GUEST EDITORIAL

SMEOS 2011 (SENSORS, MEMS AND ELECTRO-OPTIC SYSTEMS)

This special issue of the SAIEE Africa Research Journal is devoted to selected papers from the SMEOS 2011 (Sensors, MEMS and Electro-Optic Systems) Conference which was held in Berg-en-Dal, Kruger National Park, South Africa from 19 to 21 September 2011. The aim of SMEOS 2011 was to establish a forum for academia, research institutions and industry working in the field of sensors, MEMS and electro-optical systems, to share their relevant research and development ideas. Each paper presented at the conference was double-blind reviewed by at least two reviewers. Reviewers could recommend a reviewed paper to the technical chair for publication in this special issue, and a total of eleven papers eventually passed this review process.

Five of the papers discuss the integration of additional functionality into the existing CMOS technology and the vacuum packaging of some these devices. Two papers address the issue of integrating thermal MEMS devices onto a CMOS chip, a third paper investigates the wafer level packaging of CMOS MEMS devices, and two papers report on the properties and application of optical devices in standard CMOS technology.

In the paper by Maclean et al, "Optimisation of CMOS compatible microbolometer device performance", the design, simulation and characterisation of CMOS microbolometers are investigated in order to optimise the electro-thermal properties of the devices. The paper by Schoeman and du Plessis, "Characterisation of the electrical response of a novel dual element thermistor for low frequency applications", also deals with thermal elements integrated onto a CMOS chip, but in this case the interaction between two thermal elements in close proximity is being investigated. These novel devices can be used in very low frequency signal processing applications as a result of the quite long thermal time constants. In the paper by Versteeg et al, "Wafer level packaging with wedge seal method", a novel vacuum tight seal referred to as the wedge seal method is proposed. The seal consists of a silicon wedge forced into a pliable material (typically a metal) that is attached to the component wafer. The wedge-seal addresses some of the requirements of CMOS micro-bolometer packaging in that it provides a vacuum tight seal at low temperatures with tolerance to surface finish and topography.

The paper by Bogalecki et al, "Spectral measurement and analysis of silicon CMOS light sources", reports on the spectral emission from CMOS based pn junctions. Junctions in reverse bias, forward bias and in punch-through conditions are investigated, with the surprising result that a fairly wide range of wavelengths are emitted. It is indeed speculated that intra-conduction-band (c-c) electron (e-) transitions seem to be the dominant physical mechanism responsible for the wide spectrum in the avalanching and punch-through devices. One application that has already been achieved is an all silicon optical data transmission system, reported by Goosen et al in their paper "CMOS avalanche electroluminescence applications – microdisplay and high speed data communication". This paper describes a 10 Mb/s optical data link using CMOS light sources, with a bit error rate better than 10-12. Since the devices also emit in the visible range, a novel CMOS dot matrix microdisplay is also described.

Three papers report on research activities at the National

Metrology Institute of South Africa (NMISA). The first paper by Burger et al, "Ab initio frequency measurement and characterisation of frequency doubled fibre laser utilised for precision oscillators" describes the first ab initio measurement of an unknown optical frequency utilising the Ti:Sapphire-laser based optical frequency comb at NMISA. It is shown that this measurement methodology can be successfully applied to specifically relatively noisy lasers. In a second paper by Burger et al, "Methodology for in situ characterisation of a highly birefringent photonic crystal fibre for supercontinuum generation", a novel methodology for precisely determining the eigenaxes and effective twist of a solid-core polarisation maintaining fibre with a slightly elliptical effective core in an experimental setup with an ultrashort pulse laser is presented. In a third paper by Kritzinger et al, "Digital design of broadband long-period fibre gratings by an inverse scattering algorithm with flip-flop optimisation", a discrete inverse scattering method, known as layer-peeling, is used to synthesise a LPFG (long-period fibre grating) from a desired complex spectrum by a direct solution of the coupled-mode equations, while simultaneously determining the physical properties of the layered structure. Possible applications are also discussed where optimised broadband LPFGs could be utilised in the field of telecommunications and sensing.

A very interesting paper by Hugo et al originating from the Materials Science and Manufacturing Division at the CSIR, "A lensless, automated microscope for disease diagnostics", presents a digital in-line holographic microscope (DIHM) platform to be used with image processing and classification algorithms to provide a low cost, portable and automated microscope. Initial results show that the images obtained using the DIHM platform are similar to those obtained using a conventional bright field microscope. This work will be targeted towards the implementation of an automated full blood count, which could provide resource limited areas with improved healthcare facilities and reduced diagnosis times at a low cost.

Strever et al in their paper "Optical and thermal applications in grapevine (vitis vinifera l.) research – an overview and some novel approaches" describe various optical and thermal applications in grapevine research to quantify the light and temperature regime around a grape bunch. Techniques include temperature measurement techniques (thermocouples and thermal imaging) as well as methods to quantify light quantity (hemispherical photography) as well as light quality (spectro-radiometric applications) around a grape bunch.

The last paper deals with infrared detectors. In the paper "Growth and characterisation of InAs photodetectors for MWIR applications" by Wagener et al, the development of InAs photodiode structures grown by metal-organic vapour phase epitaxy and processed using conventional photolithography techniques are discussed. Due to the narrow band gap of these materials, the detectivity of the devices is often limited by the junction leakage currents. Various contributions to the leakage current and photo-response have been analysed.

Prof. Monuko du Plessis
Guest Editor

OPTIMISATION OF CMOS COMPATIBLE MICROBOLOMETER DEVICE PERFORMANCE

W. Maclean, M. du Plessis and J. Schoeman*

* *Carl and Emily Fuchs Institute for Microelectronics, Dept. of Electrical, Electronic & Computer Engineering, Corner of University Road and Lynnwood Road, University of Pretoria, Pretoria 0002, South Africa. E-mail: johan.schoeman@eng.up.ac.za*

Abstract: Uncooled IR (infrared) microbolometer performance is greatly affected by the thermal properties associated with the structural layout of each design. Equations are derived in this article which make use of basic structural dimensions to predict the expected thermal conductance and thermal capacitance of a microbolometer device. These equations enable a microbolometer designer to determine the estimated thermal time constant of a design without performing complicated analytical calculations for each layer in the design. Calculation results shown indicate the effect structural changes have on the thermal time constant of microbolometer devices. These changes aid microbolometer designers in adjusting the layout of the device to change the thermal time constant to the desired value. Structural deviations that occur during manufacturing of microbolometers are calculated and the possible causes are discussed.

Key words: Uncooled infrared bolometer, thermal properties, thermal performance prediction.

1. INTRODUCTION

Microbolometer device fabrication has been substantially simplified with the advances in MEMS (Micro-electromechanical structures), such as thermally isolating micro structures created with a complementary metal oxide semiconductor (CMOS) process [1]. These devices are manufactured by performing various processing techniques after the CMOS process has been completed. These post CMOS processing techniques are performed at a low temperature, to maintain the integrity of the underlying CMOS circuit.

Different microbolometer designs are used to determine the relationship between structural differences and the thermal performance of these devices. Specifically designed test devices will be used to isolate structural differences and determine the effect each structural difference has on the thermal performance. These changes in performance are used to derive equations that can be used to predict the thermal properties of a microbolometer design based on basic structural dimensions [2].

Experimental measurements performed on manufactured microbolometer devices reveal material property and structural differences from expected design values. These differences are calculated and possible causes for these differences are considered.

2. DESIGN OF MICROBOLOMETER TEST DEVICES

The microbolometer device shown in figure 1, illustrates the basic layout of a microbolometer with all structural elements indicated. This article will investigate performance changes with alterations to the support leg width and length. Test devices which isolate a single structural change are used. In one set the support leg width is kept constant and only the length is changed to determine the change in thermal performance associated. Another set of devices are then used where the support leg length is kept constant and only the width is changed. The range of support leg widths used are between 8 μm and 20 μm, and support leg lengths are between 89 μm and 169 μm.
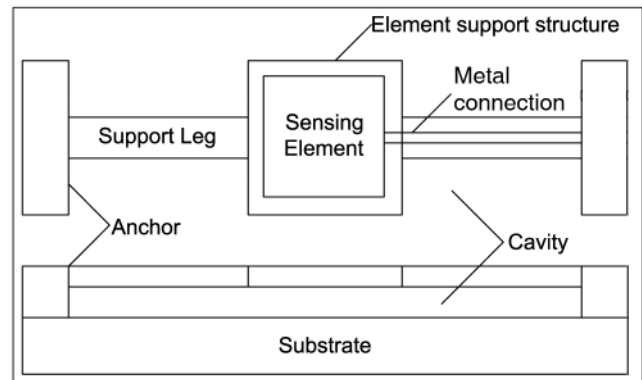
Figure 1: Basic structure of IR bolometer

Test structures used for the varying width test make use of an "I" bolometer design as shown in figure 1, bolometers used for the varying length test make use of an "L" design. The "L" design improves the shape of the microbolometer for integration in a FPA (Focal Plane Array) by keeping the device square whilst increasing the length of the support legs. FPA's are used to produce a two dimensional image from multiple microbolometers arranged in an array, connected to a readout circuit [3].

3. THERMAL PROPERTY CALCULATIONS

Thermal conductance defines the rate at which heat is transferred from the bolometer membrane throughout the

Figure 2: Material layers used for test devices

Figure 3: Thermal conductance of devices with support legs of 30 μm long

device and to the anchors which operate as heat sinks for the microbolometer device. The thermal conductance of a microbolometer can be calculated by using e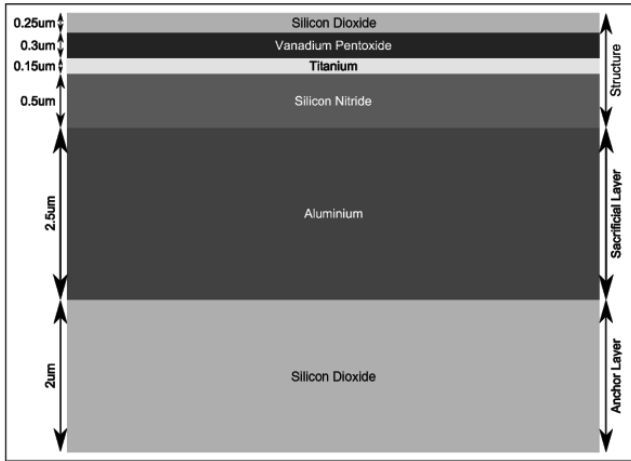quation 1, where the three terms represent the physical thermal conductance, radiation thermal conductance and gaseos thermal conductance respectively [4,5]. Devices packaged in a vacuum will have no conductance caused by gaseos thermal conductance and therefore equation 1 has to be changed accordingly.

$$G = \lambda \left(\frac{WT}{L} \right) + 4\delta\varepsilon A(Tm)^3 + \lambda_{air}\frac{A}{d}, \qquad (1)$$

where:

λ = the thermal conductivity in W/mK
W = the width of the support legs in m
T = the thickness of the material in m
L = the length of the support legs in m
δ = the Boltzmann constant
ε = the effective emissivity of the material
A = the area of the entire device in m^2
Tm = the temperature in K
λ_{air} = the thermal conductivity of air in W/mK
d = the distance from membrane structure to cavity bottom in m

The gaseos thermal conductance calculations assume that the heat loss through the air above the membrane structure to the device package is negligibly small and therefore only the heat loss through the air beneath the structure is considered. Calculation performed by using equation 1 must be done for each layer used in the design as shown in figure 2, and then added together to calculate the total thermal conductance of the bolometer device.

Devices used for the support leg width test all have a membrane area of 900 μm^2 and support legs of 30 μm long. The width of their support legs are changed in steps of 4 μm as shown in figure 3, where an increase in support leg
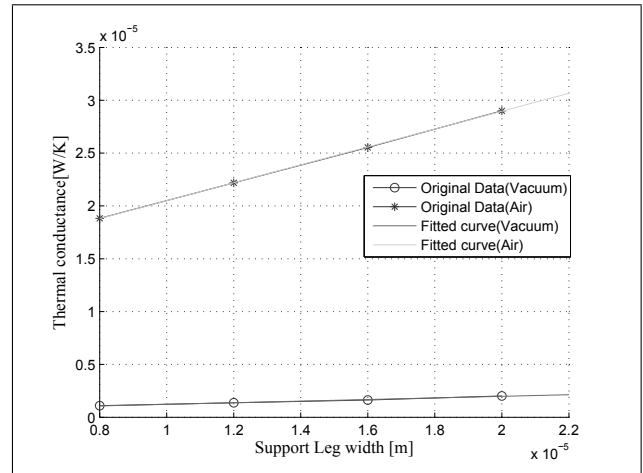
width produced an increase in thermal conductance. This increase in thermal conductance is caused by the increase in the thermal connection between the membrane and the anchors which increases the heat transfer speed.
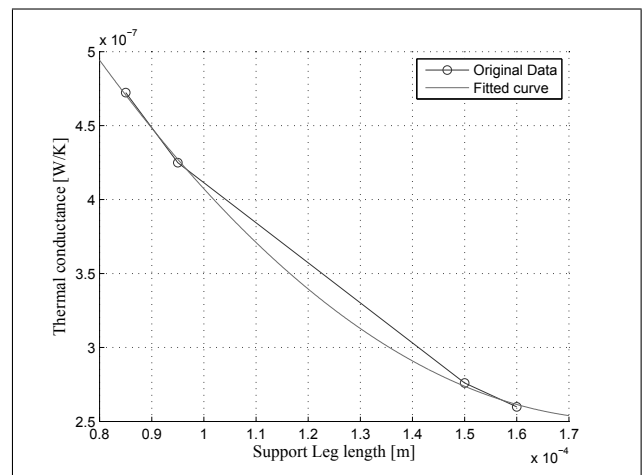
Figure 4: Thermal conductance of devices with support legs of 10 μm wide packaged in vacuum

Devices used for the support leg length test have a membrane area of 1600 μm^2 and 10 μm wide support legs. The change in thermal conductance for test devices, with different lengths of support legs, packaged in vacuum and in air are shown in figure 4 and 5, respectively. Results indicate that in order to reduce the thermal conductance, the support legs must be shortened for devices packaged in air and lengthened for devices packaged in a vacuum. This difference is due to the large contribution of gaseos thermal conductance through the air beneath the support legs to the bottom of the cavity of the devices packaged in air.

Thermal capacitance is a measure of the microbolometer's heat retention ability. The thermal capacitance is only affected by the volume of the microbolometer device and therefore it can be assumed that when the total
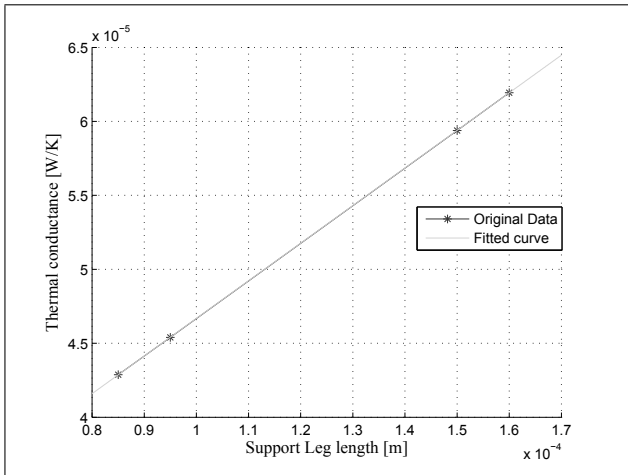
Figure 5: Thermal conductance of devices with support legs of 10 μm wide packaged in air
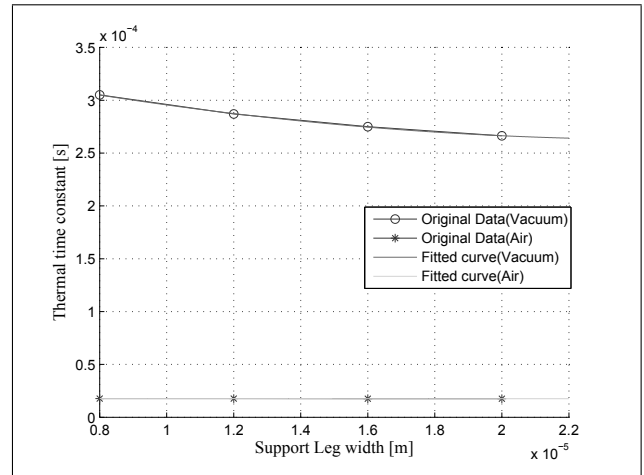

Figure 6: Thermal time constant of devices with a membrane area of 900 μm^2

volume of the microbolometer is increased the thermal capacitance is also increased. Equation 2 is used to calculate the thermal capacitance of each layer of the microbolometer device. The thermal time constant can be calculated with equation 3 by using the total thermal capacitance and thermal conductance values calculated [4, 5]. The thermal capacitance and the thermal conductance are used together to determine the thermal time constant of a microbolometer device. This thermal time constant determines the maximum speed at which the microbolometer can operate. The readout circuit can limit the operating speed further but is not capable of reducing the thermal time constant beyond this value without altering the design layout.

$$H = V\rho c, \tag{2}$$

where:

V = the volume of the device m^3
ρ = the density in g/m^3
c = the specific heat in J/gK

$$\tau = \frac{H_{Total}}{G_{Total}}, \tag{3}$$

where:

H_{Total} = the total thermal capacitance of all the layers in J/K
G_{Total} = the total thermal conductance of all the layers in W/K

Figure 6 shows the change in thermal time constant with a change in support leg width for the test devices used in the thermal conductance and thermal capacitance calculations above. Figure 7 and 8 show the thermal time constant with a change in support leg length for test devices used above packaged in a vacuum and in air, respectively.
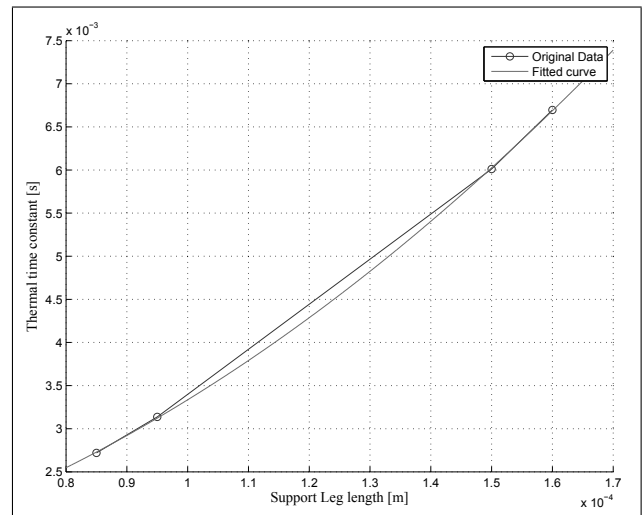

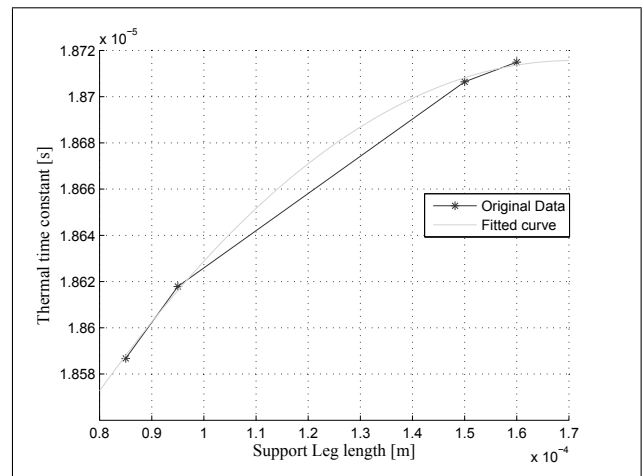Figure 7: Thermal time constant of devices with support leg widths of 10 μm packaged in vacuum

Figure 8: Thermal time constant of devices with support leg widths of 10 μm packaged in air

4. PREDICTION AND OPTIMISATION OF THERMAL PROPERTIES

Microbolometers are designed to have a specific time constant. This time constant can be adjusted to improve the performance of a device by altering its structural layout [6]. The previous sections show simulation results for devices where different structural dimensions are varied to isolate the change in thermal time constant to a specific structural dimension.

Results shown in the previous sections indicate that the time constant for bolometer devices can be lengthened by increasing the length of the support legs, or alternatively reducing the width. By using the results shown in figures 3 to 5 and the thermal capacitance values calculated with equation 2 for the same devices, equations 4 to 6 can be derived.   These equations can be used to calculated the expected thermal conductance and thermal capacitance values for devices based on basic structural dimensions.   These predicted thermal conductance and thermal capacitance values are then used to estimate the thermal time constant of a device by using equation 3.

$$
\begin{aligned}
G_{Air} &= 0.187 L_{Leg} + 1.28 \times 10^4 A_{Mem} \\
&\quad + 0.85 W_{Leg} - 6.812 \times 10^{-6}
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
G_{Vac} &= 53.3 L_{Leg}^2 - 0.023 L_{Leg} + 0.076 W_{Leg} \\
&\quad - 16.37 A_{Mem} + 1.1586 \times 10^{-6}
\end{aligned} \tag{5}
$$

$$
H = 0.483 W_{Leg} L_{Leg} + 0.528 A_{Mem} \tag{6}
$$

where:

L_{Leg} = the length of the support legs in m
W_{Leg} = the width of the support legs in m
A_{Mem} = the area of the membrane structure in m^2

Equations 4 to 6 make use of the basic structural dimensions to estimate the thermal performance of a microbolometer device. These equations are only valid for devices manufactured with the layers shown in figure 2 with Vanadium Oxide used as sensing material.

5. THERMAL PROPERTY VERIFICATION OF MANUFACTURED DEVICES

Manufactured microbolometers differ slightly from their designs due to variations in material thickness and device deformation.   Deformation of a bolometer device usually causes the distance between the device membrane and the cavity bottom to change, this affects the thermal conductance of devices packaged in air.

Manufactured test devices can be used to calculate the actual distance between the membrane and the cavity bottom if deformation is suspected.   Figure 9 shows the thermal conductance calculated from experimental measurements of test devices with different membrane sizes [4, 7].
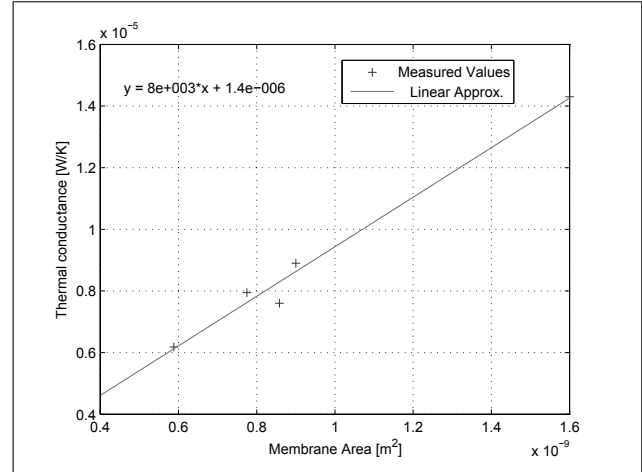

Figure 9: Thermal conductance with a change in membrane area

The actual distance between the membrane and the cavity bottom can be calculated by using the slope of the linear curve fitting shown in figure 9 and equation 7 below.

$$
G_{atm} = \lambda_{air} \frac{A}{d}, \tag{7}
$$

where:

λ_{air} = the thermal conductance of air in $W/m.K$
A = the area of the sensing membrane in m^2
d = the distance between the membrane and the cavity bottom in m

Equation 7 shows that the slope of the linear curve shown in figure 9 represents the ratio $\frac{\lambda_{air}}{d}$, where λ_{air} is a known constant. It can therefore be calculated that the actual distance between the sensing membrane and the cavity bottom is $3.2\mu m$ and not $2\mu m$ as the design indicates in figure 2. This deformation could be as a result of internal stress between the layers used in the design which caused the membrane to bend upwards after the sacrificial layer was removed, or due to the increased temperature required during the final steps of manufacturing.

Material thickness and property variations can change the performance of the actual device and make performance predictions difficult.   The thermal conductance of microbolometer devices packaged in a vacuum can be used to derive a constant which is independent of material thermal properties and layer thickness, limited to a specific design process. Figure 10 shows the thermal conductance of microbolometer devices as a function of the number of squares in their support legs. The number of squares in

a support leg is calculated by dividing the length of the support leg by its width. This shows the amount of equally sized square blocks needed to form all the support legs of a microbolometer device. These square blocks are made up of the layers used for the specific design process and are equal in thickness.
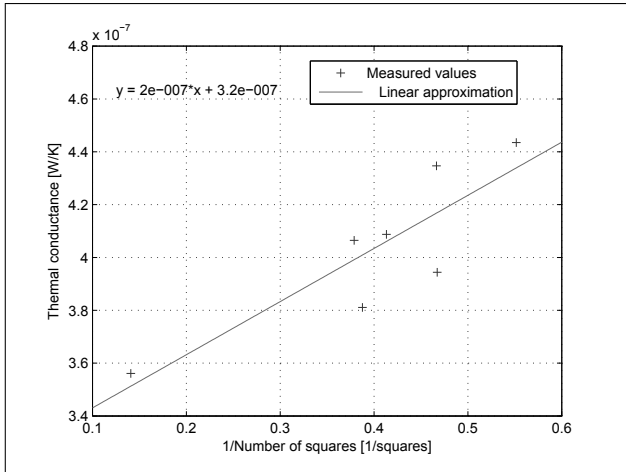

Figure 10: Thermal conductance with a change in the number of squares in the support legs

The physical thermal conductance shown in equation 8, is the major contributing factor of thermal conductance for devices packaged in a vacuum. The slope of the fitted curve shown in figure 10 is defined as the thermal conductance per square for the range of manufactured microbolometers. Substituting this newly derived thermal conductance per square value $(2 \times 10^{-7} \frac{W/K}{\#Squares})$ in equation 8 for $\frac{L}{W}$, leaves only two unknown variables, the thickness and thermal conductance of the materials in each layer. These variables are identical as shown in figure 2 for all the devices manufactured with the same process and can therefore be excluded.

$$G_{physical} = \lambda \left(\frac{WT}{L} \right), \qquad (8)$$

where:

λ = the thermal conductance of each material in $W/m.K$
W = the width of the support legs in m
T = the thickness of each material m
L = the length of the support legs in m

The thermal conductance per square can therefore be used to estimate the thermal conductance of any microbolometer device packaged in a vacuum based only on the amount of squares in their support legs as shown in equation 9.

$$G_{physical} = \left(\frac{G_{Squares}}{\#Squares} \right) + C, \qquad (9)$$

$G_{squares}$ = the thermal conductance per square in $\frac{W/K}{squares}$
$\#Squares$ = the amount of squares in the support legs in *amount*
C = the minimum thermal conductance practically achievable in W/K

6. CONCLUSION

The thermal performance calculation of a microbolometer design can be simplified by using equations 4, 5 and 6 to predict the thermal conductance and thermal capacitance prior to building a sophisticated model or performing complicated calculations. These values can be used to calculate the expected thermal time constant and adjust the structural dimensions of the bolometer device to obtain the desired performance values. Calculation results used to derive the above mentioned equations can also be used to aid in making changes to the structure of the microbolometer to achieve the required thermal time constant.

Manufactured devices perform different from simulations due to unexpected structural deformations and process variations. Calculations shown are used to determine the amount of structural deformation of microbolometer devices, by using measured thermal properties. Equations are derived that can be used to simplify the estimation of thermal conduction calculations by eliminating the use of process dependant variables that are common to all devices. These calculations are limited to a set of devices manufactured on the same substrate or with the same processing steps.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] S. Gilmartin, D. Collins, D. Bain, W. Lane, O. Korostynska, A. Arshak, E. Hynes, B. McCarthy, and S. Newcomb, "Uncooled IR nanobolometers fabricated by electron beam lithography and a MEMS/CMOS process," in *8th IEEE Conference on Nanotechnology 2008*. IEEE, 2008, pp. 131–134.

[2] Y. Tsujino, "An approach for the performance analysis of an uncooled infrared bolometer imager," *Infrared Physics and Technology*, vol. 53, pp. 50–60, 2010.

[3] A. Rogalski, "Optical detectors for focal plane arrays," *Opto-electronics review*, vol. 12, pp. 221–245, 2004.

[4] P. Eriksson, J. Andersson, and G. Stemme, "Thermal characterization of surface-micro machined silicon nitride membranes for thermal infrared detectors," *Journal of Microelectromechanical Systems*, vol. 6, pp. 55–61, 1997.

[5] F. Niklaus, A. Decharat, C. Jansson, and G. Stemme, "Performance model for uncooled infrared bolometer arrays and performance prediction of bolometers operating at atmospheric pressure," *Infrared Physics and Technology*, vol. 51, pp. 168–177, 2008.

[6] H. Najafabadi, A. Asgari, M. Kalafi, and K. Khalili, "An analytical model for detectivity prediction of uncooled bolometer considering all thermal phenomena effects," *Procedia Engineering*, vol. 8, pp. 280–285, 2011.

[7] J.-S. Shie, Y.-M. Chen, and B. Chou, "Characterization and modelling of metal-film microbolometers," *Journal of Microelectromechanical Systems*, vol. 5, pp. 298–306, 1996.

CHARACTERISATION OF THE ELECTRICAL RESPONSE OF A NOVEL DUAL ELEMENT THERMISTOR FOR LOW FREQUENCY APPLICATIONS

J. Schoeman and M. du Plessis

Carl and Emily Fuchs Institute for Microelectronics (CEFIM), University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa. E-mail: johan.schoeman@eng.up.ac.za

Abstract: This work is aimed at characterising the DC electrical response of a temperature sensitive microbolometer device. The contribution lies with the choice and the structure of the device, a novel bolometer infrared sensing structure consisting of dual sensing elements that are thermally very closely coupled on a single membrane supporting structure. A mathematical model is presented to characterise the behaviour of the device resistance and conductance for a given biasing current. A modified experiment of a well published non-optical method exploiting the normally unwanted Joule heating of a device when biased with a large direct current is employed for the experimental verification and validation of the theoretical model. The measured results indicate that the proposed model approximates the measured results well. Although some deviation occurs, this is to be expected and discussed.

Keywords: bolometer, electrical characterisation, MEMS, thermal microdetector, thermo isolator

1. INTRODUCTION

Bolometers are thermal microdetectors that can operate at room temperature without the need for cooling to cryogenic levels. Furthermore, they are relatively inexpensive to manufacture. It comes as no surprise that many diverse applications have been developed from this technology, including surveillance applications like night vision, enemy surveillance and border control, as well as fingerprint scanning, pollution and fire detection, spectroscopy, and medical applications like non-contact inflammation and infection detection [1–4]. It is clear that most of the current emphasis for these devices is placed in using the bolometer in the infrared (IR) range.

This work, however, is aimed at investigating and characterising the electrical response of a temperature sensitive microbolometer device (developed initially as an uncooled bolometric infrared sensing element) at much lower frequencies, enabling novel approaches to many signal processing applications. Albeit that the characterisation of the electrical properties is not new, with various DC and AC methods introduced by Shie *et al* [5–7] and others [8, 9], this characterisation is normally performed with the extraction of a signal proportional to the IR incident radiation in mind, while minimising the Joule heating caused by biasing. However, it is necessary to heat up the membrane sufficiently by applying a high power signal in order to achieve the appropriate non-linear response to demonstrate our device functionality.

The contribution of this work lies with the choice and the structure of the device. We propose a novel bolometer infrared sensing structure that consists of dual sensing elements. These elements are thermally very closely coupled on a single supporting membrane structure, as presented in Section 2. A mathematical model is derived to compensate for the new variables that are introduced in Section 3. This model is then compared to measured results from a manufactured device in Section 5, where it was found that the measured results corresponds very well with the expected theoretical behaviour proposed by the mathematical model of the device.

2. DEVICE DESCRIPTION

The device investigated in this work has a similar structure to traditional bolometer devices, but differs in that it consists of two metal resistive elements that have been layed out in such a way to ensure high thermal coupling between the two elements. The mask design of the titanium and sacrificial aluminium cantilever device prototype consisted of four mask definitions. These define the windows in the nitride layer, the lift-off mask for metal deposition, the gold pattern for interconnect and the etch of the sacrificial layer. It was decided to use a titanium metal film resistor. The fabrication has been conducted at temperatures below 500 °C to allow future post processing onto existing CMOS readout circuitry. Typical CMOS values of the material thermal parameters as reported by [10] were used for the device design. Figure 1 shows a photograph of the above device after the manufacturing process.

3. DEVICE CHARACTERISATION

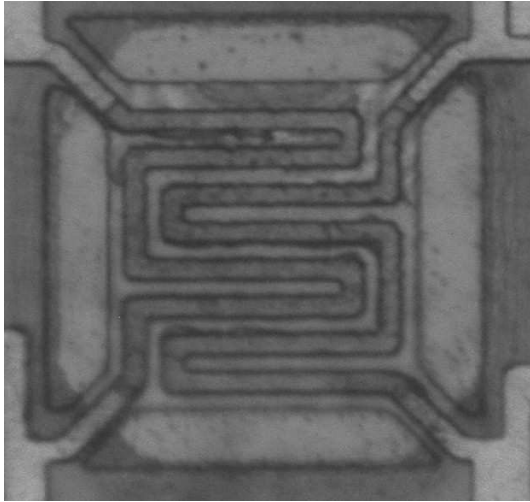It is well known that the resistance of the microbolometer

Figure 1: Photograph of the dual element bolometer device

is dependent on the difference in the active region (membrane) temperature T, or the average membrane temperature $\bar{T}$ and a chosen or reference temperature, T_0 [7]. We will assume that both the ambient temperature, T_a, and the substrate temperature, T_s, are equivalent. The microbolometer resistance of a thin-film metal device is then given as

$$R_B = R_{B0}[1 + \alpha_0(\bar{T} - T_0)] = \frac{V_B}{I_B} \qquad (1)$$

with R_{B0} the reference resistance at temperature T_0 and α_0 the associated temperature coefficient of resistance (TCR). Parameter extraction of the device resistance is done by calculating the ratio of the measured voltage and current.

