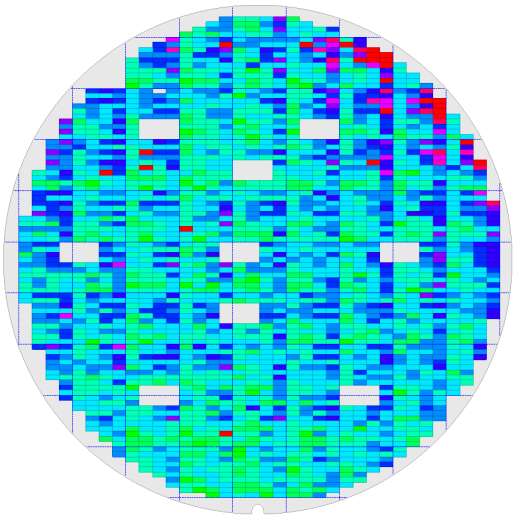


**Figure 186 - Wafer map of the EEPROM bit cell current test**

Figure 189 shows 5 wafer locations that were affected the most by the shift in the bit cell current. Just for clarification that this new procedure would not allow potential fast bits into the field, 5x16 devices were stressed with the specification + guardband write cycles, these wafers baked for 24 hours and the retention test executed. All devices passed both the endurance and retention tests. Due to this successful experiment, the test program was released into production, for high volume manufacture, for all of the EEPROM products running at ams AG.

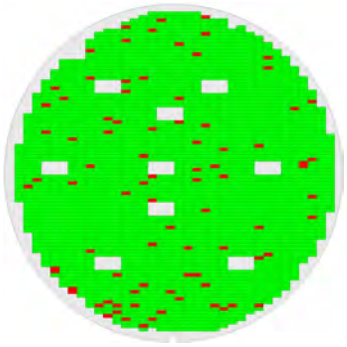


**Figure 187 - Wafer map of fast EEPROM bit cells tested with the optimized fast bit test (red dots).**

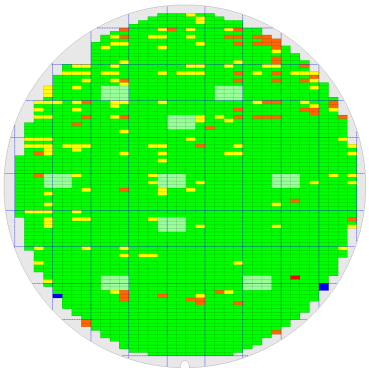


**Figure 188 - Wafer map obtained with optimized production test program.**

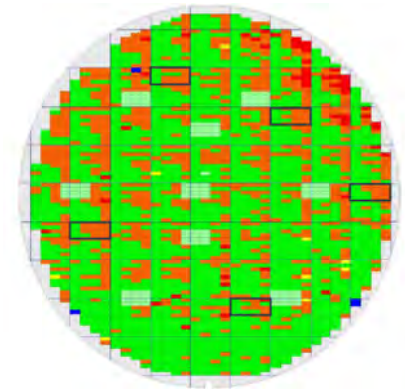


**Figure 189 - Wafer map of an EEPROM showing abnormal fail pattern**

## 8.6 EEPROM bit cell defect types

Two different types of defect can be assumed based on measured results.

### 8.6.1 Impurities within oxide layer

This causes a bit cell failure after a few program and erase cycles, this is a catastrophic fault and is mandatory to be caught. These devices die within a few cycles and are caught with the standard program fast bit test. These devices are always located on the outer rim of the wafer and are thought to be caused by micro debris falling from the spacer wafers that are placed above and below the wafer lot in the furnace. By replacing these cheap spacer wafers for every lot rather than reusing them significantly reduces the defect rate attributed to this phenomena. This phenomenon was also confirmed between two independent foundries used by ams using the same IP and process technology.

### 8.6.2 Trapped charge

Trapped charge within the bitcell causes an NBTI failure mechanism, which leads to early lifetime failure of the bitcell through endurance cycling, compared to a standard cell. This failures is also caught by the standard test program described, however, cycling these devices generally only results in an endurance failure after 750K cycles at 150 degrees. This is over 10X the guaranteed specification therefore quality of this specification is not in question and could be dropped from the screening procedure for low endurance products. However, for products up to and including a rating of 1Million cycles this standard screening test would need to be in place.

## 8.7 Test/Defect Coverage Discussion

For EEPROM, the test program is not only looking for device functionality as we usually look for in devices such as RF devices. What we are looking for is to guarantee that a device can fulfil a certain number of programming and erase events and when the device is programmed that the data stored within the device will retain the data for a number of years. If an EEPROM process has the capability of withstanding 500K write cycles and the spec is only 50K, the there is a 10x guard band on this spec. What we are interested in is from a quality perspective, that we are not at risk of a maverick event causing this guard band to reduce significantly to endanger the 50K spec. This is a different definition of coverage as we are looking at foundry control and latent defects.

These specs can only be guaranteed by doing a full product qualification with a released test program. In this case, this was the test program described in Chapter 8.4.1. Being able to capture all endurance failures with this program and detect shifts in bitcell currents for an EEPROM that had initially a 200K cycles endurance performance gave a confidence that the test coverage was 100 %. At the time of producing the cost optimised version of the EEPROM test program, chapter 8.4.3, the IP had been in production for 5 years with 300 billion bitcells tested without any RMA received from any automotive customer. Hence, it can be stated that on a bitcell level that with the no field returns in 300 billion bitcells tested that we have measured test defect coverage of 3.33 ppt.

Once the cost optimised test program was released, Chapter 8.4.3, no changes were made to the test methodology for another 3 years, with an additional 500 billion bitcells tested, without any returns being received from any automotive customers. Using the same arguments as above, we can say the defect coverage of this test program is 2.00 ppt or 1.25 ppt combined. With the yield enhanced test program, chapter 8.5 and additional 100 billion bitcells were tested with no additional returns received. Hence, combined we have a measured 1.11 ppt quality level or defect coverage.

By improving the yield of a product other than by decreasing limit but by adapting the test program in a way that best suits the devices architecture not only increases yield but also increases the quality level as outliers in the tail of the Gaussian distribution are removed. Even if these devices were shown to pass qualification, it is better to remove them as they do not contribute any major yield loss but the confidence in the delivered material is greater. However, by doing such exercises in quality and foundry control, a much deeper understanding of the device quality can be gained and any foundry improvement and excursions can easily be tracked and hence feedback to the foundry immediate.

When the EEPROM IP was developed and released to production, the foundry could produce bit cells that could only endure a maximum of 200K cycles. Now with process improvements and increased knowledge of how defects affect the EEPROM bit cells ams AG is looking at using the same EEPROM for datalogger devices that required a guaranteed 500K to 1000K cycles. Using the knowledge developed about defect coverage and data retention capabilities ams AG is able to deploy such technologies in the future.

It is now understood that capability to guarantee 1 Million endurance cycles lies in being able to detect the max difference current shift between a partially fully program programmed bit cell. If the delta between these two currents is not sufficiently large enough then it is guaranteed that the endurance will be less than 1 Million cycles and possibly be only 750K, this is know this from experimentation and characterisation, (Figure 190), of fast bit devices found in production that would have not been shipped to current customers. However, this type of defect coverage would not be at all possible without the possibility of analysing a primarily digital chip in an analogue form and this would not be possible without the test modes designed for high quality test purposes. Patents are filled for these technologies and this ensures that ams AG remain the best in class EEPROM AEC-Q100 Grade 0 quality and with a cost effective solution.

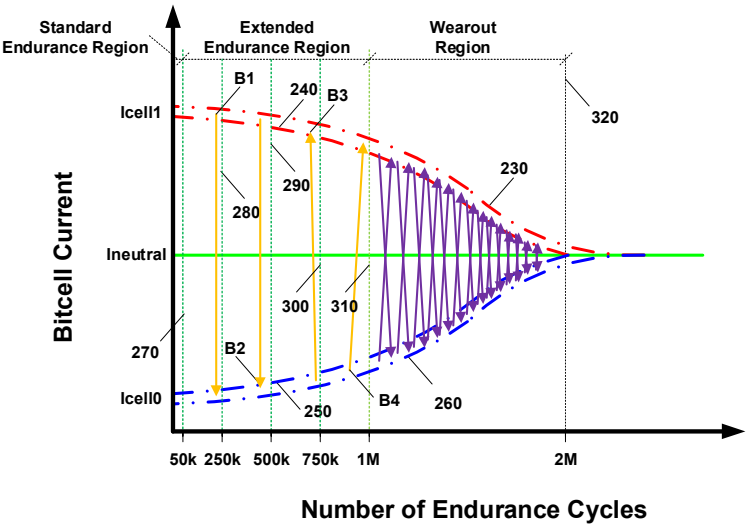



**Figure 190 - Endurance behavior of NVMs based on Fowler-Nordheim tunneling**

## 8.7.1 Test/Defect Coverage Qualification

After successfully passing the NVM screening, only random distributed intrinsic failures will be activated during the lifetime of an EEPROM bitcell [71]; however, it is extremely difficult to catch such failures.

Making an in-depth analysis of the intrinsic failures, it can be seen that there are only a few failures within one memory plane that are activated before the NVM enters the wear out region and the bit cells starts to collapse at a high rate. The beginning of the wear out region indicates the end of the bathtub curve and the bit cells start to fail. If there are only a few bit cell within one memory array that will activate during life then a way of predicting the number of endurance cycles the bitcell can take is needed to ensure the

outgoing quality level. The key to solve this issue is a screening routine that can predict the number of endurance cycles of each bit cell in the memory array and keep the test time under control.

The program and erase state of NVMs based on Fowler-Nordheim tunnelling can be measured in two ways. The threshold voltage or drain to source current can be used to judge if a bit cell is efficiently programmed or erased.

The screening procedure developed by Mielke [69] is based on the fact that weak NVM bit cells can be erased faster compared to standard bit cells.

A closer look to the endurance behaviour of NVMs based on Fowler-Nordheim tunnelling shows that after a correct infant mortality screening, only a few bit cells inside the memory have a potential risk to fail before the NVM reaches the wear out region, Figure 190. Depending on the used bit cell technology, it is possible to measure the erased and programed state of a NVM bit cell either by the threshold voltage of the bit cell or by looking at the drain to source current of the bit cell. A programmed bit cell delivers a drain to source current in the range of Icell1max and Icell1min and an erased bit cell a current between Icell0max and Icell0min. As an area efficient NVM is required, the focus is on the bit cell size because this is the main area contributor. The size of the bit cells is pushed to the process limits so as to ensure the most efficient bit cells in terms of area. A bit cell current distribution of the Icell current for program and erase across the memory plane will be observed, the range of the cell current is caused by process variations and will change from wafer to wafer and lot to lot. Therefore an adaptive method is required to take these variations into account rather than having a screening method.

From Figure 190 it can be seen that 50k endurance cycles can be safely screened and secured with standard screening procedures with an acceptable yield loss and test time. To extend the number of endurance cycles to screen for 125k endurance cycles the algorithm must be capable to find all weak bits in the extended endurance range for example B1, B2, B3 and B4, Figure 190. Such weak bits have no hard failure mechanism like short to ground or supply, the failure modes can be trapped charges in the tunnel oxide which are activated during lifetime. The failures close to the maximum allowed number of endurance cycles in the extended endurance range are the most difficult to find. As discussed earlier, the program and erase bit cells have a bit cell current distributions caused by mismatch and process variation. To push out safely the

maximum number of endurance cycles the screening procedure must be able to catch all infant mortality and weak bits in the extended endurance range. An activated weak bit can lead to a stuck at high (B3, B4) or stuck at low (B1, B2). That means depending on the technology and failure mode a fast erase bit is not necessarily a fast programing bit cell or vice versus. So it is essential to prove and screen both program and erase state.

To implement an efficient screening method it is necessary to know when the NVM bit cell will fail caused by endurance cycling and the exact location of the bit cell.

It is useful to define so called cycle classes which will give us a prediction on how many endurance cycles a NVM bit cell can withstand. Modern NVM's with standard screening procedures are capable of 50k endurance cycles at 150deg Celsius. To extend this number of write cycles for example to 125k we define 3 cycle classes 75k, 100k and 125k. As shown in Figure 190 most of the NVM bit cells of the state of the art NVM process can handle the 125k write cycles. Only a few bits are not able to achieve this high number. These bit cells are of interest. Depending on the defect, those bit cells can endure from 50k to 125k endurance cycles. In Figure 190 four examples are shown, failure B1 is activated close to 60k endurance cycles and produces a stuck at "low". B2 shows the same failure mode and is activated in the range of 85k endurance cycles. B3 generates a stuck at "high" failure close to 100k endurance cycles and B4 shows the same failure near 125k. All four failures must be safely detected by the screening procedure to enable a correct replacement of the weak bit cells so the NVM can fulfil 125k write cycles.

Cycle class statistics of the endurance performance must be generated to get a better understanding of the failure modes of the technology and how they are activated. Generating such statistics is a huge effort because endurance cycling of NVMs is very time consuming. It can be found that the endurance behaviour of NVM follows a Gaussian distribution, Figure 149. Most of the bit cells of a mature NVM process are able to withstand a high number of endurance cycles depending on the used technology the value is varying. As shown in Figure 190 the probability of a bitcell being able to endure a high number or endurance cycles of mature NVM technologies is high. The exact numbers strongly depends on the technology and the intrinsic failure rate of the production line, Figure 195. After a successful infant mortality screening, the likelihood of failure B1 (< 75k endurance cycles) is very low as the number of endurance cycles are increasing the likelihood of getting a failure (B2, B3 and B4). Out of the distribution

the maximum number of endurance cycles, and the cycle classes of a NVM technology can be extracted from the characterisation data. The number of cycle classes depends on the sensitivity of the screening algorithm.

