Colophon

Text
Participants and organization workshop.

Cover photo
'Physicists assemble pieces of the ATLAS detector for the Large Hadron Collider'.
Image courtesy Argonne National Laboratory.

Design and production
Stichting FOM, Utrecht, the Netherlands.

Contact
Isabel Poyck
Foundation FOM
PO Box 3021
3502 GA Utrecht
Phone: +31 (0)30 600 12 21
Email: isabel.poyck@fom.nl

Thanks to
The organisation is particularly grateful for the excellent service and facilities of the Lorentz Center, the effort of the senior researchers (from the preparation phase onwards) and the enthusiastic contribution of all participants during the week.

The workshop 'Physics with Industry' was organised by Marcel Bartels and Pieter de Witte of the Foundation for Fundamental Research on Matter (FOM) and Floor Paauw and Marjan Fretz of the Technology Foundation STW in collaboration with Sietske Kroon of the Lorentz Center. The event was funded by the Lorentz Center (which is partly funded by FOM and STW) and the participating companies. FOM is part of the Netherlands Organisation for Scientific Research (NWO). STW is funded by NWO and the Dutch Ministry of Economic Affairs, Agriculture and Innovation.

Foreword

It is our great pleasure to present to you the proceedings of the third 'Physics with Industry' workshop that was organised by the Foundation FOM and Technology Foundation STW at the Lorentz Center in the Netherlands. The main aim of the 'Physics with Industry workshop' is to obtain creative solutions for industrial problems and to bring (young) physicists in contact with industrial R&D.

The first 'Physics with Industry workshop' was organized in 2010 and was inspired by the 'Mathematics with Industry' workshops, which have regularly been organised by the 'International Study Group Mathematics with Industry' since 1968. As well as enabling excellent scientific research, both FOM and STW focus on contributing to the Dutch knowledge economy, for example through public-private research collaborations and the training of young scientists. The 'Physics with Industry workshop' is therefore a natural extension of FOM's and STW's ambition to help companies and to inspire (young) physicists.

59 scientists participated in the workshop 2012, ranging from PhD students to professors. These scientists spent a week working in groups on five industrial problems, which were selected by a programme committee from proposals put forward by industry. Following an introduction to the various problems by the companies on Monday, the participants worked on these in groups for the rest of the week. On Friday, the groups presented their findings to the companies.

Besides the scientific outcomes, the workshop also resulted in new public private contacts that may lead to future collaborations. A novelty in 2012 was that one company (NXP) filed a patent based on the results of this workshop. Participants were mostly driven by the shear pleasure of applying their physics knowledge to new problems, the desire to enrich their scientific network and the interest in gaining hands on experience with industrial R&D processes. Companies benefited from the scientific input they received and participating in the workshop enlarged their academic network.

These proceedings provide an overview of the scientific results obtained during the third 'Physics with Industry' workshop. We hope you enjoy reading it!

Wim van Saarloos  
Director FOM

Eppo Bruins  
Director STW
Table of Contents

Janssen Precision Engineering ................................................................................................ 3
MicroDish ................................................................................................................................. 19
NXP Semiconductors .............................................................................................................. 37
PamGene International .......................................................................................................... 63
Shell Projects & Technology ................................................................................................. 81

List of participants................................................................................................................... 99
Introduction

The third workshop Physics with Industry was organized in 2012 by the Foundation FOM and Technology Foundation STW at the Lorentz Center at Leiden, the Netherlands.

The five industrial problems discussed during the week were collected via an open call for proposals in spring 2012. A programme committee selected the five 'best problems' for the workshop. The selection criteria used by the committee were:

- it must be possible to solve the problems (or a major solution must be within reach) within one week and physics can make a clear contribution to the solution;
- it should be an urgent problem;
- the company should be willing to share detailed information.