The next parameter of importance to device characterisation is the thermal conductance G that consists of three components, i.e. the contributions of the gaseous (G_g), radiative (G_r) and solid (G_s) conductances. We may safely assume that the radiative and solid components are neglectable, as the gaseous component contributes at least ten times more to the total conductance than the combined conductance of the radiative and solid components at atmospheric pressure. Therefore, the conductance may be given simply as [11]

$$G = G_g + G_r + G_s \approx G_g \qquad (2)$$

If the thermal conductance is the dominating heat transport mechanism, it may be related to electrical parameters by means of a heat balance equation as [7, 12]

$$G(\bar{T} - T_a) = P_e + P_{IR} \qquad (3)$$

where P_e is the total electrical power dissipated by the membrane and P_{IR} is the absorbed IR radiation power. The total electrical power will consist of the combined contributions of the two closely coupled resistive elements, and will be significantly larger that the contribution of the

IR radiation power. Each of these resistive elements will also contribute to the change in membrane temperature, so Eq. (3) may be rewritten as

$$G[(\bar{T}_1 + \bar{T}_2) - T_a] = P_{e,B1} + \beta P_{e,B2} \qquad (4)$$

where $\bar{T}_1$ and $\bar{T}_2$ is the average temperature rise resulting from the joule heating caused by the respective electrical power components $P_{e,B1}$ and $P_{e,B2}$. We also introduce the coupling parameter, β, that allows for inefficient coupling of electrical power due to the device structure, especially the poor coupling in the device legs, as well as the top left and bottom right corners for this topology as in Figure 1. This also compensates for the crude approximation of $\bar{T} = \bar{T}_1 + \bar{T}_2$, as these temperature components will play a more complex role than mere linear summation. The temperature rise of the membrane caused by the second resistive element is calculated as

$$\Delta T_2 = (\bar{T}_2 - T_a) = \beta P_{e,B2}/G \qquad (5)$$

Notice that the additional thermal contribution of the second resistive element will also influence the temperature term of Eq. (1). Therefore, the resistance of the first bolometer element may be written as

$$R_{B1} = R_{B1,0}[1 + \alpha_0(\bar{T}_1 + \beta P_{e,B2}/G - T_0)] \qquad (6)$$

yielding an increase in resistance of the first element equivalent to $\beta \alpha_0 R_{B1,0} P_{e,B2}/G$. This is illustrated in Figure 2 where it was assumed that $\beta = 1$. The figure also serves to further clarify the differences in the contributions of $\bar{T}_1$ and $\bar{T}_2$ respectively.
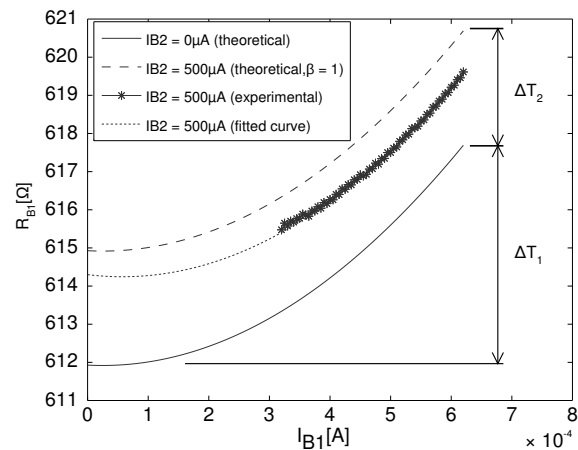

Figure 2: The effect of modulating device resistance by electric current

4.   EXPERIMENTAL SETUP

Direct knowledge of α_0 is required for successful characterisation of microbolometers. This parameter is usually determined by placing the device within an oven and measuring the device voltage at a specific reference
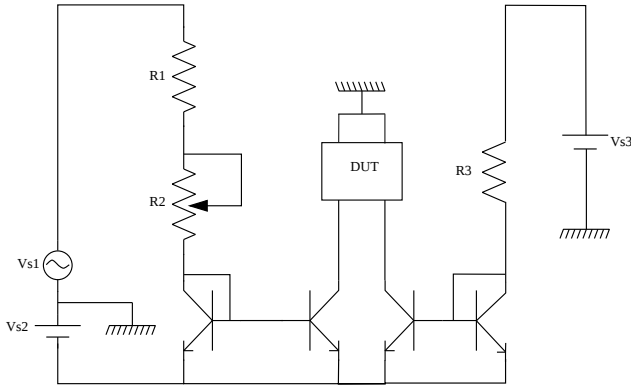
Figure 3: The implemented BJT current mirror test setup

current over a range of known temperatures [7]. From *Eq.* (2), α_0 is then determined as $\Delta R/\Delta T$. Once α_0 is known, further characterisation is possible from a DC I-V sweep, from which R_B and P_B may be determined.

Two well known methods have been published for determining the thermal conductance experimentally. The first of the two techniques require knowledge of the electrical self heating power, the device resistance and the reference parameters R_{B0} and, as mentioned earlier, α_0. R_{B0} may be determined from the I-V curve at the lowest biasing current value. Once the reference parameters have been extracted, the thermal conductance may be derived, using a simplified version of the equation presented in [7] with $T_a = T_0$, as

$$G = \frac{I_B V_B R_{B0} \alpha_0}{\frac{V_B}{I_B} - R_{B0}} \quad (7)$$

The second technique extracts the thermal conductance directly from the graph of the inverse resistance and the square of the biasing current, as discussed in [9], based on the equation

$$1/R_B = 1/R_{B0} - \alpha_0/G(I_B^2) \quad (8)$$

G may be extracted from the slope of a suitably fitted curve as $G = -\alpha/(\delta(1/R_B)/\delta(I_B^2))$.

The results of the characterisation of the first element are then used as a reference for the second part of the DC analysis. Now the second element may be stimulated by means of a controllable constant current source while the new I-V curves of the first element are measured. This has been done in incremental steps from $0\,A$ to $1\,mA$.

5. MEASURED RESULTS

The first step taken during the electrical characterisation of the device was to investigate the DC characteristics of a single element, which is then applied as a benchmark for further experimental results. All measurements were conducted with the device at atmospheric pressure. The

voltage of the device was measured with a Hewlett Packard 4155B parameter analyser by applying a DC current sweep from $0\,mA$ to $1\,mA$. The resistance can then be extracted as the ratio of the measured voltage to the applied current. Note that at this stage no current was applied to the second resistive element. This result is plotted as the bottom curve labelled $R_{B0,1} = 611.9\,\Omega$ with $I_{B2} = 0\,\mu A$ in Figure 4. It is clear that the resistance is a non-linear function of current (non-ohmic), and ranges at low currents from approximately $612\,\Omega$ to $626\,\Omega$ at $1\,mA$. A numeric approach yielded a second order polynomial curve fit given as

$$R_{B0,1} = 1.63 * 10^7 I_{B1}^2 - 800 I_{B1} + 611.93 \quad (9)$$

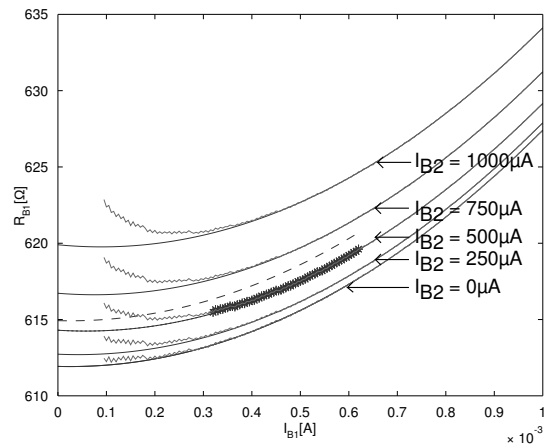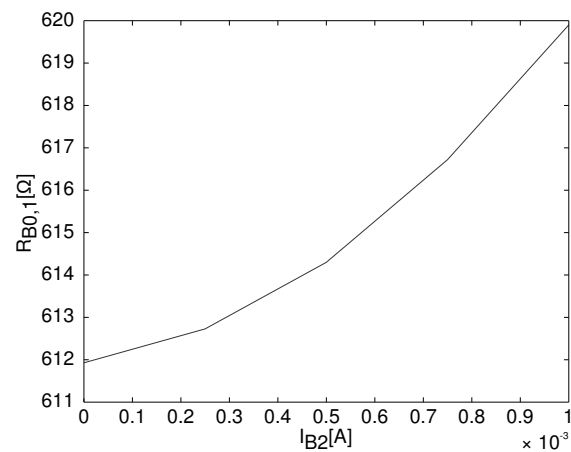Figure 4: Device current versus resistance curve for varied I_{B2} settings

Figure 5: Effect of I_{B2} on $R_{B0,1}$

Once the benchmark result has been extracted, the DC analysis can be extended by applying various controlled DC currents to the second element, and repeating the procedure discussed earlier for every control current applied. This was done for $I_{B2} = [250, 500, 750, 1000]$

μA. The experimental results are indicated, along with the benchmark result, in Figure 4. Superimposed on this figure is theoretical graph of Figure 2, where it is applicable to note again that $P_{e,B2} = I_{B2}^2 R_{B2} = 120.1 \ \mu W$ for our experimental setup at $I_{B2} = 490 \approx 500 \ \mu A$. Based on *Eq.* (6) the device should experience a $\beta P_{e,B2}/G = 5.77$ degree Celcius/Kelvin increase in temperature for 100% effective coupling ($\beta = 1$), i.e. an increase of $\beta \alpha_0 R_{B1,0} P_{e,B2}/G = 2.999 \ \Omega$ in device resistance. As seen from the experimental results of Figure 4 for $I_{B2} = 500 \ \mu W$ that $R_{B0,1} = 614.3 \ \Omega$, rather than 614.9 Ω. Using this result, we find that the average membrane temperature increased only by 4.61 degree Celcius. As mentioned previously, the losses are attributed to the less than perfect coupling in the device corners. If we assume that the thermal conductance is the same in both cases, we find that approximately 93.11 μW of the available 120.1 μW was efficiently coupled to the first device, or $\beta = 77.5\%$. The set of curve fitted equations extracted from the experimental data is given as

$$R_{B0,1} = 1.63 * 10^7 I_{B1}^2 - 800 I_{B1} + 611.93$$
$$R_{B0,1} = 1.65 * 10^7 I_{B1}^2 - 1376 I_{B1} + 612.73$$
$$R_{B0,1} = 1.68 * 10^7 I_{B1}^2 - 1923 I_{B1} + 614.30 \qquad (10)$$
$$R_{B0,1} = 1.71 * 10^7 I_{B1}^2 - 2578 I_{B1} + 616.72$$
$$R_{B0,1} = 1.73 * 10^7 I_{B1}^2 - 3102 I_{B1} + 619.90$$

but it may also be simplified by applying *Eq.* (6) to *Eq.* (9) to relate the two resistor and current values as

$$R_{B0,1} = \begin{aligned} &1.63 * 10^7 I_{B1}^2 - 800 I_{B1} + 611.93 \\ &+ 3.725 * 10^4 I_{B2}^2 R_{B2} \end{aligned} \qquad (11)$$

The effect of the control current, I_{B2}, on the resistance of the first element is plotted in Figure 5. Here we remove the effect of the device current I_{B1}.

During one of the earlier calculations we assumed that G was unaffected by the additional current, I_{B2}. One may expect that this assumption to be valid and reasonable, as the parameter is mostly dependent on material properties and surface areas which should be unaffected by biasing. However, as indicated in Figure 6 and Figure 7, it is clear that a slight increase in thermal conductance is seen as the control current is increased. This may be attributed to the fact that G is related to $\lambda_{air} A/d$, with λ_{air} the thermal conductivity of air, A the area of the device and d the distance between the membrane and the substrate. We can, therefore, conclude that a slight deformation occurs, and that the membrane to substrate distance decreases slightly due to buckling as the membrane heats up, causing the thermal conductance to increase.

6. CONCLUSION

A low-cost, CMOS compatible microbolometer process was previously established. Titanium was used as the
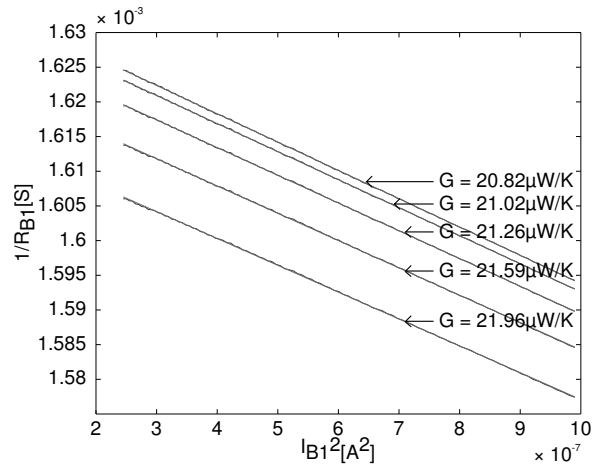

Figure 6: Thermal conductance extraction from the I_{B1}^2 vs $1/R_{B0,1}$ graph
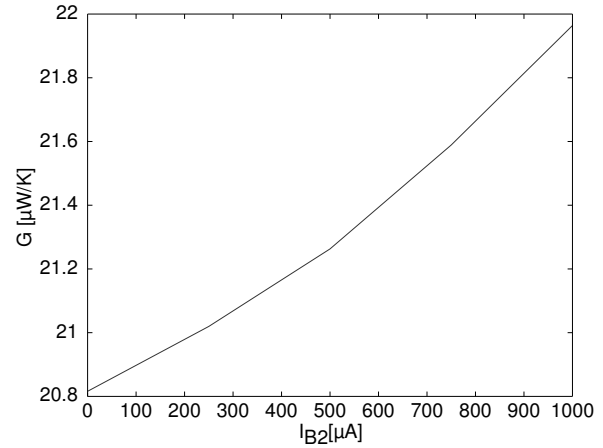
Figure 7: Effect of I_{B2} on the thermal conductance

sensitive resistive element, which is supported by silicon nitride. A novel device structure with a second resistive element was introduced. A mathematical model was then derived to include the effects of raising the membrane temperature by means of an electrical control signal via this second resistive element. A set of non-optical techniques have been reported for characterising and testing the manufactured uncooled microbolometer structures by increasing the Joule heating sufficiently for non-ohmic device operation. These techniques were employed to characterise the reference device parameters, before they were repeated to gauge the validity of the new device model. It was found that the introduced model behaves very similarly to the experimental results, albeit that a shaping factor, β was introduced to compensate for inefficient thermal coupling that has not been investigated analytically to date.

Previously published results indicated a very low frequency response in the low kilohertz range even at atmospheric

pressure, with a preliminary indication that the MEMS device has the potential to be used in novel low frequency hybrid MEMS/CMOS IC applications where traditional CMOS failed without the use of external components. The results of this work may be applied to at least two problems found with CMOS integration. As is clear from *Eq.* (6) and Figure 2, resistance modulation is achievable by means of a control current. This allows for very fine tuning of a resistor with either an on-chip or external current source. A second problem that may be addressed is that of on-chip electrical isolation. The close coupling of the thermal energy between the resistors on the same membrane enables the transfer of an AC signal on one element to the other element, as the change in membrane temperature caused by the AC signal will result in a change in the resistance of the second element. This means that the device may be used as a micro thermal isolator, transfering an electrical signal with thermal coupling while maintaining electrical isolation between its input and output. However, the exact derivation and complete AC characterisation is reserved as a seperate investigation.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] R. Wood, "High-performance infrared thermal imaging with monolithic silicon focal planes operating at room temperature," IEEE Electron Devices Meeting, Tech. Rep., December 1993.

[2] G. B. Jacobs and L. R. Snowman, "Laser techniques for air polution measurement," *IEEE Journal of quantum electronics*, vol. QE-3, no. 11, pp. 603–605, November 1967.

[3] E. F. J. Ring, "The historical development of thermometry and thermal imaging in medicine," *Journal of Medical Engineering Technology*, vol. 30, no. 4, pp. 192 – 198, July/August 2006.

[4] C. M. Travers, A. Jahanzeb, D. P. Butler, and Z. Çelik Butler, "Fabrication of semiconducting YBaCuO surface-micromachined bolometer arrays," *Journal of Mircoelectromechanical systems*, vol. 6, no. 3, pp. 271–276, September 1997.

[5] M. Ou-Yang, C.-S. Sheen, and J.-S. Shie, "Parameter extraction of resistive thermal microsensors by AC electrical method," *IEEE Transactions on Instrumentation and Measurement*, vol. 47, no. 2, pp. 403–408, April 1998.

[6] Y.-M. Chen, J.-S. Shie, and T. Hwang, "Parameter extraction of resistive thermal microsensors," *Sensors and Actuators A 55*, pp. 43–47, 1996.

[7] J.-S. Shie, Y.-M. Chen, M. Ou-Yang, and B. C. S. Chou, "Characterization and modeling of metal-film microbolometer," *Journal of Mircoelectromechanical systems*, vol. 5, no. 4, pp. 298–306, December 1996.

[8] R. Hornsey, P. Thomas, A. Savchenko, and T. Pope, "Nonoptical characterization techniques for uncooled microbolometer infrared sensors," *IEEE Transactions on electron devices*, vol. 47, no. 12, pp. 2294–2300, December 2000.

[9] P. Eriksson, J. Y. Andersson, and G. Stemme, "Thermal characterization of surface-micromachined silicon nitride membranes for thermal infrared detectors," *Journal of Mircoelectromechanical systems*, vol. 6, no. 1, pp. 55–61, March 1997.

[10] M. Von Arx, O. Paul, and H. Baltes, "Process-dependent thin-film thermal conductivities for thermal CMOS MEMS," *Journal of Microelectromechanical systems*, vol. 9, no. 1, pp. 136–145, March 2000.

[11] M. Ou-Yang and J.-S. Shie, "Measurement of effective absorptance on microbolometers," in *Instrumentation and Measurement Technology Conference, 1999. IMTC/99. Proceedings of the 16th IEEE*, vol. 1, 1999, pp. 447 –451 vol.1.

[12] F. Kohl, F. Keplinger, A. Jachimowicz, and R. Chabicovsky, "A new analytical model for detectivity prediction of resistance bolometers," in *Sensors, 2002. Proceedings of IEEE*, vol. 2, 2002, pp. 1290 – 1293 vol.2.

WAFER LEVEL PACKAGING WITH WEDGE SEAL METHOD

C. Versteeg*, J. vd Vyver** and P. van Rooyen***

* *DETEK, Denel Dynamics, Nellmapius Road, South Africa,*
E-mail: christo.versteeg@deneldynamics.co.za
** *DETEK, Denel Dynamics, Nellmapius Road, South Africa*
E-mail: james.vd.vyver@deneldynamics.co.za
*** *mHealth Inc., San Diego, CA, USA*

Abstract: Wafer level packaging could reduce the cost of MEMS based sensors through simplified processing and inexpensive components. A novel vacuum tight seal referred to as the wedge seal method is proposed in this paper. The seal consists of a silicon wedge forced into a pliable material (typically a metal) that is attached to the component wafer. The wedge-seal addresses some of the requirements of micro-bolometer packaging in that it provides a vacuum tight seal at low temperatures with tolerance to surface finish and topography. The concept was evaluated with regards to manufacturability, material suitability and performance. A demonstrator model was manufactured with Complementary Metal–Oxide–Semiconductor (CMOS) compatible processing equipment and procedures. The wedge effectively penetrated the metal base and showed promising leak rates.

Key words: Wedge seal, anisotropic etching, vacuum-tight and low temperature.

1. INTRODUCTION

Micro Electro Mechanical Systems (MEMS) are comprised of small fragile structures which typically sense mechanical movement, pressure differences or electromagnetic radiation.

Each sensor type has unique packaging requirements with regards to sensor exposure and isolation. In the case of a micro bolometer, suitable thermal isolation and sufficient Infrared Radiation (IR) exposure both affect the performance of the sensor. Thermal isolation is effectively ensured with a vacuum inside the package around the sensor and IR radiation requires a transparent medium that allows the radiation to penetrate the packaging and reach the sensor.

Many successful bolometers have been developed and deployed commercially, and due to a competitive market the unit price is under continued downward pressure. Since packaging could account for up to 75% [1] of the device costs, new low cost packaging techniques are required.

MEMS have gained much support as it builds on the well-established silicon wafer processing industry. Current packaging research focuses to build on the bulk silicon processing capabilities and to perform packaging on wafer level, before separating the devices, known as Wafer Level Packaging (WLP) or Zero Level Packaging.

Current WLP bonding techniques include: electric field assisted adhesion (anodic bonding), direct binding and joining with an intermediate layer (metal-eutectic/solder etc) [2].

Currently only plasma activated Silicon Direct Bonding (SDB) allows for a vacuum compatible joint that requires processing at temperatures below 200 °C. But this process has further limitations in that it is very sensitive to particle contamination and requires a surface roughness of less than 2 nm [2].

In this paper we propose a wedge-seal method that promises to be a low cost, low temperature (< 200 °C), Ultra High Vacuum (UHV) seal joint (Leak rates <1E-15 mbar.l/s) that can accommodate surface roughness variations in the order of 1 μm. The manufacturing requirements are also compatible with Complementary Metal–Oxide–Semiconductor (CMOS) processing equipment and procedures.

2. WEDGE SEAL DESCRIPTION

The wedge-seal joint builds on the Con Flat (CF) vacuum seal that uses two knife edges which are forced into a copper ring, with the knife edges deforming the copper ring and thus forming the seal [3]. We deployed the same principle on wafer level with a knife edge etched into a silicon wafer and a base wafer that was coated with a metal layer. The lid with the knife edge was then forced down onto the base with the metal layer so that a vacuum joint was formed (see Figure 1).
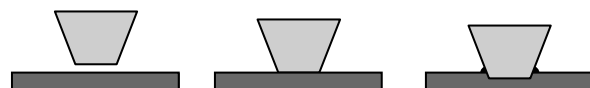

Figure 1: Wedge forming seal with metal base

3.   IMPLEMENTATION

This concept was evaluated by processing a lid with a wedge seal protruding from the surface and a flat base covered with indium.

The wedge component of the seal is formed from a Silicon wafer that has a <100> crystal orientation and by using an anisotropic wet etch (see Figure 2). The seal layout must be composed of straight lines due to this process with proper corner compensation structures.
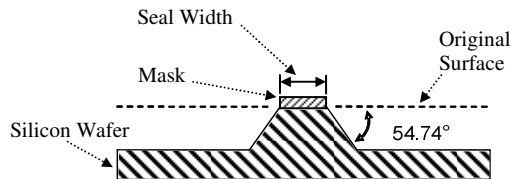

Figure 2: Anisotropically etched edge features

Photolithographic processes are used to define a mask for a nitride passivation layer, during plasma etching. The nitride then provides the masking material for the anisotropic etch. The photolithography limited the width of the seal face and the minimum width obtained was 16 um. The width could be significantly reduced by using reticles and a wafer stepper. The anisotropic etch consists of potassium hydroxide (KOH), water and isopropyl alcohol (IPA). The wedges were etched until they protruded a height of 27 μm (see Figure 3).
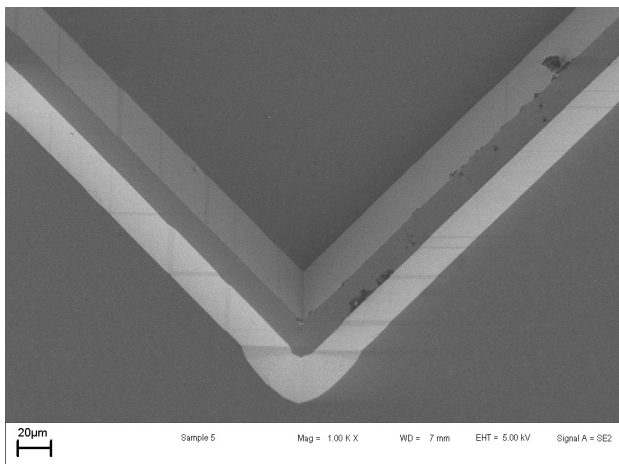

Figure 3: Scanning Electron Microscope (SEM) image of wedge structure

The base layer should provide a pliable interface to the wedge, which needs to deform the material to conform to the shape of the wedge and thereby create an intimate joint. Since the pliable base layer is exposed to the vacuum inside the cavity it must be stable at very low pressures and thus a metal is preferred. Various metals are suitable for this application but only those compatible with wafer processing are considered for the proposed wedge-seal. Some examples include aluminium, indium and gold. For our experiments a layer of indium was used to test the concept since it is very ductile at ambient. This was deposited with an evaporation process to a thickness of 1.5 µm.

The metal wedge joint has numerous factors that influence its design and processing. The wedge itself could be varied in its depth and seal width. The wall angles are fixed due to the anisotropic etching characteristics. The seal width depends on the processing capability and should be as thin as possible.

The metal base has two major properties to consider – these are the thickness of the layer and the type of material. The thickness should be at a minimum to save costs. The material itself has thermal (melting point) and mechanical characteristics (yield stress and ductility). The melting point of the material must be considered when reviewing all the post processes. If this temperature is exceeded the seal will be destroyed. The bonding process itself has three main variables namely; force, temperature and duration, apart from all the basic mechanical requirements of flatness and parallelism.

The ideal wedge seal characteristic is one where sufficient material has been displaced along the entire wedge section to form a vacuum tight seal. The wedge should however not push through the metal layer so as to damage the underlying structures or touch the base wafer and possibly destroy the seal integrity.

Metal-wedge joint processing involves the contact of the wedge seal surface with the metal layer and forcing the wedge into the metal layer so that the layer is permanently deformed (see Figure 4). Material ductility is a function of temperature and lowers until the material reaches its melting point where all mechanical strength is lost [5]. An increase in temperature should reduce the force required to penetrate the base material and increase the impression depth.
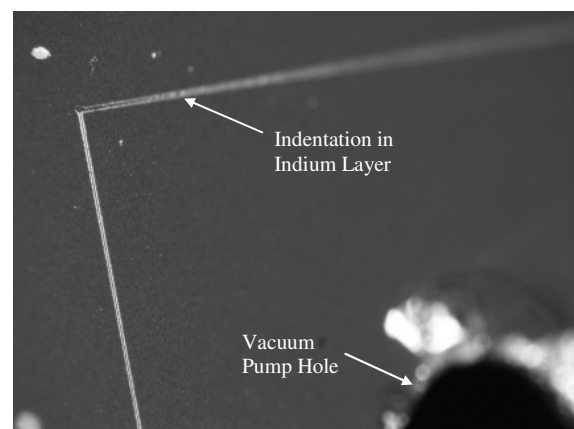

Figure 4: Indented indium surface, showing hole for vacuum pumping

As the force is increased the pressure is raised above the material yield strength, the material deforms and yields sideways as it makes way for the wedge. The material pushed away is further deformed by the slanted wedge sides which forces it further to the sides. This ensures that the seal surface and the sides make intimate contact. The vacuum joint is then established between the newly exposed metallic material and the wedge itself.

4.   SEAL TEST SETUP

The wedge seal performance was based on the hermiticity of the joint between the lid and the base wafer. The joint should be vacuum tight which is verified with a helium leak tester.

The connection to the leak test equipment requires a hole in the silicon base wafer which can be drilled with either a tungsten-carbide or diamond-impregnated drill bit. Both drill processes induce micro-cracks in the single crystal structure and they were etched to remove stress concentrations (see Figure 5). After the base wafer is covered with a metal layer and the lid with the wedge seal joined, the aluminium pump post is attached with vacuum grease and the test performed.
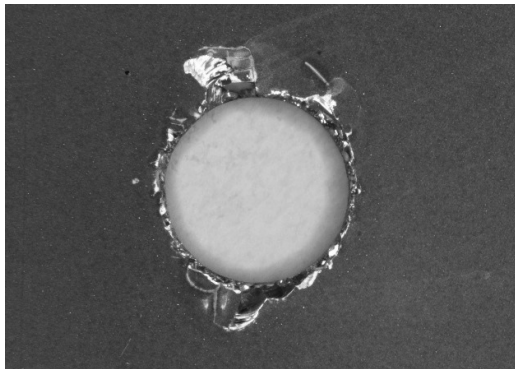
Figure 5: Hole drilled in silicon wafer

The wedge seal was defined on the lid wafer as a 1 x 1 cm square which was then bonded over the hole in the base wafer to create a cavity with an access portal. The bonding of the lid and the base was performed in a flip chip bonding machine. This allowed for the accurate and consistent bonding of the lid to the base wafer by controlling; the lid and base temperature, the force applied and the duration of the wait, ramp-up, hold and ramp-down phases. The flip chip bonding machine system calibration procedure ensures parallelism before each bonding run. After the lid was bonded onto the base wafer, the aluminium pump post was joined to the wafer with a thin layer of vacuum grease.
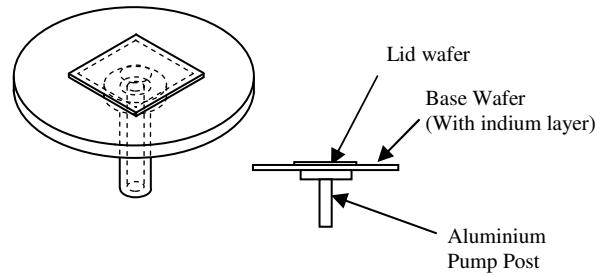
Figure 6: Wedge seal test layout

The wedge seal bonds were formed at different temperatures and forces. Initially the lids were removed to evaluate the uniformity of the indentation and to determine if the wedge had pinched through the metal layer. Depression uniformity was visually inspected under a microscope and pinching evaluated with a Near Infra-Red (NIR) light source and a camera sensitive to this bandwidth. The camera detects IR light that is not blocked by indium which should cover the entire seal length. If the camera senses IR light then the indium was totally removed or pinched so thin that it becomes transparent. Both cases constitutes as a poor bond.

Secondly after the indentation seems acceptable the packaged component is leak tested with helium to determine the hermiticity of the joint.

5.   TEST RESULTS

The 16 µm seal width showed even indentation from the wedge, this confirms that the flip chip bonder has been suitably calibrated with regards to parallelism and that the flatness of the bonding tools is acceptable.

The indentation depth could be varied significantly with different temperature and force combinations. High temperatures (> 120 °C) and high forces (>5000g) resulted in some of the seal sections pinching through with seal a width of 16 µm.

The indents of the 16 µm showed that a high force (7000 g) with a low temperature (40 °C) did not penetrate through the indium (1.5 µm thick). Wedge profiles with a seal width of greater than 16 um showed poor and uneven indentation even with forces up to 7000g. Higher forces should improve the indentation but it is not feasible.

The wedge seal hermiticity was tested after bonding (@ 40 °C, 7000 g) and showed leak rates of 1.2E -10 mbar l/s with Helium. This leak rate is close to the lowest detectable leak rate of the equipment and due to the vacuum, outgassing could become significant.

The joint has poor mechanical strength since no chemical bonds are induced.
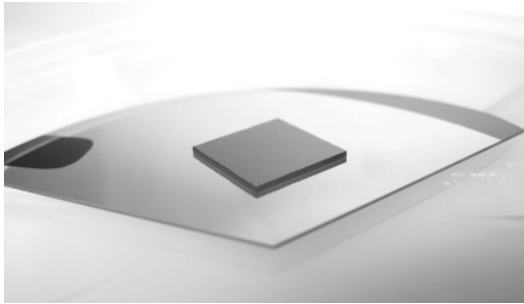
Figure 7: Bonded wedge seal lid joined to the indium covered bottom wafer.

6.　CONCLUSIONS

The wedge-seal concept was proposed as a wafer level packaging, lid to base, vacuum tight seal.

The components for this seal consists of a wedge, etched from a silicon wafer with a specific crystal orientation and a base covered with a pliable material namely indium. Both components were manufactured with CMOS compatible materials and processes. Forming the wedge seal requires mostly force and heating in between parallel and flat surfaces.

The wedge-seal performance relates to its ability to ensure that the vacuum, within the sealed cavity, isn't compromised and is quantified with a leak rate measurement.

The wedge seal components were manufactured from silicon wafers and bonded together. The silicon wedge effectively penetrated the indium base.

The vacuum was drawn through a portal in the base wafer. The leak rate test experiment showed promising results although the measurements were limited by the equipment and surface outgassing.

The wedge showed poor mechanical strength and for a sensor application it must be supported to withstand vibration and shock.

7.　REFERENCES

[1] J.H. Lau, C.K. Lee, C.S. Premachandran and Y Aibin: *Advanced MEMS Packaging*, McGraw Hill Companies Inc, USA , Chapter 7, pp. 341-342, 2010.

[2] V. Lindroos, M. Tilli, A. Lehto and T. Metooka: *Handbook of Silicon Based MEMS Material and Technologies,* Elsevier Inc, UK, First Edition, 2010.

[3] D.M. Hoffman, B. Singh, J.H. Thomas III: *Handbook of Vacuum science and technology,* Academic Press Ltd, UK, First Edition, 1998.

[4] L. Finn, R. Mulholland and G.J. Gibbon: "Design and implementation of a prototype computer based rehabilitation device for the lower extremities", *SAIEE Africa Research Journal*, Vol. 95 No. 1, pp. 29-32, March 2003.

[5] R. P. Reed, C. N. McCowan, R. E Walsh, L. A. Delgado and J. D. McColskey: "Tensile Strength and Ductility of Indium", *Material Science and Engineering A,* Volume 102, pp. 227-236, 1988

SPECTRAL MEASUREMENT AND ANALYSIS OF SILICON CMOS LIGHT SOURCES

A.W. Bogalecki*, M. du Plessis**, P.J. Venter** and C. Janse van Rensburg*

* *INSiAVA (Pty) Ltd, P. O. Box 14679, Hatfield, Pretoria, 0028, South Africa*
** *Carl and Emily Fuchs Institute for Microelectronics, Dept. of Electrical, Electronic & Computer Engineering, Corner of University Road and Lynnwood Road, University of Pretoria, Pretoria 0002, South Africa. E-mail: alfons.bogalecki@gmail.com*

Abstract: The emission spectra of *pn*-junction and punch-through (PT) carrier injection silicon (Si) CMOS light sources were measured at various current densities and temperatures. In contrast to the narrow-band forward-biased junction spectrum, that peaks around 1.1 μm (1.1 eV), the reverse-bias spectrum was found to extend from about 350 nm (3.4 eV) to about 1.7 μm (0.7 eV) covering the UV, Vis and NIR regions. Since the photon energy decreases with increasing wavelength, the significant NIR radiation implies that the quantum conversion efficiency of Si avalanche light sources is appreciably higher than previously reported. Calculating the photon flux at the emission source within the Si against photon energy allowed the deduction and quantification of the physical light emission processes with respect to silicon's electronic band structure. Intra-conduction-band (c-c) electron (e^-) transitions seem to be the dominant physical mechanism responsible for the wide avalanche spectrum.

Key words: Avalanche, CMOS, electroluminescence, light source, photon emission, silicon, spectrum, temperature

1. INTRODUCTION

1.1 Problem statement

Efficient on-chip Si light sources are desirable for wide-ranging applications that include optical data transmission, CMOS-integrated micro displays and micro-opto-electro-mechanical systems (MOEMS). Although Si electroluminescence (EL) was observed as early as 1955 [1], its high-speed capability [2] and long-term reliability [3] are established and the INSiAVA[1] group has improved the external power efficiency (EPE) of Si light sources by a factor 21 [4], relatively little is known about the physical EL processes and their dependence on factors such as current and temperature. In fact, there is still disagreement about the physical mechanisms responsible for the Si EL ([1], [3], [5] - [8]).

1.2 Background

While forward-biased Si *pn*-junctions are known to radiate strongly in a narrow band around 1.1 μm (1.12 eV), this light is due to the slow recombination of injected diffusing carriers, which limits the modulation speed of forward-biased *pn*-junction light sources to a few 100 kHz. For this reason, EL improvement research within the INSiAVA project focuses on wide-band, but fast-switching avalanching light source configurations that include *inter alia* reverse-biased *pn*-junctions and carrier-injection devices like punch- and reach-through ([9], [5]) Si light sources.

1.3 Approach

In this work, the EL spectra of two such Si light sources are investigated. Figure 1 shows the element layouts of the *pn*-junction and PT devices that were manufactured in 400 and 1760 element matrices respectively in the commercial austria micro systeme 0.35 μm four-metal CMOS process without any post-processing.
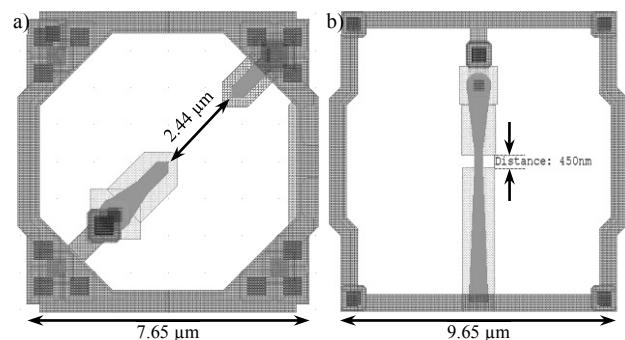
Figure 1: Element layouts of the a) n^+p-junction and b) n^+pn^+ PT CMOS light sources measured in this work.