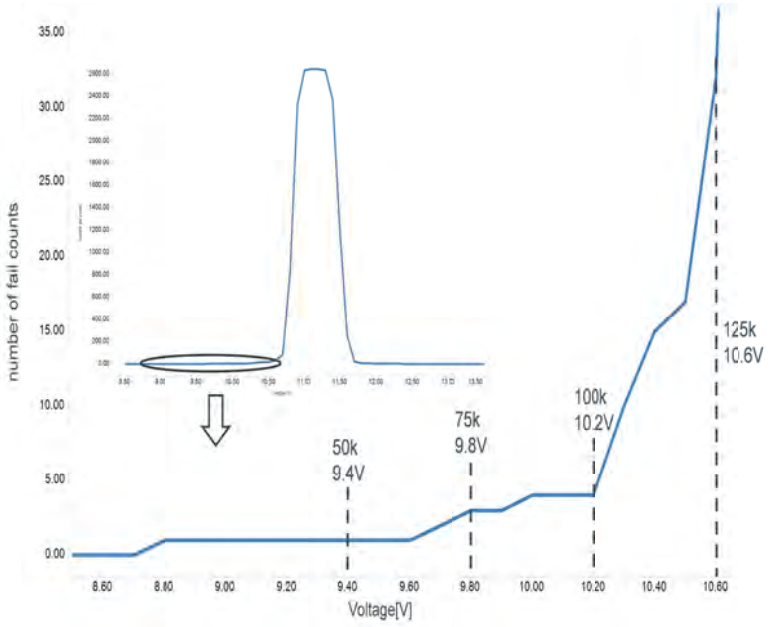



**Figure 191 - Erase Cycle Class voltage distribution**

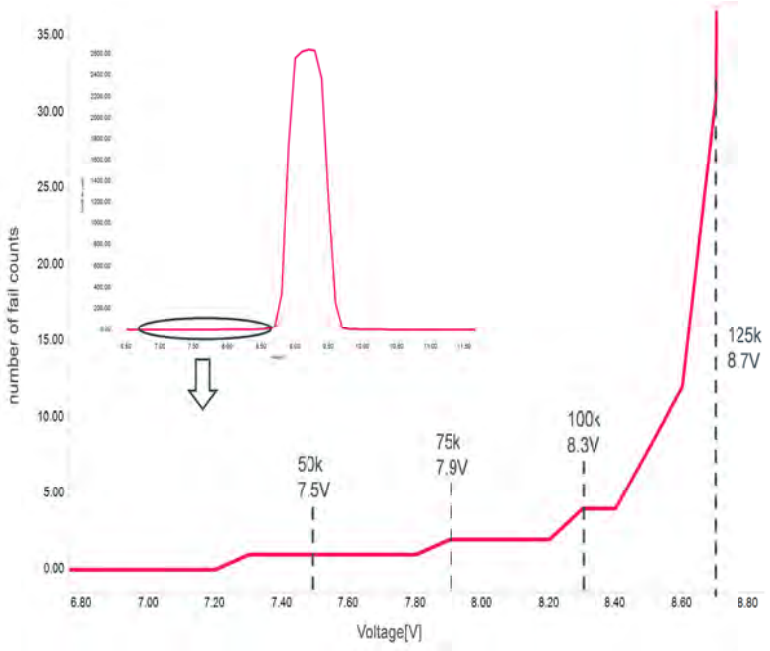



**Figure 192 - Program Cycle Class voltage distribution**

The exact voltage levels of the soft erase pulses used to judge the cycle class of each individual bit cell in the memory array depends on the used technology and can be found during characterisation, Figure 191. As stated earlier a fast erasable bit cell is not

necessarily a fast programmable bit cell therefore the screening procedure must be repeated to check for weak programed bits in the memory. A soft programming pulse is applied to the memory to check for the programming cycle classes, Figure 192. An interesting observation can be made by comparing the voltage levels for weak programming and weak erasing bits, Figure 191 and Figure 192. There is a difference between the cycle classes program and erase voltages. The reason for this behaviour is the NVM bit cell itself which is asymmetric. Resulting in the fact that the bit cell can be more easily programmed than erased. However, this all assumes that the reference current for each bit is centred, as the reference current will shift due to process corners, Figure 184 and Figure 185. This will result in causing more Program or Erase bitcell failures and as the bitcell is asymmetric, will could result in unnecessary yield loss thus delivering bit cells that could be guaranteed for higher endurance than is stated in the spec. By measuring the average current of the bitcell, the reference current can be set and hence centre the reference current for each individual die thus resulting in a centred process for all Erase and Program operations.

Observing Figure 191 & Figure 192, there is a specific voltage the bit cells will program and an erase for at when the bit cell is originally produced. This would correlate to the maximum number of endurance cycles a device will be able to endure over its lifetime. However, if we program the device with a static voltage, as described in 8.4.3, a certain charge will be stored within the bit cell that is propositional to the voltage the bit cell will program and erase. If two voltages are used to measure the resulting stored charge as a current then a delta between these two currents can be made. The measurement of the bit cell current is defined in 8.5. The bigger the delta between these two currents, the greater the distance between these two measured currents is quicker and alternative way of sweeping the voltage to measure where the bit cells will program and erase as shown in Figure 191 & Figure 192. Using this delta, Figure 193 & Figure 194, a judgment on the quality of the bit cell can be made quickly and effectively without the need for expensive and complicated voltage sweeps.

Referring to Figure 149, shows how a devices performance will change with temperature and shows the standard Gaussian distribution of any silicon performance metric. By means of characterising the program and erase voltages of an EEPROM bitcell such that the outer regions of the Gaussian distribution have been found, as described in chapter 8.4.2.

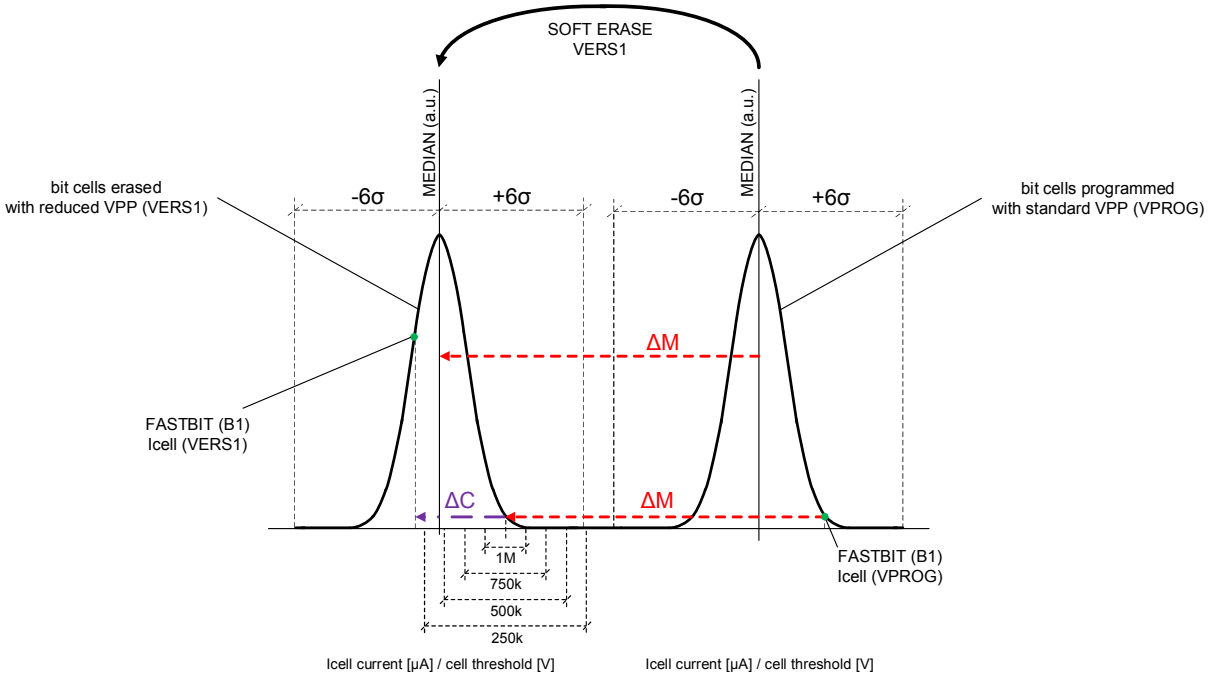



**Figure 193 - Soft erase screening based on 1 soft erase pulse**

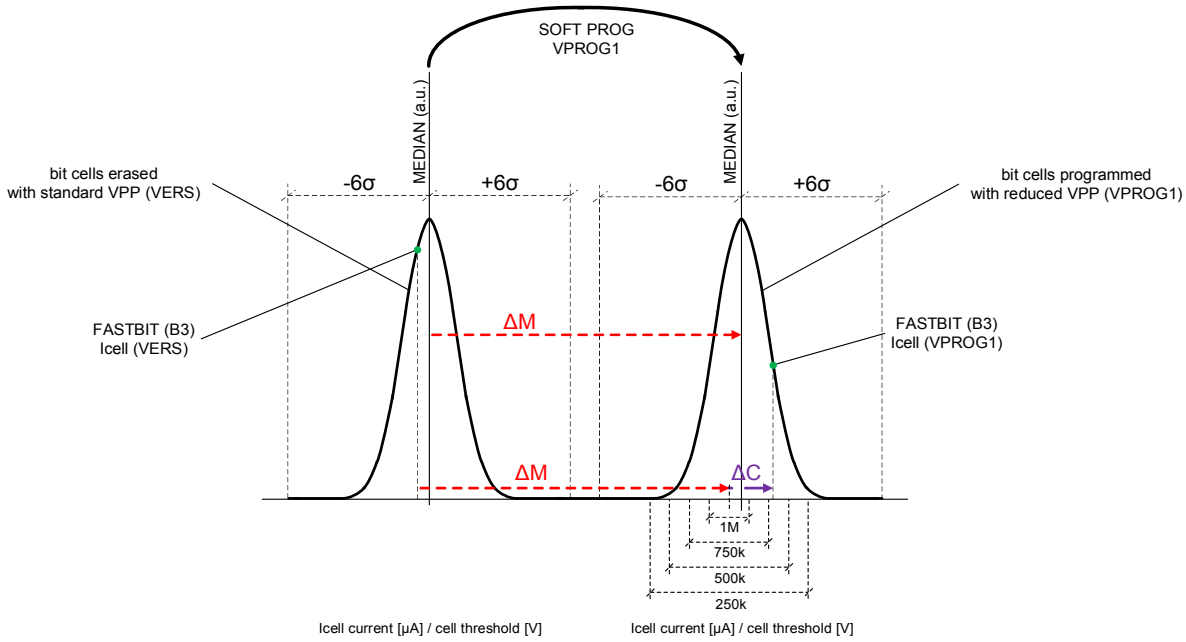



**Figure 194 - Soft program screening based on 1 soft program pulse**

Then centring the those distributions with respect to the bias current, as described in chapter 8.5.3, the bitcell can be subjected to it maximum specified write cycles or otherwise known as endurance cycles. If the program and erase bitcell current has not degraded such that a difference can still be seen between a program and erased bit then

the device is deemed to have passed the stress test or specified number of write cycles. By finding devices on the edge of the Gaussian distribution by using corner material and reduced program and erase voltages compared to what would be used in the application as described in 8.4.2, then a guard band can be established, as shown in Figure 149. This would guarantee, based on the Endurance cycling, that the bitcell is good and that the yield improvement is not a test escape even though it would have failed with previous test program revisions, this is otherwise known as over rejection.

However, the qualification of the yield improvement is not complete until the devices have been proven able to retain the programmed data for the specified retention time at a specified temperature. This qualification is done by baking a wafer that has had devices identified on the edge of the Gaussian distribution and have been endurance cycled. The bake is undertaken, as described in chapter 8.5.1, at an elevated temperature to accelerate aging process. The process of baking the wafers encourages the charge held within the bitcell to dissipate. If the bitcell cannot hold the charge held within the cell, for a specified time at a specified temperature then the data held will be corrupted and hence will not be readback correctly, therefore any comparison to data loaded into the EEPROM prior to the bake will fail readback.

Using corner material all of the above steps are whenever a change to such a test program is made to guarantee that no test escapes were generated in the any improvement exercise.

As described in chapter 8.5.3, a wafer lot was found on the edge of the process window with the characteristics of a bitcell with reduced bitcell current. Although the wafers were still in spec and meet specification, the lots were low yielding. Based on the above hypothesis, much of this yield should be recoverable with no test escapes. Based on the yield of these wafers, a yield increase from in the region of 15 % should be expected. Having implemented the centring of the bitcell current, as described in in chapter 8.5.3, a yield of 15 % was seen, Figure 188 and is in align from the defect density of the ams 0.35 um process, Figure 195. However, to prove that new test program produces devices that are specified to the actual number of Endurance loops, a lifetime test needs to be executed on a random set of parts that failed the original test program with the reference current set to a fixed value. Figure 189 shows 5 locations that were stressed to 60K endurance cycles that is the endurance specification plus a guardband of 10K cycles. All of these devices passed the Endurance test but they also

need to be shown to be able to retain the data for the specified 10 years lifetime, this was done by baking the wafers at 250 degrees for 24 hours and running the retention tests during the 2nd stage of the EEPROM test flow as described in chapter 8.5.1. All devices passed the retention tests and thus proved that the hypothesis was correct.

Figure 195 shows the defect density of the ams 0.35 um process from 2014 to 2017. This variation in the defect density not for discussion, however, what can be seen is the product yields is a direct effect due to this fluctuation in the defect density.