The committee aimed at a mix of contributions from small, medium and large companies. The committee consists of seven researchers with different backgrounds in physics:

Prof. Marileen Dogterom, FOM Institute AMOLF
Prof. Ute Ebert, Centrum Wiskunde & Informatica
Prof. Erik van der Giessen, University of Groningen
Prof. Fred MacKintosh, VU Amsterdam
Dr. Jacco Snoeijer, Twente University
Dr. Peter Steeneken, NXP Semiconductors
Prof. Lucas van Vliet, Delft University of Technology

The committee selected problems from the companies Janssen Precision Engineering, Microdish, NXP (Leuven), PamGene and Shell. Together a well balanced mix of SME’s and larger industries posing different challenges ranging from fundamental questions to more applied problems in a wide variety of industries. As soon as the five workshop problems had been selected, senior researchers from academia who are familiar with the specific subjects involved were recruited. They helped the companies to prepare their questions for the workshop and they joined the workshop week to guide the progress of the discussions.

These proceedings contain five chapters, one for each company case. Each chapter starts with a description of the case and a profile of the company, followed by a detailed description of the results obtained in the single workshop week.
Cryogenic compatible displacement sensor

Mireia Alos-Palop, Jildou Baarsma, Felipe Bernal, Victor Calado, Michiel De Dood, Huub Janssen, Akashdeep Kamra, Kees Landheer, Marco Zocca

1. Utrecht University, The Netherlands
2. Delft University of Technology, The Netherlands
3. ITP Utrecht, The Netherlands
4. AMOLF, The Netherlands
5. LION, The Netherlands
6. Janssen Precision Engineering, JPE, The Netherlands

1. Abstract

The state of the art experiments in low temperature physics require sophisticated instrumentation capable of displacement sensing with high precision and cryogenic environment compatibility. The present report discusses two such designs- μPOT design and OptoGroove design. The former is an all electrical method of position detection which is essentially a miniaturization of the classical potentiometer concept. The design involves measuring the voltage of a sliding probe on a conducting wire, which varies linearly with the position of the probe. The OptoGroove design, on the other hand, is essentially a digital optical encoder. An optical fiber is directed at a side surface of the actuating screw. The side surface has been laser engraved with a series of equally spaced parallel grooves. Time domain reflectometry allows counting the number of grooves during the motion of actuating screw, which in turn translates to the linear displacement. Both of the above techniques seem to be robust over a large temperature variation.

2. Company profile: JPE

JPE is an independent engineering group for development and realization of high-tech machinery and instruments. Especially where accurate and stable performance is involved in the sub-micron area. The company was founded by Huub Janssen in 1991 after several years of experience in the high-tech industry of companies like ASML and Philips. Nowadays, we built up a team of professionals which are able to find and implement solutions for very challenging engineering requests. JPE has gained multidisciplinary knowledge of technical issues at every level. From system level down to component level, from definition and design, up to prototyping and qualification. By following a systematic approach with high involvement, quality and efficiency are guaranteed. Mostly our projects are initiated by
physicists working on cutting edge research and technologies. For them, we develop high-end opto-mechanical solutions to be used in deep vacuum as well as cryogenic environment. Our developments typically find their way in an international market like:

- semi-conductor industry,
- astronomy and space,
- scientific experimental instruments.

Apart from turn-key projects, we also have separate R&D activities, often initiated as spin-offs of our projects. One of the successful spin-offs is the recently developed and introduced product: "PiezoKnob" which is a kind of cryogenic motor based on our own piezo technology able to position with nanometer accuracy in environments of a few degree Kelvin. This actuator is used by quantum physicists at distinguished universities like Leiden and Yale.