The n^+ in *p*-substrate junction is pointed towards a p^+ substrate contact to increase the electric field and current density at the tip. The PT device consists of two n^+ diffusions that are spaced a specific distance from each other in the *p*-substrate so that PT carrier injection can occur [9]. The large open area around the light sources is necessary to prevent reflections or shielding from the surrounding metal interconnect layers.

These two CMOS light source matrices were measured at various device currents and temperatures with a Si UV-

CMOS light source power spectra at 90 mA

Figure 2: Spectrometer-measured power spectra of the avalanche, forward-biased and PT CMOS light sources.

Vis-NIR and an InGaAs NIR spectroradiometer that covered the wavelength ranges 177 nm – 1.1 μm and 894 – 1729 nm respectively. The two resulta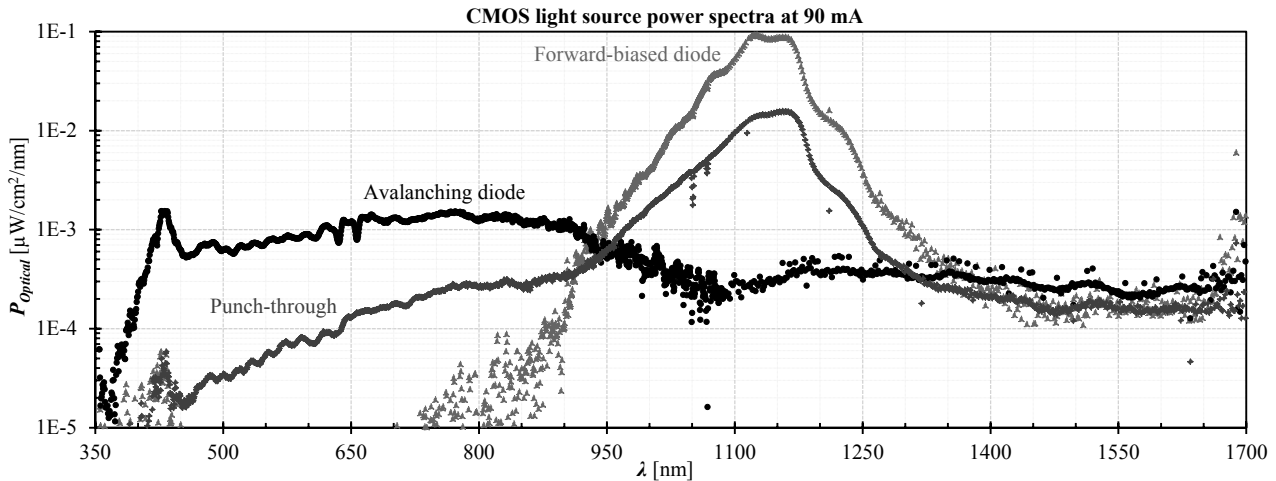nt spectra were then merged by taking into account the relative responses of the two spectroradiometers. By selectively filtering the spectroradiometer pixel noise and the wavelength-dependant Si-SiO$_2$-air stack interference and taking into account transmission losses through the Si-SiO$_2$-air stack, Si absorption and effective solid angles as seen by the spectroradiometers, the photon emission rates per photon energy at the light generation site were then calculated.

2. MEASUREMENT RESULTS

2.1 Spectral Analysis

Figure 2 shows the typical external spectra of the avalanche, forward-biased and PT light sources measured at matrix currents of 90 mA. Although these measurements were measured at room temperature, the typical IC package temperature of the light source test-chips at such currents was about 31 ºC. The spectra in Figure 2 have not been filtered for noise or the wavelength-dependant Si-SiO$_2$-air stack interference and

clearly show the strong 1.13 μm peak of the forward-biased junction and weaker wide-band emission of the avalanching junction. The avalanche spectrum shows three distinct peaks: a narrow peak at 427 nm and two wide-tailed distributions around 761 nm and 1.25 μm. The PT light source combines attenuated spectra of both the forward-biased and avalanche light sources, but still has about twice the integrated optical power of the avalanche source.

A representative example of the resultant photon generation rates of the avalanche, forward-biased and PT CMOS light sources at room temperature are shown in Figure 3. These spectra experienced identical filtering and loss compensation and clearly show the 1.09 eV peaks of the forward-biased and PT sources. Avalanche peaks are visible around 0.9, 1.03, 1.41 and 2.89 eV.

Forward-bias and PT: The 1.09 eV peaks of the forward-biased and PT sources correspond to the shortest phonon-assisted (PA) conduction to valence (c-v) transitions in the energy bands of Si (Figure 4) and allow determining the temperature at the light generation site within the Si. At 300 K this band gap of Si is about 1.12 eV.

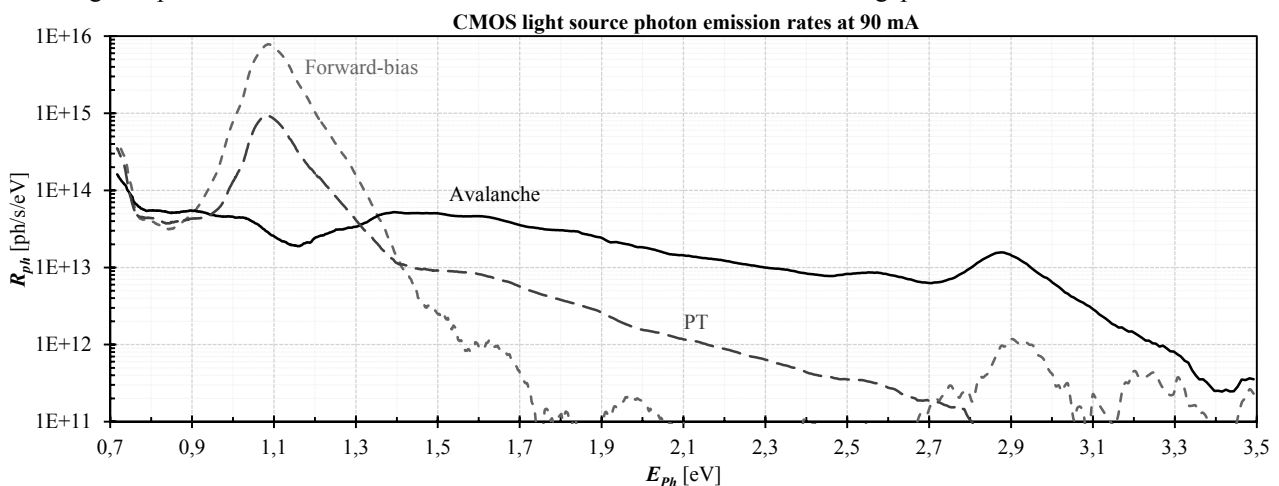**CMOS light source photon emission rates at 90 mA**

Figure 3: Calculated photon emission rate spectra of the avalanche, forward-biased and PT CMOS light sources.
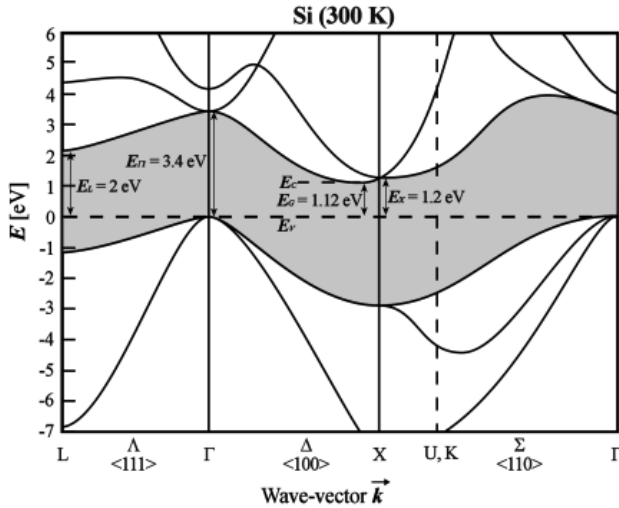
Figure 4: Si energy band diagram at 300 K [10].

The direct band gap energy of Si decreases with increasing temperature according to the experimentally determined relation ([10], pg. 15)

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \qquad (1)$$

where the fitting parameters have the values

$E_g(0)$= 1.1675 eV
α = 481.5 μeV/K
β = 645.5 K

According to (1) the temperature at the origin of the 1.1 eV peaks in the measured forward-biased and PT light sources was approximately 167 ºC.

Avalanche: Three possible radiative transitions might be responsible for the avalanche EL. Direct (non-PA) c-v transitions, which would show a emission peak at a photon energy of about 3.4 eV, PA c-v transitions, where emissions would peak around 1.12 eV plus the maximum of the e^- energy distribution in the conduction band, and c-c transitions, where the photon emission peak would coincide with the maximum e^- energy in the conduction band. One possible origin of the avalanche emission peak at 2.89 eV would therefore be direct (non-PA) c-v transitions. Considering that this gap requires an energy of about 3.4 eV at room temperature (Figure 4), the ΔE_{Ph} of at least 0.51 eV suggests a local temperature of about 1342 ºC, which seems impossibly high. If this mechanism is indeed responsible for the 2.89 eV avalanche emission peak, then raising the temperature of the device by 100 ºC should decrease the direct energy gap by about 40 meV. The temperature measurements in section 2.3 showed at most a shift of 7 meV, which seems to suggest that the spectral position of the 2.89 eV avalanche emission peak is independent of the temperature-variable band-gap and that this peak does not originate from direct c-v transitions.

Figure 5 shows simulated bulk e^- energy distributions in the conduction band for various electric field strengths. According to [11], the e^- distribution peak at 2.89 eV is due to the second maximum of density of states of electrons in the conduction band at this energy.
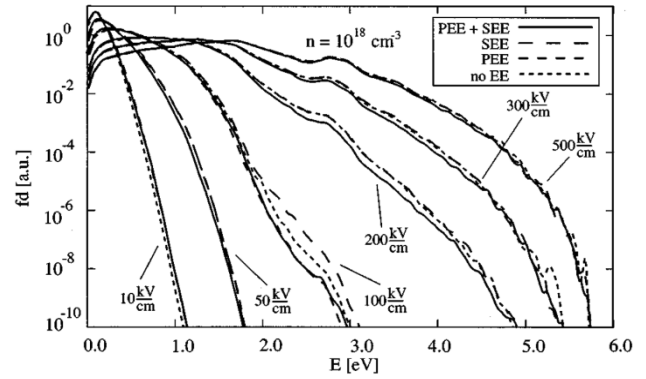

Figure 5: Electron energy distributions for different electric fields simulated by Mietzner *et al* [11].

The avalanche emission spectrum in Figure 3 with its 1.41 and 2.89 eV peaks and relative photon emission rate distribution for $E_{Ph} > 1.17$ eV agrees very well to a case where the electric field strength would be between 300 and 500 kV/cm. The measured breakdown voltage of about 9.8 V and estimated background doping of about $1.3 \cdot 10^{17}$ cm^{-3} suggests that the maximum electric field in the avalanche light source could reach strengths of up to 650 kV/cm. Although the magnitude of the electric field in the *pn*-junction device of Figure 1 is not yet known, the pointed n$^+$ diffusion is expected to increase the local electric field strength dramatically higher than 100 kV/cm. Since the n^+ diffusion is surrounded by the CMOS LOCOS field oxide, which through its curved "bird's beak" shape also shapes the diffusion vertically into a sharp tip, the electric field at the tip is enhanced even more.

Considering alternatively that the emission peak at 1.41 eV is due to hot e^- c-v transitions would imply an e^- distribution peak at 1.41 − 1.08 = 0.33 eV. Figure 5 consequently suggests an electric field strength of about 100 kV/cm, which seems too low for *pn*-junctions in reverse-bias avalanche breakdown.

The avalanche emission rate in Figure 3 drops by about 0.82 of an order of magnitude (factor 6.6) for an energy difference of about 1.05 eV between 1.41 and 2.46 eV. From Figure 5 a 400 kV/cm e^- energy distribution drops about by 0.85 of an order of magnitude (factor 7) from 1.5 eV to 2.5 eV. The 100 kV/cm distribution function drops from its peak at about 0.35 eV to 1.35 eV by more than two orders of magnitude (factor 110). The observation that the high energy tails of the avalanche emission agree very well with the e^- distribution function at about 400 kV/cm, but not at all at 100 kV/cm also reinforces the view that the avalanche emission is due to conduction to conduction (c-c) and not c-v transitions.

Similarly, the avalanche EL photon emission distributions around the peaks at about 0.91 eV, which has also been measured by *inter alia* [6] on NMOSFETs, and 1.02 eV can be explained as originating from direct inter-conduction band transitions in lower electric field regions along the flat edge of the n^+ diffusion and a distance away from the sharp diffusion corner. According to Figure 5 the average field strength away from the sharp corner is between 100 and 200 kV/cm.

The agreement between e^- energy distribution in Figure 5 and avalanche emission rate in Figure 3 seems to suggest that the electric field strength in the avalanche CMOS light source ranges between 100 kV/cm and about 400 kV/cm at the pointed diffusion corner. As also postulated by [6] and [8] the dominant avalanche light generation mechanism seems to be due to direct c-c e^- transitions.

2.2 Source current dependence

Figure 6 depicts the optical powers of the avalanche, forward-biased and PT CMOS light sources per µA integrated over the 350 – 990 nm and 990 – 1700 nm regions against varying source element current.
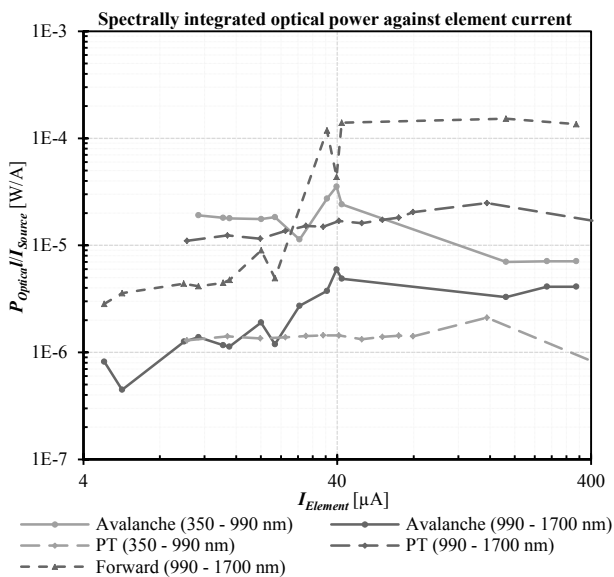

Figure 6: Measured integrated optical powers per source current against element current.

In effect, Figure 6 illustrates the effectiveness of each light source type in converting the current flowing through it into optical power. The overall trends are that the integrated optical power between 350 and 990 nm decreases with increasing source current, while the NIR radiation between 990 and 1700 nm increases with source current. The avalanche and forward-biased optical powers increase with source currents up to 40 µA after which the 350 - 900 nm radiation decreases with current while the 900 - 1700 nm optical remains relatively constant for driving currents above 40 µA/element. Although the exact reason for this phenomenon is still unknown, measurement error seems excludable as all three

independently measured data sets of the avalanching and forward-biased junction emissions show a similar maximum around 40 µA/element. The most probable reason for this behaviour is speculated to be the pointed diffusion shape of the diode CMOS light source in Figure 1 a). The 350 – 900 nm and 900 – 1700 nm PT power efficiency components as function of elemental source current exhibit similar trends with maxima around 156 µA/element.

Figure 7 shows the photon emission rate R_{Ph} per Source current I_{Source} of the avalanche CMOS light sources for four different source currents.
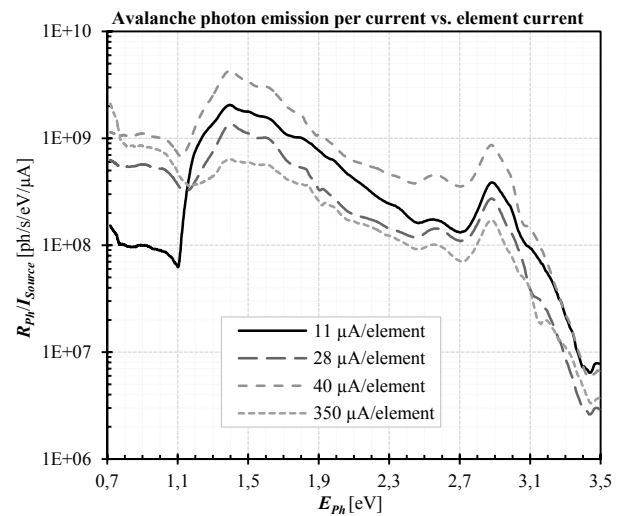

Figure 7: Measured avalanche photon emission rate R_{Ph} per source current for different source currents.

At low source currents, the avalanche light source is most efficient at emitting photons with energies between 1.17 eV and 3.2 eV with a peak at about 1.43 eV. As the source current increases, the photon emission rate above 1.17 eV decreases while photons with energies up to 1.17 eV are emitted at a growing rate. For element currents between 28 and 40 µA, photons above 1.17 eV experience an increasing emission rate. For $I_{Element} > 40$ µA the overall photon emission rate then decreases again, but this decrease is more pronounced for $E_{Ph} > 1.17$ eV.

One possible explanation for the decreasing R_{Ph} for $E_{Ph} > 1.17$ eV and increasing R_{PH} for $E_{Ph} < 1.17$ eV for $I_{Element} \leq 28$ µA is that for low currents the largest current flow and consequently largest e^- density occurs through the localized high electric field at the sharp tip of the n^+ diffusion. Forcing a higher device current increases the proportion of the current flowing across the flat diffusion edges on either side of the tip, which due to the lower electric field strength present there causes the increasing low-energy photon emission rate.

2.3 Temperature dependence

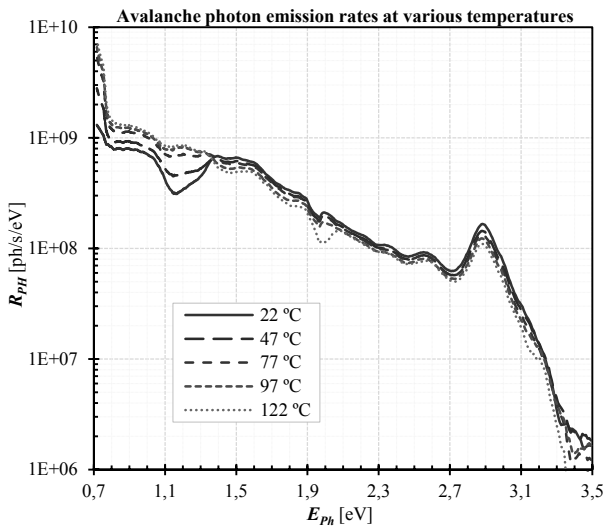*Avalanche:* Figure 8 depicts the measured temperature dependence of the avalanche photon emission rate.

Figure 8: Measured avalanche photon emission rates for five selected IC package temperatures from 22 - 122 ℃.

As illustrated in Figure 9, after being almost equal at a temperature of 22 ℃, R_{Ph} increases almost linearly with temperature for $E_{Ph} < 1.37$ eV and fills the low-temperature dip at 1.18 eV in Figure 8, while the emission rate above 1.37 eV decreases with increasing temperature. At temperatures above about 100 ℃ the rate of increase with temperature of the 0.7 – 1.37 eV component and consequently the total photon emission rate slows down and flattens off.
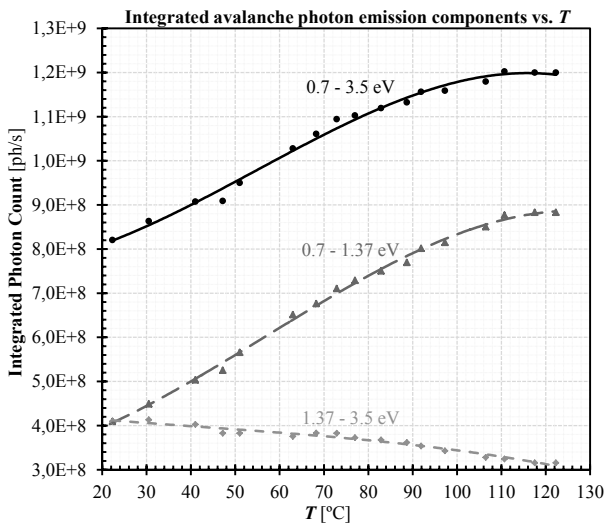

Figure 9: Integrated avalanche spectral components against temperature.

The negative temperature coefficient of $R_{Ph}(E_{Ph})$ for $E_{Ph} > 1.37$ eV was explained by [7] to be caused by factors such as decreasing impact ionization with increasing temperature. With increasing temperature, the avalanche multiplication rate decreases due to increased phonon scattering, which decreases the number of hot carriers travelling through the depletion region and consequently decreases radiative recombination. A secondary factor is

the increasing absorption of Si with increasing T. Although this theory explains the avalanche emission behaviour with temperature for $E_{Ph} > 1.37$ eV, it fails to explain the increasing EL with T for $E_{Ph} < 1.37$ eV.

Closer examination of the R_{Ph} peaks in Figure 8 reveals that neither of the peaks at 0.9 eV or 2.9 eV shift with temperature. Although it could also be an artefact of the changing R_{Ph} distribution around the peak at about 1.47 eV, it seems to be shifting by about 50 meV to higher photon energies when the temperature changes from 22 to 122 ℃.

Punch-through: Figure 10 shows the measured temperature dependence of the PT light source photon emission rate.
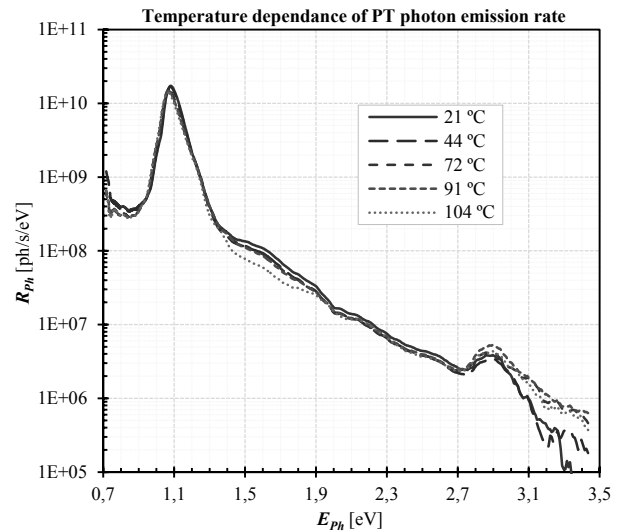
Figure 10: Measured PT photon emission rates for five selected IC package temperatures from 21 to 104 ℃.

The maximum of the PT emission spectra shifts from 1.093 eV at 21 ℃ to 1.089 eV at 104 ℃. According to (1) this shift of − 4 meV implies that the temperature at the site of light generation has increased from 110 ℃ to 163 ℃, a ΔT of 42 ℃ while the IC package temperature has increased by 83 ℃. The slight difference in temperature changes is attributed not only to unavoidable thermal gradients between the IC package and the light generation site, but also that longer temperature forcing settling periods and temperature measurements closer to the light generation site might be needed.

Although not measured, it is expected that the temperature behaviour of the 1.1 eV peak emission of the forward-biased *pn*-junction light source is similar to the abovementioned peak shift of the PT source.

Further investigation of Figure 10 reveals that while the PT photon emission rate falls with temperature for $E_{Ph} < 2.7$ eV, it increases above 2.7 eV. This behaviour is completely opposite to the avalanche emission temperature behaviour and currently unexplained.

3. CONCLUSION

Spectroradiometric measurements of avalanching, forward-biased *pn*-junction and PT CMOS light sources in the wavelength range 350 – 1700 nm (0.7 – 3.5 eV) revealed their distinctly different photon emission spectra for different current densities and temperatures. While the forward-biased *pn*-junction showed the typical strong but narrow photon emission peak around 1.1 eV due to PA c-v carrier transitions, the wide-band avalanche spectrum seems to be mainly generated by direct hot e^- c-c transitions. The width and shape of the avalanche spectrum seems to relate to the electric field strength through which the electrons drift. Introducing high electric fields by utilizing sharp diffusion corners enhances the EL in the visible wavelengths dramatically. The PT light emission spectrum consists of a scaled combination of the forward-biased and avalanche spectra and while the spectral behaviour with current density is comparable to the respective dominant emission spectra of the forward-biased and avalanching *pn*-junction light sources, its temperature behaviour is distinctly different.

In general, the NIR emission for energies in the range 0.7 – 1.17 eV (1.1 – 1.7 μm) increases with increasing device current for all CMOS light sources. The photon emission in the energy range 1.25 – 3.5 eV (350 – 990 nm) of all CMOS light sources decreases with increasing current.

While NIR emission in the range 0.7 – 1.37 eV (0.9 – 1.7 μm) of the avalanche light source increases with temperature, photon emission in the energy range 1.37 – 3.5 eV (350 – 904 nm) decrease with increasing *T*. The temperature behaviour of the PT CMOS light source seems to be inverse to the avalanche source, but with a zero temperature coefficient near $E_{Ph} \approx 2.7$ eV (459 nm).

The overall result is that for visual display and Si detector applications, avalanching *pn*-junctions are the most efficient UV-Vis-NIR Si light sources and should be operated at low currents and temperatures. For applications where Si wave-guiding or long-wavelength detectors (like InGaAs) are desired, the PT and forward-biased Si light sources are superior and should be operated at high currents and low temperatures. The decisive difference between the PT and forward-biased light sources is that while the forward-biased source is an order of magnitude more efficient than the PT source, but the PT device is electrically modulateable above a few 100 kHz. Table 1 summarizes these recommendations with respect to Si light source requirements.

Table 1. Recommended Si CMOS light source parameters as a function of requirements.

Desired		Recommended			
		Type	Geometry	$I_{Element}$	T
UV-Vis-NIR	More UV-Vis	Avalanche	Sharp	Low	Low
	More NIR		Flat/Round	High	High
NIR	> 100 kHz, less effective	PT	Flat/Round	High	Low
	More effective, ≤ 100 kHz	Forward-bias			

4. REFERENCES

[1] R. Newman, W.C. Dash, R.N. Hall and W.E. Burch: "Visible light from a Si p-n junction", *Phys. Rev.*, Vol. 98 A, pg. 1536, 1955.

[2] A. Chatterjee, B. Bhuva and R. Schrimpf: "High-speed light modulation in avalanche breakdown mode for Si diodes", *IEEE Electron Device Letters*, Vol. 25, No. 9, pg. 628, September 2004.

[3] A. Chatterjee and B. Bhuva: "Accelerated stressing and degradation mechanisms for Si-based photoemitters", *IEEE Transactions on Device and Materials Reliability*, Vol. 2, No. 3, pg. 60, 2002.

[4] A.W. Bogalecki and M. du Plessis: "Design and manufacture of quantum-confined punch through SOI light sources", *SPIE Photonics West symposium on Photonic Integration: Optoelectronic Integrated Circuits XII*, San Francisco, USA, pp. 76050B-1, January 2010.

[5] M. du Plessis, P.J. Venter and A.W. Bogalecki: "Using reach-through techniques to improve the external power efficiency of silicon CMOS light emitting devices", *SPIE Photonics West symposium on photonic integration: Silicon Photonics V*, San Francisco, USA, Paper 7606-37, January 2010.

[6] N.C. de Luna, M.F. Bailon and A.B. Tarun: "Analysis of near-IR photon emissions from 50-nm n- and p-channel Si MOSFETs", *IEEE Transactions on Electron Devices*, Vol. 52, No. 6, pp. 1211, 2005.

[7] M. Lahbabi, M. Jorio, A. Ahaitouf, M. Fliyou and E. Abarkan: "Temperature effect on electroluminescence spectra of silicon p–n junctions under avalanche breakdown condition", *Materials Science and Engineering*, Vol. B86, pp. 96–99, 2001.

[8] J. Bude: "Hot-carrier luminescence in Si", *Physical Review B*, Vol. 45, No. 11, pp. 5848, March 1992.

[9] P.J. Venter, M. du Plessis, M. Goosen, I.J. Nell and A.W. Bogalecki: "Improved efficiency of CMOS light emitters in punch-through with field oxide manipulation", *IEEE International Conference on Microelectronics 2010*, Cairo, Egypt, pg. 36, December 2010.

[10] S.M. SZE and K.K. Ng: *Physics of Semiconductor Devices*, Third Edition, John Wiley & Sons, Hoboken, New Jersey, 2007.

[11] T. Mietzner, J. Jakumeit, and U. Ravaioli: "Local iterative Monte Carlo analysis of electron–electron interaction in short-channel Si-MOSFETs", *IEEE Transactions on Electron Devices*, Vol. 48, No. 10, pg. 2326, October 2001.

CMOS AVALANCHE ELECTROLUMINESCENCE APPLICATIONS – MICRODISPLAY AND HIGH SPEED DATA COMMUNICATION

M.E. Goosen*, M. du Plessis**, P.J. Venter**, A.W. Bogalecki*, A.C. Alberts* and P. Rademeyer

* INSiAVA (Pty) Ltd, P.O. Box 14679, Hatfield, 0028, Pretoria, South Africa
E-mail: marius.goosen@up.ac.za
** Carl and Emily Fuchs Institute for Microelectronics, Dept. of Electrical, Electronic & Computer Engineering, Corner of University Road and Lynnwood Road, University of Pretoria, Pretoria 0002, South Africa

Abstract: All-CMOS silicon light sources, although not the choice semiconductor process for light generation, offer the possibility of large scale manufacturing, integration with digital and driver electronics as well as a wide operating temperature range. These advantages do however come at a cost of reduced efficiency, but offer significant cost advantages inherent when using a standard CMOS technology. This paper presents two applications of standard CMOS integrated light sources. A fully functional microdisplay utilising avalanche electroluminescence for visible light generation and implemented in a completely standard 0.35 μm CMOS technology is presented. The microdisplay has an operating temperature range of -50 to 125 °C, which cannot be achieved by competing microdisplay technologies. Utilising the same silicon light sources, a 10 Mb/s optical communication link is established operating at a BER of less than 10^{-12}. The data communication link presented in this paper constitutes the fastest all-silicon data communication link achieved thus far.

Key words: Microdisplay, Optical interconnect, silicon light emission, silicon photonics

1. INTRODUCTION

Avalanche electroluminescence in the visible spectrum from *pn*-junctions was reported for the first time in 1955 [1]. Due to the low internal quantum efficiency of indirect bandgap silicon light emitters, the main industry approach was to use III-V element as well as organic material emitters to accomplish light emission for applications ranging from long-haul optical communication links to microdisplays.

In a microdisplay market dominated by OLED and LCD technologies, a fully integrated CMOS microdisplay, fabricated in the 0.35 μm *austriamicrosystems* (AMS) CMOS process with no process modifications or post-processing, has a number of attractive advantages which could be exploited to lead to interesting applications in this environment. The advantages include the use of a robust and mature technology which is CMOS, resulting in favourable cost, configurability, wide operating temperature range and reliability. The reliability of silicon light sources has been shown through accelerated stress tests resulting in negligible intensity variations under all conditions of aging [2]. Other silicon-based microdisplays, implemented using porous silicon, were deemed quite promising for near-the-eye (NTE) applications where low luminance levels are sufficient for direct viewing [3, 4].

Short distance communication, such as chip-chip, board-board and rack-rack interconnects, require less optical power to maintain an acceptable BER. Hence low efficiency silicon light emitters become a viable solution,

although at a cost of higher power dissipation [5]. The low efficiency of the indirect bandgap silicon light emitters, if optimised, can be utilised in high speed all-silicon optical interconnects [6, 7].

Section 2 describes the silicon light emitters utilised for both the all-CMOS microdisplay as well as the high speed data optical transmission presented in Section 3 and Section 4 respectively.

2. SILICON LIGHT EMITTERS

The silicon light emitters utilised in this research are formed in bulk silicon with the creation of reverse biased *pn*-junctions. The reverse biased *pn*-junctions, operated in the avalanche breakdown region, are shaped to form a point, hence aptly named point sources, in order to enhance the local current density and increase the external power efficiency (EPE).

Light directing structures utilising the back-end-of-line (BEOL) stack direct the light generated in the bulk toward the SiO_2-air interface improving the light extraction efficiency (LEE) by an average factor of 2.18 [8]. The BEOL-stack reflectors can be designed to either focus the light into a narrow pencil-beam for coupling into an orthogonally aligned optical fibre for communication or to generate a wide beam width for a microdisplay with a wider half-power viewing angle. Additional techniques to further increase the light extraction efficiency of punch-through and reach-through devices have also been demonstrated [9], indicating the

possibility of increased optical output power and higher data rates.

The emission spectrum, ranging from 400 nm up to 1000 nm allows for both viewing applications, such as a microdisplay, and short-distance optical communication. Figure 1 illustrates the emission spectrum of a silicon light emitter. The photopic luminance function is also shown as indication of the typical spectrum that the human eye perceives under well-lit conditions.
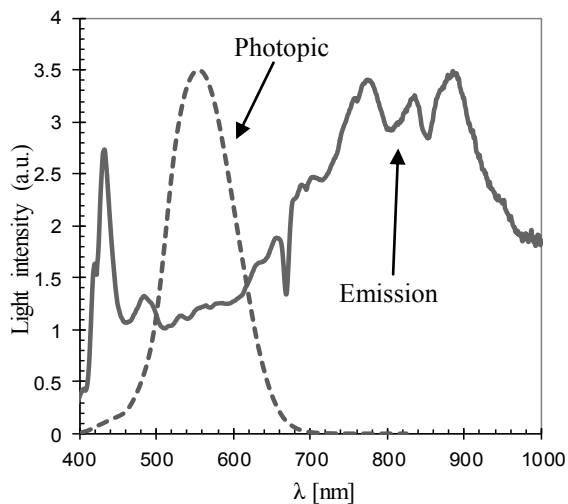

Figure 1. Emission spectrum of the silicon light source driven under avalanche conditions as well as the photopic luminosity function.

The photopic luminosity function as shown in Figure 1, is relatively narrow compared to the emission spectrum of the silicon light source. Approximately 10 % of the emitted optical power falls in the photopic region seen by the eye. For optical data communication, the majority of the spectrum is covered by commercially available silicon avalanche photodiodes (APDs), usually with peak responsivity at around 800 nm.

3.   ALL-CMOS MICRODISPLAY

3.1 Microdisplay design

Each pixel within the 8x64 pixel dot-matrix display is comprised of 30 individual point sources, each with their own BEOL reflector. The minimum pixel dimension is hence determined by the BEOL reflector aperture, which is approximately 5 μm. Each 50 μm pixel was connected to a row-select line as well as a column-select line. The column-select lines were combined into a 6-64 decoder to reduce the amount of interface interconnects. Figure 2 illustrates the scanning methodology utilised in the microdisplay.