To be able to calculate the expected yield of a device based on the defects per $cm^2$, several statistical estimates can be used. The most simple is the seeds algorithm, using the algorithm for one of the EEPROM products give a predicted yield of 99.5 % based on the die size and number of defects. However, this is a simple model and does not take into account the number of masks used within EEPROM process. Using a Bose-Einstein algorithm estimates a yield of 96 % based on the die size and 0.05 of defects per $cm^2$. This would be the worst-case yield. By plotting the yield over a certain period, and comparing it to the defect density curve, if the test is capturing real defects then the yield should follow the same curve over a similar space of time. Observing Appendix 7, the average yield of an EEPROM product can be observed. What can be seen is that the same characteristic curve of the defect density variation is present over a similar period; it is not possible to fully align the charts as the test of a lot compared to when they are produced can vary due to processing times in the FAB and expected ship dates to the customer.

As there has been a full product qualification for each step performed, a full understanding of effect of defects on each process complexity has been attained. Because no RMA has ever been received on the ams 0.35 EEPROM technology dictates that there is 100 % test coverage for the EEPROM specification of 50000 write cycles, at 150 degrees, with a data specification retention of 10 years.

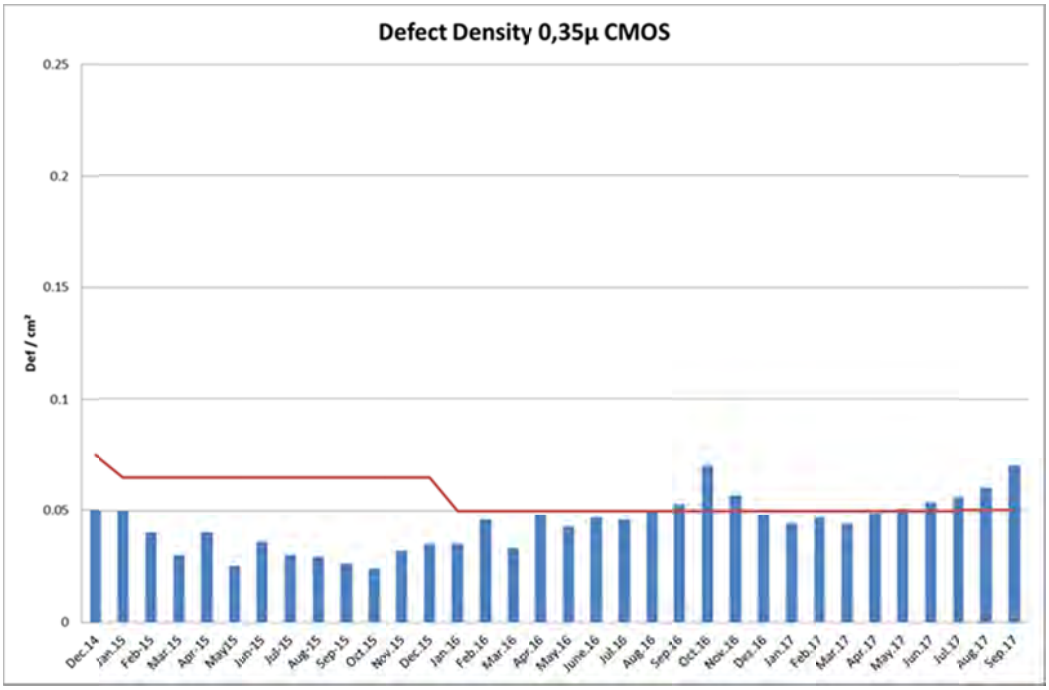



**Figure 195 – Defect Density of the ams 0.35 um CMOS process**

## 8.8 Conclusion

This chapter has shown that the program and erase threshold voltages of an EEPROM bit cell are very sensitive. As such the fast bit test methodology has to have significant attention to detail when both dealing with program and erase voltages that are used to detect potential oxide issues that would lead to an endurance failure. However to ensure a good yielding product detail has to also to paid to the reference at which the bit cell is set for the fast bit tests. A corner lot will certainly help to get a product into a safe state before starting production, however; only by running a product in high volume will the true corners of the foundry process be seen. By paying attention to the lots that do not comply with the norm and producing a thorough investigation into the root cause of maverick events, will ultimately drive the yield to the level of the expected defect density of the silicon foundry and avoid unnecessary yield loss, and at the end, increase the quality and reliability of the end-product. With such tools, a truly cost effective automotive quality grade EEPROM can be produced. To date, ams AG has shipped over 900 billion EEPROM bit cells with no returns received from any automotive customers. However, as with any process there are many ways things can be improved. It is just a case of thinking about the where the next feature or where improvements can be made.

# 9 THE FUTURE OF TEST OF LSI DEVICES – THE NEXT 10 YEARS

## 9.1 Introduction

The next 10 years will be an exciting time for the semiconductor business as a whole not just for the LSI business segment. Moore's law is ending for the digital segment and transistors in the lower geometries are now approaching the size of an atom. Where once corner wafers needed to be run to see the extremes of the process, with the new processes such as the 28nm, the whole process corner can be observed over one wafer lot. However, many of the past learnings from the use of the past main stream digital processes such as the 0.18um and 0.13um, that are becoming the defacto standard for analogue, can be leveraged for the future use of these processes in the LSI space.

One major development that is currently a focus within Universities around the world and that is gaining traction with integrated device manufacture (IDM)'s is the process of predicting wafer yields and quality related parameters based on the WAT data coming from the silicon foundry. Based on this data many parameters for the test of a product can be supplied in advance such that certain values would not need to be measured and hence a large test time saving gained from not having to make these measurements.

With the development of new standards such as the 1687.2 standard that was discussed in Chapter 2 it will be possible to extend the parts of the DFT of a chip such

that it looks like a tester instrument. This will allow test engineers to use lower costs testers that do not have expensive high end mixed signal instrumentation and use the tester software to configure the DFT to test the device in such a way that tester instrumentation currently does today. This standard will also allow ease of access to blocks that are not easily accessible today.

In the past, there was a concerted effort to replicate the IDDq of the digital world into the analogue world. However, in general this work failed due to process shifts causing unstable measurements to become meaningless and not having the tools to be able to allow easy access the individual device such that meaningful measurement of bias could be made. However, recent efforts by both the author and NXP semiconductors have produced good results that allow correlating simple current tests of separate blocks of a device to actual specification based parameters. In combination with dynamic part averaging (DPAT) techniques, this methodology will enable screening such that only centred devices will be shipped to the customer in the future. Hence ensuring any suspicious outliers will be removed from the device population thus resulting in future quality levels approaching the sub DPPB level.

A controversial topic that has gathered much attention in recent years is the discipline of actually modelling how a defect manifests itself. In the digital industry, this work was completed 30 years ago and the mass adoption of SCAN has ensured the quality levels shipped to the field of digital circuits are in the 99.99% region. Unfortunately, this was not the case for the analogue industry and up until November 2016 was the first commercially available analogue fault simulator available to the market. However, this tool, DefectSim, from Mentor Graphics, uses non-formally agreed fault models to simulate the test coverage. By using this type of tool allows the user to see where the test coverage is lacking and allows the user to use other tools such as the 1687.2 standard to design new test access to the device to allow testing of the block such that the user can reach 100% test coverage. Thus, in the future much more auto generated test scripts will be available allowing reuse of test blocks from device to device and as such allow ease of porting code from one tester to another.

The development standard 2427, is addressing the need of standardising analogue fault models. The fault models will be used to calculate the test coverage of a silicon circuit design in the future based on the test program used to test it in production. Used with tools such as DefectSim, the industry will have a common framework to work with

so that a customer will be able to compare one companies test coverage with another, with the knowledge that the same baseline is used for both. Therefore, in the future, it will not be possible to say that a company has 100 % test coverage as 100 % of the specifications are tested as it is in today's market place.

Also today, for the automotive industry, there is the requirement to carry out an ISO26262 circuit level FMEA; this is due to safety concerns and the ability to screen defects that might go undetected due to device issues not being seen unless they are in the system. ISO26262 gives a framework to do an FMEA of the device within the system and gives the user the knowledge to overcome these issues during the screening process. As the volume of devices in the consumer space gets ever increasing larger, there will be a push, in the future, for an ISO26262 style FMEA for consumer products. However, instead of being from the safety aspect, but from the point of view of having to rework the possibility of 1000's end products due to infant mortality and test escapes. Having a quality level of 1ppm for a product that ships 1 billion parts per year would mean having to rework and verify 1000 units returned from the field which clearly would be unaffordable and unacceptable to the for the end manufacture and customer of the product.

## 9.2 New Quality Related Test Procedures – Learning From digital

### 9.2.1 High Voltage Stress

One way the digital world increased the test coverage and removed a large portion of infant mortality defects was by the use of stress in the test program. By having a system such as SCAN that allowed all the flip flops to be exercised allowed the use of voltage stress to strain weak transistors such that a device that would fail in the field would fail at the production stage. Two stress schemes have been used for some time to strain devices during the production phase, [77], one being the DVS (Dynamic Voltage Stress) in Figure 196, and the EVS (Enhanced Voltage Stress) in Figure 197.

For the analogue world, there has not been any major development in this area. However the author has been active in this area of research and believes the following will become commonplace in the future so as to reduce the infant mortality rate in the field.

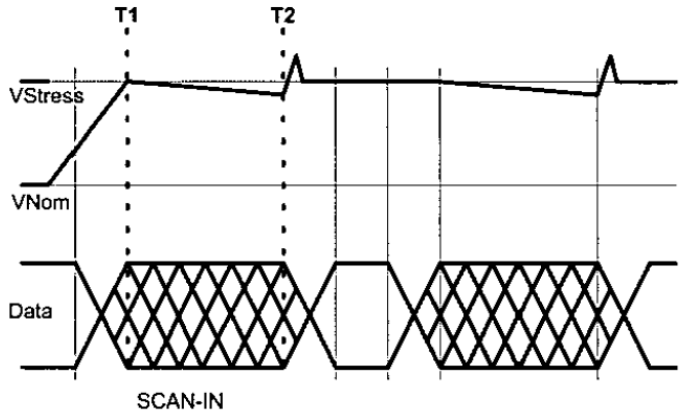



**Figure 196 – Dynamic high voltage stress digital timing diagram**

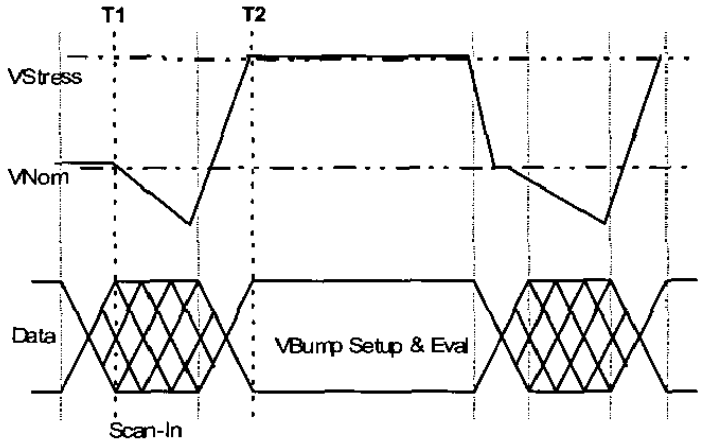



**Figure 197 – Enhanced high voltage stress digital timing diagram**

In the future, for LSI devices, to drive down the infant mortality rate, an analogue test program will encompass a stress section of the program to remove infant mortality fails. The big difference between a digital and an analogue test program is that digital is used to program the function of analogue transistors, or to setup individual analogue blocks of the device. Therefore, a way is needed to stress each block equally in a known or standardised way. If one considers the way that a device register map is configured for an analogue device, it can be seen that a matrix of addresses and register values are common throughout the world. Therefore standardising a stress program for the future is relatively easy. If for every even nibble of a register map 16#A and16#1 the odd is programmed then each adjacent block will stressed. If the inverse is then done such that the even nibble is 16#01 and 16#A for the odd then each block has been stressed equally and adjacent blocks have stressed each other. This pattern then can be used with

the voltage stress developed by the digital world 20 years ago. We will call this pattern the analogue infant mortality inducing pattern or AIMI pattern.

Using the methodology developed by the digital industry, [77], we can now produce two stress tests for the LSI devices that will become commonplace in the future.

#### 9.2.1.1 Dynamic Voltage Stress

Using the same idea as is in Figure 196, a voltage of $1.3xVDD_{max}$ would be set, the AIMI pattern runs with standard digital levels i.e. 0 $\rightarrow VDD_{max}$, then VDD reduced back to VDD. Each transistor would then be dynamically stressed at VDD = $1.3xVDD_{max}$, as the transistor switch on and off.

#### 9.2.1.2 Enhanced Voltage Stress

Using the same idea as in Figure 197, the 1st segment of the AIMI pattern would be run with VDD = $VDD_{max}$, the voltage would then be set to VDD = $1.8xVDD_{max}$, a static wait time would then be executed while the device was stressed under the static conditions. VDD would be returned to $VDD_{max}$ and the 2nd segment of the AIMI pattern executed. VDD would be returned to VDD = $1.8xVDD_{max}$ and the device stressed as second time for a static time under 2nd segment conditions.

This would lead to a static stress of all the oxides without over stressing the gate.

## 9.3 Dynamic Part Average Testing

Referring to Figure 198, it can be seen how the AEC - Q001 Rev-D standard defines how to dynamically set limits as a lot is being tester so that outliers can be removed. This ensures that any devices that do not fit within the distribution of the population of the devices are removed, as they are different for some small reason.

Using this process with the IDDq analogue measurement technique defined in 2.3.14 a very powerful approach to defect oriented test can be developed. If all the biases of a chip are seen to be good and only devices that are selected to be shipped to the customer have the same distribution with the device population, then it can be stated that the under lying quality of the part is of the highest order possible.