3. Problem description

For measuring displacements in an ambient environment several solutions are commercially available. One can think of: optical encoders, laser interferometers, capacitive or inductive sensors etc. Regrettably, only a limited number of these commercial sensors can operate in a vacuum environment and even less in a cryogenic environment. Recently developed actuator operates in such cryogenic environments, as mentioned before. This actuator, a friction based inertia drive, has a position resolution in the nanometer range combined with a stroke of several millimeters (see JPE website for detailed information). It is typically used to position optical elements etc. in a cryostat. As one can imagine, a cryostat is a closed environment and thus one cannot look what is happening. Therefore, many of our customers are looking for a position sensor which can measure the position of the actuated optical elements. Important issues concerning this displacement sensor is the limited space in cryostats, and the necessary feed-through for cabling (with low thermal impact) etc. In order to get an impression of the specifications, a list of the basic requirements as well as some extended requirements can be seen in the table below:

<table>
<thead>
<tr>
<th>Displacement sensor</th>
<th>Basic requirements</th>
<th>Extended requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental temperature</td>
<td>1 to 300 K</td>
<td></td>
</tr>
<tr>
<td>Measuring range</td>
<td>5 mm</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>1 µm</td>
<td>1 nm</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>2 µm</td>
<td>5 nm</td>
</tr>
<tr>
<td>Absolute position accuracy</td>
<td>10 µm</td>
<td>100 nm</td>
</tr>
<tr>
<td>Measuring frequency</td>
<td>5 Hz</td>
<td>5 kHz</td>
</tr>
<tr>
<td>Output</td>
<td>Digital (t.b.d.)</td>
<td></td>
</tr>
<tr>
<td>Max volume</td>
<td>10mm diameter, 30 mm</td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>The smaller the better</td>
<td></td>
</tr>
<tr>
<td>Cost price</td>
<td>1000 Euro per axis</td>
<td></td>
</tr>
<tr>
<td>Wires (feed-throughs)</td>
<td>≤ 6</td>
<td>≤ 4</td>
</tr>
<tr>
<td>Maximum dissipation within cryostat</td>
<td>≤ 50 µW @4K</td>
<td>≤ 1 µW @20 mK</td>
</tr>
</tbody>
</table>
4. Problem solving strategy

The team of researchers and the company leader analysed the problem from different perspectives. After long discussions, two main strategies are going to be developed in what follows. The first method focuses on the resistance measurement, named $\mu$POT, and the second method focuses on an optical encoder, named OptoGroove.

$\mu$POT@JPE

The $\mu$POT concept is essentially a digital potentiometer, in which the sliding contact is directly connected to the linear translation stage of the JPE PiezoKnob. A square-wave shaped contact strip is patterned on a fixed substrate, and kept at a known voltage. The measured voltage at the sliding contact is then a direct measure of the displacement.

OptoGroove@JPE

The OptoGroove concept is based on measuring the backreflected intensity of a laser that impinges on an engraved pattern on the PiezoKnob spindle. Measurement of the linear displacement is in this case indirect, but the device can be easily implemented with commercially available components.

5. Theoretical analysis: $\mu$POT

Potentiometers are the devices that offer a variable resistance which is controlled by a mechanical knob. If we reverse the idea, we get a position sensor - the resistance, and hence the measured voltage, varies with the position of a contact.

\[
V = \frac{R_1(x)}{R} V_{\text{bias}},
\]

where $R$ is the total resistance in the potentiometer circuit. The variation of $R_1$ is linear with $x$, the position coordinate. This gives us an expression for the measured voltage.

\[
V = \frac{x}{L} V_{\text{bias}} + V_{\text{offset}},
\]

where $V_{\text{offset}}$ accounts external circuit resistance between ambient state with cryogenic state and the contact resistance of the sliding contact. The idea now is to use this principle at a micrometer level. Towards this end, a circuit can be designed as shown in Fig. 1 using photo-lithography techniques. The circuit made of a conducting material is biased at a voltage $V_{\text{bias}}$. A third probe, which is rigidly connected to the stage, is positioned on the grid at $x$ in some coordinate system. The relevant signal is the voltage measured by this probe. An important thing to realize is that the resolution of this measurement is decided by the spacing between adjacent wires as shown in Fig. 2. This resolution can easily be chosen to be in the sub-micron region while fabricating the circuit. The only requirement is that the contact between the probe and the circuit should be much smaller than 1$\text{mm}$. 