A column is activated through the applied 6-bit word (only one column active at any given moment), and any of the individually addressed rows can be activated, activating and driving the chosen pixel. For sequential

scanning, the active column will move one column right or left, and the rows are used to choose the active pixels. This process continues, allowing flicker-free image formation.
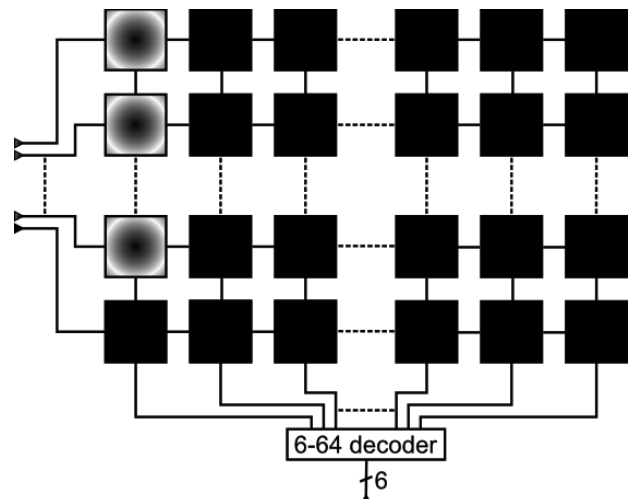
Figure 2: Illustration of the scanning methodology employed for image formation.

The half power viewing angle of the display is dependent on the type of BEOL reflector implemented on each of the point sources within the pixel. Figure 3 illustrates the 3D radiation pattern of the microdisplay.
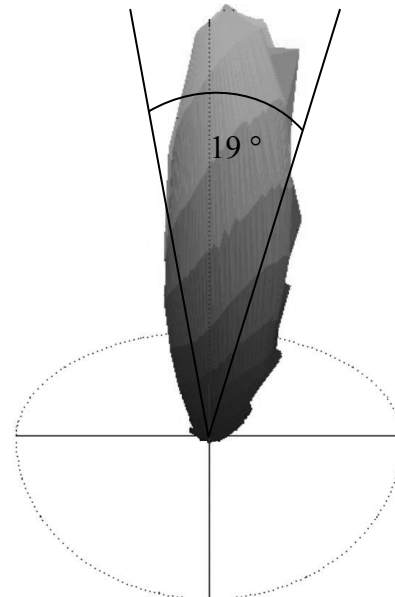
Figure 3. 3D radiation pattern of the implemented microdisplay. The measured half power viewing angle is 19 degrees.

3.2 Results

As indicated in Figure 3, a half power viewing angle of 19 degrees was achieved, although it should be noted that the viewing angle is determined by the variable BEOL reflector. The half power angle was measured using a photomultiplier tube (PMT) with a small circular aperture, in conjunction with a goniometer, rotating the

microdisplay in fixed steps over the two spherical axes around the PMT. A viewing angle of 19 degrees is ample for NTE applications where the display is held in a fixed position relative to the eye. It can also be noted that the radiation pattern is slightly slanted due to the general direction to which the planar light sources are aligned.

The prototype developed depends on external control for the rendering of characters and images. Grey scale images can be produced by either using pulse width modulation (PWM) or current controlled pixels. Since the microdisplay is formed using the emissive silicon light sources, a true black can be achieved with an inactive pixel.

Figure 4 illustrates an image formed on the all-CMOS microdisplay, followed by a grey scale bar implemented with the use of PWM.

Figure 4: Image formed on the all-CMOS microdisplay clearly showing the illuminated pixels in contrast to the true-black (off) pixels [10].

The image as shown in Figure 4 was formed without the use of an image intensifier. The CMOS display as shown has sufficient optical power in the photopic region for direct viewing with a simple magnification lens.

4. HIGH SPEED DATA TRANSMISSION

4.1 Communication link setup

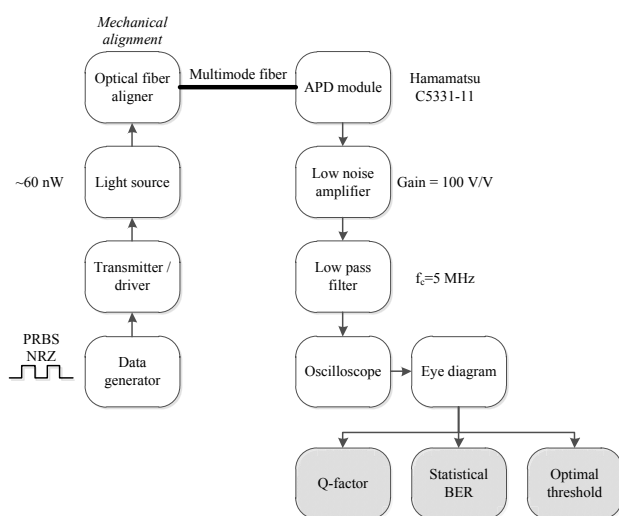Figure 5 illustrates the optical communication link setup used in the data rate demonstration.

Figure 5: High speed optical communication setup.

A data generator was used in the creation of pseudo-random NRZ sequences and applied to the transmitter/ driver. The transmitter/ driver implemented a differential driving scheme in order to reduce electromagnetic radiation due to large switching currents. One of the directly modulated light sources, capable of delivering up to 60 nW$_{peak}$ of optical power, was mechanically aligned with a low attenuation step index multimode fibre (MMF) with a core-diameter of 600 μm. The light source array was matched to the core diameter of the fibre, resulting in a total of 3520 point sources. The light source array was biased and driven for maximum extinction ratio to maximise the achieved signal to noise ratio (SNR). The MMF exhibits a low attenuation in the region spanning from 400 nm up to 2000 nm, hence in the silicon emission spectrum.

A commercial APD, *Hamamatsu 5331-11*, was used to convert back to the electrical domain before amplification and filtering. The APD has an active area diameter of 1 mm, hence no lenses are required to couple from the fibre to the photodiode. It further has a peak responsivity at 620 nm of 0.42, a transimpedance gain of 40 kΩ and a noise density referred to the input of approximately 0.5-1 pW/√Hz.

The low noise amplifier was implemented using 2 stages, each with a 10 V/V gain, and a total noise density of 20 nV/√Hz. The filter cut-off frequency was chosen as half the data rate, hence 5 MHz, suppressing additional frequency components introducing inter-symbol interference.

4.2 Results

An oscilloscope was used, as recommended by ITU-T G.976 [11] and discussed in [7], in order calculate a quality factor (Q-factor) which relates the signal power to the noise power, hence a statistical SNR, within the eye diagram. From the calculated Q-factor the statistical BER and optimal sampling threshold is determined. The eye diagram achieved after filtering is depicted in Figure 6.
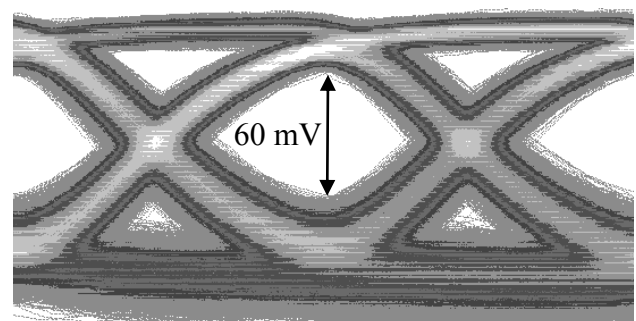

Figure 6: Eye diagram of the high speed optical communication link at a data rate of 10 Mb/s.

The increased amount of noise in the lower level of the eye diagram in Figure 6 is due to the increased amount of shot noise generation with a larger incoming optical signal. The APD utilised has an inverting gain

transimpedance amplifier, hence the low level in the eye diagram is in fact the high level of the optical signal. The statisctical information extracted from the eye diagram is depicted in Figure 7.
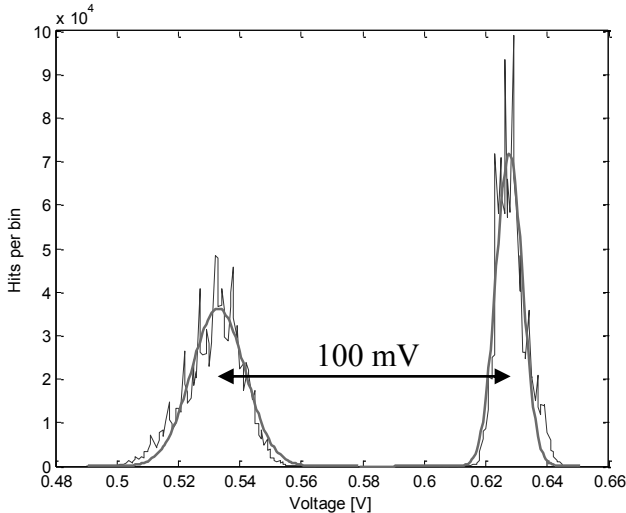

Figure 7. Histogram of the vertical eye opening.

The noise in the vertical eye amplitude follow a Gaussian distribution since there are no channel bandwidth limitations closing the eye in a deterministic fashion. Hence, a Gaussian distribution was fitted to the noise in order to optain the noise in the eye as is shown in Figure 7.

The eye diagram in conjunction with the statistical information gathered by the histogram, results in a calculated Q-factor of 7.22, an optimal sampling threshold of 0.59 V and a statistical BER of 2.64×10^{-13}. The achieved SNR for the 10 Mb/s optical data communication link equates to 26.05. It is noted that the data rate achieved is not limited by the switching capability of the silicon light sources, but rather by the achievable optical output power and the noise contribution of the APD determining the SNR and BER. Silicon light sources have been shown to switch in excess of 350 MHz, through E-O-E testing [7], and in excess of 20 GHz using streak camera techniques for emission resulting from hot carrier luminescence [12].

Another fact worth mentioning is the low LEE which is currently achieved. The LEE is defined as the percentage of light which can be directed to exit the surface. Currently, only about 1 % of the generated optical power exits the surface of the die, of which another 10 % is lost due to the fibre coupling efficiency. Hence, taking these factors into account approximately 6.6 μW is generated on the CMOS die.

The previous fastest reported all-CMOS optical communication link achieved a data transfer rate of 1 Mb/s at a BER of 10^{-14} [7]. In order to achieve a data transfer rate of 10 Mb/s the amount of input optical power has to scale with the square root of the bandwidth increase. Hence, a factor 3.3 improvement is necessary

for scaling from 1 Mb/s to 10 Mb/s. The BER curves at the different data rates are illustrated in Figure 8. The demonstrated data transfer rates are indicated by encircled points.
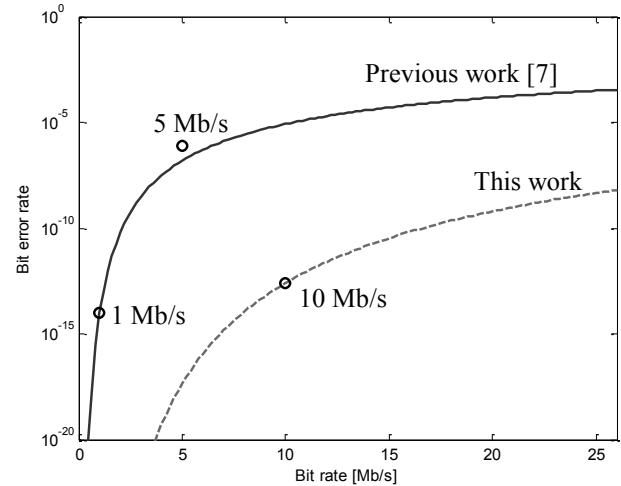

Figure 8. Theoretical BER versus bit rate curves. Solid line – results presented in [7]. Dashed line – result presented in this paper.

5. CONCLUSION

The electrical performance parameters of the prototype all-CMOS microdisplay is presented in Table 1.

Table 1. CMOS dot matrix display electrical characteristics

Parameter	Typical	Units
Si source breakdown voltage	9.2	V
Supply voltage	12	V
Current per pixel	5.1	mA
Maximum current (all pixels on)	41	mA
Power consumption (all pixels on)	0.492	W
Refresh rate	42.8	Hz
Pixel pitch	50	μm
Display active area	1.32	mm^2
Half power viewing angle	19	°

The cost benefit and flexibility coupled with the robustness of the mature technology which is CMOS, may lead to interesting applications in the microdisplay environment. One interesting possibility is the combination of an optical communication link and a microdisplay, since the silicon light sources are inherently fast switching devices [13]. With an increase in device efficiency, the power consumption can be reduced while maintaining the same luminance. Although not comparable to OLED and LCD technologies in terms of power consumption and resolution, the wide operating temperature range and implementation flexibility should allow all-CMOS microdisplays to be applied in certain niche markets.

An all-silicon optical communication link was demonstrated operating at a data transfer rate of 10 Mb/s

while maintaining a statistical BER of less than 10^{-12}, a tenfold increase in data transfer rates over previous reported all-silicon communication links [7]. The light sources were implemented in standard 0.35 μm CMOS process, with no post-processing, obtaining results exceeding other more exotic silicon-based technologies [5]. With increased EPE and LEE of the silicon light emitters, multi-Gb/s optical interconnects for short-haul communications remain feasible and attractive.

6.   ACKNOWLEDGMENTS

7.   REFERENCES

[1] R. Newman, "Visible light from a silicon p-n junction", Phys. Rev., 100(2), pp. 700–703, 1955.

[2] A. Chatterjee and B. Bhuva, "Accelerated Stressing and Degradation Mechanisms for Si-Based Photoemitters", IEEE Trans. on Device and Materials Reliability, 2(3), pp. 60-64, 2002.

[3] A. Smirnov, A. Berezovik, P. Poznyak, V. Labunov, S. Lazarouk, "Silicon based LED microdisplays: The experience of design and manufacturing," Proc. of SPIE Vol. 6637, XV International Symposium on Advanced Display Technologies, 663703, 2007.

[4] P. Jaguiro, P. Katsuba, S. Lazarouk, M. Farmer, A. Smirnov, "Si-based emissive microdisplays," Physica E 41, pp. 927–930, 2009.

[5] S. Sayil, "Avalanche breakdown in silicon devices for contactless logic testing and optical interconnect", Analog Integrated Circuits and Signal Processing, 56(3), pp. 213-221, 2008.

[6] M. du Plessis, H. Aharoni and L.W. Snyman, "Silicon LEDs fabricated in standard VLSI technology as components for all silicon monolithic integrated optoelectronic systems", IEEE Journal on selected topics in Quantum Electronics, 8(6), pp. 1412-1419, 2002.

[7] M.E. Goosen, P.J. Venter, M. du Plessis, I.J. Nell, A.W. Bogalecki and P. Rademeyer, "High-speed CMOS optical communication using silicon light emitters", Paper 7944-32, Session 9, Proc. SPIE 7942, SPIE Photonics West, San Francisco, USA, 26 January 2011.

[8] A.W. Bogalecki, M. du Plessis, P.J. Venter, M.E. Goosen and I.J. Nell, "Integrated optical light directing structures in CMOS to improve light extraction efficiency", Proc. of the International Conference on Microelectronics, Cairo, pp. 168-171, 19-22 Dec. 2011.

[9] P.J. Venter, M. du Plessis, I.J. Nell, A.W. Bogalecki, and M.E. Goosen, "Improved efficiency of CMOS light emitters in punch through with field oxide manipulation", Proc. of the International Conference on Microelectronics, Cairo, pp. 36-39, 19-22 Dec. 2011.

[10] P.J. Venter, A.W. Bogalecki, M. du Plessis, M.E. Goosen and I.J. Nell, "CMOS dot matrix microdisplay", Proceedings of the Advances in Display Technologies conference, Proc. SPIE 7956, Paper 79560Y; SPIE Photonics West, San Francisco, California, USA, 22-27 January 2011.

[11] ITU-T Recommendation G.976, "Digital transmission systems – Digital sections and digital line system – Optical fibre submarine cable systems", Series G: Transmission Systems and Media, Digital Systems and Networks, 1997.

[12] A. Chatterjee, B. Bhuva and R. Schrimpf, "High speed light modulation in avalanche breakdown mode for Si diodes", IEEE Electron Device Letters, 25(9), pp. 628-630, 2004.

[13] J.J.D. McKendry, R.P. Green, A.E. Kelly, Z. Gong, B.G.D. Massoubre, E. Gu, M.D. Dawson, "High-speed visible light communications using individual pixels in a micro light-emitting diode array", IEEE Photonics Technology Letters, 22(18), 2010.

AB INITIO FREQUENCY MEASUREMENT AND CHARACTERISATION OF FREQUENCY DOUBLED FIBRE LASER UTILISED FOR PRECISION OSCILLATORS

J.P. Burger*, C. Matthee* and R. Kritzinger*

**National Metrology Institute of South Africa (NMISA), Private Bag X34, Lynnwood Ridge, 0040, South Africa. E-mail: JBurger@nmisa.org*

Abstract: An *ab initio* measurement of a free running, frequency doubled erbium fibre laser is made with an optical frequency comb in concert with an internally calibrated wavemeter. The measurement is validated via a Monte Carlo uncertainty analysis. The operation and characteristics of a new high performance optical metrology source for an optical frequency standard is also verified in the process.

Key words: Optical frequency standards, optical frequency comb, metrology, photonics.

1.   INTRODUCTION

The National Metrology Institute of South Africa (NMISA) is starting to develop precision atomically-referenced optical oscillators for metrology and technological applications. Stabilised optical oscillators offer high precision, exceeding those of traditional microwave standards like the common caesium (Cs) and hydrogen standards, and also have application in ultrahigh stability RF oscillators when down converted with an optical frequency comb [1]. Optical frequency combs can be seen as an optoelectronic gearing system to convert the optical oscillators with frequencies around typically 150 THz to 600 THz to RF frequencies in the 10's of MHz range (or vice versa), while retaining the same fractional uncertainty of the frequency during the process. It can also be used to optically synthesise a frequency comb from a microwave reference for measurement and other uses, with basically the same fractional frequency accuracy, as the microwave reference.

Optical oscillators can use inherently stable lasers, to access transitions in atoms for referencing/stabilising to provide simple, but highly stable molecular clocks (like iodine stabilized green laser based on Nd:YAG or Yb-fibre technology[1, 2]). Furthermore more exquisite optical clocks utilise lasers that access ultrahigh Q transitions in atoms/ions to provide fractional frequency uncertainties $<10^{-17}$ [3]. Furthermore such oscillators offer significant advantages in distribution of frequency (RF or optical) over low loss fibre [4] or high performance antenna arrays [5], and have many other industrial and reference measurement applications.

The NMISA is developing its optical frequency systems on a robust and compact optical telecommunications fibre technology platform coupled with advanced electronics. In the last decade high performance single polarisation single frequency fibre lasers have become available in two forms, namely distributed feedback (DFB) lasers [6] and distributed Bragg reflector (DBR) lasers [7] that both use in-fibre Bragg grating technology. These lasers can interface with components originally developed for the optical fibre telecommunications field, which are usually quite robust, and include modulators, couplers, amplifiers, Faraday rotators and in-fibre filters. All of these components are spliced together in systems with a semi-automated fusion splicer, and results in a manufacturing process quite similar to the soldering together of electronic components on a circuit board. Even atomic cells can now be built in optical fibre [8]. This fibre and advanced electronics technology platform therefore provides the opportunity to build very robust, compact, transportable and potentially cheap systems for certain niche markets, which are sensitive to these attributes.

The work described in this paper is concerned with measurements on some parts of the optical metrology source subsystem that is currently being developed. The optical metrology source subsystem contains many parts, but two of the important parts are a DBR fibre laser which is frequency doubled in an external nonlinear waveguide that gives a near-infrared line. Some of the power and wavelength characteristics of this source are presented. Furthermore the frequency characteristics that is measurable on a Ti:Sapphire (Titanium doped Sapphire) laser optical frequency comb together with a wavemeter is shown. In the process it is shown that absolute optical frequency measurements can now be undertaken at the NMISA, potentially down to the ~Hz level (with the proper microwave reference) at optical frequencies of ~400 THz, without prior knowledge of the optical frequency, as an input to the process.

2.   METHODOLOGY

Optical frequency combs [9] are now the standard way to link optical frequency measurements with established microwave references. Nonetheless there are now only three labs that operate such combs in the Southern Hemisphere (Brazil, Australia and South Africa) according to our knowledge, due to the relatively high complexity and cost. New fibre combs [10] have now started to make the technology quite robust though.

The comb is synthesized via a mode-locked laser and uses a highly nonlinear photonic crystal fibre for spectral broadening to reach the required spectral bandwidth of an octave. The measurement of an unknown laser (unit under test or UUT) with frequency f_{UUT} is done by detecting a beat with a beat frequency f_b against an Nth comb element of the optical frequency comb with frequency

$$f_N = N f_{rep} \pm f_0 \tag{1}$$

where N is the comb element number, f_{rep} is the laser repetition rate and f_o is the offset frequency. f_{rep} is changed via the cavity mirror spacing on the laser's Fabry Perot cavity and f_o via the laser's intracavity dispersion via translation of thin glass prisms inside the cavity. If all the abovementioned quantities are known and measured the UUT's frequency is therefore simply determined to be

$$f_{UUT} = N f_{rep} \pm f_o \pm f_b \; . \tag{2}$$

Therefore five parameters have to be known to determine the UUT's frequency, that is N, f_{rep}, f_o and f_b and the two signs of the last numbers. In practice N is the hardest to determine and is the point of discussion in the next 3 sections.

2.1 Resolving comb elements with comb only: General remarks

Resolving the unknown comb number N in such optical frequency measurements is still challenging due to a typical close comb element spacing of 200 MHz at optical frequencies. In order to avoid ambiguity in the measurement of N an implicit frequency pre-knowledge of the laser frequency below <0.25 parts per million (ppm) is implied at a wavelength of λ~800 nm, when a single direct measurement is undertaken. There has been some work to increase the laser repetition rates in both solid state and fibre lasers via shorter laser cavities [11], high harmonic modelocking [12] or frequency filtering in Fabry Perot cavities [13], but repetition rates around the ~200 MHz is still dominating for a range of practical reasons.

In practise f_o and f_{rep} is controlled via proportional-integral (PI) feedback control electronics in the Menlo Systems comb (model FC-8004), to ensure f_b is within the detection bandwidth of bandpass electronics in the system (a few MHz at ~30 MHz). The abovementioned frequency comb, which is also utilised at the NMISA has a repetition rate of around 200 MHz, and both the repetition rate and offset frequency are referenced to a Cs frequency standard (Agilent 5071A). f_o is fixed to an absolute value of 20 MHz. This frequency fixture limit on f_0 and f_b and the adjustment limitations of f_o and f_{rep}

therefore constrains the operable parameter space of the instrument, but with an advantage of sensitive detection due to component optimization. Alignment of ultrashort pulses in propagation vector, polarization, wavelength, and time domains is necessary to detect the offset beat f_o. Typically the UUT vs. comb beat is quite weak because it involves the mixing of an external laser (UUT) with typically ≤1 mW power (else the avalanche photodetector will saturate) with a comb element that has only a few tens of nW in power. Nevertheless, with careful alignment the latter UUT vs. comb beat will result in a ~ 30 dB signal to noise ratio that is typically necessary for successful frequency counting. Furthermore the sign of f_b is determined by slightly changing the repetition rate, while observing the change in beat frequency, f_b. The sign of f_o is changed by adjusting the intracavity dispersion.

2.2 Procedures for resolving comb elements with comb

A "simple" procedure that could theoretically resolve the number N, which has been attempted at the NMISA, is detailed here to give an idea of some of the difficulties that are encountered in resolving the exact frequency. This procedure works by keeping the comb element number N constant, but tuning f_{rep} so that f_o can alternate between positive and negative.

Assume that the frequency of the unknown laser, f_{UUT}, stays constant between performing two sets of measurements; one set with the offset frequency, f_0, tuned to be negative and the second with the offset frequency positive. During the measurements, care was taken to keep the comb element, N, the same for both measurements.

$$f_{UUT} = N f_{rep1} - f_{01} + f_{b1} \tag{3}$$

$$f_{UUT} = N f_{rep2} + f_{02} + f_{b2} \tag{4}$$

Since both the laser frequency and the comb element are unknown, one must be eliminated from the two equations. By setting Equations 3 and 4 equal, and adding uncertainty contributors (assuming dominant uncertainties are coming from the UUT vs. comb beat notes) one obtains

$$N + \Delta(N) = \frac{(f_{01} + f_{02} + f_{b2} - f_{b1} + \Delta(f_{b2} - f_{b1}))}{f_{rep1} - f_{rep2}} \; , \tag{5}$$

where $\Delta(x)$ denotes the uncertainty in a number x. But the uncertainties in f_{b2} and f_{b1} have the same Gaussian-like distribution, and therefore uncertainty in the difference is $\Delta(f_{b2} - f_{b1}) = \sqrt{2}\,\Delta(f_{b2}) = \sqrt{2}\,\Delta(f_{b1}) = \sqrt{2}\,\Delta(f_b)$ [14]. From Equation 5 the uncertainty in N is therefore

$$\Delta(N) = \frac{\sqrt{2}\,\Delta(f_b)}{f_{rep1} - f_{rep2}} , \qquad (6)$$

where the difference between the two repetition frequencies $f_{rep1} - f_{rep2}$ is small; typically ~ 20 Hz. To obtain $\Delta(N) < 0.5$ with 95% confidence, the uncertainty (standard deviation) in individual beat frequency measurements must be < 4 Hz. Some of the lasers, as discussed in this paper have uncertainty even up to the MHz level on the timescales of the measurements, and therefore this methodology becomes unfeasible for a large number of lasers utilised within metrology environments. It should be noted that a two-step method has been suggested and tested by others for measurement of lasers with up to 10's of kHz of uncertainty [15], but this still falls short for really noisy lasers (some of which might be quite accurate though).

2.3 Method for measurements at the NMISA

So far these measurement procedures that use the comb only for frequency information could not be fully implemented at NMISA, especially as lasers with relatively large uncertainties were being measured. These uncertainties stem from processes that include frequency modulation of the UUT that is used to lock the laser to an atomic/molecular cell (like on older national length standards), or other intrinsic noise or drift of the lasers. Nevertheless accurate average values are needed on some of these sources as some of them are even used in real metrological setups (for example for dimension) where traceability to primary time standards is necessary. Therefore some initial rough measurement is needed, especially for the lasers that have been used so far, to obtain an estimate on the comb mode number for the subsequent frequency comb measurement.

Therefore a wavemeter with an internal laser reference is utilised at the NMISA for performing a premeasurement. Previous attempts at using a non-referenced Fabry Perot wavemeter were unsuccessful; even these meters drift over time and smaller than 0.25 ppm absolute accuracy is needed at all times. This need is now fulfilled by utilizing a 0.2 ppm absolute accuracy wavemeter from Bristol Instrument (series 621). This wavemeter has an internal stabilized Helium Neon laser for continuous self-calibration. It was necessary to show though that the wavemeter could indeed resolve the actual comb element numbers in a statistically significant way to have confidence in the measurement process and effectively validating it. A new metrology source (as under development at NMISA) was utilised in the process, and these measurements are discussed in the next section.

3. MEASUREMENTS ON DOUBLED LASER

The experimental setup is shown in Figure 1 below. The modelocked laser system produces ~50 fs slightly chirped pulses, which are injected in highly nonlinear fibre to generate combs spanning more than an octave. There are two separate supercontinuum combs – one is passed through the nonlinear interferometer to generate f_o and the other is used for measuring f_b. The two separate but coherent combs enable one to separately optimise the magnitude of the offset frequency beat and UUT vs. comb beat. The device under test was a single longitudinal mode DBR optical fibre laser (NP Photonics series RFLM) made in erbium fibre that is fully fibrerised with polarisation maintaining (PM) fibre. This source is injected into a frequency doubling waveguide after an attenuator. The connectors A and B were disconnected at various stages of the experiment to monitor power on calibrated InGaAs or Si power measurement heads (Ophir), to observe the spectra on a scanning grating spectrometer (Anritsu) or to measure the wavelength (Bristol Instruments 621).
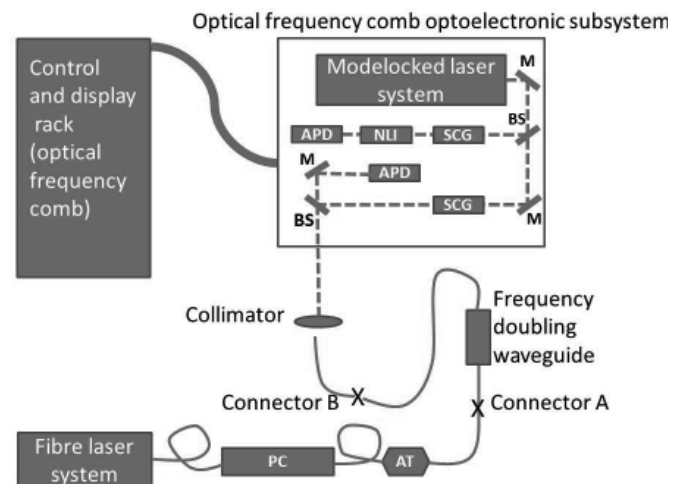

Figure 1: Simplified schematic of experimental setup. The abbreviations are as follows: PC=Polarization controller/adapter, AT=Manual attenuator, BS=Beam splitter, M=Mirror, APD=Avalanche Photodetector, SCG=Supercontinuum Generator (consisting of launching and collimating optomechanics and highly nonlinear fibre), NLI=Nonlinear Interferometer

3.1 Characteristics of frequency doubling waveguides

A custom-made periodically poled lithium niobate (PPLN) doped waveguide with MgO (the MgO increases the photorefractive damage resistance) has been used to frequency double the light from the fibre laser. The waveguide has been pigtailed with appropriate PM fibres. Such MgO:PPLN waveguide devices have been found to be convenient and efficient frequency doubling devices [16]. Such nonlinear waveguides are also used in systems for laser stabilisation to atoms [17]. The poling period is optimized for the particular laser utilised in this specific case. The waveguide device is able to generate more than

10 mW of output power for a source input power of ~140 mW (measured at the ends of the pigtails). The input wavelength is around ~1550 nm and the doubled output at half the input wavelength. This doubled wavelength is weakly visible, and therefore also convenient for length. One also expected a quadratic dependence of output power on input power as shown in the measurement of Figure 2.
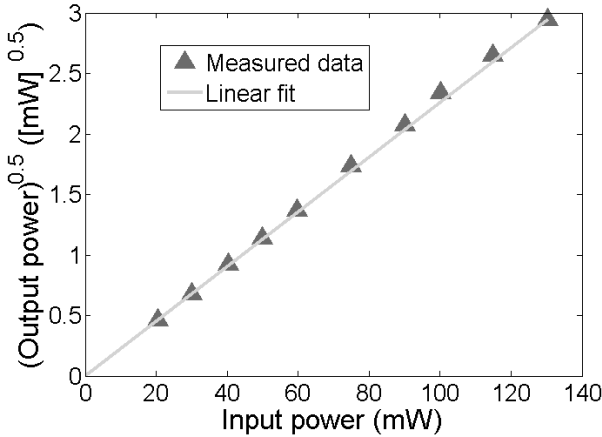

Figure 2: Output power characteristics of the MgO:PPLN waveguide, when the temperature has been optimized

Furthermore the waveguide was determined to have a wavelength acceptance bandwidth (full-width at half maximum) of ~0.05 nm. The temperature tuning characteristic is quite critical and is shown in Figure 3, which displays the nearly sinc-type shape expected from phase matching considerations.
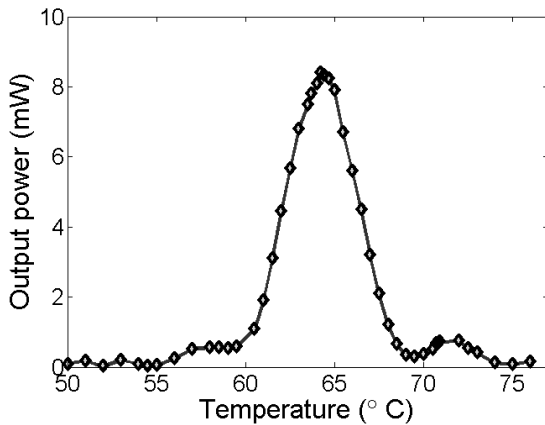

Figure 3: Temperature tuning characteristics of the MgO:PPLN waveguide

3.2 Measurements of laser against optical frequency combs

The frequency doubler places the telecommunications wavelength of the DBR fibre laser within the operational frequency range of the Ti:Sapphire laser optical frequency comb. The doubler therefore links the telecommunication range with such a visible comb that is

also utilised at the NMISA and elsewhere for visible length standard characterisation and calibration.

Three separate measurements of the UUT were done over short periods of time, and the averaged frequencies recorded on the comb every 1 second. The Allan deviations [19] of these records are shown in Figure 4; some periodicity in the UUT frequency is revealed in these and the actual frequency plots. The fractional frequency deviation over the measurement times were always less than 5×10^{-9}.
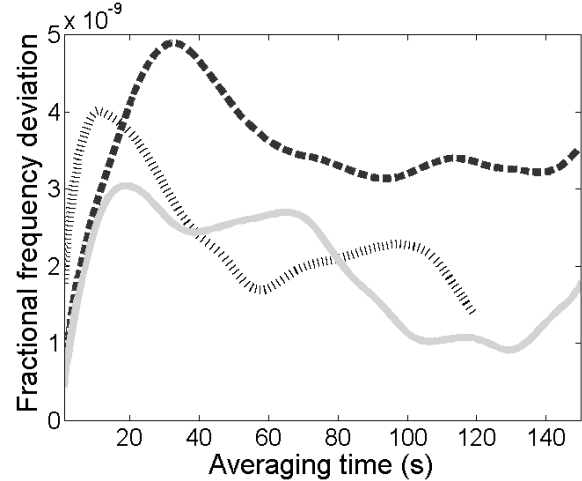

Figure 4: Overlapping Allan deviation of the UUT

Furthermore, connector B was also loosened and the fibre also intermittently connected to the wavemeter. The timerecord was also collected for the wavemeter, and an average value calculated. The corresponding sequential measurement of the laser frequency as recorded on the wavemeter and the frequency comb were always less than five minutes apart in time. The mode number was then calculated as

$$N = \left[\frac{\left(\langle f_{UUT} \rangle \pm \langle f_b \rangle \pm f_0 \right)}{\langle f_{rep} \rangle} \right] \quad , \quad (7)$$

where the large square brackets signifies rounding and the triangular brackets averaging. f_{UUT} is determined from the wavemeter reading. The appropriate signs that were measured in each case were also applied. The offset from integer numbers (δN) were also calculated. The largest absolute value of δN was ~0.16, which means all estimated mode numbers were apparently quite accurate. The apparent accuracy is verified in the next paragraph.