Referring to Figure 199, it can be seen on the left hand side of the wafer that there is an area shown where there are issues. The dies that have no black outline are devices that would have failed the original test program with the static limits. By the use of DPAT, [78],

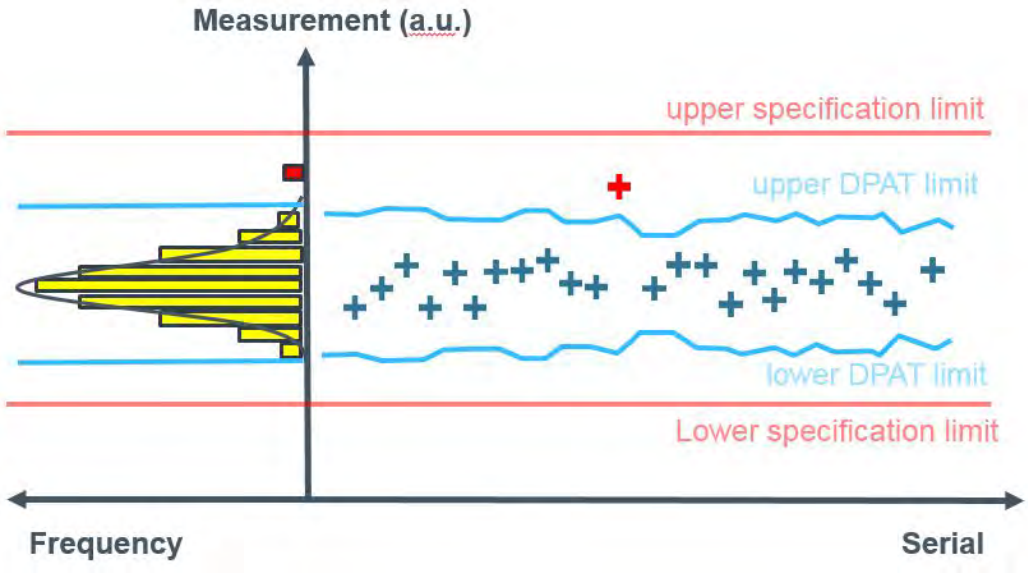



**Figure 198 – Diagram of how DPAT dynamically changes test limits to remove outliers**

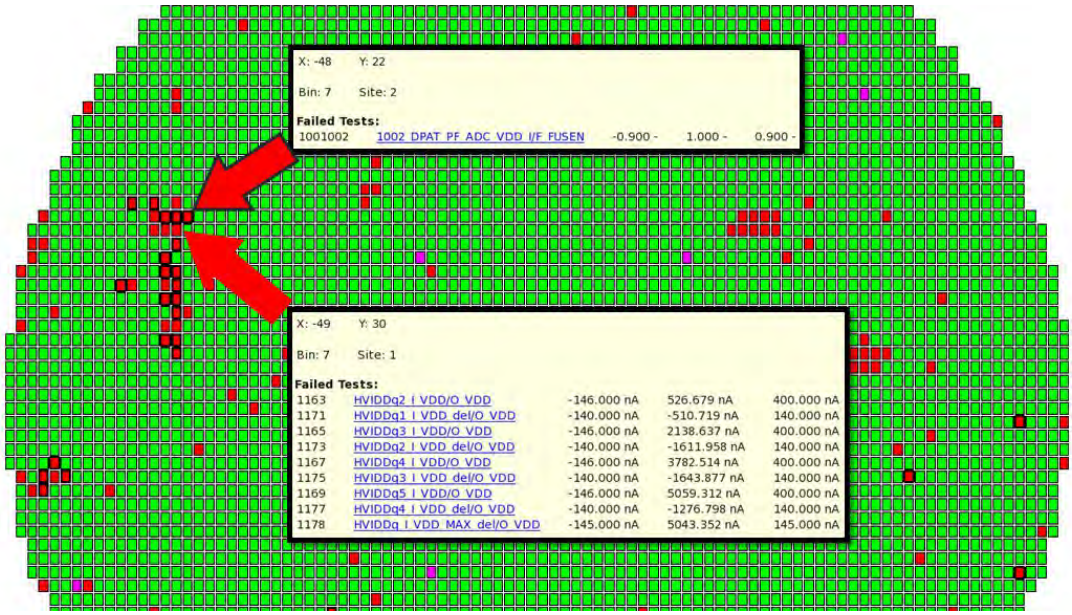



**Figure 199 – Power of DPAT to remove suspicious outliers**

the dies marked in black were shown to be abnormal. With the addition on the newly detected die, it can be seen that there is a clear stripe. This was a known issue within the silicon foundry. However, only if the issue were large enough would it have been detected at wafer sort. This is due to the need to have wide limits to take into account of process variation. By using DPAT, [78], these problem areas were identified easily and avoided shipping potential infant mortality failures into the field, thus saving the end customer much time consuming work in the future if the parts were to be returned.

## 9.4 Using Scribe Line Monitoring Data for Test

In the semiconductor market, there is a term called big data, in the authors opinion this should be called data cemetery. There is so much data available to engineers that simply goes un-utilized. Take for instance the wafer acceptance test data of the silicon foundry, here there are process monitors that are placed in the scribe lines of a wafer and these process monitors are measured to guarantee that the wafer meets a certain specification and that the wafers have been produced correctly. The potential uses of this data are astronomical but only until recently has this area of research been explored.

There are two areas that this research will affect the future of LSI devices.

### 9.4.1 WAT Data to Quality Related Parameter Correlation

Referring to Figure 200, one can see a set of wafer maps, [79,80], one is a reconstructed wafer map of the WAT as though it was measured per die and the other is gradient of a test parameter that has a spatial pattern correlation to the WAT parameter(s). As the listed WAT parameters and wafer sort items are all quality related metrics, if a correlation exists then a method for detecting quality related silicon foundry excursions could be detected prior to wafer sort. This would result in the ability to pre-judge if a more extensive test program is needed to test those wafers or just to scrap the lot before it even reaches test, thus enabling a much faster turnaround on problem lots from the silicon foundry. This would require an extra process step before the wafers are sent to test, however this type of system could easily be handled by software that is already installed at silicon foundries and will be available in future systems.

| Match | WS Item(s) Name | Pattern | WAT Parameters Name | Pattern | Similarity $\gamma$ | Comprehensible Correlations* |
|---|---|---|---|---|---|---|
| 1 | VoltslstErase, TailErase/ProgramFunc, TailFunctionProg, BasicFunctional, EnduranceRead, RetentionRead/ICell |  | CWET, CW2, CW2I, CW1, CW1MEF, TWC, JET1, QPMZ1, QPMZ2, QPMZ3 |  | 0.78 | High |
| 2 | SATrip_ICellD0, SATrip_ICellD1, ..., SATrip_ICellD15 |  | CWET, CW2, CW2I, CW1, CW1MEF, TWC, JET1, QPMZ1, QPMZ2, QPMZ3 |  | 0.75 | Median |
| 3 | VoltslstProgram |  | TMFBL1, WU1, WU2 |  | 0.58 | Low |
| 4 | VoltslstProgram |  | CWHPY1, CWHPY2, HBNNB, JET2, LQ1, LQ2, UDP1, UPD2 |  | 0.53 | High |
| 5 | MarginHighIRef, MarginLowIRef |  | SEJG1, SEJG2, SEJG3, XFGG1, XFGG2 |  | 0.47 | None |

**Figure 200 – WAT to Wafer Sort data correlation**

### 9.4.2 WAT Data to Yield Correlation

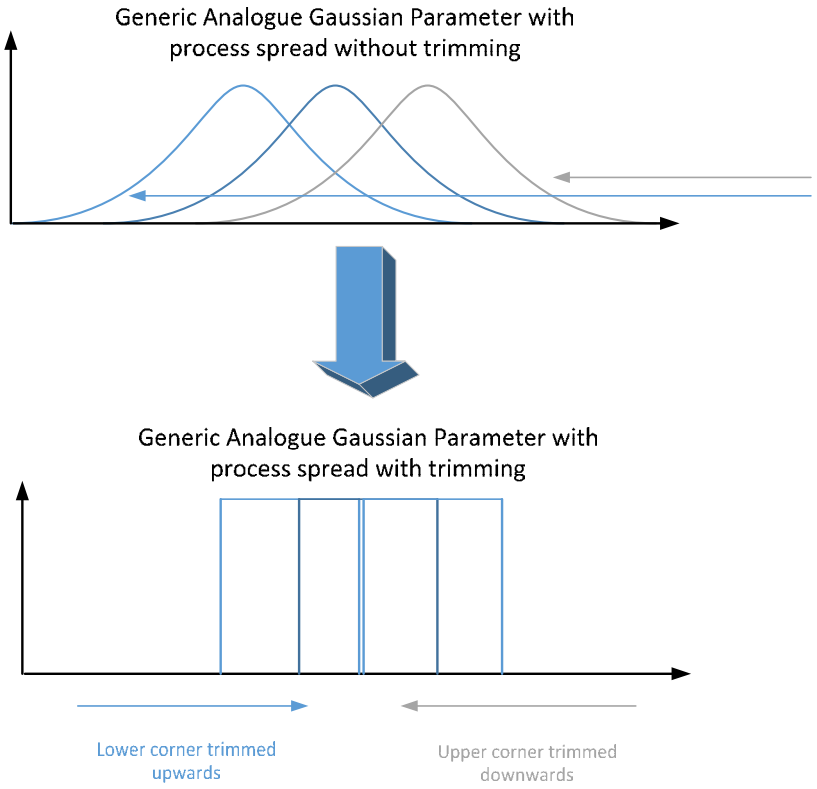
Using the same process as described in 9.4.1, but using all the test parameters and yield as a training. The WAT test data can be used to predict the yield of a product, [82]. This has several consequences. As a yield value based on actual semiconductor physical data and correlation has been performed, the yield does not match at the wafer sort or final test stage an immediate issue can be raised. Either being a quality related indictor or yield simply being lower than expected. Either way, risks in delivering bad product or the possibility of having to little yield compared to the expected and hence reduction in revenue can be avoided. These tools will be available in the future in existing software suites already used by semiconductor professionals.

### 9.4.3 WAT Data to Analogue Trimming Parameter Correlation

As LSI devices become more complex and have to fulfill increasing extraneous specifications, more need for trimming is becoming necessary. 10 years ago maybe the bandgap voltage was trimmed whereas today there can be over 20 parameters being trimmed, ranging from oscillator frequencies to reference currents and voltages. This is becoming ever more important for such applications as high resolution, low voltage range ADC's as the error in the reference can easily be larger than the resolution step of the ADC and hence if the device were not trimmed accurately this would result in the ADC not meeting specification.

Any analogue trim parameter will have the standard Gaussian distribution, that will change over the process corners (see Figure 201). By trimming the analogue parameter, the edges of the Gaussian distribution are moved into the center and hence a rectangular distribution is produced. Depending on where on the process curve the device was produced will result in an average trim code (see Figure 201), for the device. Therefore, each silicon lot of material produced would have a different average trim code requirement that would depend on the distribution. In today's trim routines, at best, a binary or at worst, a linear search from min to max, would be executed. As each iteration requires an analogue measurement and a settling time, the amount of test time for an LSI product today can be as much as 80 % of the test time.

In the future, it will be possible, by having the WAT data being fed directly into the test program, to estimate the starting trim codes, [81], for each silicon lot such that a




For a centered process most devices would not need to be trimmed. Trimming is needed to bring in the edges of the Gaussian into the centre of the distribution. However most test programs start from a fixed point and make iterations to find the best trim code

By using WAT data the centre of the distribution for analogue parameters for the lot can be estimated and a pre-judged trim code starting point can be calculated for each parameter. Hence large test time savings are possible due to not having to cycle through all trim code possibilities

**Figure 201 – WAT to analogue parameter estimation for predicting optimum trim codes**

maximum of two iterations will be needed to find the best trim code for each individual lot, therefore potentially reducing the test time by as much 60 %.

In addition, in combination with Section 2.3.14, in the future, trimming will probably be the only specification based metric that is tested in production as all other parameters will be tested based on a defect-based test of the bias current and voltage of the block.

## 9.5 1687.2 – Future of DFT and Test Access of an LSI device

As described in Chapter 2, a new standard is being formed that will tackle the standardization of test access and control of design for test (DFT) circuitry within a device. By having the ability to program the control of devices functions over the SCAN interface a hierarchal approach can be taken. This would allow canned methods to be developed one time for each block, and hence allow reuse of each test code segment so that it can be ported from test program to test program, and even tester platform to tester platform. In addition, as the test instruments function in the device can be described, drivers can be written such that a mixed signal block of an LSI device can

be made to look like an instrument of an LSI tester. This would allow the tester to talk to the device with a set of digital patterns as it is today. The state of the art would be advanced by measuring the on chip response with its own instrument but rather it looking like the tester measured the response so that all the standard logging and DPAT routines can be executed as they are today so that quality is maintained. This is very beneficial as using an on chip DAC or ADC prevents distortions or corruption of a signal as it is brought into or out of the device to the tester. Therefore maintaining the integrity of the signal, and potentially avoiding problematic guardbands and correlation errors due to loadboard parasitics. This would be achieved with a standardized test bus that would be inserted to an LSI, as a SCAN chain is inserted today. See Figure 202. This standard allows extending a tester's programming language to control embedded instruments described by IEEE 1687 [83], in the same way that the programming language controls the tester's instrumentation. For example, consider a DUT (an IC, or an IC on a board) that contains a DAC accessible via IEEE 1687. A test engineer of the future would write a simple program on the tester that reads the ICL and PDL for the DAC, and set the DAC's input to a given digital code based on a requested voltage. The test program would look like:

```
VI.ForceVoltage("MyEmbeddedDAC",3V);
VI.Connect("AnalogTestPin");
VI.SourceCurrent("AnalogTestPin",-100uA);
MultisiteFloat meas = VI.MeasureVoltage("AnalogTestPin",RANGE_4V);
```

**`MyEmbeddedDac`** represents a core that is accessed by IEEE 1687, and **`AnalogTestPin`** is a tester V/I source-and-measure instrument that can be connected to a DUT pin through the interface board. The requested 3 V output would be translated into a DAC input code by the software.