- 5 -
Figure 1. The basic scheme of the μPOT.

Figure 2. The accuracy of the measurement is relatively insensitive to the size of the probe. (a) For a small contact area, only one wire is connected to the probe. For a larger contact area, more than one wires are shorted and connected to the probe. The voltage measured is still the voltage of the first wire in contact. Hence the resolution is still the spacing between the consecutive wires. (b) Equivalent circuit depiction of the two systems. For a thicker probe, $R'$ is the resistance which is shorted by the probe. This still does not affect the measurement as long as $R' \ll R_1 + R_2$. 
5.1 Measurement Scheme

The systems where the position sensor is meant to be used do not have a very rapidly varying position. Hence, we can choose to measure the position in bursts and not continuously. As mentioned earlier, the probe is attached to the stage whose position is to be measured. When the stage moves, the probe slides on the circuit. One measurement cycle is described below step-wise along with the basic specifications of the measurement instruments.

- The bias voltage to the circuit is switched on. The voltage steady state is reached depending upon the bandwidth of the connections.
- 2ms after the bias voltage is switched on, the analog voltage read by the voltmeter is sampled and converted to digital data. The voltmeter required needs to have a bandwidth larger than 1kHz and an input impedance larger than 1MΩ. A 14 bit analog to digital converter is required to provide enough dynamics range in the measurement electronics to resolve a step of sub-micron size with a 5mm travel range. Both these components are rather standard and base range.
- Immediately after the sampling, the circuit bias is turned off.

As per the requirement of the design, we need to measure @ 5 Hz. Hence each 2ms cycle is repeated every 200ms.

5.2 Main Features

- The measurement scheme gives us the absolute position.
- The required temperature range of operation is large - 1K – 300 K. This requires a measurement scheme which is insensitive to temperature variations, which the measurement scheme under consideration is. This is because the measurement inherently depends on the ratio between two resistances (See Eq. (1)).
- The design requires only two extra wires going into the system - one of them to bias the circuit and another connected to the probe. This is true for one axis. For each extra axis, we need one one extra wire for only the probe as the same wire can be used to bias all the potentiometers.

5.3 Circuit design

First step is to determine the range of resistance we can have. The upper limit is imposed by the input impedance of any voltmeter (~ 1 MΩ). The lower limit is imposed by two requirements - the total resistance should be larger than a typical circuit and contact resistance, and the power dissipation should be no more than ~ 1mW at low temperatures. A dc voltage bias of 1V, limits the the potentiometer resistance down to 1 kΩ. We design the potentiometer resistance in the range of 1kΩ < R < 1MΩ.

We propose using metal or alloy wiring by standard optical lithography on a Si/SiO2 wafer with taking the following material requirements in consideration:

- Non-magnetic and non-superconducting
- Chemically stable (no oxidation)
- nearly constant conductivity in 1-300 K range
When using gold the following dimensional parameters are taken: Electrodes of 300 nm in width and 900 nm in depth, yielding a resistance per unit length of $740\, \Omega\, m^{-1}$ at 1 K [1]. The width of the meandering is taken to be 1 mm and a length of 5 mm. The electrode pitch is twice the electrode thickness of 300 nm. This sets the resolution limit to 600 nm. The number of winding then becomes 8300, which can be covered by the number of voltage intervals in a 14 bit analog to digital converter. The total resistance range is $\sim 6 - 700\, k\Omega$ from cryogenic temperatures to room temperature. The maximum continuous power dissipation with 1 V dc bias is $160\, \mu W$. The average power dissipated with 2ms-200ms pulsed cycle will be 100 times lower: $1\, \mu W$. We propose that an alloy such as AuPd could be used to reduce the resistance range [3].