3.3 Statistical interpretation of measurement and validation of measurement process

The frequency deviations of the three separate runs were combined to calculate a histogram of actual frequency deviation. The histogram was fitted with a Gaussian probability density function (PDF) for subsequent calculations. It was assumed furthermore that the

extracted standard deviation was applicable to a 1 s measurement period, as is used by the electronics of the comb. The instability originating from the Cs clock that supplies referencing to the comb was accounted for by assigning a standard deviation of 2×10^{-11} to it within a 1 s averaging (measurement) time. One can now calculate a numerical PDF for δN (with symbol $\delta\widetilde{N}$) when the Cs clock and laser instability and wavemeter accuracy are taken into account. In such a calculation we define the PDF of the offset from integer comb numbers as

$$\delta\widetilde{N} \equiv \left(\frac{(F_{UUT} \pm F_B \pm f_0)}{F_{rep}}\right) - \left\langle\frac{(F_{UUT} \pm F_B \pm f_0)}{F_{rep}}\right\rangle. \qquad (8)$$

where the second term in triangular brackets is an average over infinite time, that results in an integer number. F_{UUT} is a random variable with square PDF [18] with an average of $\langle f_{UUT} \rangle$ and standard deviation determined by the manufacturer's specification on the wavemeter. F_B is a random variable with a Gaussian PDF, an average of $\langle f_b \rangle$ and a standard deviation determined from the Gaussian fit described above. Lastly, F_{rep} is a random variable with a Gaussian PDF, an average of $\langle f_{rep} \rangle$ and an uncertainty dictated by the Cs clock. The offset frequency also has a negligible uncertainty (<1 Hz) that is not taken into account. A Monte Carlo calculation utilising ten million pseudo-random numbers for each PDF then results in the histogram shown in Figure 6 for a 1 s measurement.
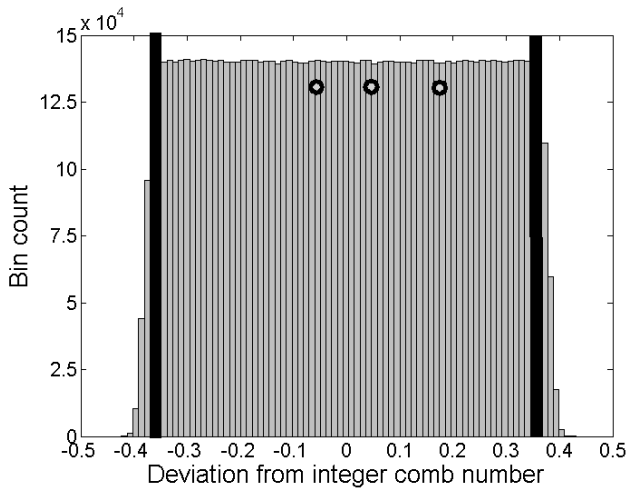

Figure 6: The numerical histogram for offset from integer comb numbers. The round dots represent averaged measured (offset) numbers. The two thick black vertical lines represent the 95% confidence interval.

4. DISCUSSION

It is clear that the wavemeter can give measurements of frequency that nearly coincides with integer comb numbers. All the measurements of comb number lies within the 95% confidence interval of the PDF of the offset from the integer mode number. One can also deduce from Figure 6 that there is a >99% probability to

always get $|\delta N|<0.5$, and therefore to resolve the comb number N. Therefore, there is a high certainty that the presented methodology is working correctly in measuring frequency, because the wavemeter is used to assign real numbers to the "photonic vernier's tickmarks". The wavemeter can do this as it can measure with <0.2 ppm absolute accuracy, which allows it to operate below the ~200 MHz ambiguity range of the optical frequency comb. Furthermore the wavemeter has been calibrated by the manufacturer against other absolute standards, so that the absolute wavelength of the source can also be assured to be correct. It can also be extrapolated that it would be possible to take *absolute* frequency readings at the ~170 Hz level for more stable lasers oscillating at hundreds of THz given the ~7.5 kHz uncertainty of the comb in 1 s, and when averaging is done over a time period >½ hour (limited by the Cs reference clock). This measurement capability can of course be improved with improved references like hydrogen masers or Cs fountain clocks.

5. CONCLUSION

This paper related the first *ab initio* measurement of an unknown optical frequency utilising the Ti:Sapphire-laser based optical frequency comb at NMISA. This measurement methodology can be successfully applied to relatively noisy lasers specifically. This methodology required the use of an internally calibrated high precision wavemeter. The measurement method was shown to be statistically sound via a Monte Carlo analysis. This successful measurement is in contrast to comb-only-methods, which might end up with large uncertainties on the laser frequency values due to some uncertainty in the precise comb element number (*N*) value. Measurement methods that utilise the comb only, given repetition rates of ~ 200 MHz is not usable for more "noisy" lasers. Future measurement will include an extra beamsplitter and free space to fibre optic coupler, so that the comb and wavemeter measurement can take place at the same time.

The optical source was shown to be able to link the optical telecommunications range, with the near-infrared range coinciding with the input operational range of 500 nm - 1200 nm of the optical frequency comb. Because of the visual visibility of the beam it can be used in length applications. When the optical source is stabilised it will also be usable as an ultrastable oscillator for a range of time and frequency applications. In the future these improved characteristics will be measured with the same methodology as presented here, except that the Cs clock might be augmented with a hydrogen clock, to be able to measure the correspondingly higher stability.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] J.L.Hall, J. Ye, S. Diddams, L-S. Ma, S.T. Cundiff and D.T. Jones: "Ultrasensitive spectroscopy, the ultrastable Lasers, the ultrafast Lasers, and the seriously nonlinear fiber: a new alliance for physics and metrology", *IEEE Journal of Quantum Electronics*, Vol. 37, No. 12, December 2001.

[2] J-P Wallerand: "A frequency doubled amplified-fiber laser for molecular iodine spectroscopy near 515 nm", *Digest: 2004 Conference on Precision Electromagnetics*, London, pp. 50-51, June 2004.

[3] C.W. Chou, D.B. Hume, J.C.J. Koelemeij, D.J. Wineland and T. Rosenband: "Frequency comparison of two high accuracy Al^+ optical clocks", *Physical Review Letters*, Vol. 104, No. 7, pp.070802-070806, February 2010.

[4] G. Grosche, O. Terra, K. Predehl, R. Holzwarth, B. Lipphardt, F. Vogt, U. Sterr, and H. Schnatz, "Optical frequency transfer via 146 km fiber link with 10^{-19} relative accuracy", *Optics Letters*, Vol. 34, No. 15, pp .2270-2272, August 2009.

[5] J.-F. Cliché and B. Shillue: "Precision Timing Control for Radioastronomy: Maintaining Femtosecond Synchronization in the Atacama Large Millimeter Array", *IEEE Control Systems Magazine*, pp.19-26, February 2006.

[6] L.B. Fu, M. Ibsen, P.W. Turner , D.J. Richardson, D. N. Payne: "Keyed axis single-polarization all-fibre DFB laser", *Electronics Letters*, Vol. 38, No. 24, pp. 1537-1539, November 2002.

[7] Ch. Spiegelberg, J.Geng, Y. Hu, T. Luo, Y. Kaneda, J. Wang, W. Li, M.Brutsch, S. Hocde, M. Chen, J. Babico, K. Barry, W. Eaton, M. Blake, D. Eigen, I. Song, S. Jiang: "Compact 100 mW fiber laser with 2 kHz linewidth" , *Proceedings: Optical Fiber Communications Conference 2003*, Vol. 3, Paper PD45, pp. 1-3, February 2004.

[8] A. D. Slepkov, A. R. Bhagwat, V. Venkataraman, P.Londero, and A. L. Gaeta, "Generation of large alkali vapor densities inside bare hollow-core photonic band-gap fibers", *Optics Express*, Vol. 16, No. 23, pp. 18976-18983, November 2008.

[9] J. Ye and S.T. Cundiff (eds): *Femtosecond Optical Frequency Comb Technology: Principle, Operation and Applications*, Springer Science and Business Media, New York, 2005.

[10] B. R. Washburn, S.A. Diddams, N.R. Newbury, J.W. Nicholson, M.F. Yan and C.G. Jørgensen:

"Phase-locked, erbium-fiber-laser-based frequency comb in the near infrared", *Optics Letters*, Vol. 29, No. 3, pp.250-252, February 2004.

[11] H. Byun, D. Pudo, J. Chen, E.P. Ippen, and F. X. Kärtner: "High-repetition-rate, 491 MHz, femtosecond fiber laser with low timing jitter", *Optics Letters*, Vol. 33, No. 19, pp. 2221-2223, October 2008.

[12] S. Zhou, D.G.Ouzonov and F.M. Wise: "Passive harmonic mode-locking of a soliton Yb fiber laser at repetition rates to 1.5 GHz", *Optics Letters*, Vol. 31, No. 8 , pp. 1041-1043, April 2006.

[13] J.W. Sickler: *High repetition rate mode-locked erbium-doped fiber lasers with complete electric field control,* Thesis (Ph.D.)--Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Engineering, 2008.

[14] E.W. Weisstein, "Normal Difference Distribution." *MathWorld-Website:*http://mathworld.wolfram.com/

[15] J.Zhang, Z.H.Lu, Y.H.Wang, T.Liu, A. Stejskal, Y.N. Zhao, Z.H.Lu, L.J.Wang, andR. Dumke:"Absolute Mode Number Determination of Frequency Combs", *Proceedings: 2007 IEEE Frequency Control Symposium,* pp. 104-106, 2007.

[16] M.M. Fejer: "Nonlinear frequency conversion in periodically-poled ferroelectric waveguides", D.B. Ostrowsky and R. Reisnisch (eds.), *Guided Wave Nonlinear Optics*, pp.133-145, Kluwer Academic Publishers, Netherlands, 1992.

[17] C. Svelto, F. Ferrario, A. Arie, M. A. Arbore and M. M. Fejer: "Frequency Stabilization of a Novel 1.5 µm Er–Yb Bulk Laser to a 39 K Sub-Doppler Line at 770.1 nm", *IEEE Journal of Quantum Electronics*, Vol. 37, No .4, pp. 505-510, April 2001.

[18] Working Group 1 of the Joint Committee for Guides in Metrology, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (JCGM 100:2008, GUM 1995 with minor corrections),* Bureau International des Poids et Mesures, Paris, pp.11-14, September 2008.

[19]D. W. Allan: "Statistics of Atomic Frequency Standards", *Proceedings of the IEEE*, Vol.54, No.2, pp. 221-230, February 1966.

METHODOLOGY FOR *IN SITU* CHARACTERISATION OF A HIGHLY BIREFRINGENT PHOTONIC CRYSTAL FIBRE FOR SUPERCONTINUUM GENERATION

J. P. Burger*, A. Ben Salem**, R. Cherif** and M. Zghal**

* *National Metrology Institute of South Africa (NMISA), Private Bag X34, Lynnwood Ridge, 0040, South Africa. E-mail: JBurger@nmisa.org*
** *Cirta'Com Laboratory, Engineering School of Communication of Tunis (Sup'Com), University of Carthage, Ghazala Technopark, Ariana 2083, Tunisia*

Abstract: A novel methodology for precisely determining the eigenaxes and effective twist of a solid-core polarisation maintaining fibre with a slightly elliptical effective core in an experimental setup with an ultrashort pulse laser is presented. This geometrical identification relies on experimental modal analysis and by utilising an incoherent ultrashort optical pulse fragment measurement and modelling procedure, and is a prerequisite for precise measurements and characterisation of this class of fibres. This orienting method is applied in the study of low threshold and temporally and polarisation stable supercontinuum generation in ~2m length of photonic crystal fibre, utilizing a modelocked Ti:Sapphire laser.

Key words: Polarisation maintaining fibre, orientation measurement, supercontinuum, photonics

1. INTRODUCTION

Polarisation-maintaining (PM) optical fibre has some physical provision for maintaining the state of linearly polarised light traveling through a single-mode core as is needed in various applications. Alignment and angular position determination of such fibre is important in the manufacturing, characterisation and utilisation of the fibre [1]. A number of methods are mentioned in literature for orienting such fibres. The first consists of applying a compressive stress to the side of a PM fibre and observing a change in the fibre's polarisation crosstalk or extinction ratio [2]. One can furthermore use imaging [3, 4] (which might utilise phase contrast) and some derivatives of this are used on modern fusion splicers. Side illumination with a laser and resultant scattering can also be used [5]. Other methods like monitoring the ellipticity of polarisation (see [6] and references therein) cannot discern between the fast and slow axis and are of more limited value. Much higher birefringence and stronger confinement are available in newer photonic crystal fibres (PCFs) [7] as compared to the traditional solid glass (PM) fibres. The latter attributes make PCFs attractive for a range of applications, including supercontinuum (SC) generation via nonlinear spectral broadening of ultrashort optical pulses [8-13]. Advantages of using birefringent PCF include the fact that all the spectral components exhibit the same stable linear polarisation and a reduced power requirement for SC generation [14]. Stable polarisation becomes important in especially optical frequency measurement and divider systems based on optical frequency combs [15]. Optical comb systems have been shown to be high precision frequency dividers for atomically stabilised optical oscillators (i.e. lasers) to generate high precision radio frequency (RF) clocking. Co-aligned and stable polarisation in the comb and the polarised laser (under test) are important in ensuring amplitude stability and low noise in the RF beat signal.

Despite all this previously mentioned work, there is to our knowledge, no published work, on the precise orientational measurement of PM-PCFs, as needed in characterisation and utilisation of such fibres. This is important especially when longer fibre strands are used, and twist is present in the fibre. The previously mentioned methods utilising imaging [3, 4] becomes hard at the high magnification needed for PCFs, and side illumination [5] becomes essentially impossible to decipher due to the large number of airholes, that can essentially block the view of the central area. The elasto-optic method [2] requires special piezoelectric contacts on the fibre, and it is not clear how stresses will propagate to the core of the PCF as opposed to solid glass fibres. Typically light is simply injected into the PCF using precision mechanical stages, and it is impossible to position high magnification microscopes in the same setup for viewing the fibre end-faces which have airholes of the order of ~100 to 500 nm in size. Therefore imaging for determining fibre orientation is impossible *in* the setup where the fibre is utilised, and moving the whole setup to a high magnification microscope is also difficult and would require special chucks for the fibre ends for mounting in a special microscope adapter.

A novel method for rapidly determining the orientation of a high birefringence PM-PCF in a setup where the fibre is simply put by hand into V-grooves is presented. This method would be especially applicable to the testing of these fibres in a manufacturing or research environment, where for example the nonlinearity also needs to be extracted. The method is based on modal analysis and measurement of the polarisation response with an ultrashort pulse probing laser. The abovementioned orientational information extracted by the new method and electromagnetic modelling was also utilised in the nonlinear optics study of the photonic crystal fibre that was investigated for generating octave

spanning supercontinua at *low* energy input (~100pJ) from a modelocked Ti:Sapphire laser. Such energy levels are now in the same energy range as those originating from high repetition rate fibre oscillators that are being developed [15-18] (of which nearly fully spliced versions could be made). These high repetition rate sources might offer *ab initio* measurement of precision optical frequency standards, without the need for a wavemeter due to the large longitudinal mode spacing. The type of fibre that was used in this study could potentially be used with such fibre sources, for stable low energy threshold SC.

2.   EXPERIMENTAL SETUP

The experimental setup consisted of a modelocked linearly polarised Ti:Sapphire laser (it can run in continuous wave mode as well) that is coupled to a PM-PCF via a precision stage and a single objective lens. The laser produced ~56 fs (full width at half maximum) long pulses at 202 MHz repetition rate. The input polarisation into the fibre was controlled with a half-wave plate. At the output of the fibre the light was collimated with a lens and sent through an analyser (polariser). The output of the analyser could be sent to a wavelength insensitive power detector or optical spectrum analysers. The fibre mounting simply consisted of V-grooves in steel, with magnets pressing down on the fibre in the V-grooves. A fixed length of PM-PCF (2m) wound around a glass beaker was used in all the experiments.

3.   MODAL MEASUREMENTS

2.1 Fibre analysis

Scanning electron microscopy (SEM) and finite element modelling was utilised in the present case on a commercial small core PCF with birefringence ~7×10^{-4} and a nonlinear coefficient of ~2.5×10^{-20} m^2.W^{-1} to determine the theoretical modes. Most fibres are manufactured to have fully rotationally symmetric modes, but with highly nonlinear and high birefringence PM-PCF, as used here, it is possible to observe pronounced geometric elongation in the modes. The core of the fibre is shown in Figure 1, and it was determined that both polarisation modes are elongated along the slow axis, due to presence of two larger holes along the fast axis. Therefore measurements of the near and far-field were undertaken to determine the output axes in the setup.
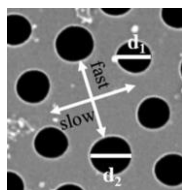

Figure 1: A SEM micrograph of the core region of the PM-PCF. The distance between the two large airholes centre is approximately 2 μm.

2.2 Near-field measurement

The near-field was imaged with a high numerical aperture (NA) aspheric lens onto a charge coupled device focal plane array (CCD) as shown in Figure 2. Such a near-field image (screen capture) is shown in Figure 3, together with measured cross-sections and is rotated by 90 degrees (for the correct view), to compensate for the 90 degree mounting of the array in the laboratory. From the image it can already be seen that the fast axis is nearly perpendicular with the optical table's surface, because of the horizontal elongation of the beam. The determination of axes was done by calculating second order moments (σ^2_x and σ^2_y) along straight lines going through the numerical mass centres of the images, when the images were rotated (Figure 4). The *x*-axis is parallel to the laboratory optical table and the *y*-axis perpendicular. An analysis of Figure 4 indicated the angle of the fast axis to be -1°. The near-field image is very sensitive to focusing and therefore the far-field was also examined to check the correctness of this measurement.
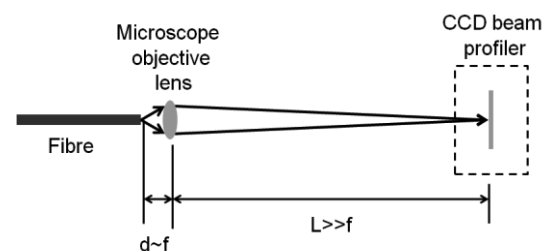

Figure 2: The experimental setup utilised for measuring the near-field. The microscope objective is placed at a distance **d** nearly equal to the focal length **f** from the fibre; and a much enlarged image is formed at a distance **L>>f** onto a focal plane array.
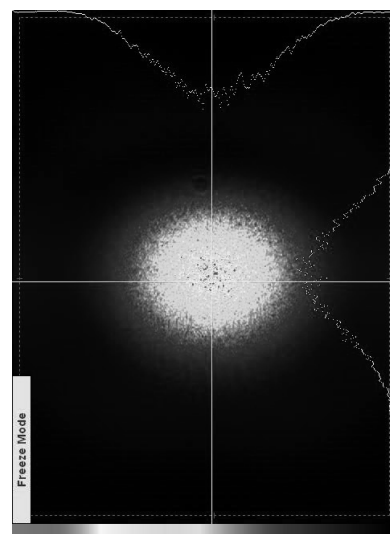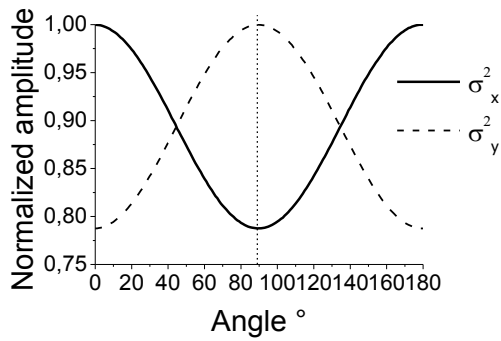

Figure 3: Near-field image of the fibre mode.

Figure 4: Second moments of the image in Figure 3.

2.2 Far-field measurement

The far-field was projected onto a diffusing screen in the focal point of a large high NA lens, when the lens is less than a focal length away from the fibre as shown in Figure 5. A CCD camera was then used to take an image of the diffusing screen (Figure 6) and the second moments calculated (Figure 7). The results indicated that the major axis of the image was at 89°.
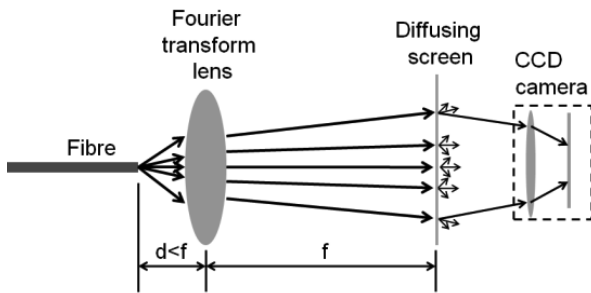

Figure 5: Setup utilised for far-field measurement. f is the focal length of the Fourier transform lens.
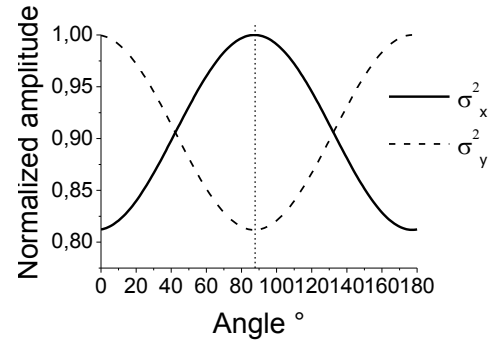

Figure 6: Image of the far-field.

Figure 7: Second moments of the far-field.

2.3 Discussion

The far-field orientation was determined to be exactly as expected; it exhibited a a 90° degree rotational difference w.r.t. the near-field, and therefore the near and far-field measurements are consistent in showing a -1° orientation of the fast axis at the fibre output. The input side of the fibre still had an unknown orientation, which could be determined through polarisation measurements as shown in the next section.

4.　POLARISATION RESPONSE WITH ULTRASHORT PULSE INJECTION

4.1 Model

The fibre is probed with ultrashort pulses and it was shown by ultrashort pulse propagation modelling [13] that the polarisation mode dispersion was such that two temporally separated orthogonally polarised pulse fragments were created in the first few cm of the fibre. Under such circumstances the pulse fragments were incoherent at the output of the fibre. Nonetheless, a Jones matrix analysis of the system can be undertaken, if new Jones projection matrices are introduced [20]. This analysis predicts an output intensity dependence on polarisation (polarisation response) as follows:

$$
\begin{aligned}
I_{out} \propto &\cos^2(\Psi)\cos^2(\Psi-\varepsilon) \times \\
&\left\{\begin{array}{l}\cos^2(\kappa)\sin^2(\kappa) \\ -2\sin(\zeta)\cos(\zeta)\times\left[\cos^3(\kappa)\sin(\kappa)+\cos(\kappa)\sin^3(\kappa)\right]+ \\ \cos^2(\zeta)\cos^4(\kappa)+\sin^2(\zeta)\cos^4(\kappa)\end{array}\right\} \\
+ \ &\sin^2(\Psi)\sin^2(\Psi-\varepsilon) \times \\
&\left\{\begin{array}{l}\cos^2(\kappa)\sin^2(\kappa) \\ -2\sin(\zeta)\cos(\zeta)\times\left[\cos^3(\kappa)\sin(\kappa)+\cos(\kappa)\sin^3(\kappa)\right]+ \\ \sin^2(\zeta)\cos^4(\kappa)+\cos^2(\zeta)\cos^4(\kappa)\end{array}\right\}
\end{aligned}
$$

(1)

The angular symbols are defined in Table 1.

Table 1: Definition of angles.

Angle symbol	Definition
ψ	Input polarisation angle
ε	Angle of fast axis of the PCF at the input side of the PCF.
η	Angle of fast axis of the PCF at the output side of the PCF.
ζ	Twist angle in order to give the fast axis orientation at the input side of the PCF ($\zeta = \eta - \varepsilon$).
κ	Angle of the polarisation axis of the polarizer

In order to demonstrate the sensitivity of the output response to the fibre rotation, two polarisation response surfaces are calculated for firstly $\varepsilon = 5°$ and $\eta = 6°$ (Figure 8) and then $\varepsilon = 5°$ and $\eta = 8°$ (Figure 9), i.e., only a 2° difference in output angle/twist. The clear sensitivity of the response to the system angles makes the intensity dependence on the polarisation ideal for determining the fibre twist.
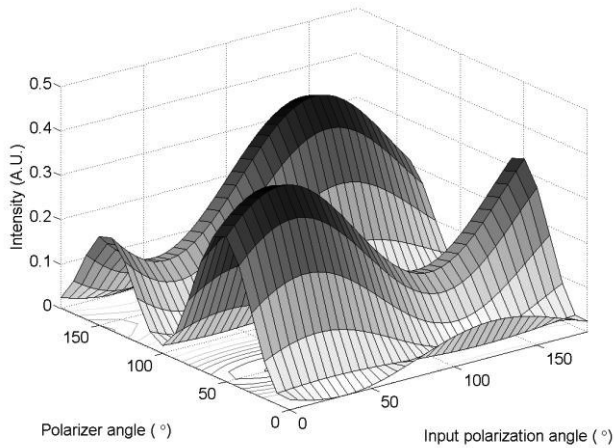
Figure 8: Theoretical polarisation response for $\varepsilon = 5°$ and $\eta = 6°$.
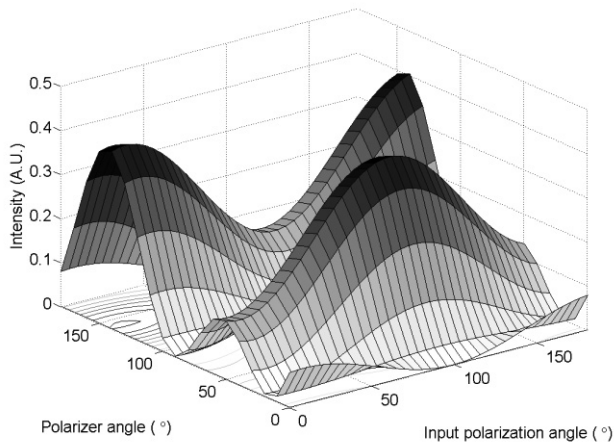
Figure 9: Theoretical polarisation response for $\varepsilon = 5°$ and $\eta = 8°$.

4.2 Measurements

The polarisation response methodology was applied to the fibre utilised in the study. A measurement of the polarisation response is shown in Figure 10. The graph was surface-fitted as a function of fibre twist to yield a single solution: a fibre twist angle of 85.2°. The theoretical response is shown in Figure 11.
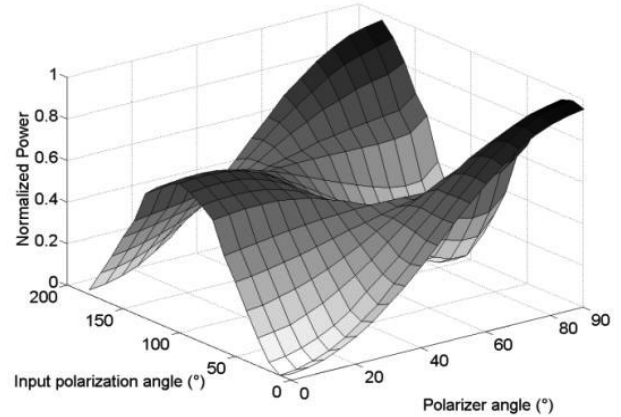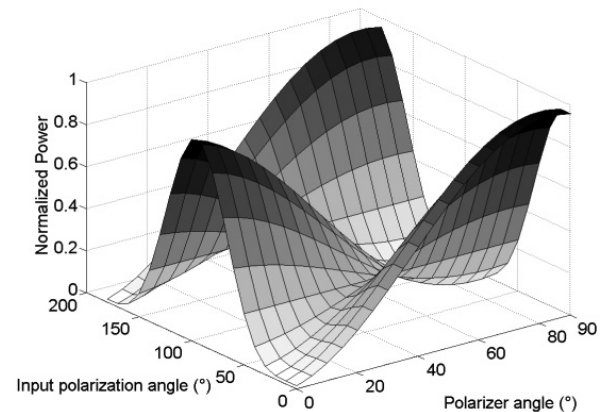

Figure 10: Experimental polarisation response.

Figure 11: Theoretical polarisation response.

4.3 Discussion

The shapes shown in Figures 10 and 11 are very similar and smaller, unmodelled effects are responsible for some differences. Firstly, the model assumes perfect polarisation extinction is possible in the fibre. The fibre exhibited an imperfect polarisation extinction of ~ -15 dB and the underlying mechanisms implied by this could possibly explain discrepancies. The unaccounted-for mechanisms include the fact that launching with a strongly focused beam implies deviation from the linear polarisation input [21]. Furthermore eigenmodes are in fact hybrid [22], and even more so in strongly guided structures such as the fibre used here. Furthermore, birefringent fibres are known to have coupling between polarisation modes due to inevitable non-uniformity along the length of fibre [22]. Lastly the laser beam from

the modelocked laser is slightly elliptical with some substructure, and this geometric structure is also expected to play a role in the polarisation at the focus where the fibre is situated.

5.   APPLICATION IN SC STUDIES

Polarisation resolved studies were also undertaken on the same fibre after the axes were determined. Pulses with energy of ~100 pJ per pulse were injected into the fibre. Spectra were recorded as a function of input and output polarisation (for steps of 5º degree rotation) and some of these are shown in Figures 12-14 with the angles referenced with respect to the fast axis at the input. The spectra as shown are contour maps of the logarithm of spectral density. White signifies the highest intensity. The slight smearing coupled with blotchiness is because of the relatively coarse 5º sampling interval and interpolation that MATLAB's contour plot algorithm utilizes. The main features of the SC can be clearly seen though. The nature of supercontinuum generation is complex [8], and in-depth discussions are beyond the scope of this paper. The main features can be easily recognized from the graphs that are shown here.
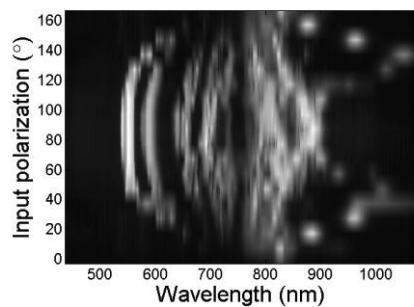

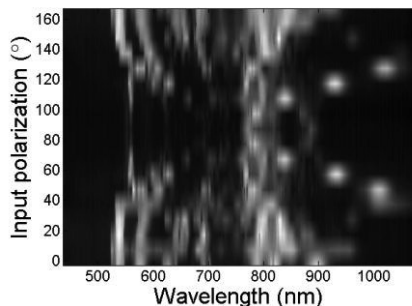Figure 12: Output spectra recorded on fast axis.

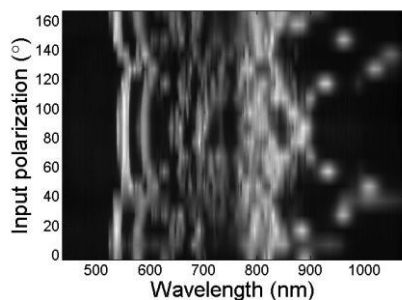Figure 13: Output spectra recorded on slow axis.

Figure 14: Output spectra recorded at 45º w.r.t. the fast axis.

Figure 12 shows the spectral evolution as a function of injected power - starting from zero power (at input polarisation of 0º). The laser spectrum is initially centred at 820 nm, and the measurement at 0º input polarisation is mostly polarisation leakage, with little nonlinear signature. A series of dots starts forming to the right as the input polarisation angle is increased (this should be a streak, but is not finely sampled enough here). These dots represent a redshifted Raman soliton (others can also just be seen forming at higher power corresponding to close to 90º polarisation). The fibre initially supports an N-th order soliton for the first section [10], if one assumes that all the light is falling above the zero-dispersion wavelength. Because of higher order dispersion and nonlinear effects, this higher order soliton decays into redshifted fundamental solitons along with blue shifted nonsolitonic radiation. The curved waves on the right hand side of Figure 12 are literally the spectral signatures of the dispersive waves (blue shifted nonsolitonic radiation), and the blue shift is clearly power dependent. Some radiation was recorded up to 1400 nm (not shown here), and more than an octave of bandwidth is spanned, even though the spectrum was not flat. The spectra were very stable on timescales exceeding a month. Figure 13 is essentially the same as Figure 12, except that the whole image is effectively 90º phase shifted. This represents the spectra along the slow axis, for a range of input powers (the lowest input power would correspond to 90º input polarisation). Figure 14's measurement essentially shows a mixture of Figure 12 and 13 as it represents the radiation at 45º w.r.t. both the slow and fast axis. In the experiment leading to Figure 14 half of the radiation in Figure 12 and Figure 13 are detected simultaneously. Interestingly, a larger blue shift is seen for one of the polarisation's dispersive waves than the other's and could be attributable to the fact that zero-dispersion wavelength for the two polarizations are predicted to be 15 nm apart, and close or even inside the pump spectrum, leading to different degrees of soliton-effect compression of the pump pulse for the two polarizations.

6.   CONCLUSION

A novel method for measuring the orientation of PM-PCF at both the input and output ends have been presented. The method requires the fibre to be simply mounted in V-grooves by hand and furthermore needs only a CCD imager, power meter, inexpensive lenses and a waveplate and polarizer. A short pulse modelocked Ti:Sapphire laser was utilised in the measurement. It would also be straightforward to also measure the birefringence in the same setup, by measuring the separation of two pulse fragments (at lower energy than shown here to avoid most fibre nonlinearities) with a second-harmonic autocorrelator [23] (usually available in an ultrafast optics laboratory). Small discrepancies were measured between the theoretical and modelled polarisation response, and was discussed. The latter discussion illustrates that there is further scope, to study smaller polarisation effects in PM-PCF, which could be significant in some niche

applications. SC studies were also undertaken on the same fibre as used in the orientation measurement. The measurements showed greater than an octave of bandwidth of SC by using less than 100 pJ of input energy and a relatively long length of fibre. The spectra were also very stable in time, possibly making this class of fibre useful for use in optical combs based on high repetition rate modelocked fibre lasers, if high coherence could also be demonstrated.

7.   ACKNOWLEDGEMENTS

8.   REFERENCES

[1] J. Noda, K. Okamoto and Y. Sasaki: "Polarization Maintaining Fibers and Their Applications", *Journal of Lightwave Technology*, Vol. 4, No. 8, pp. 1071-1089, August 1986.

[2] S. L. A. Carrara, B.Y. Kim and H. J. Shaw: "Elasto-optic alignment of birefringent axes in polarization-holding optical fiber", *Optics Letters*, Vol. 11, No.7, pp. 470-472, July 1986.

[3] Fujikura Ltd. product bulletin #88112000 on the FSM-20 PM, p. 2, 1990.

[4] Y. Hu, J. Tan, Z. Chen, C. Zeng, S. Chang and Z. Meng: "Phase contrast method for measurement of birefringent axes orientation of polarization-maintaining fiber", *Proceedings of the International Society for Optical Engineering*, Vol. 2895, pp. 350-354, 1996.

[5] J. B. Aniano: "System for determining birefringent axes in polarization-maintaining fiber", *U.S. Patent 5317575*, May 1994

[6] D. de Oliveira Maionchi, W. Campos and J. Frejlich: "Angular alignment of a polarization-maintaining optical fiber", *Optical Engineering*, Vol. 40, No. 7, pp. 1260-1264, July 2001.

[7] K. Suzuki, H. Kubota, S. Kawanishi, M. Tanaka and M. Fujita: "Optical properties of a low-loss polarization-maintaining photonic crystal fiber", *Optics Express*, Vol. 9, No. 13, pp. 676-680, December 2001.

[8] J. Dudley, G. Genty and S. Coen: "Supercontinuum generation in photonic crystal fiber", *Reviews of Modern Physics*, Vol. 78, pp. 1135-1184, October-December 2006.

[9] M. Lehtonen, G. Genty, and H. Ludvigsen and M. Kaivola: "Supercontinuum generation in a highly birefringent microstructured fiber", *Applied Physics Letters*, Vol. 82, No. 14, pp. 2197-2199, April 2003.

[10] A.Proulx, J-M.Ménard, N.Hô, J.M. Laniel, R.Vallée and C. Paré : "Intensity and polarisation dependences of the supercontinuum generation in birefringent and highly nonlinear microstructured fibers", *Optics Express*, Vol. 11, No. 25, pp. 3338-3345, December 2003.