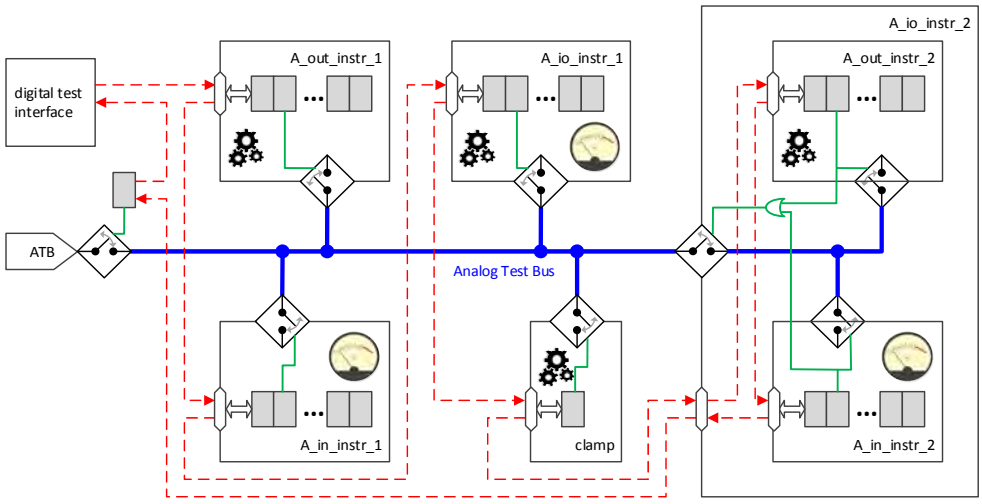



**Figure 202 – 1687.2 A standardised analogue test bus insertion**

In the future, this will allow a seamless transition in a test program between the embedded DAC sourcing a voltage inside the device and a measurement by a tester instrument on the outside of the device.

## 9.6 2427 - A Standard for Analogue Defect Modelling and Simulation

This proposed standard will define simulation models of manufacturing defects that have been observed within integrated circuits (ICs). All defects should be detected by manufacturing tests of analogue, digital, and mixed-signal circuits, but the portion of all possible defects that can be detected, or "covered", in practice depends on many factors, which this standard will consider. This standard will consider redundancy, since most analogue circuits have redundancy and defect tolerance, intentional or not. This standard will not consider combinations of variations that could result in a circuit failing to meet all its specifications.

This standard will focus on defects in analogue functions [84]. In this context, an "analogue function" means a function that has input, internal, or output signals with meaningful values in a defined continuous range, and the function has at least one tested performance that is sufficiently non-deterministic that its test has upper or lower limits.

The primary purpose of this standard will be to allow people to communicate information about defect coverage in a way that allows prediction of costs and the need for action. This standard will also guide the simulation of defects to ensure that the defect models achieve an optimum trade-off between cost-effective simulation times and modelling defect behaviour seen in real circuits.

This standard will also support defect-oriented testing (DOT), which means applying tests, specification-based or alternatives, to achieve a target coverage of a set of potential defects. Applying a sine wave to test total harmonic distortion (THD) is an example specification-based test that can achieve very high defect coverage of circuitry through which the signal passes. Applying DC or a square wave to a filter is an example of an alternative test that might achieve lower coverage than a sine wave for a single path, but which might be applied more cheaply to more paths simultaneously to achieve higher defect coverage of a circuit containing many signal paths.

A second purpose of this standard will be able to facilitate a more efficient simulation of defects in analogue circuits. Simulation time with sufficient accuracy is the major

challenge in analogue simulation, and estimating analogue defect coverage increases the challenge because there may be thousands or millions of potential defects to consider in an analogue circuit within an IC.

A third purpose of this standard will be to allow estimation of defective parts per million (DPPM) to facilitate trade-offs between cost of test, time to market, and quality. The portion of all potential defects that tests must detect in a circuit will depend on the likelihood of the defects occurring, the consequences of undetected defects in the intended application for the IC, and the likelihood that a defect that occurs has a costly consequence. It is also possible that some defects do not affect a circuit's datasheet specifications but might later do so, and therefore they pose a reliability risk; this standard will consider coverage of these defects also.

A fourth purpose of this standard will be to facilitate better design for test (DFT) and test generation. Many DFT and test techniques have been developed which, despite cost advantages, are not used in practice because their impact on defect coverage is questioned due to the lack of a well-understood and silicon-corroborated analogue defect models. By allowing reliable comparisons of DFT and test techniques, automation and quality improvements will becomes much more feasible with this standard in the future.

By using 2427 in combination with 1687.2, it will become possible in the future to produce a tool that will be able to automate analogue test insertion and measure the resulting test coverage, such as we have in today's marketplace with SCAN based digital tests. Therefore, this will eventually become known as Analogue SCAN.

## 9.7 ISO26262 – Functional Safety and Future of LSI devices

Functional safety features form an integral part of each automotive product development phase, ranging from the specification, to design, implementation, integration, verification, validation, and production release. The standard ISO 26262 is an adaptation of the functional safety standard IEC 61508 for automotive electric/electronic systems. ISO 26262 defines functional safety for automotive equipment applicable throughout the lifecycle of all automotive electronic and electrical safety-related systems.

Like its parent standard, IEC 61508, ISO 26262 is a risk-based safety standard, where the risk of hazardous operational situations is qualitatively assessed and safety measures

are defined to avoid or control systematic failures and to detect or control random hardware failures, or mitigate their effects.

In summary for LSI devices, this means that a FMEA should be performed from an application point of view to see the impact of how a defect on the application board would affect the LSI device and how a defect in the LSI device would affect the application. It is entirely possible that a defect with in an LSI device would not become detectable unless the chip was used in the application, [85]. This standard should help identify those risks and help designers build in better DFT or add more functional safety systems to a device to mitigate those risks.

Currently, FMEA's are done very rarely for consumer devices, and if done they are executed without very much attention to detail as the consequences are not so dramatic if there were to be an issue. However, at the time of writing, a paradigm shift is being seen in the consumer electronics space. High volume customers are starting to request larger qualification lots than that of automotive customers. This is a result of needing to be confident that a device that will go into high volume production will be producible across all process corners and that any major defects from these runs detectable during the qualification. This is in fact manual version what ISO26262 is trying to achieve. Therefore, in the future, terms such as high reliability, automotive etc will cease to exist, as all semiconductors manufactured will need to be made to the same exacting standards, be it from a safety point of view or from a security of supply point of view. At the end, the two requirements are the same in that the end users requires products with a defect density equal to zero i.e. dppb $< 0$ or defect tolerance to be a part of all device functions of the future.

# 10 FUTURE WORK

## 10.1 EEPROM

Although the author has a strong history in analogue/mixed signal test, his recent works have been concentrated mainly in the reliability and quality of EEPROM technologies. With the authors first patent application filling being successful will allow the author to discuss these advancements in testing of EEPROM's much more easily.

The basis of the patent was to protect the intellectual property of the test solution developed to scrutinise the individual bit cell currents of an EEPROM. Using this data, a set of tests can be developed to check the program and erase currents at different voltages. Analysing the stored current within the bit cell and the change between the two program and erase voltages used, one can build a picture of the silicon impurities that will be crucial to the operating lifetime, and reliability of each bit cell within the EEPROM.

As with every research undertaking, a mountain of data was produced to verify the hypotheses that were generated in this research undertaking. As usual, the research findings were not as one would have totally expected and as a result created more questions than were originally asked.

As part of this research, the resulting findings will go some way into predicting wear out mechanisms of semiconductor devices. We believe that by endurance cycling an EEPROM, the extended bias conditions of a semiconductor device are replicated but in a much more tightly controlled and predictable manner. As such, we believe will be able to create an Analogue NBTI model using this data. This will hopefully, extend the state of the art as research into wear out mechanisms and NBTI have so far been limited to digital circuits in advanced nodes. There are many analogue nodes, 0.13um, 0.18um,

0.35um, which also require this research, as automotive analogue technologies use these process nodes extensively.

Also, as well as performing extended endurance cycling of these EEPROM's, experiments with retention bakes were also made. The standard spec for an automotive EEPROM is to retain data for 10 years at 150 degree and as part of the test solution, a bake a rescreen should be performed to screen out any retention bake failures. What was found was possibly a way to identify retention failures without doing a bake of the silicon.

This will form the largest part of the author's future work in the coming years and as a result hopefully the most reliable and best performing EEPROM available to the automotive market.

## 10.2 Analogue Mixed/Signal

The author would like to try and experiment with a phase switched discrete stepped chirp, the creation of such a signal is theoretically possible, however, one potential obstacle could be the maximum waveform size that one can load into the AWG at any one time. This could be an interesting work although the use of such a waveform is yet to be determined.

The ams AG 0.35 um optical process is receiving much attention due to the excellent performance it exhibits. Due to this, there is much interest in this process as a semiconductor full service foundry. As such, there is much interest in pulsed current tests. Unfortunately, the basic DC source from LTX is not so easy to use when performing pulsed current tests and as such, the author is looking into using a standard DC source to develop a loadboard instrument solution.

An extension of this is to investigate current steering DAC's so that an onboard AWG could be created that can deliver sinusoidal current into virtual ground so that it can be used for device testing. Another solution that is envisaged is to use high performance of the tester AWG and use one of the voltage current converters that are currently offered. Both solutions have drawbacks, the DAC solution has current glitches that results in a non-continuous waveform and the voltage to current converter although offers a smoother waveform is restricted to milliamp ranges when microamps are required.

## 10.3 Design for Test

Although the author has extensive experience in analogue/mixed signal test, he does not believe his future is restricted to this discipline and as such wants to expand significantly into the DFT field. As discussed in Chapter 2.7, the author is the leader and founding member of an industrial working group focusing on revolutionising the analogue test world.

Apart of the author's future work will be to deliver an extension to IEEE 1687 standard that will address analogue test access to a device. A huge part of this work is to develop the Analog PDL description language to mimic what a tester would do at device level to be able to use a devices infrastructure to source stimuli to the device from the device and/or preform measurements of the devices response to such on chip stimuli. As such, a paper was presented at ITC17 on a topic that details four use cases where the new standard would be applied to the test of analogue/mixed signal devices, [86].

As a founding member of the working group that is producing the extension to 1687 standard, one of the self-defined tasks is to develop a common framework for defining analogue fault models, 2427. One of the issues that have been faced by industry was the lack of models and a generic process that had corners models to allow simulation of a circuit without the need of NDA's with silicon foundries. This has restricted the sharing of data related to defect simulation and test coverage. The author has gained permission to share the basic 0.35 um CMOS process of ams AG with a set of benchmark circuits. In conjunction with Mentor Graphics, we have produced the world's 1st set of benchmark with a process model that can be used to simulate a circuit across corners such that meaningful test coverage can estimated. As such, a paper was presented at ITC17 that details the background of the work, and where the work package can be found, such that an individual or organisation (academic or industrial) can use it, [87].

The author will use and continue to develop these new standards as part of his work to bring testing of analogue devices into the 21st Century. This will drive defect rate to less and ppb and ensure the safe operation of automotive devices.

# 11 CONCLUSION

Although the semiconductor industry is at a junction and there is a paradigm shift occurring in the business, there is no reason that the semiconductor test development personnel cannot go the extra mile and deliver test solutions that are leading edge but also deliver the required quality, reliability, and throughput of a semiconductor in the 21st Century.

The demonstrated solutions to five problems experienced in the past were developed with simple common sense approaches that required a simple understanding or appreciation of the issues at hand. If the author was content, with having lower quality, reliability and higher test times then certainly these issues would not have been investigated and been resolved but at a serious risk of losing the business that came with it. However, it can also be noted that the author has seen an increased interest from current and potential customers of making test solutions with the authors test team because of the research he and his team are making in the area of mixed signal test development.

There is certainly much more work to be done in this area but this will take another paradigm shift in the industry in that all the elements of producing a semiconductor will work together as a whole rather than several separate entities. However, beforehand, the ATE industry could do much more to produce simple methods that can be re-used, as the author has done himself, but available to the whole customer base of the ATE provider. This would ensure that the most cost effective and reliable solutions are available to the collective. It would also be of much benefit to the semiconductor industry, maybe not the ATE industry, if the ATE industry would invest more effort into making current hardware solutions better to the customer base however by using simple software solution. This would affect the ATE industries profit and loss, as its

customer base would need to buy less of the ATE industries product. A solution to this could be a licensing option that gave the user access to special modes such as phase switching enabling better performance of its lower-end products.

For now, with the combination of all of the described approaches a higher yield, better quality, and reliability, faster test times and at the same time as guaranteeing harder specifications with lower end instruments are feasible.

As all with all development in this world, only rarely are big state of the art developments made by a single individual. This can be said for this entire dissertation, ideas come from all aspects of life, some even take decades to mature before they are even acted on or even for the opportunity to arise for those ideas to even to be tried tested. As this dissertation is built from several papers that the author has written with other colleagues it was worth to split out the individual contributions from the author, this has been addressed in Appendix 8.

# 12 REFERENCES

[1] Peter Sarson, "ACR BER correlation to ATE for a COFDM VHF RX", 20th International Mixed-Signal Testing Workshop, pp.1 - 4, Paris, France (2015).

[2] Peter Sarson, "IDDq Analogue," Emerging Test Strategies Session: presented at European Test Symp., May 2014.

[3] Peter Sarson, Automotive Session: "Test Development : Getting to 0ppm," presented at VLSI Test Symp., April 2016.

[4] Peter Sarson ,Constantinos Xanthopoulos, "Automotive Alternative Test", presented at VLSI Test Symp., April 2017.

[5] Salvador Mir, "Mixed-signal/RF design-for-test:principles and advanced techniques", 2014 19th IEEE European Test Symposium (ETS), Tutorial A.