[11] Z. Zhu and T. G. Brown: "Polarization properties of supercontinuum spectra generated in birefringent photonic crystal fibers," *Journal of the Optical Society of America B*, Vol. 21, No. 2, pp. 249-257, February 2004.

[12] Z.Zhu and T.G. Brown: "Experimental studies of polarization properties of supercontinua generated in a birefringent photonic crystal fiber", *Optics Express*, Vol. 12, No. 5, pp. 791-796, March 2004.

[13] A. Ben Salem, R. Cherif, M. Zghal, and J. Burger: "Highly Birefringent Photonic Crystal Fiber for Coherent Infrared Supercontinuum Generation", *Proceedings: Progress in Electromagnetics Research Symposium*, Marrakesh, Morocco, pp. 1247-1251, March 2011.

[14] P. Blandin, F. Druon, M. Hanna, S. L. Fort, C. Lesvigne, V. Couderc, P. Leproux, A. Tonello and P. Georges: "Picosecond polarized supercontinuum generation controlled by intermodal four-wave mixing for fluorescence lifetime imaging microscopy," *Optics Express*, Vol. 16, No. 23, pp. 18844-18849, November 2008.

[15] J. Ye and S.T. Cundiff (eds), *Femtosecond Optical Frequency Comb Technology: Principle, Operation and Applications*, Springer Science and Business Media, New York, 2005.

[16] S.Zhou, D.G. Ouzonov and F.M. Wise: "Passive harmonic mode-locking of a soliton Yb fiber laser at repetition rates to 1.5 GHz", *Optics Letters*, Vol. 31, No. 8, pp. 1041-1042, April 2006.

[17] J. Chen, J. W. Sickler, E. P. Ippen, and F. X. Kärtner: "High Repetition Rate, Low Jitter, Fundamentally Mode-Locked Soliton Er-Fiber Laser," *Conference paper: Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies*, OSA Technical Digest Series (CD) (Optical Society of America), paper CThHH3, May 2007.

[18] H. Byun, D. Pudo, J. Chen, E.P. Ippen, and F. X. Kärtner: "High-repetition-rate, 491 MHz, femtosecond fiber laser with low timing jitter", *Optics Letters*, Vol. 33, No. 19, pp. 2221-2223, October 2008.

[19] A. Martinez and S. Yamashita: "Multi-gigahertz repetition rate passively modelocked fiber lasers using carbon nanotubes", *Optics Express*, Vol. 19, No. 7, pp. 6155-6163 , March 2011.

[20] A. Ben Salem, Rim Cherif, M.Zghal and J.P.Burger: "Insights into the polarisation behaviour of a long birefringent photonic crystal fibre under low energy ultrashort pulse excitation", *Proceedings: 2011 IEEE Africon Conference*, Livingstone, Zambia, September 2011.

[21] J.Lekner: "Polarisation of tightly focused laser beams", *Journal of Optics A: Pure and Applied Optics* ,Vol. 5, No. 1, pp. 6–14, January 2003.

[22] T. A. Eftimov and W. J. Bock: "Analysis of the Polarisation Behavior of Hybrid Modes in Highly Birefringent Fibers", *Journal of Lightwave Technology*, Vol. 16, No. 6, pp. 998-1005, June 1998.

[23] J.P.Burger, W.H. Steier and S. Dubovitsky: " The energy-limiting characteristics of a polarisation-maintaining Sagnac interferometer with an intraloop compressively strained quantum-well saturable absorber", *Journal of Lightwave Technology*, Vol.20, No. 8, pp. 1382–1387, August 2002.

DIGITAL DESIGN OF BROADBAND LONG-PERIOD FIBRE GRATINGS BY AN INVERSE SCATTERING ALGORITHM WITH FLIP-FLOP OPTIMISATION

R. Kritzinger*, J. Burger*, J. Meyer and P. L. Swart†**

* *Optical Frequency Standards and Photonics Technology Group, R&D Core, National Metrology Institute of South Africa, Private Bag X34, Lynnwood Ridge, Pretoria, 0040, South Africa.*
E-mail: rkritzinger@nmisa.org
** *Department of Electrical and Electronic Engineering Science, Photonics Research Group, University of Johannesburg, P. O. Box 524, Auckland Park, 2006, South Africa.*
E-mail: johanm@uj.ac.za
† Deceased

Abstract: A discrete inverse scattering method, known as layer-peeling, is used to synthesise a LPFG (long-period fibre grating) from a desired complex spectrum by a direct solution of the coupled-mode equations, while simultaneously determining the physical properties of the layered structure. The physical properties of the grating structure are determined in a recursive layer-by-layer manner by using causality arguments to design LPFGs exhibiting a flat-top spectral profile. The results obtained from the layer-peeling method are optimised using the flip-flop method to ease the fabrication process. We found that the layer-peeling method has the highest performance and executes in the least amount of time. A discussion of possible applications where optimised broadband LPFGs could be utilised in the field of telecommunications and sensing demonstrates the importance of the results.

Keywords: Long-period fibre grating, complex spectrum, layer-peeling, flip-flop method.

1. INTRODUCTION

Over the years, it has become important to design appropriate optical filters, for example LPFGs, to achieve a desired spectral response. LPFGs are transmission-type spectral filters that can be used in various applications with their guided-to-cladding mode power exchange, for example as gain equalisers for erbium-doped fibre amplifiers (EDFAs) [1], as channel routers in optical add-drop multiplexers (OADMs) [2] or as sensors [3].

LPFG synthesis entails the derivation of the physical properties of a transmission filter from a desired spectral response [4, 5]. The grating synthesis problem is by no means trivial, especially compared to the well-known problem of computing the spectrum directly from a specific grating structure [6]. When designing fibre gratings, care should be taken to strictly monitor the complexity of the index modulation of the grating structure, such that it can be practically realised in the optical fibre core during the grating fabrication process, which was one of the aims of this research as well.

It has been shown that LPFG structures (e.g. exhibiting a flat-top spectral profile) can be reconstructed fast and efficiently from a desired complex spectrum using the discrete layer-peeling (DLP) method [4, 5]. The physical properties of a grating structure are calculated in a recursive layer-by-layer manner by using causality arguments. This method is stable and has a low algorithmic complexity that scales as $O(M^2)$. Genetic algorithms (GAs) and variational optimisation can also used to synthesise LPFGs [7, 8]. However, these methods have a low algorithmic efficiency and a slow convergence when complex LPFGs are designed. In this paper, the DLP method and flip-flop optimisation method are used to design LPFGs to achieve a desired spectral response. The flip-flop method has been known to be effective in designing interference coatings and rugate filters [9]. The flip-flop method is used to optimise a DLP synthesised LPFG to exhibit either a low or high index change at each grating period to ease the grating fabrication process. Possible applications are also discussed where synthesised LPFGs could be utilised.

2. BACKGROUND

2.1 Grating synthesis by layer-peeling algorithm

During the synthesis of an LPFG, the grating structure and physical properties are derived from a desired spectral profile utilising the transfer matrix model [4-6]

$$\begin{bmatrix} R_j(\delta) \\ S_j(\delta) \end{bmatrix} = G_j^{LPFG} \begin{bmatrix} R_{j-1}(\delta) \\ S_{j-1}(\delta) \end{bmatrix} \tag{1}$$

$$G_j^{LPFG} = \begin{bmatrix} S_{LPFG1} - i(\delta/\gamma_L)S_{LPFG2} & i(q/\gamma_L)S_{LPFG2} \\ -i(q^*/\gamma_L)S_{LPFG2} & S_{LPFG1} + i(\delta/\gamma_L)S_{LPFG2} \end{bmatrix} \tag{2}$$

Where:

δ = detuning parameter

$\gamma_L \equiv (q^2 + \delta^2)^{1/2}$

$S_{\text{LPFG1}} = \cos(\gamma_L \Delta z)$

$S_{\text{LPFG2}} = \sin(\gamma_L \Delta z)$

$R_M(\delta)$ and $S_M(\delta)$ are the amplitudes of the mode fields traversing through a section j of length Δz, where the coupling coefficient $q(j\Delta z)$ is unique for each section j, and $\rho_j = (-q_j^{*}/|q_j|)\tanh(|q|\Delta z)$ [4, 5]. When using the DLP algorithm, the LPFG structure is divided into M layers separated by a distance Δz, and the main aim is to obtain the strength of the instantaneous scattering points ρ_M, given a valid pair of transmission $R_M(\delta)$ and cross-coupling $S_M(\delta)$ spectra [4]. The spectral fields, $R_M(\delta)$ and $S_M(\delta)$, are periodic with a period of $\delta_w = \pi/\Delta z$. The detuning parameter for uniform LPFGs is defined as $\delta \equiv (1/2)(\beta_a - \beta_b^{\mu}) - (\pi/\Lambda)$, where β_a is the propagation constant of the core mode, β_b^{μ} is the propagation constant of the μth cladding mode and Λ is the grating period [1]. The resonant wavelength of an LPFG structure is defined as $\lambda_{\mathrm{LPFG}} = \Delta n_{\mathrm{eff}}\Lambda$, where Δn_{eff} denotes the difference between the effective refractive indices of the core and the cladding [1]. By letting all the coupling take place at a single point, i.e. $|q| \to \infty$, while the product $q\Delta z$ remains constant, the distributed coupling process can be separated into two parts that consists of replacing the transfer-matrix matrix G_j^{LPFG} by a product of two transfer matrices, G_Δ^{LPFG} (pure propagation between instantaneous scattering points in LPFG) and G_ρ^{LPFG} (mode-coupling in the jth section) [4].

$$G_\rho^{LPFG} = \left(1+|\rho|^2\right)^{-1/2}\begin{bmatrix} 1 & \rho \\ -\rho^{*} & 1 \end{bmatrix} \qquad (3)$$

$$G_\Delta^{LPFG} = \begin{bmatrix} \exp(i\delta\Delta z) & 0 \\ 0 & \exp(-i\delta\Delta z) \end{bmatrix} \qquad (4)$$

Where:

$1 \le j \le M$

The discrete coupling ratio in the LPFG is then simply expressed as $\rho_M = s_M(0)/r_M(0)$, where the time-domain coefficients are defined as [4]: $R_j(\delta) = \sum r_j(\tau)\exp(i2\Delta z\delta\tau)$ and $S_j(\delta) = \sum s_j(\tau)\exp(i2\Delta z\delta\tau)$, where $\tau = 0,1,\ldots,M$. Since the value of ρ_M is known at the final layer, we can now remove this layer and obtain the impulse responses of layer $M-1$ as follows [4, 5]:

$$r_{M-1} = \left[r_M(\tau) + \rho_M^{*} s_M(\tau)\right]/\sqrt{|\rho_M|^2+1} \qquad (5)$$

$$s_{M-1}(\tau-1) = \left[s_M(\tau) - \rho_M r_M(\tau)\right]/\sqrt{|\rho_M|^2+1} \qquad (6)$$

The DLP procedure of reconstructing a LPFG from a complex spectrum profile is summarised in [4, 5, 8].

2.2 Grating synthesis by flip-flop optimisation method

For many years, the flip-flop method has been used extensively for the synthesis of thin-film structures (with practically realisable refractive index profiles) [9, 10]. The flip-flop method was effective in obtaining an equivalent thin-film structure exhibiting either a low or high refractive index at all layers (exhibiting equal thickness). In this paper, results are presented where the flip-flop method was used to synthesise an LPFG originally designed using the DLP method. The results obtained from the flip-flop method are intended to be used for the fabrication of an LPFG in single-mode fibre using a point-by-point fabrication technique [11]. The flip-flop method is described as follows:

1. Divide the LPFG to exhibit equal grating periods. Since the grating was originally designed by using the DLP method, the grating does not need to be divided to exhibit equal grating periods. Take note of the π phase shifts at each zero point of the refractive index profile of the synthesised LPFG;
2. Assign initial refractive index values to each grating period using a suitable starting index profile design;
3. Evaluate a fitness function that represents performance of the starting index profile design;
4. Move through the synthesised LPFG and change the state (i.e. refractive index) of each grating period one-by-one. The fitness function has to be calculated for each flipped state. A grating period can exhibit either a low or high refractive index value. The superior fitness function value for a particular flipped state is retained and used in subsequent calculations;
5. If the fitness function improved after evaluating all grating periods in a single pass, step 4 is repeated, otherwise the synthesis process is complete.

3. NUMERICAL RESULTS

3.1 Design of a broadband LPFG using the DLP method

In this section an LPFG is reconstructed from a complex spectrum profile. The cross-coupled power spectrum has a flat-top, nearly rectangular passband described by a "Super-Gaussian" function [4]:

$$S_M(\delta) = \sqrt{T} \times \exp\left[-\left(\delta/\delta_{pb}\right)^{20}\right] \qquad (7)$$

where the maximum cross-coupled power (T) in the passband is -0.46 dB. The passband's full width at half-maximum (δ_{pb}) is 25 nm and the resonant wavelength (λ_{LPFG}) is 1562 nm. The grating chirp is 8.3 nm/cm. The effective refractive index difference (Δn_{eff}) between the guided mode and fifth-order cladding mode is 3.4×10^{-3}. The reference grating period (Λ) equals 459 µm. The grating length (L_{LPFG}) is 30 mm and the number of layers used to reconstruct the LPFG is 34, where $M = L_{\mathrm{LPFG}}/\Delta z$. During the simulation, the number of wavelengths (X) is equal to 1000. The power transmitted by the core mode and cladding mode are shown in Figure 1. The transmission spectrum is kept above -11 dB, exhibits a full-width half maximum (FWHM) bandwidth of 24.9 nm, and has no sidelobes. The ripple variation of the

transmission spectrum and cross-coupled spectrum are 0 dB and 1.2×10^{-3} dB, respectively. Figure 2 shows the coupling coefficient and index modulation of the synthesised LPFG. The LPFG has uniform grating periods and a refractive index profile that exhibits a sinc-like envelope, where there exists a π phase shift at each minimum point [4, 5]. Zhang et al. observed similar results after the DLP method was used to reconstruct an LPFG [5]. An error function is used for the measurement of the difference between the synthesised and desired cross-coupling spectrum of an LPFG. The error function is used to evaluate the performance of the synthesis method and is expressed as: $d\{S_{\text{SYN}}, S_{\text{DES}}\} = \sum (|S_{\text{SYN,k}}| - |S_{\text{DES,k}}|)^2$, where $S_{\text{SYN,k}}$ and $S_{\text{DES,k}}$ are the kth value of the synthesised and desired cross-coupling spectrum, respectively [8]. The temporal performance (R_s) metric is commonly used for comparing the performance of different algorithms for the solution of the same problem [12]. Temporal performance is commonly expressed in solutions per second (sol/s) and attains a maximum when the algorithm's run-time is at a minimum [12]. The reconstruction of the LPFG took 768 milliseconds using MATLAB® on a 2.2 GHz Athlon K7 computer. The DLP algorithm reproduced the original cross-coupling power spectrum with an absolute error of almost 0%. The maximum error in computing the cross-coupled power in the simulation is of the order of 10^{-4} dB. The absolute
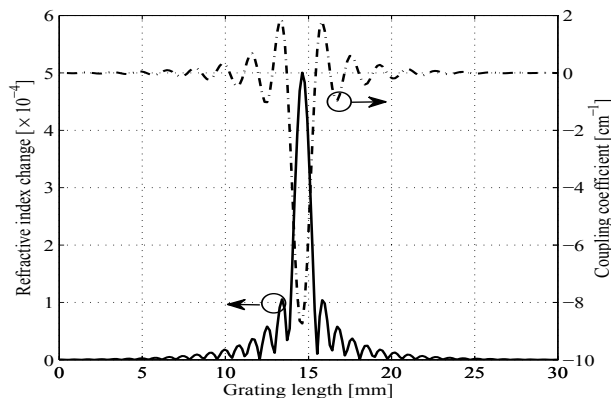
error is calculated in the wavelength region, 1489 to 1635 nm. The DLP algorithm produced 655.5 KFLOPS within 768 ms of execution time. The temporal performance of the layer-peeling algorithm was calculated at 1.3 sol/s.

In practice the synthesised index modulation profile illustrated in Figure 2 will be difficult to implement due to its complex profile. Most often, it is difficult to develop a fabrication system that delivers the required accuracy [11]. The synthesised refractive index profile illustrated in Figure 2 is segmented using the grating period (i.e. 459 μm) to consider practical fabrication issues. Unfortunately, the LPFG structure, considering practical fabrication issues, differs from the original synthesised LPFG in that the transmission spectrum does not exhibit a perfect flat-top profile and in that the spectrum bandwidth decreases.

3.2 Design of a broadband LPFG using the flip-flop optimisation method

The refractive index change profile (considering fabrication issues) can be simplified even further to ease the LPFG fabrication process by using the flip-flop method. The aim is to retain the grating periods as they are, but to change the index change values to exhibit either a low or high index change at each grating period – essentially a "digital" design using only two refractive index values. Figure 3 illustrates the transmission spectrum obtained for the core mode and cladding mode after implementing the flip-flop method in conjunction with the DLP method. The results were obtained after three iterations of the flip-flop method in less than 80 seconds. In Figure 3, the transmission loss of the flip-flop synthesised LPFG exhibits a transmission loss greater than > 6 dB within the stopband. The index change profile of the flip-flop synthesised LPFG illustrated in Figure 4 will ease the LPFG fabrication process. New positions exist for π phase shifts along an LPFG, when the flip-flop method is implemented. Figure 5 illustrates that the dispersion for the optimised filter is small (4 ps/nm) when the LPFG is reconstructed using the DLP
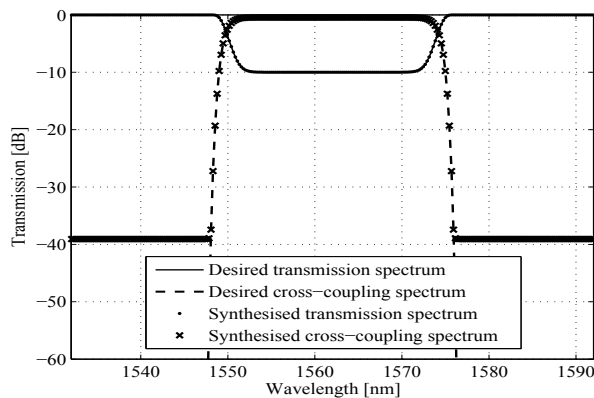

Figure 1: Spectrum results associated with the synthesised LPFG compared to results of desired LPFG

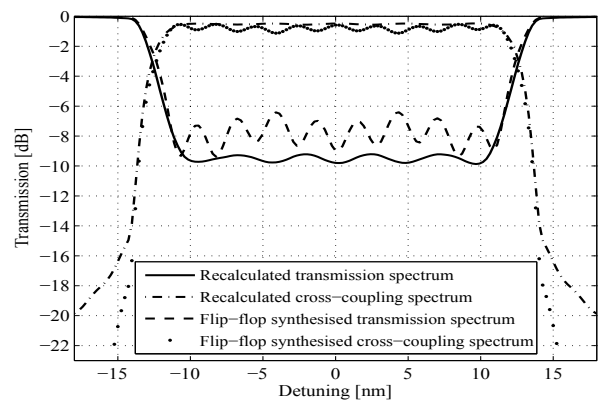Figure 2: Refractive index modulation and coupling coefficient of synthesised LPFG

Figure 3: Spectrum results associated with the flip-flop synthesised LPFG compared to spectra of DLP synthesised LPFG with fabrication issues considered
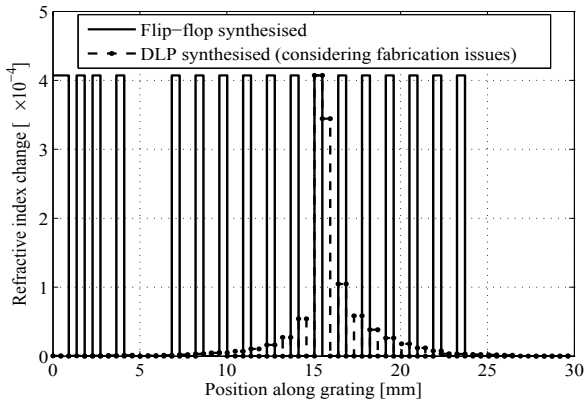
Figure 4: LPFG index modulation profile results obtained using the practical DLP method and flip-flop method
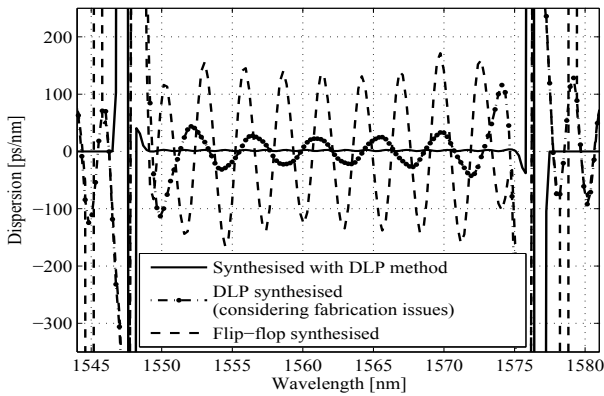

Figure 5: Predicted dispersion of optimised LPFG obtained using the DLP method and flip-flop method

method, but increases when the filter structure results are adjusted for practical implementation, especially when the flip-flop method is used in conjunction with the DLP method. The flip-flop method does not do a good job of reproducing the transmission and cross-coupling spectrum of the LPFG synthesised using the DLP method, resulting in an absolute error of ~4%. The maximum error in computing the transmission power in the flip-flop simulation is of the order of 10^{-1} dB. The flip-flop algorithm produced 72.1 KFLOPS within 78 seconds of execution time. The temporal performance of the flip-flop algorithm was calculated at 1.3×10^{-2} sol/s.

3.3 Discussion

In this paper, LPFGs have been synthesised using the DLP method and flip-flop method. These synthesis methods were used to reconstruct an LPFG structure from a complex spectral profile. The DLP method (discussed in Section 3.1) reconstructed an LPFG efficiently, resulting in an absolute error of almost 0% and a temporal performance of 1.3 sol/s. A high-precision LPFG fabrication system would be required to induce the synthesised refractive index change profile in an optical fibre. The flip-flop method (discussed in Section 3.2) could not reconstruct an LPFG efficiently. The flip-flop

method exhibits an absolute error of ~4% and a temporal performance of 1.3×10^{-2} sol/s. The flip-flop method produced a much simpler refractive index change profile compared to the DLP method, but still exhibited π phase shifts at discrete points in the synthesised refractive index change profile. The temporal performance metric was used to compare the performance of different algorithms for the solution of the same problem. It was calculated that the DLP method has the highest performance and executes in the least amount of time.

4.   LPFG FABRICATION AND RESULTS

The experimental set-up discussed in [11] was used to fabricate the broadband LPFG. The grating was fabricated in photosensitive single-mode fibre using a TEM_{01*} - mode CO_2 laser (Edinburgh Instruments PL2-M) by implementing the point-by-point fabrication method. Figure 6 illustrates the refractive index change profile of the synthesised LPFG considering practical fabrication issues, as well as the index change profile implemented during the grating fabrication process. The index modulation profile implemented during LPFG fabrication is narrow and exhibit a maximum index change of 3.8×10^{-4} near the centre of the grating. The index change profile implemented during LPFG
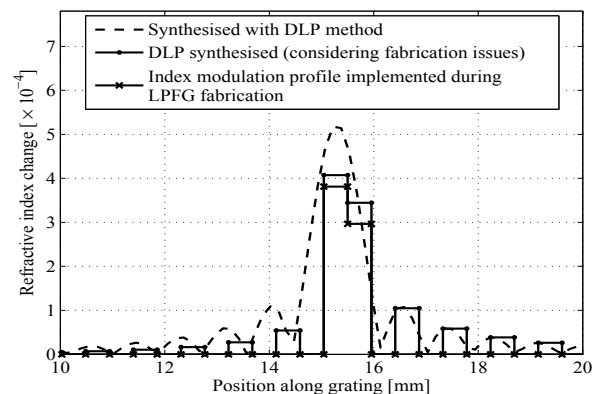

Figure 6: Index change profile implemented during fabrication of complex LPFG designed with DLP method
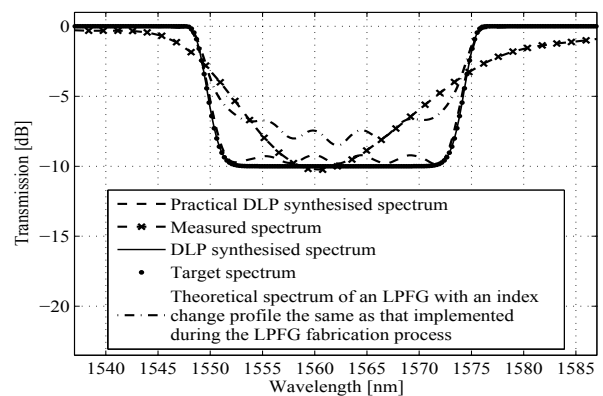

Figure 7: Spectra simulated and measured after fabricating complex LPFG designed with DLP method

fabrication (line-cross graph) can resemble the index change profile considering fabrication issues (line-dot graph) more closely by decreasing shutter exposure time. Figure 7 illustrates the transmission spectrum measured after the complex LPFG was fabricated. The measured spectrum did not exhibit a flat-top, but exhibits minimal sidelobes and a FHWM stopband bandwidth close to the target spectrum (dotted graph). The measured spectrum exhibits a FWHM stopband bandwidth of ~27 nm and a maximum transmission loss of 10.3 dB at ~1561 nm. Figure 7 shows that the measured transmission spectrum (dash-cross graph) differs from the theoretical transmission spectrum (dash-dot graph). The reason for this could be attributed to grating imperfections or accumulation of numerical errors and phase errors when the DLP algorithm is implemented during the simulation of the LPFG transmission spectrum (using index change profile implemented during LPFG fabrication process). The greater the difference between the original DLP synthesised index change profile (dashed graph of Figure 6) and the practical index change profile (line-cross graph of Figure 6), the greater the grating phase error will be. The grating fabrication set-up in [11] does place a limit on the accuracy that can be achieved in manufacturing LPFGs designed with the DLP method. The LPFG experimental results obtained with the DLP algorithm were not good enough, since the LPFG transmission spectrum did not resemble a flat-top type spectral profile. However, the measured spectrum did exhibit a high transmission loss and broad bandwidth in the stopband. The theoretical results obtained from the flip-flop method were not considered in experiments, because the simulation results obtained from the DLP algorithm was used by the flip-flop method to simplify the index change profile. It was expected that if no decent experimental results were obtained with the DLP algorithm, better experimental results would not be obtained with the aid of the flip-flop method either, unless a more accurate LPFG fabrication set-up is used. To our knowledge, this was the first time a point-by-point fabrication method was used to manufacture a DLP synthesised LPFG.

5.   APPLICATIONS

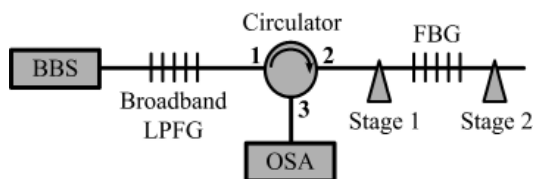*5.1 Bragg grating sensor interrogated by an LPFG*

Figure 8: Schematic diagram of FBG sensor system with LPFG as interrogation element

Figure 8 illustrates a system that operates on the principle of using a broadband LPFG to interrogate a fibre Bragg grating (FBG) strain sensor, which is based on strain
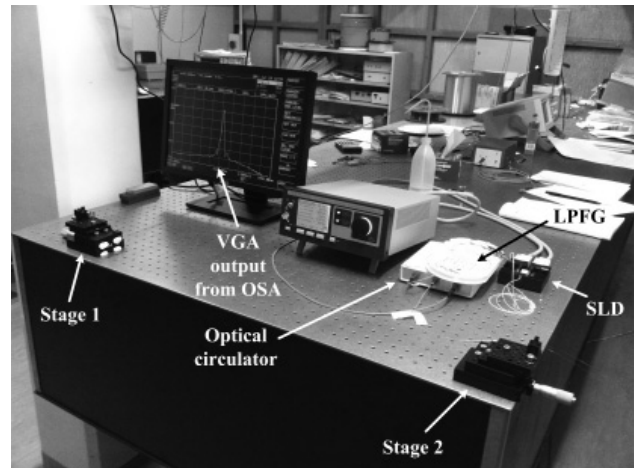

Figure 9: Illustration of the experimental set-up for fibre grating sensor interrogation

related optical intensity measurements. The experimental set-up consisted of a superluminicent-diode (SLD) light source, broadband LPFG, optical circulator, two translation stages, a narrowband FBG and an optical spectrum analyser (Ando AQ6315B). Each translation stage contains one fibre clamp. The distance between the fibre clamps is ~1 m before the fibre is strained. Figure 9 illustrates an image captured of the experimental set-up for fibre grating sensor interrogation. The LPFG length is 40 mm and exhibits a 1567.7 nm resonant wavelength (λ_{LPFG}) when mode-coupling is performed to the fifth-order cladding mode. The broadband LPFG is used as an edge filter that yields an almost linear relationship between the wavelength shift of the FBG and the detected light intensity at different strain-induced values [13]. The FBG length is 5 mm and exhibits a resonant wavelength of 1546.4 nm. A FBG is a reflection filter that reflects light at a specific resonant wavelength [6]. During the experiments the resonant wavelength (λ_{FBG}) of the FBG is located at the negative slope side of the LPFG curve (i.e. intensity-descending side of LPFG transmission band) as shown in Figure 10. The reflected FBG power decrease when λ_{FBG} shifts to longer wavelengths. Figure 11 and 12 illustrates the shift in λ_{FBG} and decrease in light intensity
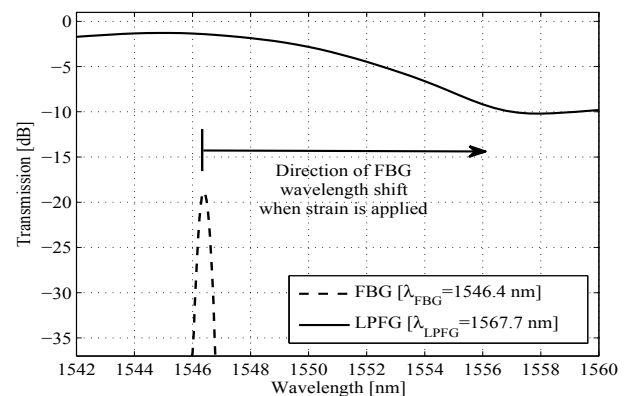

Figure 10: Spectrum of FBG strain sensor and LPFG employed as edge filter
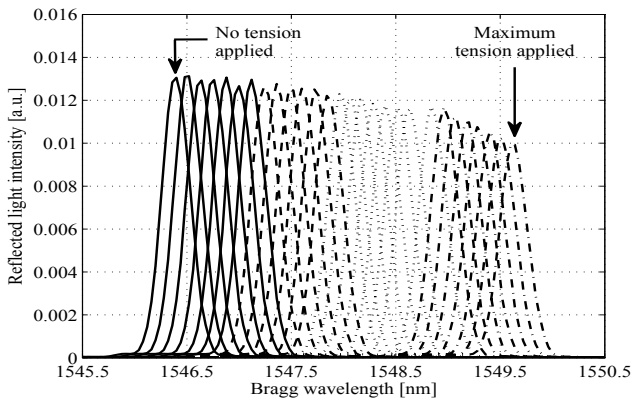
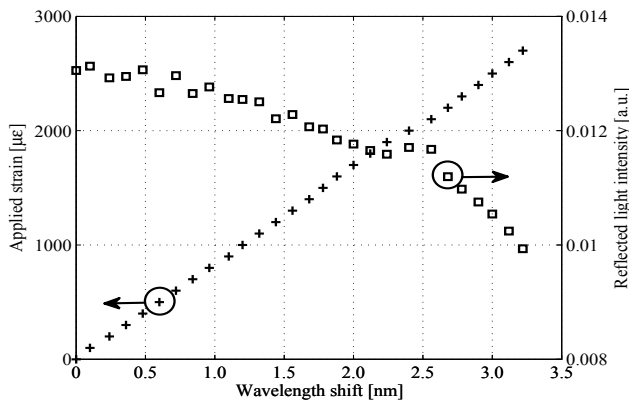Figure 11: FBG resonant wavelength versus reflected light intensity

Figure 12: FBG wavelength shift versus applied strain and reflected light intensity

by varying applied strain from 0 to 2700 $\mu\epsilon$. The FBG is strained by moving the second translation stage as shown in Figure 8. The FBG produced satisfactory results when it was placed under tension, resulting in a maximum wavelength shift of 3.2 nm at a translation stage displacement of 2.7 mm. Compared to other experiments where narrowband LPFGs are utilised [13, 14], the broadband LPFG used in this particular experiment exhibits steeper side skirts, which results in the decrease in light intensity to be more rapid when strain is applied to the FBG. Using a LPFG exhibiting steep side skirts in a FBG strain sensing system can be useful in environments where small changes in fibre strain and significant changes in optical light intensity need to be monitored. To our knowledge, this is the first time results are presented where a broadband CO_2-laser-induced LPFG is used to interrogate a FBG strain sensor.

5.2   Wavelength-tuneable OADM using broadband LPFGs and a narrowband FBG

Figure 13 illustrates an OADM utilising a pair of broadband LPFGs and a narrowband FBG, designed for operation in a dense wavelength-division multiplexing (DWDM) network. Evanescent field coupling in an
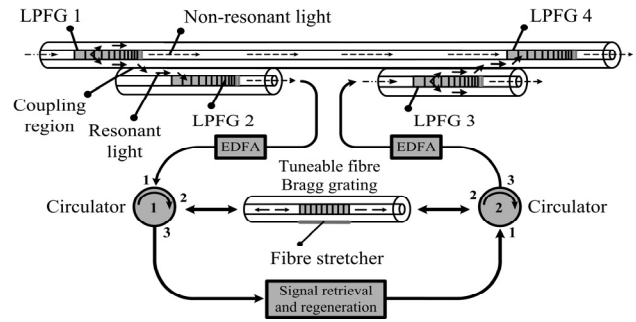

Figure 13: Broadband ADM designed for WDM network

optical fibre-based wavelength-selective coupler (WSC) allow broadband light to be transferred between two fibres containing non-uniform LPFGs [15]. Since the multiplexing part of the OADM illustrated in Figure 13 is the mirror image of the de-multiplexing part, only results obtained for the de-multiplexing part are presented. Figure 14 illustrates the experimental set-up of the de-multiplexing part of the OADM. A wavelength-tuneable laser source (Agilent 81600B) was used to transmit a specific wavelength channel at 1 mW through the WSC. An EDFA was used to restore the power level of the output signal obtained from the WSC to its original level. A tuneable FBG (with 37.2% peak reflectivity) is used to select a specific wavelength channel from the broadband light routed through the WSC. The resonant wavelength of the FBG is 1549.3 nm when no fibre strain is applied. The results shown in Figure 15 were obtained from an optical spectrum analyser (Yokogawa AQ6317C). Figure 13 illustrates that the non-resonant wavelengths travelling beyond the FBG are routed to the multiplexing part of the OADM. All signals exiting the second circulator are pre-amplified before multiplexed on the original fibre link.
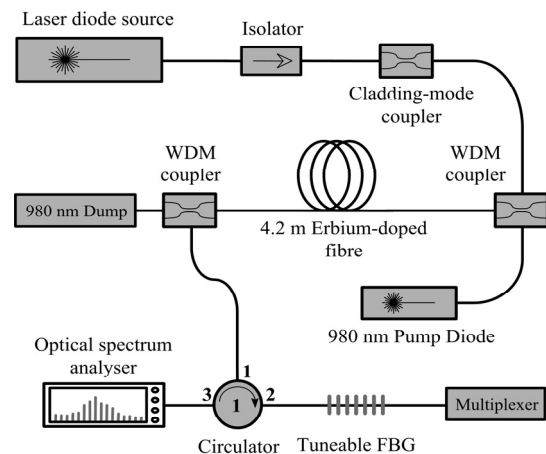

Figure 14: Schematic diagram of experimental set-up for the de-multiplexing part of the OADM

6.   CONCLUSION

We have shown that the DLP method can be used successfully to reconstruct broadband LPFGs from a complex spectrum. The DLP method has the highest performance and executes in the least amount of time, but
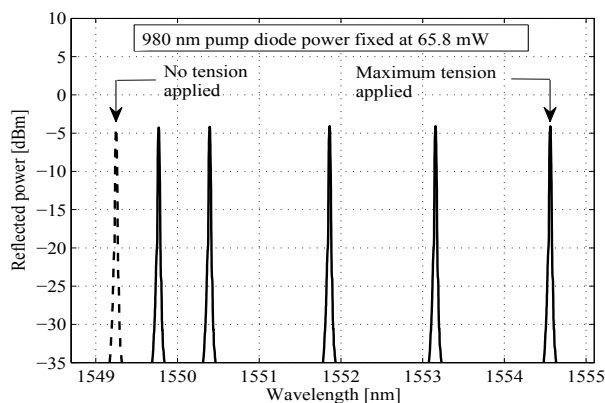
Figure 15: Power spectra of injected wavelength channels

results in index modulation profiles that are difficult to implement in practise. When the flip-flop method is used in conjunction with the DLP method, much simpler LPFG index modulation profiles are obtained, which may be implemented in a LPFG fabrication process. The flip-flop method does not require a complex initial design, and has a rapid convergence. The application of using broadband LPFGs in the field of telecoms and sensing was experimentally demonstrated through the discussion of possible applications. Precise control of the induced index change in the fibre core is of utmost importance in order to manufacture LPFGs that resemble the target spectrum more accurately. The utilisation of the modified DLP and flip-flop synthesis methods in conjunction with an automated fabrication system shows great promise for the manufacture of high-performance LPFGs for application in various industries in the future.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] A. M. Vengsarkar, J. R. Pedrazzani, J. B. Judkins, P. J. Lemaire, S. Bergano, and C. R. Davidson: "Long-period fiber grating based gain equalizers", *Optics Letters*, Vol. 21, No. 5, pp. 336-338, 1996.

[2] Y. G. Han, S. B. Lee, C. S. Kim, and M. Y. Jeong: "Tuneable optical add–drop multiplexer based on long-period fiber gratings for coarse wavelength division multiplexing systems", *Optics Letters.*, Vol. 31, No. 6, pp. 703–705, 2006.

[3] S. W. James and R. P. Tatam: "Optical fibre long-period grating sensors: characteristics and application", *Measurement Science and Technology*, Vol. 14, pp. R49-61, 2003.

[4] J. Brenne and J. Skaar: "Design of grating-assisted co-directional couplers with discrete inverse-scattering algorithms", *Journal of Lightwave Technology.*, Vol. 21, No. 1, pp. 254-263, 2003.

[5] J. Zhang, P. Shum, S. Y. Li, N. Q. Ngo, X. P. Cheng, and J. H. Ng: "Design and fabrication of flat-band long-period grating", *IEEE Photonics Technology Letters*, Vol. 15, No. 11, pp. 1558–1560, 2003.

[6] T. Erdogan: "Fiber grating spectra", *Journal of Lightwave Technology.*, Vol. 15, No. 8, pp. 1277-1294, 1997.

[7] G. W. Chern and L. A. Wang: "Design of binary long-period fiber grating filters by the inverse-scattering method with genetic algorithm optimization", *Journal of the Optical Society of America A*, Vol. 19, No. 4, pp. 772-780, 2002.

[8] R. Kritzinger and J. Meyer: "Design and fabrication of novel broadband long-period fiber gratings using synthesis techniques", *Journal of Lightwave Technology*, Vol. 29, No. 8, pp. 1077-1084, 2011.

[9] W. H. Southwell, "Coating design using very high- and low-index layers", *Applied Optics*, Vol. 24, No. 4, pp. 457-460, 1985.

[10] J. A. Dobrowolski, "Comparison of the Fourier transform and flip-flop thin-film synthesis methods", *Applied Optics.*, Vol. 25, No. 12, pp. 1966-1972, 1986.

[11] R. Kritzinger, D. Schmieder and A. Booysen: "Azimuthally symmetric long-period fibre grating fabrication with a TEM_{01} -mode CO_2 laser", *Measurement Science and Technology.*, Vol. 20, pp. 034004-, 2009.

[12] R. W. Hockney: *The science of computer benchmarking*, SIAM, Philadelphia, chapter 2, 1997.

[13] P. Saidi Reddy, R. L. N. Sai Prasad, K. Srimannarayana, M. Sai Shankar, D. Sen Gupta: "A novel method for high temperature measurements using fiber Bragg grating sensor", *Optica Applicata.*, Vol. 40, No. 3, pp. 685-692, 2010.

[14] R. W. Fallon, L. Zhang, L. A. Everall, J. A. R. Williams and I. Bennion: "All-fibre optical sensing system: Bragg grating sensor interrogated by a long-period grating", *Measurement Science and Technology.*, Vol. 9, pp. 1969-1973, 1998.

[15] R. Kritzinger and A. Booysen: "Wavelength-tuneable add/drop multiplexer using broadband transmission filters and a narrowband reflection filter", *Proceedings: SPIE*, Vol. 7004, No. 70042O, 2008.

A LENSLESS, AUTOMATED MICROSCOPE FOR DISEASE DIAGNOSTICS

S. Hugo*, T. Naidoo*, H. Swart*, S. Potgieter*, P. van Rooyen* and K. Land*

* Materials Science and Manufacturing, CSIR, PO Box 395, Pretoria, South Africa
E-mail: shugo@csir.co.za

Abstract: Optical microscopy is widely accepted as one of the gold standards in disease diagnosis. However, factors such as cost and the need for a trained eye limit the prevalence of such equipment, particularly in resource-limited areas such as rural clinics. Lensless microscopy, which is based on principles of digital holography, has illustrated the possibility of using simple and cheap optical components combined with software algorithms to implement microscope platforms. We present a digital in-line holographic microscope (DIHM) platform to be used with image processing and classification algorithms to provide a low cost, portable and automated microscope. Initial results show that the images obtained using the DIHM platform are similar to those obtained using a conventional bright field microscope. Applications of this work are targeted towards the implementation of a full blood count, which could provide resource-limited areas with improved healthcare facilities and diagnosis times.

Key words: Microscopy, point-of-care disease diagnostics, digital in-line holographic microscopy.

1. INTRODUCTION

1.1 Background

The concept of holography dates back to 1948, with its name derived from the Greek words "holos" and "graphein", meaning entire and to write, respectively [1]. A hologram is a recorded interference pattern between a reference wave and a wave field scattered from an object. Traditionally, holographic plates were used as recording mechanisms to capture the hologram. The recorded holographic image is then illuminated with the reference wave to obtain a reconstructed object wave, which is the same as the original object wave.

Throughout the last few decades, significant developments in the field of holography have been made, particularly with the direct recording of Fresnel holograms, made possible with charge-coupled devices (CCDs) [2]. This method enables full digital recording and processing of holograms, without having to use photographic recording processes. The technique has been labelled direct holography, and more recently, digital holography.

In digital holography, the interference between a reference wave and a scattered wave from an object are captured at the surface of a CCD, or other similar light-sensitive electronic device. The resulting hologram is thus electronically recorded and stored. Numerical reconstruction of the object from the electronically recorded hologram can then be performed. Numerical reconstruction can be implemented using approximations of diffraction theory calculations, e.g. the Fresnel-Kirchoff integral.

Digital holography has seen many improvements and has been applied to a variety of measurement tasks, including shape measurement and deformation analysis, encryption of information, etc. [1]. Included in these applications is the use of digital holography in imaging, and in particular microscopy [3 - 5], providing microscope systems with a large field of view compared to conventional bright field microscopes.

1.2 Motivation and objectives

Digital holography demonstrates the implementation of microscopy methods using simple recording devices (e.g. CCDs), without the need for specialised optical components such as lenses. This illustrates the possibility of using simple and cheap optical components with software algorithms to create a low cost and compact microscope platform.

In addition to the advantages of digital holography for microscopy applications in terms of cost and complexity, digital holography also provides advantages over conventional microscope systems in terms of field of view and depth of field. Digital holography allows for images with a large field of view to be captured, where the field of view is proportional to the size of the digital sensor. In terms of depth of field, conventional microscopes provide a focussed image at a single plane, whereas a hologram contains the complete three-dimensional information of the optical wavefront. This allows for reconstruction of the object image at any focal plane.

Optical microscopy is an established field with vast applications. Among the most important applications is that it is considered one of the gold standards for medical diagnosis, with use in advanced laboratories through to rural clinics. However, conventional optical microscopy makes use of specialised optical components and requires a trained operator.

Implementation of a low cost microscope by means of digital holography techniques could enable microscopy to become a more accessible tool, particularly in resource limited areas. By combining this platform with algorithms for identification of a specific disease to be diagnosed from a biological sample, an automated microscope with a large field of view can be realised, making analysis more efficient. In addition, the reduced complexity of the optics of the digital holography microscopy system would allow for cheaper, more compact and mobile equipment to be developed compared with many traditional microscope systems, potentially providing a point-of-care disease diagnostic solution ideally suited to resource limited settings.

The development of effective point-of-care disease diagnostics, particularly for developing countries, is rapidly becoming a more prominent area of research [6, 7]. The potential impact of such diagnostic tools is high, as there is a global increase in non-communicable and infectious disease, and diagnostics remain expensive and unavailable in areas that are in the greatest need of medical assistance [8].

Applications of this work as a disease diagnostic tool could provide resource limited areas with improved healthcare facilities and reduced diagnosis times without requiring skilled personnel on site. This would be particularly useful and novel in a local South African context, where existing medical diagnostics are generally expensive and thus inaccessible to the majority of the population, with a limited number of trained personnel to operate the diagnostic tools. Other developing countries across the globe, from Africa to Asia, would also benefit from the realisation of a disease diagnostic as presented here, as would developed countries, where point-of-care and home-test medical diagnostics for chronic diseases could be realised.

2. PRINCIPLES OF HOLOGRAPHY

Fundamental principles of holography include propagation and interference of light waves, which can be explained using scalar diffraction theory [1, 9].

There are different holography set-ups (off-axis, in-line etc.), many of which have the same theoretical principles [10]. In-line holography with spherical waves is the simplest realisation of the holographic method [3] and will form the basis of the work discussed in this paper.

2.1 Hologram generation

The generation of a hologram is illustrated in figure 1. Mathematically, the reference wave, E_R, with real amplitude, A_R, and phase, φ_R, is described by

$$E_R(x,y) = A_R(x,y)\exp(j\varphi_R(x,y)) \qquad (1)$$

while the object wave, E_O, with real amplitude, A_O, and phase, φ_O, is described by

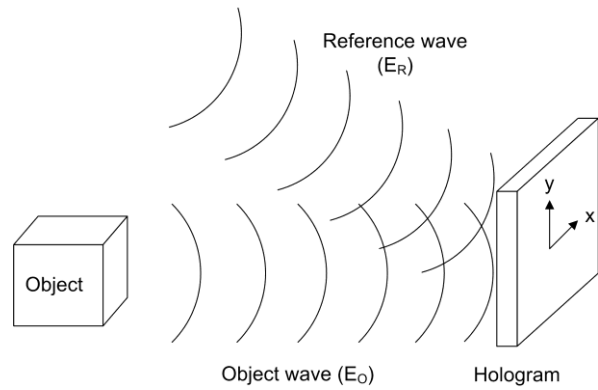$$E_O(x,y) = A_O(x,y)\exp(i\varphi_O(x,y)) \qquad (2)$$

Figure 1: Illustration of hologram generation

The interference of the reference and object waves at the surface of the recording medium has an intensity, $I(x,y)$, that represents the hologram. This can be calculated by

$$I(x,y) = |E_O(x,y) + E_R(x,y)|^2 \qquad (3)$$

$$= E_O(x,y)E_O*(x,y) + E_R(x,y)E_R*(x,y)$$
$$+ E_O(x,y)E_R*(x,y) + E_R(x,y)E_O*(x,y)$$

2.2 Image reconstruction from hologram

A hologram of an object that is situated at a distance z from the recording device (e.g. CCD) can be used to reconstruct a virtual image of the object at the position of the original object. A real image of the object is also formed at a distance z from the recording device, but in the opposite direction of the recording device. This is illustrated in figure 2. For reconstruction of the undistorted real image at the position previously occupied by the object during recording, the conjugate reference beam is used to illuminate the hologram and obtain a reconstructed image of the object.

For image reconstruction, the diffraction of a light wave at an aperture is calculated. In this case, the hologram is the aperture, and the light wave is the conjugate reference beam, which is assumed to be perpendicular to the hologram.

This is the in-line digital holography set-up. The diffraction can be described quantitatively by the Fresnel-Kirchoff integral in equations (4) to (6). Equation (4) forms the basis for numerical hologram reconstruction. It should be noted that the spherical reference wave is assumed to be far enough from the object to have planar

properties, and thus the conjugate reference beam is equivalent to the reference beam: $E_R* = E_R \equiv A_R$.
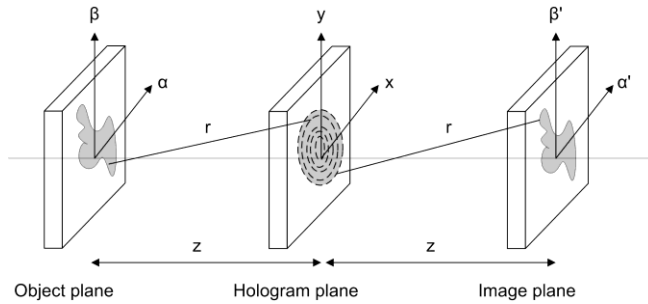

Figure 2: Illustration of coordinate system for hologram reconstruction. Image adapted from [1]

$$I(\alpha',\beta') = \frac{j}{\lambda} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) E_R(x,y) \frac{e^{(-jkr')}}{r'} dxdy \qquad (4)$$

with

$$r' = \sqrt{(x-\alpha')^2 + (y-\beta')^2 + z^2} \qquad (5)$$

and

$$k = \frac{2\pi}{\lambda} \qquad (6)$$

2.3 Numerical reconstruction

Numerical approximations can be implemented to calculate the Fresnel-Kirchoff integral for reconstruction of the image. A commonly used technique is the Fresnel approximation, which is used in this work as it is valid for microscopy applications where the object dimensions are small compared to the distance z.

Equation (5) can be approximated by a Taylor series [1, 9], which enables equation (4) to be modified to the Fresnel approximation or Fresnel transformation as given by equation (7). The double integral can be recognised as a Fourier transform, simplifying the numerical calculation for image reconstruction.

$$I(\alpha',\beta') = \frac{j}{\lambda z} \exp(-jkz) \exp(-j\frac{\pi}{\lambda z}(\alpha'^2 + \beta'^2))$$

$$\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[E_R(x,y) h(x,y) \exp(-j\frac{\pi}{\lambda z}(x^2 + y^2)) \right. \qquad (7)$$

$$\left. \exp(-j\frac{\pi}{\lambda z}(x\alpha' + y\beta'))dxdy \right]$$

3. DIGITAL IN-LINE HOLOGRAPHIC MICROSCOPE (DIHM) PLATFORM

The digital in-line holographic microscope (DIHM) platform that was implemented makes use of a laser diode light source, an aperture, and a complimentary metal oxide semiconductor (CMOS) image sensor.

A digital hologram of the object being investigated is captured by the image sensor. Image reconstruction is then performed on the hologram to provide an image of the object with a large field of view. CMOS sensors, such as those found in mobile cameras, are readily available and can be utilised for digital microscopy applications, providing a lensless system that is low in cost.

To construct the DIHM platform, interchangeable light sources are used. An infrared laser diode (808 nm) (L808P010,Thorlabs) and a blue laser diode (408 nm) (ML320G2, Thorlabs) have been used. An aperture of 50 μm was used to improve the coherence of the light. The object to be imaged was mounted on a glass microscope slide with dimensions 76 mm x 26 mm x 1 mm. The slide contained either a printed microscopic image, or a stained and fixed blood film, as prepared by pathologist laboratories. The image sensor used was a 1/2.5-Inch 5MP CMOS digital image sensor (MT9P031, Aptina) with a 2.2 μm x 2.2 μm pixel size. The DIHM platform was set up in such a way that the distance d between the aperture at the light source and the object was 200 mm to ensure a planar wave at the object plane. The distance z between the object and the image sensor was set at 2 mm. Figure 3 provides a simple illustration of the DIHM platform set-up with the various components and parameters of the system and figure 4 shows the physical experimental set-up that was manufactured.
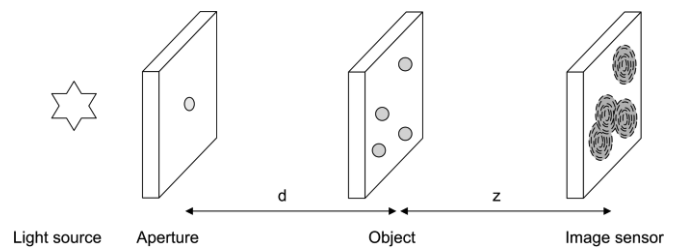

Figure 3: Illustration of DIHM platform components and parameters

4. SIMULATIONS

An algorithm was implemented in Matlab to simulate the wave propagation and hologram formation as generated by the physical DIHM platform. The simulation of the wave propagation parameterizes and closely models the physical experimental platform and enables diffraction patterns and holograms of arbitrary objects of a microscopic scale to be generated. This provides a simulation tool that enables parallel testing and optimization of the physical DIHM platform parameters.
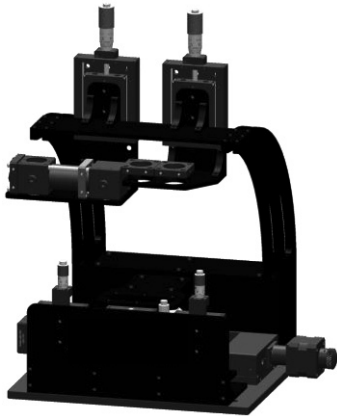
Figure 4: Mechanical design of experimental DIHM platform

5. RESULTS

5.1 Simulation results

The set of images in figures 5 to 7 show results of a forward wave propagation simulation using an artificial object. The artificial object has a diameter of 50 μm with smaller irregular shapes inside, similar to the structure of a white blood cell found in human blood.

The model simulates the propagation of a defined light source through the object to create a diffraction pattern (not shown) and a hologram of the object as shown in figure 6. The hologram is then used to reconstruct an image of the original object and the result is shown in figure 7.

The reconstruction algorithm uses as an input either an image of a hologram produced by the wave propagation simulation, as in figure 7, or a hologram image produced by the experimental DIHM platform (figures 8 and 9). The reconstruction algorithm then generates an in focus image of the original object.

5.2 Experimental DIHM results

Once simulations of hologram generation and image reconstruction were verified for an artificial object, holograms were recorded using the DIHM platform. As an initial test, a positive 1951 United States Air Force (USAF) Wheel Pattern Test Target slide (R3L1S4P, Thorlabs) was used as the object, with lines and numbering in varying sizes, with the smallest feature sizes around 2 μm. The infrared laser was used as the light source, with the set-up as described in section 3.
Using these parameters, figure 8(a) shows the digital hologram of the central area of the USAF slide, recorded by the CMOS sensor on the DIHM platform.
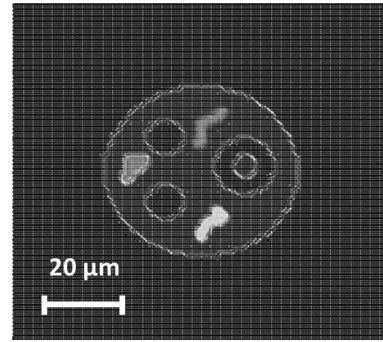

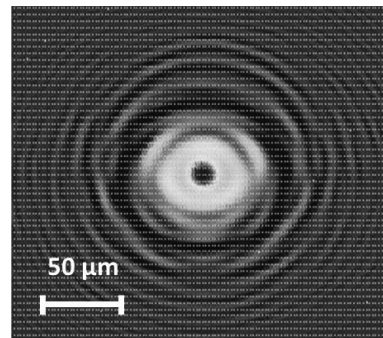Figure 5: Image of the original simulated object

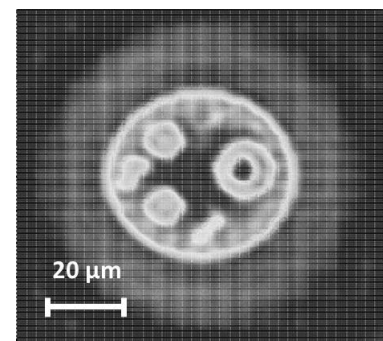Figure 6: Image of the projected hologram

Figure 7: Image of the object reconstruction from the hologram

The digital hologram was then used as an input to the image reconstruction algorithm. The algorithm first performs pre-processing of the hologram image by means of a Laplacian filter to enhance the contrast of the hologram. The reconstructed USAF slide image is shown in figure 8(b).

To test the abilities of the DIHM further, blood smear slides were imaged. Experiments were performed with both the infrared and blue laser diodes.
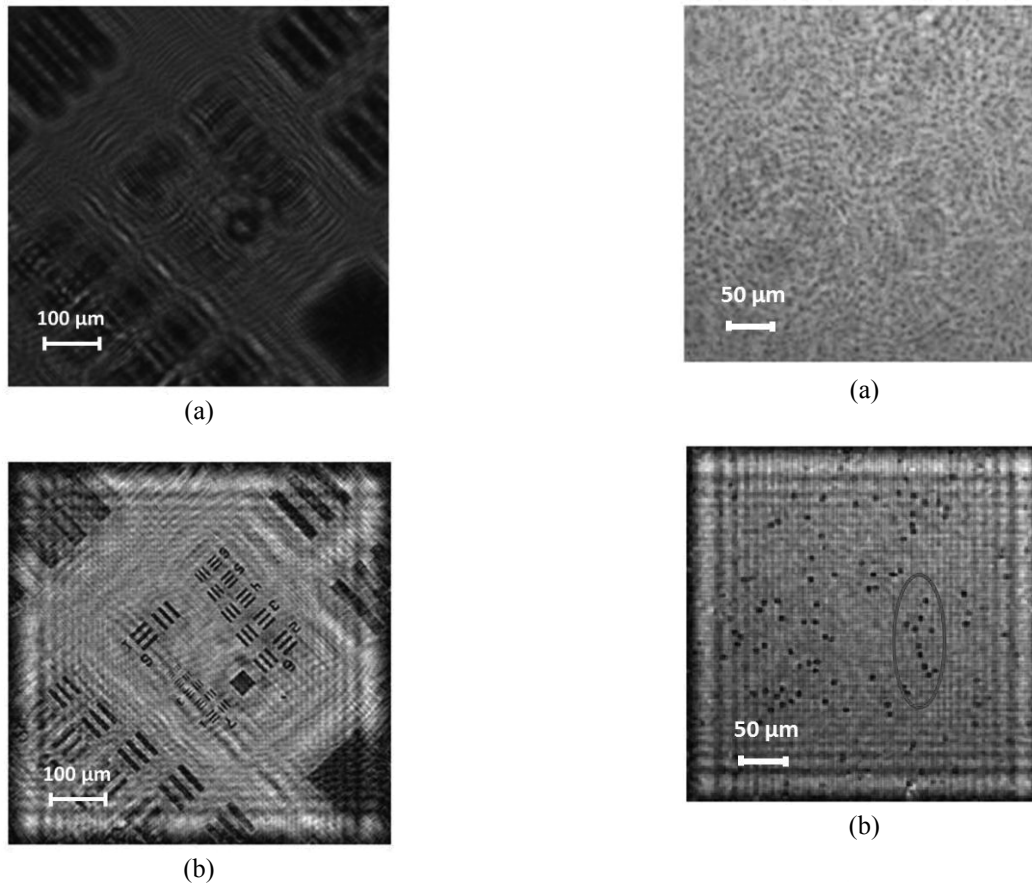
(a)

(b)

Figure 8: (a) Hologram of USAF slide obtained using the DIHM platform, and (b) reconstructed image of USAF slide using pre-processed hologram as input to the reconstruction algorithm

A hologram of a small area of a blood film slide that was obtained using a blue laser diode, and other parameters as described in section 3, is shown in figure 9(a). The corresponding reconstructed image is shown in figure 9(b), with a comparison to an image of the same area of the blood film obtained using a conventional bright field microscope with 400X magnification. The circled areas in figure 9(b) and (c) assist in highlighting corresponding areas in the two images.

The abrupt changes at the edges due to the finite sample size of the hologram results in noise at the borders of the reconstructed images in figures 8(b) and 9(b). This can be addressed by using windowing functions. However, initial results obtained indicated that the reconstructed image was of a higher quality when using the cropped images as compared with using standard windowing functions. More advanced and customized windowing techniques could be investigated and implemented to improve the reconstructed image quality.

The blue light source provided clearer results for imaging red blood cells, which are more prevalent than white blood cells in a blood film. This suggests that information from different light sources could be combined for optimal image reconstruction results.
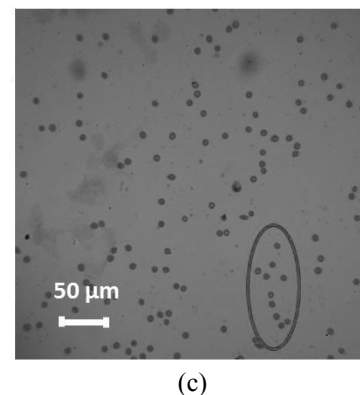

(a)

(b)

(c)

Figure 9: (a) Hologram of blood smear obtained from DIHM, (b) reconstructed image of blood smear, and (c) comparison to image of blood smear obtained using conventional bright field microscope (400 X).

6. DISCUSSION

Initial results for both the simulated and experimental DIHM system show the successful implementation of both the physical DIHM platform as well as the reconstruction algorithm.

The implementation of the wave propagation simulations allowed for diffraction patterns and holograms to be generated as a parameter optimisation tool for the DIHM platform, and also enabled the simulations to be verified against existing diffraction theory literature. By using a simple circular aperture through which a light source propagates, the Fraunhofer diffraction pattern or Airy

pattern produced was recorded. The maximum and minimum values along a cross section of the simulated diffraction pattern compared well to expected theoretical values [9] as shown in table 1.

Normalized x values simulated	Normalised x values [9]	Maximum or minimum
0	0	Max
1.187	1.22	Min
1.661	1.635	Max
2.215	2.233	Min
2.69	2.679	Max
3.244	3.238	Min
3.718	3.699	Max

Table 1: Locations of maxima and minima of simulated and theoretical circular aperture diffraction patterns

The diffraction patterns obtained using the wave propagation simulations indicated the successful implementation of the diffraction theory. This shows that the simulations can be used as a tool in investigating diffraction pattern and hologram pattern generation and the parameters required to achieve this.

In addition, since reconstruction of the original image (either simulated or real) was successful using a hologram as an input, this shows that the reconstruction algorithm was successfully implemented.

The reconstruction algorithm provided successful image reconstruction for both simulated artificial objects, as well as real objects where the holograms were recorded using the physical DIHM system. This shows that the physical implementation of the DIHM platform was successfully carried out.

Initial tests to investigate the reconstruction of images of blood cells from a recorded hologram of a blood smear slide have been shown to be successful with a high resolution. Red blood cells are typically 6 – 8 um in size, and it is clear that the DIHM can resolve these cells (figure 9(b)).

Potential improvements to the DIHM system can be investigated and implemented in terms of physical and image processing techniques. An off-axis set-up as opposed to the in-line digital holography set-up would improve the quality of the reconstructed image, as the reference wave and object waves can be spatially separated and analysed independently. Improvements in the image quality could then further be obtained by using a digital sensor with a smaller pixel size.

The finite sample size or sub-section of the entire hologram used in the image processing of this work does affect the reconstructed image quality. However, [1] demonstrates that the sub-image contains enough of the

hologram information to adequately reconstruct the image. As the sub-image becomes smaller, the quality of the reconstructed image will deteriorate. The optimal size of the sub-section of the hologram could be investigated in future.

A number of image processing techniques have been implemented in this work, including the implementation of a pre-processing Laplacian filter to enhance the contrast of the hologram image before performing image reconstruction. This high frequency filter was found to have the greatest effect on the reconstructed image quality out of all the processing techniques investigated. Smoothing functions were also implemented to soften the noise in the background of the image.

Other image processing techniques investigated included interpolation, stretching, rescaling, and histogram equalisation of the image [11]. The techniques initially investigated were found to have a minimal effect on the quality of the reconstructed image. Further investigation of pre- and post-processing techniques can be carried out in future work to determine the improvements to the reconstructed image quality.

In addition to investigating blood cells and other biological elements for medical imaging, a potential application of the digital holography techniques implemented in this work would be the inspection of micro-defects, which would find use in various manufacturing fields.

7. CONCLUSION

Initial results show that the DIHM system produces similar images to those obtained with a conventional bright field microscope. Using this successful lensless imaging system as a basis, the system can be further developed to perform automated identification and counting of the objects being investigated. For example, an initial application of this work will be targeted towards the implementation of an automated full blood count, which could provide resource limited areas with improved healthcare facilities and reduced diagnosis times at a low cost.

8. REFERENCES

[1] U. Schnars and W. Jüeptner: *Digital Holography: Digital Hologram Recording, Numerical Reconstruction, and Related Techniques,* Springer-Verlag, Germany, first edition, chapter 1-3 pp. 1-69, 2005.

[2] U. Schnars and W. Jüptner: "Direct recording of holograms by a CCD-target and numerical Reconstruction", *Applied Optics*, Vol. 33 No. 2, pp. 179-181, 1994.

[3] J. Garcia-Sucerquia, W. Xu, S. Jericho, P. Klages, M. Jericho and H. Kreuzer: "Digital in-line holographic microscopy", *Applied Optics*, Vol. 45 No. 5, pp. 836-850, February 2006.

[4] A. Ozcan and U. Demirci: "Ultra wide-field lens-free monitoring of cells on-chip", *Lab on a Chip*, Vol. 8, pp. 98-106, 2008.

[5] W. Bishara, T.-W. Su, A.F. Coskun and A. Ozcan: "Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution", *Optics Express*, Vol. 18, No. 11, pp. 11181-11191, May 2010.

[6] W. Lee, Y.-G. Kim, B. Chung, U. Demirci and A. Khademhosseini: "Nano/Microfluidics for diagnosis of infectious diseases in developing countries", *Advanced Drug Delivery Reviews*, Vo. 62, pp. 449-457, 2010.

[7] P. Yager, G.J. Domingo and J. Gerdes: "Point-of-care diagnostics for global health", *Annual Review of Biomedical Engineering*, Vol 10, pp. 107–144, 2008.

[8] "Mobile Applications on Health and Learning", *United Nations Department of Economic and Social Affairs, Division for Public Administration and Development Management, Compendium of ICT Applications on Electronic Government* , New York, Vol. 1., 2007.

[9] J.W. Goodman: *Introduction to Fourier Optics*, McGraw- Hill, Inc., second edition, chapter 3-4 pp. 32-89, 1996.

[10] M. Kim: "Principles and techniques of digital holographic microscopy", *SPIE Reviews*, Vol. 1, pp. 1 – 50, 2010.

[11] R.C. Gonzalez and R.E. Woods*: Digital Image Processing*, Prentice-Hall, Inc., second edition, chapter 3-5 pp.75-281, 2002.

OPTICAL AND THERMAL APPLICATIONS IN GRAPEVINE (*VITIS VINIFERA* L.) RESEARCH – AN OVERVIEW AND SOME NOVEL APPROACHES

A. E. Strever*, D. Bezuidenhout**, R. Zorer***, T. Moffat*, J. J. Hunter**** & *

* *Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa, E-mail: aestr@sun.ac.za*
** *Software Engineering, DynaSoft, 8 Schreiner Close, Strand, 7140, South Africa.*
*** *GIS and Remote Sensing Unit, Biodiversity and Molecular Ecology Department – DBEM, IASMA Research and Innovation Centre, Fondazione Edmund Mach, Trento, Italy.*
**** *ARC Infruitec-Nietvoorbij, Private Bag X5026, 7599 Stellenbosch, South Africa.*

Abstract: In this article, some optical and thermal applications in grapevine research are reviewed and methods to quantify the light and temperature regime around a grape bunch are discussed. This includes temperature measurement techniques (thermocouples and thermal imaging) as well as methods to quantify light quantity (hemispherical photography) as well as light quality (spectroradiometric applications) around a grape bunch. Available methods for real-time quantification of grapevine canopy size and density for application in variable rate technology sprayers are discussed, and a novel and simple approach of using opto-electronic sensors for quantification of grapevine canopy thickness and density is presented. Some scientific as well as practical applications of these individual techniques are discussed, along with their potential integration to improve knowledge of the grape bunch and canopy interaction with the environment.

Key words: Optical, thermal, *Vitis vinifera*, hemispherical photography, gap fraction.

1. INTRODUCTION

The wine industry plays an important role in South-Africa, growing in 2008 to an annual contribution to the economy of R26.2-billion or 2.2% of the gross domestic product (GDP) [1]. Wine and grapevine-related research contributes significantly to this thriving industry. Wine research is closely related to the monitoring of the grapevine and its environment. Novel techniques for quantifying and understanding the environment around a grape bunch are needed in research, especially due to the variable nature of a vineyard and its product [2, 3]. When considering sensor applications in grapevine research, an important consideration is the spatial and temporal scale at which these measurements are conducted, in and around the vineyard. On a macro-scale we can consider the interpolation of land surface temperature data from satellites using geographical information systems [4, 5]. This approach is also dependent on a sensor network for calibration purposes and can include satellite/aerial remote sensing techniques. On a meso-scale (on-farm or between-vineyard scale) we mostly deal with remote sensing technology (which can also be performed on proximal or canopy scale) as well as weather stations that are used to make meteorological observations. On a micro scale, measurements are made within or around the grapevine canopy, within or around a bunch or even within or around a single berry or leaf (also referred to as nano-scale measurements). If an analogy is to be made between a plant and sensors, it could be reasoned that a plant is an array of sensors able to react to light quantity and quality on a spatial and temporal basis, leading to physiological reactions affecting its growth, reproductive functionality and commercial quality. A complicating factor when using sensors in grapevine research is that we are dealing with effects of light quality and quantity on leaves and fruit in a spatially complex and temporally changing environment. The ideal sensor arrangement would be able to quantify, for instance in an imaging spectrometer arrangement, light quantity and quality over time in various positions within the canopy. Unfortunately, such an arrangement would be extremely expensive, especially if replication is required. For the moment at least, we can rely on sensors and modelling to get close to this goal, but certainly electronics and sensor technology could help in studying these interactions in a more efficient way, be it in terms of lower cost or higher resolution of the data generated. Another issue in using sensors in grapevine research is that sometimes care is not taken to understand the calibration conditions and measuring limits of the instruments. This is especially true when conducting light measurements, where the calibration of for instance pyranometers is done in high light environments comparable to the conditions that the sensor would encounter under clear skies. Measurements are then done inside canopies under low light conditions with different wavelength distribution and it is accepted that the calibration is still valid. While we are looking for improved solutions, this paper aims to expose the electronic engineering world to some of the applications and challenges of grapevine (and certainly also other fruit) research related to light and temperature

applications in the hope of improving the scope of current, and the development of novel technology. The importance of optical/thermal sensors in current research can be narrowed down for the purpose of this paper to: a) improving sustainable viticultural practices (such as the efficiency of spray application), b) assessing the adaptation of the plant to a change in its environment (i.e. climate change and cultivar adaptation) and c) using sensors to support general grapevine management/decision making.