[6] N. B. Carvalho and J. C. Pedro, "Multi-Tone Intermodulation Distortion of 3rd Order Microwave Circuits", 1999 IEEE Int. Microwave Theory and Tech.Symp. Digest, pp.763-766, Anaheim, Jun. 1999

[7] Leticia Aladren; Paloma Garcia-Ducar; Jesus de Mingo; Cesar Sanchez-Perez; Pedro Luis Carro ,"Behavioral Power Amplifier Modeling and Digital Predistorter Design with a Chirp Excitation Signal", 2011 IEEE 73rd Vehicular Technology Conference (VTC Spring), Pages: 1 – 5, 2011

[8] S. Boyd,"Multitone Signals with Low Crest Factor", IEEE Trans. Circuits Syst., vol. CAS-33, no. 10, pp.1018-1022, Oct. 1986.

[9] M. R. Schroeder, "Synthesis of Low-Peak-Factor Signals and inary Sequences With Low Autocorrelation", IEEE Trans. Information theory, pp.85-89, 1970.

[10] B. R. Petersen, D. D. Falconer, "Suppression of adjacent-channel, cochannel, and intersymbol interference by equalizers and linear combiners" Communications, IEEE Transactions on, Year: 1994, Volume: 42, Issue: 12 Pages: 3109 – 3118

[11] W. Allen, D. Bailey, S. Demidenko, V. Piuri, "Test chirp signal generation using spectral warping," Electronic Design, Test and Applications, 2002. The First IEEE International Workshop on

[12] A. Mishra, M. Soma, “A Time-Domain Method for Pseudo-Spectral Characterization” VLSI Test Symposium, 2008. VTS 2008. 26th IEEE

[13] H. Kitayoshi, S. Sumida, K. Shirakawa, S. Takeshita, “DSP Synthesized Signal Source for Analog Testing Stimulus and New Test Method.”, Proceedings of the IEEE International Test Conference, pp.435-440, 1985.

[14] Swen Müller, Paulo Massarani, “Transfer-Function Measurement with Sweeps”, Journal of the Audio Engineering Society, Jun 2001, Vol 49, No 6, pp 443-471.

[15] D.G.Long, “Phase unwrapping for multidimensional rational and finite-length sequences” Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on

[16] A. V. Oppenheim, R. W. Schafer and J. Buck., Digital Signal Processing, Prentice-Hall, 1999.

[17] P. Sarson, “RF Filter Characterization using a chirp” Design & Test Symposium (IDT), 2014 9th International

[18] X. G. Xia, “Discrete chirp-Fourier transform and its application to chirp rate estimation”, Signal Processing, IEEE Transactions on, Year: 2000, Volume: 48, Issue: 11 Pages 3122 – 3133

[19] R. Xuan., T. Haihong, X. Jian, S.Jia, L.Weiping, “Long-time coherent integration detection of weak manoeuvring target via integration algorithm, improved axis rotation discrete chirp-Fourier transform”, IET Radar, Sonar & Navigation, Year: 2015, Volume: 9, Issue: 7 Pages: 917 – 926

[20] P. Sarson, “Group delay filter measurement using a chirp”, 2016 21th IEEE European Test Symposium (ETS) Pages: 1 - 2

[21] W. P. M. Allen, D. G. Bailey, S. N. Demidenko, V. Piuri, “Analysis and application of digital spectral warping in analog and mixed-signal testing” , Reliability, IEEE Transaction on, Volume 52, Issue 4 Pages 444 – 45747

[22] W. Allen, D. Bailey , S. Demidenko, V. Piuri, “Test chirp signal generation using spectral warping”, Electronic Design, Test and Applications, 2002. Proceedings. The First IEEE International Workshop on

[23] S. Demidenkol, V. Piuri, “Using spectral warping for instrumentation and measurements in mixed-signal testing” , Instrumentation and Measurement Technology Conference, 1999. IMTC/99. Proceedings of the 16th IEEE, Volume 3 Pages 1547 – 1552

[24] Nobuharu Aoshima, “Computer-generated pulse signal applied for sound measurement”, The Journal of the Acoustical Society of America, Volume 69, Issue 5, 1981.

[25] C.J. Clark; A.A. Moulthrop; M.S. Muha; “Transmission response measurements of frequency-translating devices using a vector network analyzer”, IEEE Transactions on Microwave Theory and Techniques, ( Volume: 44, Issue: 12, Dec 1996 )

[26] Joel Dunsmore, "A new calibration method for mixer delay measurements that requires no calibration mixer", Microwave Conference (EuMC), 2011 41st European

[27] Jonathan Scott ; Michael Hoy ," Group-Delay Measurement of Frequency-Converting Devices Using a Comb Generator ", IEEE Transactions on Instrumentation and Measurement ( Volume: 59, Issue: 11, Nov. 2010 )

[28] Thilo Bednorz, "Group delay and phase measurements at converters and multistage converters without LO access", Microwaves, Communications, Antennas and Electronics Systems, 2009. COMCAS 2009. IEEE International Conference on

[29] P. Sarson, "Test Time Efficient Group Delay Filter Characterization Technique Using a Discrete Chirped Excitation Signal", 2016 IEEE International Test Conference (ITC) ,Year: 2016, Pages: 1 – 6

[30] Peter Sarson, "An ATE Filter Characterization ToolKit Using a Discrete Chirped Excitation Signal as Stimulus", Journal of Electronic Testing: Theory and Applications, Springer (2017), DOI : 10.1007/s10836-016-5633-x

[31] B Kanmani, "The 'Phase-Reversal' in DSB-SC: A Comment", Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th.

[32] Mini Circuits BPF-B48+ Datasheet , https://ww2.minicircuits.com/pdfs/BPF-B48+.pdf

[33] K. Wakabayashi, K.Kato, T. Yamada, O. Kobayashi, H. Kobayashi, F. Abe, K.Niitsu "Low-Distortion Sinewave Generation Method," J Electron Test 28:641–651 (Oct. 2012).

[34] F. Abe, Y. Kobayashi, K. Sawada, K. Kato, O. Kobayashi, H. Kobayashi, "Low-Distortion Signal Generation for ADC Testing", IEEE International Test Conference, Seattle, WA (Oct. 2014).

[35] K. Wakabayashi, T. Yamada, S. Uemori. O. Kobayashi, H. Kobayashi, K. Niitsu, H. Miyashita, S. Kishigami, K. Rikino, Y. Yano, T. Gake, "Low-Distortion Single-Tone and Two-Tone Sinewave Generation Algorithms Using an Arbitrary Waveform Generator", IEEE International Mixed-Signals, Sensors and Systems Workshop, Santa Barbara, CA (May 2011).

[36] M. J. Barragan, G. Leger, D. Vazquez, and A. Rueda, "On-chip Sinusoidal Signal Generation with Harmonic Cancelation for Analog and Mixed-signal BIST Applications," Analog Integrated Circuit and Signal Processing, vol. 82, no. 1, pp. 67–79 (Jan. 2015).

[37] M. M. Elsayed and E. Sanchez-Sinencio, "A Low THD, Low Power, High Output-Swing Time-Mode-Based Tunable Oscillator Via Digital Harmonic-Cancellation Technique," IEEE Journal of Solid-State Circuits, vol. 45, no. 5, pp. 1061–1071 (May 2010).

[38] B. Karthik Vasan, S. K. Sudani, D. J. Chen, & Geiger, R. L. “Low-Distortion Sine Wave Generation Using a Novel Harmonic Cancellation Technique”, IEEE Transactions on Circuits and Systems I:Regular Papers, 60(5), 1122–1134 (2013)

[39] B. Dufort, G. W. Roberts, “On-chip Analog Signal Generation for Mixed-Signal Built-In Self-Test”, IEEE Journal of Solid-State Circuits, 34(3), 318–330 (2013).

[40] A. Maeda, “A Method to Generate a Very Low Distortion High Frequency Sine Waveform Using an AWG”, IEEE International Test Conference (Oct. 2008).

[41] K. Kato, F. Abe, K. Wakabayashi, C. Gao, T. Yamada, H. Kobayashi, O. Kobayashi, K. Niitsu, “Two-Tone Signal Generation for Communication Application ADC Testing”, The 21st IEEE Asian Test Symposium, Niigata, Japan (2012).

[42] S. Shibuya, Y. Kobayashi, H. Kobayashi, “High-Frequency Low- Distortion Signal Generation Algorithm with Arbitrary Waveform Generator”, IEEE 11th International Conference on ASIC, Chengdu, China (Nov. 2015).

[43] T. Komuro, S. Sobukawa, H. Sakayori, M. Kono, H. Kobayashi, “Total Harmonic Distortion Measurement System for Electronic Devices up to 100MHz with Remarkable Sensitivity”, IEEE Trans. on Instrumentation and Measurement, vol. 56, no. 6, pp. 2360 - 2368 (Dec. 2007).

[44] Peter Sarson; Haruo Kobayashi, “Using distortion shaping technique to equalize ADC THD performance between ATEs”, 2016 IEEE 21st International Mixed-Signal Testing Workshop (IMSTW) Pages: 1 – 6

[45] F. Ye, H. Wang, T. Yi, Z. Hong, “A 14-bit 2-GS/s DAC With a Programmable Interpolation Filter”, 2015 IEEE 11th International Conference on ASIC, pp 1-4, Chengdu, China (Nov. 2015).

[46] Q. Ha, F. Ye, C. Chen, X. Zhu, M. Wang, Y. Lin, N. Li, J. Ren, “A 4-channel 8-bit 650-MSample/s DAC With Interpolation Filter for Embedded Application”, IEEE 9th International Conference on ASIC, pp. 492 – 495, Xiamen, China (Nov. 2011).

[47] M. R. Sadeghifar, J. J. Wikner, “A Higher Nyquist-Range DAC Employing Sinusoidal Interpolation”, NORCHIP, 2010, pp. 1 – 4

[48] R. E. Crochiere, L. R. Rabiner, Multirate Digital Signal Processing, Prentice-Hall (1983).

[49] P. Sarson, S. Shibuya, T. Yanagida, H. Kobayashi, “A Technique for Dynamic Range Improvement of Intermodulation Distortion Products for an Interpolating DAC-based Arbitrary Waveform Generator Using a Phase Switching Algorithm”, IEEE VLSI Test Symposium (April, 2017)

[50] W. Jiang , S. Zhao, Q. Tan, D. Liang, X. Li , Z. Zhu, “A Time-Domain Compensation Model with High IMD3 Suppression for Microwave Photonic Links”, IEEE International Conference on Electronic Information and Communication Technology (August 2016)

[51] Engelke, P., Polian, I. ; Renovell, M. ; Kundu, S. ; Seshadri, B. ; Becker, B., "On Detection of Resistive Bridging Defects by Low-Temperature and Low-Voltage Testing," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, Volume 27 Issue 2 Date Feb. 08, Pages 327 – 338.

[52] Engelke, P., Polian, I. ; Renovell, M. ; Kundu, S. ; Seshadri, B. ; Becker, B., "The Pros and Cons of Very-Low-Voltage Testing," An Analysis Based on Resistive Short Defects. 2014 IEEE 32th VLSI Test Symposium (VTS), pages 171–178.

[53] Majhi, J. ; Azimane, M. ; Gronthoud, G. ; Lousberg, M. ; Eichenberger, S. ; Bowen, F. "Memory Testing Improvements Through Different Stress Conditions" ESSCIRC 2005, pages 299-302

[54] Shen, ST. ; Liu, C. ; Ma, EH. ; Cheng, IC. ; Chien-Mo, J. "Reliability Screening of a-Si TFT Circuits: Very-Low Voltage and IDDq Testing" Journal of Display Technology, Volume:6 , Issue: 12 Date Aug. 10, Pages 592-600

[55] J. Kim ; K. Lee, "Three-transistor one-time programmable (OTP) ROM cell array using standard CMOS gate oxide antifuse" , IEEE Electron Device Letters , Volume:24 , Issue: 9 , Sept. 2003, Pages: 589-591

[56] P. Candelier ; N. Villani ; J. -P. Schoellkopf ; P. Mortini, "One time programmable drift antifuse cell reliability" , Reliability Physics Symposium, 2000. Proceedings. 38th Annual 2000 IEEE International, Pages:169-173

[57] Qian, X. ; Han, C. ; Singh, A. "Detection of Gate-Oxide Defects with Timing Tests at Reduced Power Supply" 2012 IEEE 30th VLSI Test Symposium (VTS), pages 120-126

[58] Ackaert, J. ; de Backer, E. ; Lowe, A. ; Yao, T. ; Goessens, C. ; Greenwood, B. ; Verpoort, P. "Improvement of Non-Volatile Memory Tunnel Oxide Robustness and Integrity by Design Optimization of the Memory Cell" Integrated Circuit Design and Technology, 2015

[59] Peter Sarson, "Very low supply voltage room temperature test to screen low temperature soft blown fuse fails which result in a resistive bridge", 2016 17th International Symposium on Quality Electronic Design (ISQED) Pages: 135 - 139

[60] J. Pathak, M. Thomas, J. Payne, G. Schatzberger, A. Wiesner, F. Leisenberger, E.Wachmann, and M. Schrems, "Embedded Non-Volatile memory modules for low voltage high temperature applications," Non-Volatile Semiconductor Memory Workshop, Feb. 2006, pp. 27–30.

[61] Mammen Thomas, J. Pathak, Jim Payne, Friedrich Leisenberger, Ewald Wachmann, Gregor Schatzberger, Andreas Wiesner, Martin Schrems. "A Non-Volatile Embedded Memory for High Temperature Automotive and High-Retention Applilcations" ISQED 2006.

[62] Mammen Thomas, J. Pathak, Jim Payne, Friedrich Leisenberger, Ewald Wachmann, Gregor Schatzberger, Andreas Wiesner, Martin Schrems. "Development of a Non-Volatile Memory Meeting Automotive Temperature &

Reliability Requirements and Results" IMAPS International Conference and Exhibition on High Temperature Electronics (HiTEC 2006).

[63] M. Lenzlinger and E. Snow, "Fowler-nordheim tunneling into thermally grown SiO2" IEEE Tech. Dig., pp. 273–283, 1969.