2.   BUNCH LIGHT/TEMPERATURE QUANTIFICATION

2.1 Bunch temperature measurements

Thermocouples: Several types of thermocouples and logging devices are available, with their main advantage in grapevine research being the ability to measure temperature inside an organ/environment. It is therefore mostly used to measure soil, bunch or berry temperatures in microclimate studies. Thin-type thermocouples are 0.5 mm in diameter and 3.3 mm in length, compared to the standard type thermocouples that are 2.5 mm in diameter and 8 mm in length (Figure 1). The thinner thermocouples can be adapted for use with the same type of datalogger than the thicker type, and have obvious benefits when monitoring for instance grape berries, as the entry wound is minimal and the distance into the berry can be adjusted for different measurements.

Figure 1: Thin (left) and thick (right) type thermocouples that can be used in grape berry temperature assessment.

Thermal imaging is used in grapevine studies for the detection of stress responses in grapevines under different irrigation regimes [6] as well as to detect pre-symptomatic increase or decrease in temperature in grapevine leaves infected with *Plasmopara viticola* [7]. It offers an advantage above conventional infrared thermometry, in the sense that it is possible to account for target background effects, but has the disadvantage that more extensive processing of the signal is required to achieve this. Another advantage over the use of thermocouples for bunch/berry temperature measurement is that it can be used to quantify target variability in temperature. Low-cost thermal imaging techniques are now within reach of most researchers in plant science. In Figure 2 a field set-up with a Raz-IR Nano thermal imager (SPI, Las Vegas, USA) is shown, and in Figure 3 bunch surface temperature differences between a shaded or exposed bunch, as well as temperature variability on the bunch are noticeable. The importance of assessing not

only bunch mean temperature, but also variability in berry temperature is shown in Figure 4, where it is evident that just one hour change in irradiation and ambient temperature can affect bunch surface temperature variability.

Figure 2: Thermal imager set up for measurements of grape bunch temperatures.
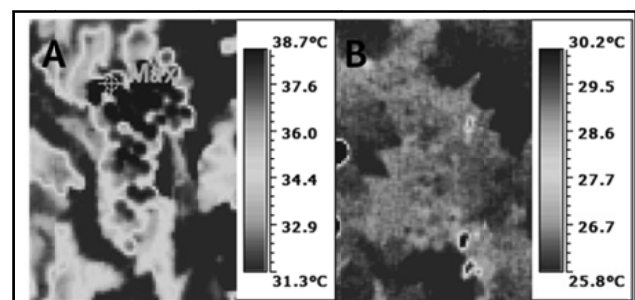

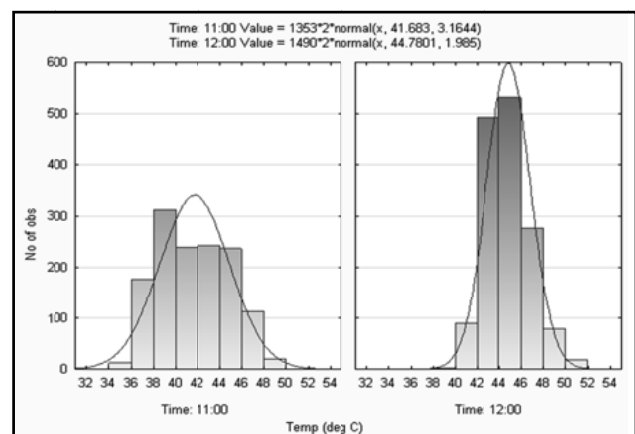Figure 3: Thermograph of the eastern (A) *versus* western (B) side of a bunch in a north-south oriented row

Figure 4: Histogram of "temperature pixels" assessed on the same bunch at 11:00 and 12:00 on the same day.

2.2 Bunch light quantity/quality measurements

Hemispherical photography: While ambient temperature is of importance to grape ripening processes, assessing fruit exposure to sunlight is also important due to the potential effect of radiation on berry temperature and composition, with potential effect on grape quality. Differences in temperature between ambient air and exposed fruit increase as solar radiation increase and

wind speed decrease [8]. The light microclimate above bunches can be computed by means of hemispherical pictures (Figure 5) in order to assess the effect of row direction, canopy gap fraction and trellis height during different times of the day, or days within a growing season on the quantity of light received. It is possible to use a special lens (Fisheye Converter FC-E8 Nikon Corp., Tokyo, Japan) and camera (Coolpix 995, Nikon Corp., Tokyo, Japan) combination, and analyse the results using Gap Light Analyser software ver. 2·0 [9] in order to compute canopy gap fraction at 10-degree zenith and azimuth resolution as well as the annual percentage of incident direct and diffuse radiation.
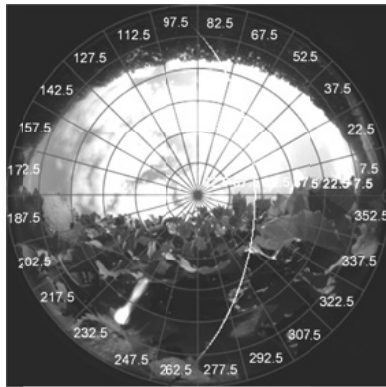

Figure 5: Hemispherical photograph taken in a north-south oriented vineyard row, with coordinate grid indicated. Sun path is indicated by the curved line across the grid.

Spectroradiometry can be used at all scale levels (aerial, satellite, canopy, leaf or bunch) to assess the spectral distribution of direct or diffuse light and also light reflected, transmitted or absorbed by different canopy elements. While pyranometers or quantum radiometers can aid in assessing total photosynthetically active radiation penetrating into a canopy, hemispherical photography can aid in modelling total radiation over time, as well as its potential reaction to a change in latitude/longitude as well as row direction. It is, however, also relevant to know how light penetrating into a canopy is affected in terms of spectral distribution, as the quality of light intercepted by leaves or fruit can affect their functioning through photon receptors able to differentiate between different wavelengths of light [10, 11].

3. CANOPY SIZE/DENSITY QUANTIFICATION

The gap fraction as well as leaf layer number (LLN) measurements are relevant in canopy microclimate studies, but also when considering light, air as well as spray penetration into a grapevine canopy. These parameters can be measured by way of point quadrat analysis [12], but this is a laborious method which is also prone to human error. It is also possible to quantify (by way of using a ruler or photographic techniques) the sunfleck pattern underneath a vineyard canopy [12], but this is also time-of-day and row direction dependent.

Opto-electronic sensors therefore offer possibilities to simplify/automate these types of measurements.

3.1 Opto-electronic sensors
Several electronic sensors are available to aid in vineyard canopy size/density quantification utilising different technologies and requiring different types of signal processing. Wavelength-specific sensors are available, similar to those deployed on satellite/aerial platforms, but they can also be deployed at a proximal level (i.e. the Greenseeker™). These sensors use the red and infrared wavebands to assess plant health. Photographic techniques have also been used to assess canopy condition [13], but this approach requires extensive processing to be useful.

We developed a simple gap fraction sensor that uses laser diodes and photo transistors to assess canopy transmittance in beams perpendicular to the surface of the canopy (Figure 6 and Figure 7). The setup is very simple and inexpensive, but needs some alignment prior to operation (Figure 8). Also the test system has been designed to run on a pre-installed wire over the grapevine canopy, which is not optimal for extensive measurements. The test system was, however, designed as a proof of concept, and could be adapted to be mounted on a spray device after mechanical stabilisation, which will also address the beam centring issue. The goals of designing such a simple setup was to a) minimise cost, b) minimise signal processing and to c) mimic a beam, or then potentially a spray droplet's ability to penetrate through the canopy. The logic is that in forced spray systems, penetration through the canopy would potentially lead to wastage, and spray nozzles could be adapted to lower application where potential penetration could occur through the whole canopy.

As the penetration of light as well as spray droplets through a canopy would be dependent on canopy thickness as well, an experiment was done by mounting two optical (infrared) and one ultrasonic low-cost distance sensor onto the setup (Figure 9).

The test run was performed on canopies where the grapevines were subjected to canopy manipulation by way of early shoot removal (Figure 10). Half of the shoots were removed for three grapevines, leaving three vines in-between unaltered as controls.

Some results of a trial run are shown in Figure 11, from which it is evident that the ultrasonic beam performed quite well compared to the hand measurements and gap fraction results, but the optical (infrared) distance sensors were erratic. The shoot-thinned grapevines showed a much higher gap fraction than the unmodified vines.
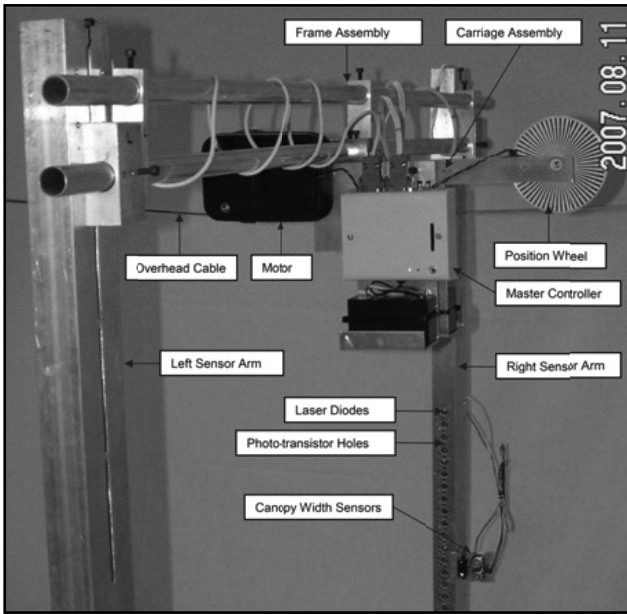
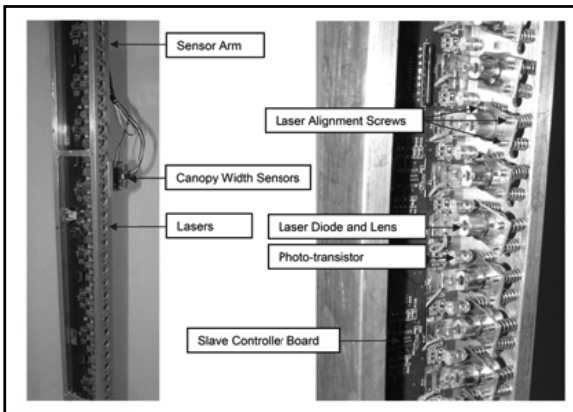Figure 6: Canopy gap fraction analyser components.

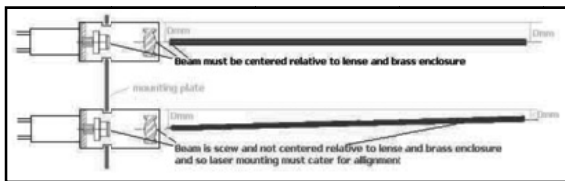Figure 7: Detail of canopy gap analyser laser beam and photo transistor setup inside the sensor arms.

Figure 8: Beam centring to align it to the photo transistor holes.
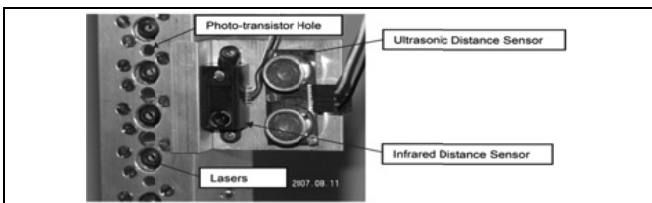

Figure 9: Ultrasonic and infrared (optical) distance sensors installed on the laser beam track to assess canopy thickness.

Figure 10: Unmodified (left) or early shoot thinned (right) grapevines used in test setup.
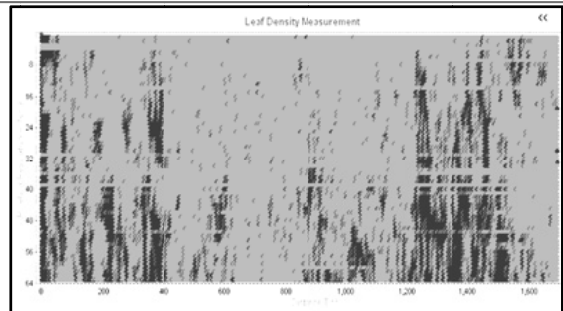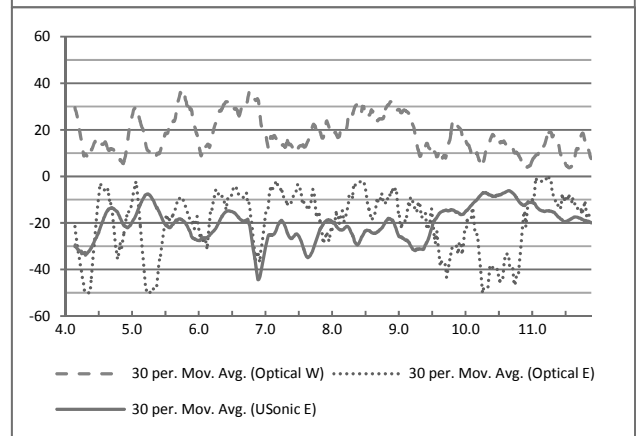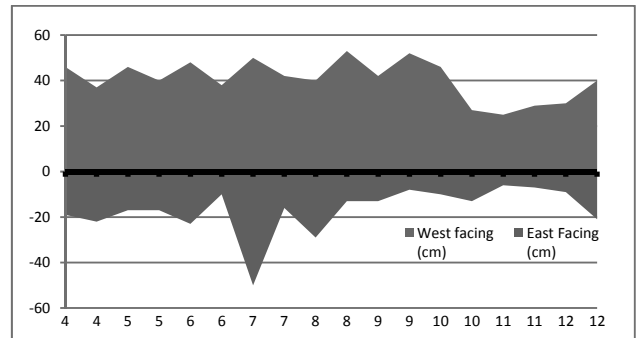

Figure 11: Top – hand measurements of canopy thickness; middle – ultrasonic and optical measurement of canopy thickness; bottom – gap fraction map showing gaps in darker pixels. The numbers on the x-axis correspond to grapevines 4, 5 and 6 (shoot thinned), grapevines 7, 8 and 9 (unmodified) and grapevines 10, 11 and 12 (shoot thinned).

4. SENSOR APPLICATIONS IN SUSTAINABLE VITICULTURE

One example of a macro/meso scale application of sensor technology is the initiative "Grapelook", co-funded by the Department of Agriculture, Western Cape, South Africa, and the European space agency (ESA). It is executed by WaterWatch (Netherlands) and the

University of Kwa-Zulu Natal in South Africa, with the aims of increasing water use efficiency in vineyards, promoting sustainable optimal resource utilization, reducing input costs (i.e. fertilizers) and protecting the environment [14]. The system provides an online interface where grape producers can access information such as leaf area index, crop coefficients, aspect, actual evapotranspiration, evapotranspiration deficit and biomass water use efficiency on maps utilising a Google Earth™ image as background.

An example of sensor applications for sustainable agriculture/viticulture on the meso-scale is the use of variable rate spraying. To this end, the method of assessing canopy dimensions/density real-time is of interest. Several options are available, including multi-band type sensors, photographic techniques, ultrasonic or infrared distance sensors, laser gap fraction analysis or LIDAR. Even though the gap fraction analysis system looks promising to use as a low-cost alternative, canopy profiling through the use of LIDAR would probably still offer the best way to combine canopy profiling and gap fraction analysis into one sensor. It has already shown promise for use in spray technology [15]. Preliminary results from some of these studies demonstrate reduction in spray drift, reduction in cost of application (savings on chemicals, but also less refill cycles) and therefore also reduced impact on the environment [16].

5.   CONCLUSIONS AND FUTURE WORK

In combination, the technologies that are described in this paper can be powerful drivers to model berry temperature, as already proposed [17]. Furthermore, ways to profile and describe a grapevine canopy and therefore the bunch environment better and with more sampling points (such as when using LIDAR) can help us to better understand and quantify the variability that is often measured in grape temperature or light interception, and perform better measurements to compliment research on the microclimate level.

It is also necessary to find ways to describe the grapevine canopy light and temperature environment better, in order to add value to regional/macro scale analyses such as those performed when satellite data are interpolated. This could improve current available decision making tools by incorporating meso- and micro-scale variability in temperature/light into existing models and to set up new models for the future.

6.   REFERENCES

[1]   Anonymous, "South-African wine industry 'driving jobs, GDP'", www.southafrica.info, 2010.

[2]   A. Strever, "A study of within-vineyard variability with conventional and remote sensing technology.", MSc(Agric)Viticulture Thesis, Stellenbosch University, 2003.

[3]   J. J. Hunter, E. Archer, and C. G. Volschenk, "Vineyard management for environment valorisation", in *Eighth International Terroir Zoning Congress*, Soave, (Verona), Italy, 2010, pp. 7-3 to 7-15.

[4]   R. Zorer, L. Delucchi, M. Neteler, and G. Nicolili, "Development of a GRASS-GIS application for the characterization of vineyards in the Province of Trento (Poster)", in *VIII International Terroir Congress* Soave, (Verona), Italy, 2010.

[5]   R. Zorer, D. Rocchini, L. Delucchi, F. Zottele, F. Meggio, and M. Neteler, "Use of multi-annual MODIS land surface temperature data for the characterisation of the heat requirements for grapevine varieties", in *Multitemp 2011*, 2011.

[6]   O. M. Grant, Å. u. Tronina, H. G. Jones, and M. M. Chaves, "Exploring thermal imaging variables for the detection of stress responses in grapevine under different irrigation regimes", *Journal of Experimental Botany,* vol. 58, pp. 815-825, 03/01 2007.

[7]   M. Stoll, H. R. Schultz, and B. Berkelmann-Loehnertz, "Exploring the sensitivity of thermal imaging for Plasmopara viticola pathogen detection in grapevines under different water status", *Functional Plant Biology,* vol. 35, pp. 281-288, 2008.

[8]   R. Zorer, T. Cobelli, T. Tomasi, L. Zulini, and M. Bertamini, "Effect of temperature and light availability on ripening of *Vitis vinifera* L. cv. Chardonnay.", in *XIV International GESCO Viticulture Congress*, Geisenheim, Germany, 2005, pp. 319-325.

[9]   G. W. Frazer, C. D. Canham, and K. P. Lertzman, "Gap Light Analyzer (GLA), Version 2.0: Imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs, user manual and program documentation". : Simon Fraser University, Burnaby, British Columbia & the Institute of Ecosystem Studies, Millbrook, New York, 1999.

[10]   R. E. Smart, S. M. Smith, and R. V. Winchester, "Light Quality and Quantity Effects on Fruit Ripening for Cabernet Sauvignon", *American Journal of Enology and Viticulture,* vol. 39, pp. 250-258, 1988.

[11]   H. Smith and G. C. Whitelam, "Phytochrome, a family of photoreceptors with multiple physiological roles", *Plant, Cell & Environment,* vol. 13, pp. 695-707, 1990.

[12]   R. Smart and M. Robinson, "Sunlight into wine: A handbook for winegrape canopy management". Adelaide: Winetitles, 1991.

[13]   H. Bjurström and J. Svensson, "Assessment of grapevine vigour using image processing", Master's thesis, Linköping University, 2002.

[14]   C. Postimus, "The intelligent pixel - helping the farmer: water use efficiency", *SABI magazine,* vol. 3, pp. 10-11, 2011.

[15]   J. Llorens, E. Gil, J. Llop, and A. Escolà, "Ultrasonic and LIDAR Sensors for Electronic

Canopy Characterization in Vineyards: Advances to Improve Pesticide Application Methods", *Sensors,* vol. 11, pp. 2177-2194, 2011.

[16]   J. Llorens, E. Gil, J. Llop, and A. Escolà, "Variable rate dosing in precision viticulture: Use of electronic devices to improve application efficiency", *Crop Protection,* vol. 29, pp. 239-248, 2010.

[17]   G. Cola, O. Failla, and L. Mariani, "BerryTone—A simulation model for the daily course of grape berry temperature", *Agricultural and Forest Meteorology,* vol. 149, pp. 1215-1228, 2009.

GROWTH AND CHARACTERISATION OF INAS PHOTODETECTORS FOR MWIR APPLICATIONS

M.C. Wagener*, V. Wagener and J.R. Botha

Department of Physics, Nelson Mandela Metropolitan University, P.O. Box 77000, Port Elizabeth, 6031, South Africa. * E-mail: Magnus.Wagener@nmmu.ac.za

Abstract: This paper reports on the development of InAs photodiodes by the Department of Physics at Nelson Mandela Metropolitan University. The device structures have been grown by metal-organic vapour phase epitaxy and processed using conventional photolithography techniques. Due to the narrow band gap of these materials, the detectivity of the devices are often limited by the junction leakage currents associated with avalanche multiplication and trap assisted tunnelling and generation effects, as well as surface related conduction channels. The various contributions to the leakage current and photo-response have been analysed, and correlated to the material and electrical characteristics of the device structures grown.

Keywords: InAs, photodiode, photoresponse, MOVPE

1. INTRODUCTION

The $InAs_{1-x}Sb_x$ material system is suitable for detection in the highly strategic 3 µm to 5 µm range of the infrared spectrum. As a prelude to the development of the InAsSb photodiodes, the electrical and structural characteristics of epitaxial InAs films have been characterised in detail. In particular, the accurate determination of the p-type and background doping of InAs offers various challenges due to the large contribution by the surface of the material to the measured conductivity [1]. Various electrical characterisation techniques have consequently been developed at Nelson Mandela Metropolitan University (NMMU) in order to optimise the device structure [2-4]. The basic material characterisation work leading up to device processing has consequently greatly contributed to development of InAs photodetectors with response characteristics comparable to commercial devices.

2. EXPERIMENTAL DETAILS

The homoepitaxial InAs diode structure was grown by metalorganic vapour phase epitaxy using a laboratory scale, horizontal Thomas Swan reactor operating at atmospheric pressure. The layers were grown on p^+ InAs substrate misoriented by 2 degrees from (001) towards (111)B. Liquid TMIn and TBAs were used as indium and arsenic precursors, respectively. The layers were grown at 600°C using a group V to group III vapour ratio of 10. The diode structure used is depicted in Fig. 1. A p^+ buffer layer was grown directly on the InAs substrate, followed by a 2 µm thick undoped layer. The p-type doping was achieved using trimethylcadmium due to the relatively low diffusion coefficient of cadmium [5]. Horikoshi *et al.* estimated the effective diffusion coefficient D_{eff} of Cd by measuring the junction depth following various anneals, with $D_{eff} \approx 6 \times 10^{-12}$ cm^2/s at 600°C [5].

The hole density was determined by a thermoelectric characterisation technique developed at NMMU [2]. A maximum doping density of 7×10^{17} cm^{-3} could be achieved. The undoped InAs is characteristically n-type, with a highly compensated carrier density in the medium to low-10^{15} cm^{-3} range [4]. The background carrier density was shown to decrease with increasing growth temperatures [4]. The growth temperature was however limited by the Cd diffusion profile required. The emitter was formed by a Sn-doped layer with a carrier density of 5×10^{17} cm^{-3}. The n^+-n-p doping sequence was used to avoid the strong residual doping effect that persists after Sn doping. Using Sn prior to the growth of the undoped layer would consequently have raised the electron carrier density significantly above the background doping level. The Ohmic contacts were formed by depositing Au/Ti contacts through a photoresist mask, with the photodiode isolated by a two-step mesa etch. The diameter of the circular photodiodes and metal contracts were 500 µm and 200 µm, respectively. No anti-reflective coating or surface passivation treatment was used.
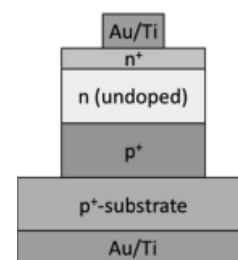

Figure 1: Schematic of the n^+-n-p photodiode structure grown by MOVPE.

The photo-response of the diode was obtained using a 100 W tungsten-halogen light source passed through a ¼-m monochromator and appropriate high pass filters. The photocurrent was measured using a Keithley preamplifier and a phase sensitive lock-in amplifier. The incident flux density was calibrated using a commercial InGaAs photodiode and a pyroelectric sensor.

3.   RESULTS AND DISCUSSION

The diode structure depicted in Fig.1 has the junction placed between the undoped layer and the p$^+$ buffer layer. This type of buried junction differs from the more conventional p$^+$-n structure, which is typically formed by diffusion/implantation of a p-type dopant into undoped InAs substrate. The characteristic inversion layer that is observed at the surface of p-type InAs [2] is, however, expected to impede minority carrier (hole) collection in a p$^+$-n photodiode, with the hole current relying on tunnelling between the surface layer and the p$^+$ emitter. Due to the high doping and defect density of bulk material, however, hole extraction from the back-contact is expected to be more efficient. Since the surface layer is also expected to result in a higher surface recombination velocity for a p-type front layer, the photodiode structure proposed in Fig. 1 is thought to be more suitable for InAs.

Considering the relatively high growth temperature used, the diffusion of cadmium from the heavily doped p-type buffer layer into the undoped layer has been considered. Since the diffusion front remains below the growth surface, the role of vacancies on the diffusion profile was neglected [6]. The Cd profile was therefore assumed to be described by a single complementary error-function. Depending on the growth rate and the thickness of the undoped layer, the position of the Cd diffusion edge/junction could therefore be well controlled [5]. Figure 2 depicts the calculated concentration profiles for the Cd doped base layer and the tin doped emitter. In the case of the device described in this paper, the junction is expected to lie at approximately 200 nm below the n$^+$ emitter, thereby ensuring that the undoped region is fully depleted.
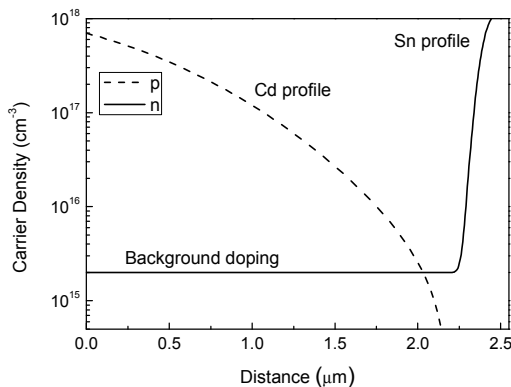

Figure 2: Simulated diffusion profiles for cadmium and tin. The abscissa denotes the distance from the intended p-doped layer interface. The background doping density has been estimated from variable temperature Hall Effect measurements [4].

The room temperature current-voltage characteristic of the photodiode, in the dark, is shown in Fig. 3. The diode exhibits rectifying behaviour with a rectifying ratio of two orders of magnitude. The reverse current follows a cube-root voltage dependence, indicative of space-charge generation within a graded junction [7]. The capacitance-voltage measurements performed on the diode at 77 K (Fig. 3 inset), also revealed a $1/C^3$ voltage dependence, indicative of a graded junction [7].
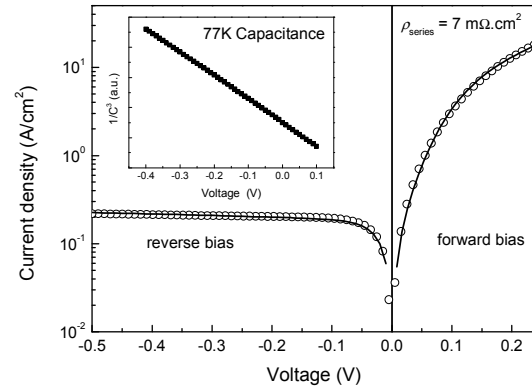

Figure 3: Room temperature dark current-voltage characteristic of the InAs photodiode manufactured at NMMU. The solid line represents the simulated behaviour. The forward bias current is dominated by band-to-band recombination, with the reverse bias current attributed to generation within the space-charge region of the diode. The inset shows the $1/C^3$ voltage dependence of the 77 K capacitance-voltage characteristics.

The forward bias characteristics appear to be ideal, which suggests that the recombination current within the diode is dominated by radiative band-to-band recombination, with a negligible contribution from Shockley-Read-Hall recombination centres. The series resistance of ρ_{series} = 7 mΩ.cm^2 was estimated from the high injection region of the forward bias measurements.

In the case of generation-recombination processes dominating diode characteristics, the differential resistance-area product is described by:

$$R_0 A = \frac{2kT\tau_0}{e^2 n_i W},\qquad(1)$$

where W is the depletion width, τ_0 the effective generation lifetime, and the other parameters have their usual meaning. Equation (1) is derived by differentiating the well know relation describing the generation-recombination process in a p-n junction [7]. The measured resistance-area product of $R_0 A$ = 0.16 Ω.cm^2 therefore relates to an effective lifetime of τ_0 = 8 ns. The resistance-area product is an important performance parameter, with the measured room temperature value comparable to commercial devices [6].

The photo-response of the InAs photodiode is depicted in Fig.4. The photodiode has a maximum sensitivity of 0.75 A/W at 2800 nm, decreasing sharply at the InAs absorption edge. The various contributions to the photocurrent have been determined by simulating the

spectral response associated with the emitter (n$^+$-region), the depletion region and the p-type base [8]. It is evident that a significant contribution is attributed to generation within the emitter, with the depletion region representing a smaller component to the overall response. The photo-response of the diode was also independent of the applied reverse bias, thus validating the proposed doping profile described in Fig. 2.
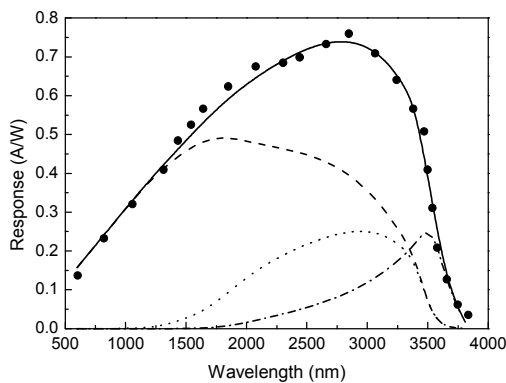

Figure 4: Room temperature spectral response of the InAs photodiode. The simulated response (—) has contributions from the emitter (- - -), base (-.-.-) and space-charge (····) region of the diode.

Figure 5 depicts the internal quantum efficiency (QE) calculated from the total response simulated in Fig. 4. Due to the Cd diffusion profile, the space-charge region of the photodiode is positioned near the surface, with the shallow n$^+$ emitter acting as a front-surface field (FSF) for holes. As a result, the FSF reduces the surface recombination velocity, thereby increasing the short-wavelength QE of the photodiode. The surface recombination could be estimated from the measured photo-response, with the increased short-wavelength response associated with a maximum surface recombination velocity of 10^5 cm/s.

The Cd doping profile is also expected to produce a back-surface field for electrons, enhancing the photo-response associated with the base of the diode. The buried diode structure, however, exhibits a reduced QE at the peak response wavelength. The negative slope of the QE below the peak response wavelength is indicative of collection from the base of the photodiode. The relatively narrow space-charge region of 200 nm obtained from 77 K capacitance measurements is likely to be the main contributing factor. Carrier recombination within the p-type region of the diode is also a likely contributor to the overall decrease in the peak response. A recent paper [4] presented deep level transient spectroscopy (DLTS) measurements performed on InAs diodes fabricated at NMMU. The measurements revealed a band of *electron* traps within InAs. These levels were found to be situated near mid-gap and related to extended defects. These electron traps are likely to impact device performance, with the trapping behaviour determined by the charge nature of the defects. Based on the DLTS results, the

extended defects are expected to be completely filled within the n-type region of the diode, whilst trapping minority carriers in the p-type region. Since no *minority* carrier (hole) deep level transients were observed, it is likely that extended defects are more detrimental within p-type regions of the device. As a result, the peak response of n$^+$-n-p diode structures is expected to be more sensitive to the crystalline quality.
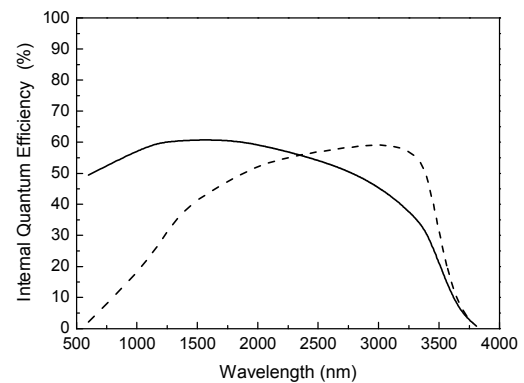

Figure 5: Comparison of the room temperature internal quantum efficiency of an InAs n$^+$-n-p photodiode fabricated at NMMU (—) and a Hamamatsu (part number: P10090) p-n photodiode (- - -).

In conclusion, the good rectifying and photo-response characteristics of an n$^+$-n-p photodiode have been demonstrated. The front-surface field associated with the n$^+$-n emitter also significantly enhanced the short-wavelength response of the photodiode. The more conventional p$^+$-n diode structure would therefore be more appropriate for applications requiring a narrow spectral band-width. A relatively low peak response was however observed and attributed to non-optimum junction characteristics. It is anticipated that the optimisation of the Cd diffusion profile, in particular the depletion width, would further improve the photodiode response.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1] M. Noguchi, K. Hirakawa, and T. Ikoma: "Intrinsic electron accumulation layers on reconstructed clean InAs (100) surfaces", *Physical Review Letters*, Vol. 66, 2243, 1991.

[2] M. C. Wagener, V. Wagener, and J. R. Botha: "Thermoelectric evaluation of the dopant density of

p-type InAs", *Applied Physics Letters*, Vol. 94, p. 262106, 2009.

[3] M. C. Wagener, V. Wagener, and J. R. Botha: "Hall and thermoelectric evaluation of p-type InAs", *Physica B*, Vol. 404, pp. 5038-5041, 2009.

[4] M. C. Wagener, V. Wagener, and J. R. Botha: "Growth temperature dependence of the background doping in MOVPE-grown InAs" *Journal of Crystal Growth*, doi: 10.1016/j.jcrysgro.2011.09.008

[5] Y. Horikoshi, H. Saito, and Y. Takanashi: "Simultanious diffusion of zinc and cadmium into InAs", *Japanese Journal of Applied Physics*, Vol. 20 No. 2, pp. 437-438, 1981.

[6] V. Tetyorkin, A. Sukach, and A. Tkachuk: *Advances in Photodiodes*, InTech, Croatia, Chapter 20, pp. 427-446, March 2011.

[7] S. M. Sze: *Physics of Semiconductor Devices*, John Wiley & Sons, Inc., USA, second edition, pp. 74-92, 1981.

[8] R. H. Bube: *Photoelectronic properties of semiconductors*, Cambridge University Press, USA, first edition, pp. 244-255, 1992.