[64] Sarson, P.; Schatzberger, G.; Seitz, R., "Automotive EEPROM Qualification and Cost Optimization", 22nd Asian Test Symposium, 2013, p105-106.

[65] D. Sweetman; "Guardbanding VLSI EEPROM test programs", VLSI Test Symposium, 1991. 'Chip-to-System Test Concerns for the 90's', Digest of Papers, Pages: 155 – 160

[66] D. Wilkie, M. Hensen, "A study of EEPROM endurance correlation with wafer level reliability data", Integrated Reliability Workshop, 1995. Final Report., International; Pages: 55 – 60

[67] J.Plantier; H. Aziza ; J. M. Portal ; C. Reliaud, "Investigation of EEPROM memories reliability during endurance and retention tests"

[68] Friedrich Leisenberger, Schatzberger Gregor, "Efficient Contact Screening of Compact NVMs for High Reliability Automotive Applications" IMSTW, 2015

[69] Neal R. Mielke Patent US 4,963,825 "Methode of Screening EPROM-related Devices for Endurance Failure"

[70] J.Fellner, G.Schatzberger, A.Wiesner, "Fully Automatical Test and Qualification System for a High Endurance Embedded EEPROM Module"; 2008 IEEE International Integrated Reliability Workshop Final Report, Pages: 118-120

[71] Muroke, P. "Flash Memory Field Failure Mechanisms", 44th Annual International Reliability Physiscs Symposium, San Jose 2006

[72] Johannes Fellner, Gregor Schatzberger Patent US 8,189,409 B2 "Readout Circuit for Rewriteable Memories and Readout Methode for Same"

[73] Young-Bog Park , Schroder, D.K. "Degradation of thin tunnel gate oxide under constant Fowler-Nordheim current stress for a flash EEPROM Electron Devices", in Electron Devices, IEEE Transactions on , vol.45 , no.6, June 1998

[74] M.Thomas J. Pathak ; J. Payne ; F. Leisenberger ; E. Wachmann ; G. Schatzberger ; A. Wiesner ; M. Schrems, "A non-volatile embedded memory for high temperature automotive and high-retention applications", 7th International Symposium on Quality Electronic Design (ISQED'06), Pages

[75] F. P. Leisenberger, Gregor Schatzberger, "An Efficient Contact Screening Method and its Application to High-Reliability Non-Volatile Memories", Journal of Electronic Testing 2016, 1573-0727, p 1-12

[76] G. Schatzberger ; F. P Leisenberger ; P. Sarson, "Yield improvement of an EEPROM for automotive applications while maintaining high reliability", 2016 IEEE 34th VLSI Test Symposium (VTS), Pages: 1-6.

[77] M. Quach; Tuan Pham; T. Figal; B. Kopitzke; P. O'Neill, " Wafer-level defect-based testing using enhanced voltage stress and statistical test data evaluation", Test Conference, 2002. Proceedings. International

[78] Christian Streitwieser, "Real-time test data acquisition and data processing enabling closed loop control systems for adaptive test", 2016 IEEE International Symposium on Circuits and Systems (ISCAS)

[79] Fan Lin; Chun-Kai Hsu; Kwang-Ting Cheng, "AdaTest: An efficient statistical test framework for test escape screening", 2015 IEEE International Test Conference (ITC), Pages: 1 – 8

[80] CK, Hsu, P. Sarson, G. Schatzberger, F. Leisenberger, J. Carulli, S. Siddhartha, KT Cheng, "Variation and Failure Characterization Through Pattern Classification of Test Data From Multiple Test Stages", IEEE 47th International Test Conference, Fort Worth, TX (Nov. 2016)

[81] Chen-Yong Cher; Yiorgos Makris; C. Thibeault; Alan J. Drake, "Innovative practices session 7C: Self-calibration & trimming", 2013 IEEE 31st VLSI Test Symposium (VTS), Pages: 1 - 1

[82] Ali Ahmadi; Ke Huang; Amit Nahar; Bob Orr; Michael Pas; John M. Carulli; Yiorgos Makris, "Yield prognosis for fab-to-fab product migration", 2015 IEEE 33rd VLSI Test Symposium (VTS), Pages: 1 - 6

[83] Stephen Sunter; J. -F. Côté; Jeff Rearick, "Streaming Access to ADCs and DACs for Mixed-Signal ATPG", IEEE Design & Test, Year: 2016, Volume: 33, Issue: 6, Pages: 38 - 45,

[84] Stephen Sunter; Krzysztof Jurga; Andrew Laidler , "Using Mixed-Signal Defect Simulation to Close the Loop Between Design and Test", IEEE Transactions on Circuits and Systems I: Regular Papers, Year: 2016, Volume: 63, Issue: 12, Pages: 2313 - 2322,

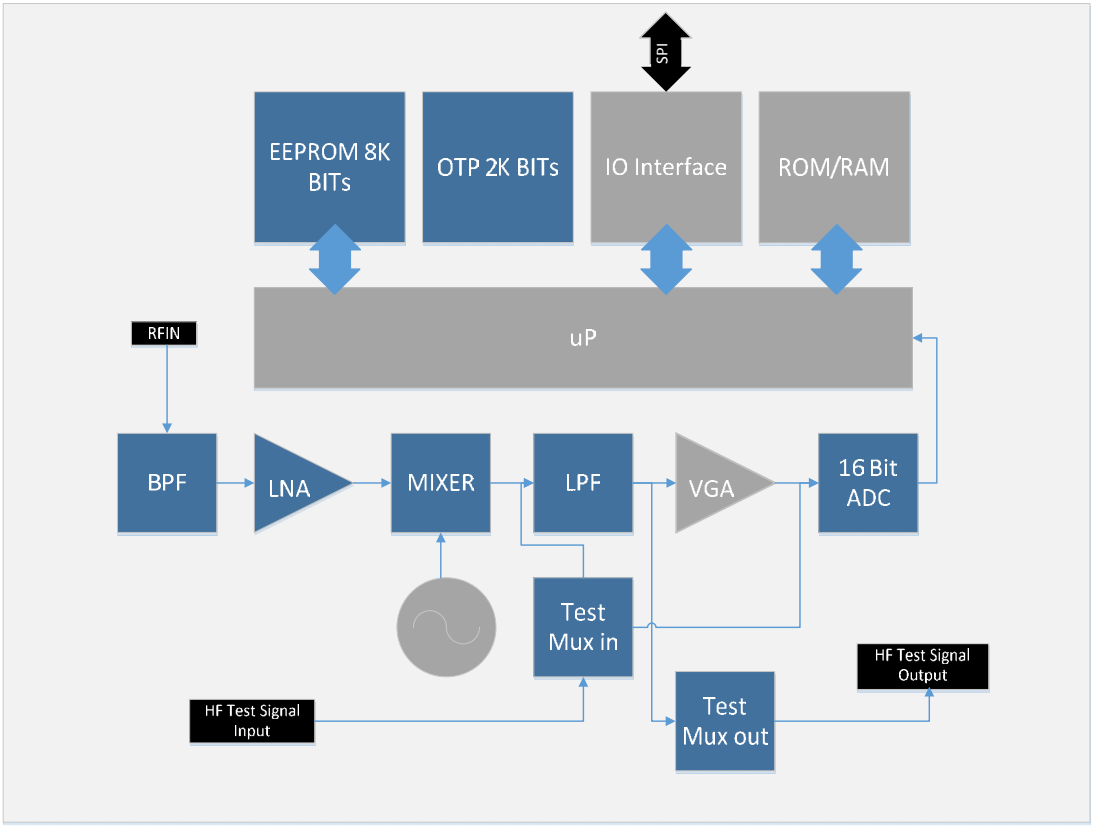
[85] Bagalini, E., Sini, J., Sonza Reorda, M., Violante, M., Klimesch, H., Sarson, P. "An automatic approach to perform the verification of hardware designs according to the ISO26262 functional safety standard". In Test Symposium (LATS), 2017 18th Latin-American. IEEE.

[86] P. Sarson, J Rearick, "Use Models for Extending IEEE 1687 to Analog Test", IEEE 48th International Test Conference, Fort Worth, TX (Nov. 2017)

[87] S. Sunter, P,Sarson "A/MS Benchmark Circuits for Comparing Fault Simulation, DFT, and Test Generation Methods", IEEE 48th International Test Conference, Fort Worth, TX (Nov. 2017)

# 13 APPENDICES

# APPENDIX 1 - RF ANALOGUE MIXED SIGNAL BLOCK DIAGRAM




This device is a figment of my imagination that I made up to use as a vehicle to tell a story for this dissertation. This device has all the elements for innovation that are described within the chapters of this dissertation and the blocks are shown in blue.

The test mux innovation is described in chapter 2.7, the Low Pass and Band Pass Filter testing innovation is described in chapter 3 & 4, and the Analogue to Digital Converter testing is described in chapter 5 & 6, the Non-Volatile Memory testing innovation is described in chapter 7, EEPROM testing innovation is described in chapter 8.

# APPENDIX 2 - PRODUCTION TESTER


This is a Fusion MX LSI tester, originally manufactured by LTX Corporation now called Xcerra. All the test solutions described in this dissertation were developed using this test platform.

This test platform has all the required instrumentation to test an RF, High Power, Mixed Signal, Large Scale Integration device. The review of this tester is described in chapter 2.2.

# APPENDIX 3 – IDDQ ANALOGUE DATALOG

1200 -0.150 uA 0.800 uA F 5.030 uA O_3V_VCC IpowAna_VCC_1

7000 0.213 mA 0.447 mA F 8.615 mA O_3V_VCC IpowAna_VCC_5

7001 1.157 mA 1.643 mA F 9.806 mA O_3V_VCC IpowAna_VCC_6

7002 6.103 mA 7.048 mA F 14.918 mA O_3V_VCC IpowAna_VCC_7

7003 8.070 mA 9.330 mA F 17.006 mA O_3V_VCC IpowAna_VCC_8

7004 8.270 mA 9.530 mA F 17.182 mA O_3V_VCC IpowAna_VCC_9

7005 11.095 mA 12.805 mA F 20.176 mA O_3V_VCC IpowAna_VCC_10

7006 11.195 mA 12.905 mA F 20.325 mA O_3V_VCC IpowAna_VCC_11

7007 11.505 mA 13.395 mA F 20.709 mA O_3V_VCC IpowAna_VCC_12

7008 13.620 mA 15.780 mA F 22.920 mA O_3V_VCC IpowAna_VCC_13

7009 14.425 mA 16.675 mA F 23.769 mA O_3V_VCC IpowAna_VCC_14

7010 16.545 mA 19.155 mA F 26.032 mA O_3V_VCC IpowAna_VCC_15

7011 16.340 mA 18.860 mA F 25.848 mA O_3V_VCC IpowAna_VCC_16

7012 19.165 mA 22.135 mA F 28.735 mA O_3V_VCC IpowAna_VCC_17

7013 19.570 mA 22.630 mA F 29.209 mA O_3V_VCC IpowAna_VCC_18

7014 19.970 mA 23.030 mA F 29.555 mA O_3V_VCC IpowAna_VCC_19

7015 19.970 mA 23.030 mA F 29.512 mA O_3V_VCC IpowAna_VCC_20

7016 20.680 mA 23.920 mA F 30.313 mA O_3V_VCC IpowAna_VCC_21

7017 22.995 mA 26.505 mA F 32.749 mA O_3V_VCC IpowAna_VCC_22

7018 31.775 mA 36.725 mA F 41.579 mA O_3V_VCC IpowAna_VCC_23

7019 32.075 mA 37.025 mA F 41.876 mA O_3V_VCC IpowAna_VCC_24

7020 33.185 mA 38.315 mA F 41.922 mA O_3V_VCC IpowAna_VCC_25

7119 1.109 mA 1.442 mA F 0.047 mA O_3V_VCC IpowAna_VCC_24_25_del

7120 5.896 mA 7.444 mA F 0.095 mA O_3V_VCC IpowAna_VCC_25_26_del

4420 -34.950 dBm -25.050 dBm -31.590 dBm D_IF1_OUT P1CH1

4430 -73.200 dBm -58.800 dBm F -105.389 dBm D_IF1_OUT P2CH1

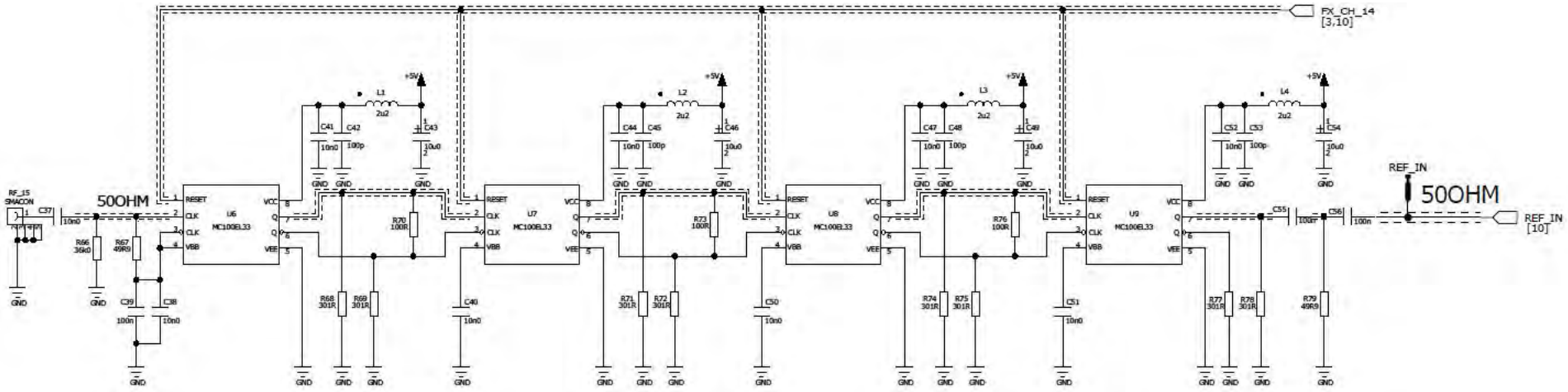
4440 -61.200 dBm -46.800 dBm F -100.087 dBm D_IF1_OUT P3CH1

4450 -110.000 dBm -47.500 dBm -102.260 dBm D_IF1_OUT P4CH1

| | | | | | | |
|---|---|---|---|---|---|---|
| 4460 | 11.230 dB | 13.570 dB | F | 5.302 dB | D_IF1_OUT | GRF_CH1 |
| 4470 | 59.500 dB | 68.500 dB | F | 98.799 dB | D_IF1_OUT | GIF_CH1 |
| 4480 | 23.000 dB | 100.000 dB | | 27.173 dB | D_IF1_OUT | GIMR_CH1 |
| 4490 | 84.100 dB | 103.900 dB | F | 78.712 dB | D_IF1_OUT | GT_CH1 |
| 4500 | -34.950 dBm | -25.050 dBm | | -31.529 dBm | D_IF1_OUT | P1CH1_R |
| 4510 | -73.200 dBm | -58.800 dBm | F | -105.506 dBm | D_IF1_OUT | P2CH1_R |
| 4520 | -61.200 dBm | -46.800 dBm | F | -101.924 dBm | D_IF1_OUT | P3CH1_R |
| 4530 | -110.000 dBm | -47.500 dBm | | -105.383 dBm | D_IF1_OUT | P4CH1_R |
| 4540 | 11.230 dB | 13.570 dB | F | 3.582 dB | D_IF1_OUT | GRF_CH1_R |
| 4550 | 59.500 dB | 68.500 dB | F | 98.977 dB | D_IF1_OUT | GIF_CH1_R |
| 4560 | 23.000 dB | 100.000 dB | | 28.459 dB | D_IF1_OUT | GIMR_CH1_R |
| 4570 | 84.100 dB | 103.900 dB | F | 77.053 dB | D_IF1_OUT | GT_CH1_R |

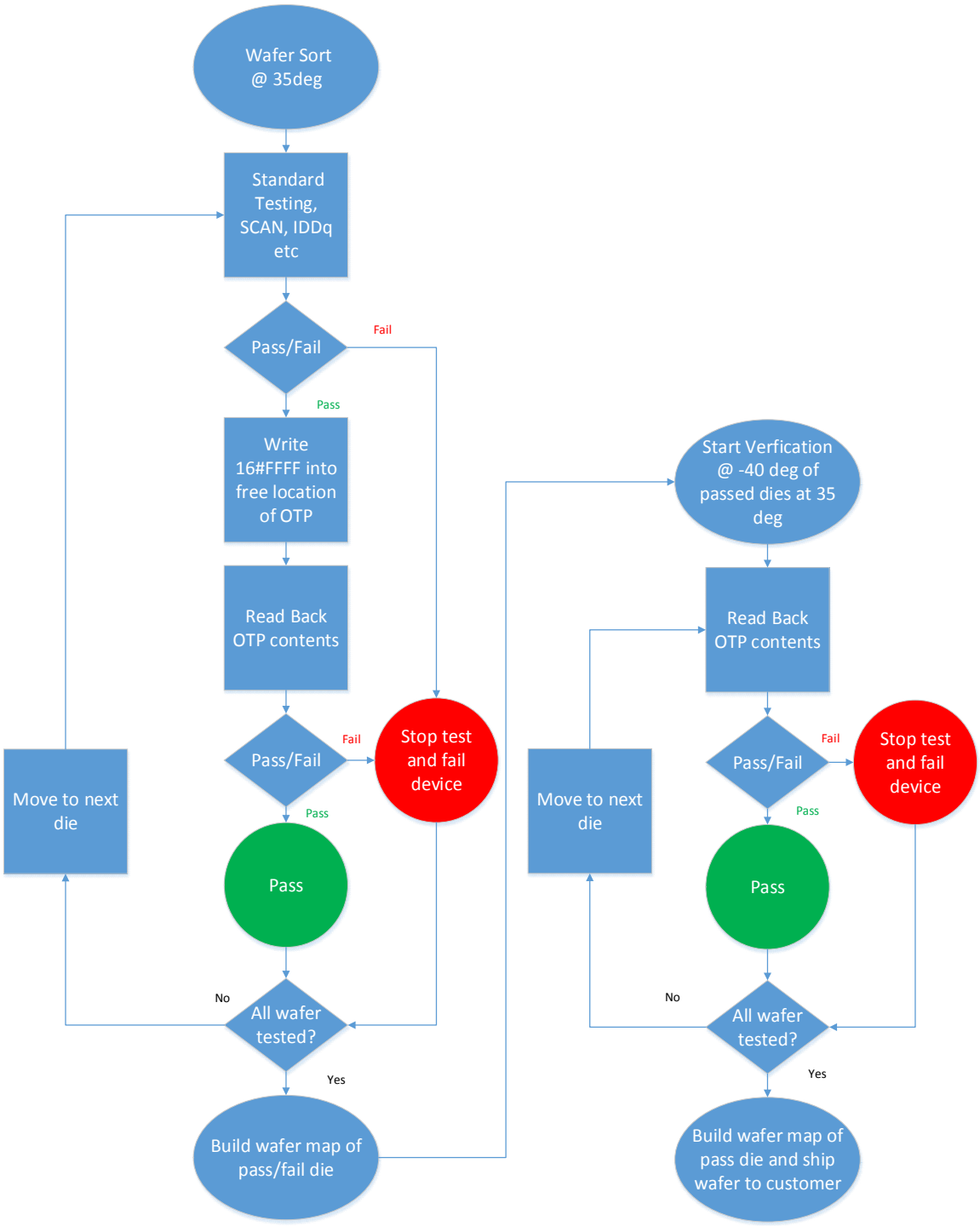
This is a dlog of a faulty RF device, using the standard method employed within the semiconductor industry; this device would have been detected using an expensive RF test methodology. What can be seen is that a matrix of current tests can be used to test the device more thoroughly as a way to better diagnose yield issues and can be used for yield analysis and aid investigations in case of future process issues. Using this method for testing a mixed signal device can increase the quality level as latent defects that would not be detected by RF tests can be seen by small changes in the difference currents between blocks of a device, see Appendix 1 and 2.3.14. Also as a set of current tests are much quicker to test than RF tests the result is a much quicker test program. As the test program no longer needs to measure RF signals, a much cheaper tester could be used and the multisite count can be significantly increased resulting in a much more cost efficient test process.

# APPENDIX 4 – ECL DIVIDER REFERENCE CLOCK CIRCUITRY




This circuit produces a low phase noise clock based, which is a function of the frequency of RF generator divided by 256. This circuit generates a 10MHz low phase noise reference clock that can be used for a RF applications where the LO needs to be phase locked to the tester with a superior performance to any standard resource available on a standard Mixed Signal tester. The device used is the MC100EL33.

# APPENDIX 5 - OTP PRODUCTION TEST FLOW



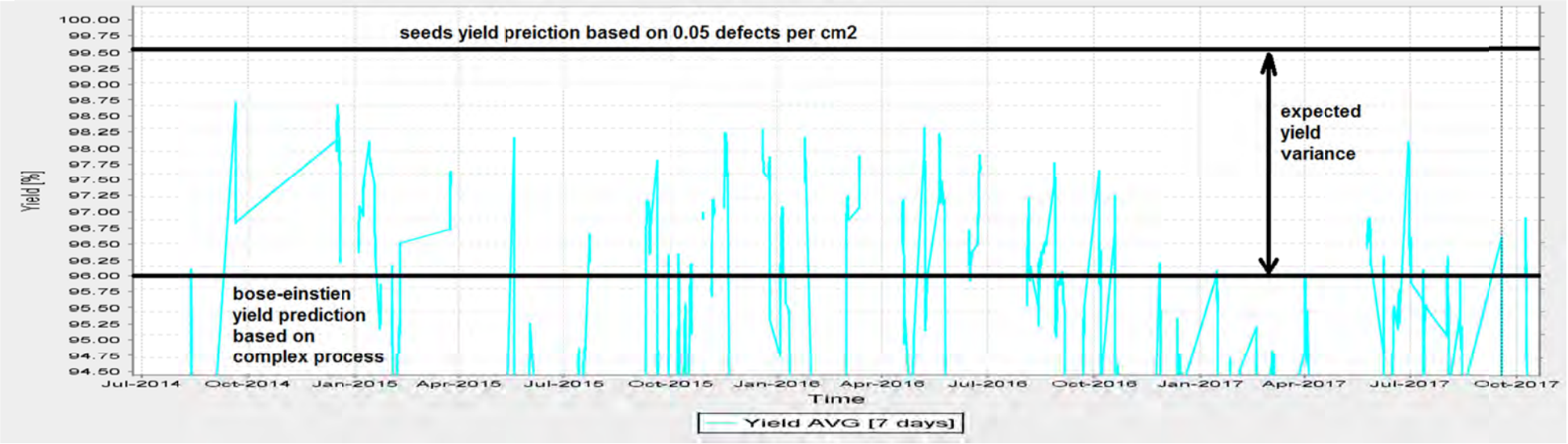

The above flow describes the flow used in testing Non-Volatile Memory described in chapter 1.

# APPENDIX 6 - PRODUCTION PROBER AT -40 DEGREES


This is an Electroglas Prober that is used to test silicon wafers before they are sliced up and packaged. This prober is special in that it has an option that allows the prober to test at minus temperatures.

This is an actual setup of the wafer test of the device described in chapter 7, the wafer map of a tested wafer can be observed on the operator interface along with the temperature of the prober.

# APPENDIX 7 – YIELD OF EEPROM PRODUCT FROM AMS 0.35 UM SILICON FOUNDRY




This chart shows the yield varying because of the varying defect density levels in the ams 0.35 um silicon foundry. As real defects are being measured then no test escapes should be expected. The yield is in alignment with the expected yield calculated based on the different statistical models used for predicting silicon yield based on die size and defects per $cm^2$.

# APPENDIX 8 – AUTHORS INDIVIDUAL CONTRIBUTIONS

1. Produced a method to test and characterise filters using a chirp at both baseband and RF that can reused easily within a semiconductor company that has its own tester infrastructure. Using the software and hardware IP allows measuring and characterising gain and phase parameters with a simple set of instructions that significantly reduces the time to debug, release, and deploy a working test program with good repeatability and reproducibility. This results in a very fast time to market that was not previously possible.
   The author has built the infrastructure with the ams AG workspace to use stepped chirps and coached the Intellectual Property team on how to build discrete chirps. Using both stimuli, the author has developed an algorithm to test and characterise mixed signal blocks for phase and time parameters such that an in-depth breakdown of frequency-time-phase profile of a system can be produced resulting in very accurate group delay measurements of R and mixed signal blocks.

2. Industrialized the phase switching algorithms for harmonic and Intermodulation Distortion suppression techniques from Gunma University and applied it to an ATE systems AWG. The result is the possibility to extend the useable range of an existing installed base of ATE AWG's. This saves the cost of having to buy new tester instrumentation when the current hardware is marginally worse than required.
   The author has shown both with simulation and direct measurement on ATE test systems that not only can the harmonic space be made more spectrally pure by supressing harmonic but also by reducing harmonic components by indirectly reducing other components. The key finding being that suppressing the $4^{th}$ order harmonic or intermodulation distortion can have a large impact on reducing odd order components and thus having a positive impact on the whole spectrum of in-band components.
   The author also showed the sensitivity of an Interpolating DAC architecture to the phase switching algorithm and has shown that the Interpolating DAC architecture is sensitive to the phase of a signal. By exploiting this phenomenon,

the author has shown that using phase switching can produce extremely high suppression of intermodulation products and harmonics.

3. Developed a new screening technique for antifuse OTP's that results in being able to deliver automotive grade quality devices, using a consumer grade design. The author has shown how too effectively screen badly burnt fuses of an antifuse technology that did not exhibit the full burn cycle due either defects or badly processed wafers. A correlation between negative temperature screening at typical voltage, to room temperature at low voltage, demonstrated that it was feasible to find failures that would only been seen at negative temperature to be screened effectively at room temperature. The author demonstrated with such a screening that if only a room temperature screening would have been implemented, that the defect rate in the field would be 3 ppm.

4. By using the DFT infrastructure developed by ams for EEPROM technologies, it was possible to develop a leading edge screening solution that both correlated to laboratory characterization tests and the ATE but in a test time efficient manner. By using the same DFT, it was also possible to develop a test routine that supplied a device specific reference. This enabled a test program that is extremely robust and is immune to extreme process variations seen in a silicon foundry.
   The author has analysed multiple wafer lots of different EEPROM products of ams, has through a thorough analysis of program and erase voltages, and measurement of the bitcell current has increased the yield and quality of ams EEPROM products as well as driving down test time and test cost. The author decreased the test time by implementing a production style test program, which was derived from a characterization test program that had been running in production for some years during the ramp up phase.
   The author verified the quality increase and confirmed that no test escape occurred by a combination of endurance cycling of marginal parts, to the maximum specified number of write cycles. The author then confirmed that those devices that were on the edge of the distribution still maintained the data

storage after a bake of the wafers that simulated the lifetime of the device at the given specified temperature.

5. The Author founded an Industrial Working Group to advance the state of the art of DFT of Analogue/Mixed Signal devices. This group has now founded two official IEEE Working Groups called “P1687.2, Standard for Describing Analog Test Access and Control” and “P2427, Standard for Analog Defect Modelling and Coverage”. As part of these groups, the author has contributed significantly to both the draft standards and has supplied silicon IP blocks based on the ams 0.35 um process to be used for the development of both standards. This would be in the form of using the blocks to produce different access mechanisms so as to be able to stimulate the different nodes of the devices and thus using the analogue fault models to gauge the fault coverage based on the added access.

---