### 5.4 Contact between probe and sample

Moving parts are always the weakest link of any electronic system. In our system, we have a moving probe that measures the voltage. Hence the probe beckons special attention. In any mechanical system, rolling motion is preferred over sliding motion. One natural question arises - “Can we have a rolling probe contact in stead of a sliding one?” The answer is “In principle, yes.” Figure 3 depicts the two possibilities we have. The rolling contact however makes thing a bit more complicated in terms of manufacturing cost and process. Hence it might be useful to test the system with a sliding contact and find out the lifetime of such a system. Only if the lifetime turns out to be too small, one can try to implement a rolling contact. Another argument in favor of sliding contact is that it has virtually no weight, whereas the rolling contact with a wheel has considerable weight which gives an orientation dependent contact force. Hence, we choose to work with a sliding contact in this design.

There are some parameters which should be taken into consideration while designing the probe. Most important of these are the contact force and contact radius. Our system requires a contact between the probe and the sample which is strong enough to have a low contact impedance, small enough to have a contact radius much less than 1mm, and weak enough that it leads to minimal abrasion during the sliding. A typical number which satisfies these requirements is $0.1N$ of contact force. Here, we model the contact and calculate the parameters of interest.

Any contact between a flat surface and a needle–like object can be modelled as a contact between an infinite plane and a sphere. This is a standard problem called a Hertz contact which can easily found in literature [2]. An elastic sphere of radius $R$ indents an elastic half-space to depth $d$, and thus creates a contact area of radius (Fig. 4)

$$a = \sqrt{Rd},$$

(3)
The applied force $F$ is related to the displacement $d$ by

$$F = \frac{4}{3} E^* \sqrt{RD^{3/2}}, \quad (4)$$

where

$$\frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}, \quad (5)$$

and $E_1, E_2$ are the elastic moduli and $\nu_1, \nu_2$ the Poisson’s ratios associated with each body.

Using typical values of elastic moduli, we calculate the dependence of contact radius ($a$) on radius of the sphere ($R$) and the contact force ($F$). The results are shown in Fig. 5. For reasonable numbers of sphere diameter and contact force, we obtain a contact radius in few micrometers, which is within our requirements. We choose the nominal value of 0.1N contact force and 1mm probe tip diameter.

The probe is connected to a Leaf Spring (henceforth referred to as just ‘spring’) which is used to fix the average value of contact force. This design is insensitive to any small contraction (or expansion) due to temperature variation. Any change in lengths leads to a small change in the contact force exerted by the spring thereby leading to a small change in contact radius. Our design, however, does not depend on the actual value of contact radius, as long as the latter is much smaller than 1mm.

Now we consider the problem of the shape of the leaf spring, see Fig. 6. Supposing a spring force of 0.1N on the surface, the compression of the SiO$_2$ is about $u_x = 0.5\mu m$ and contact length in sliding direction is a few microns as has been calculated before. We assume a Young’s modulus of $E = 110GPa$, which corresponds to Titanium. With simple mechanics, we
now calculate the thickness \((t)\) of the leaf spring strip. We set the width \((b)\) of the leaf spring to 2\(\text{mm}\) and the length \((L)\) to 5\(\text{mm}\). We start with the simple Hooke’s law for a spring:

\[
F_x = C_x u_x.
\]  

(6)

The spring constant can be calculated as follows:

\[
u_x = \frac{F_x}{C_x},
\]  

(7)

thus, one can write:

\[
C_x = \frac{3}{2} Eb \left(\frac{t}{L}\right)^3.
\]  

(8)

Using values above mentioned, we obtain a thickness of 0.8\(\text{mm}\) for the leaf spring strip. The mass of the spring is negligible.

One final check to be performed is the heat production due to the sliding contact. We know the contact force is about 0.1\(\text{N}\), assuming a friction coefficient of 0.3 gives the heat dissipation as 30\(\mu\text{J/mm}^{-1}\). This is very small as compared to the heat generated by the actuator.

### 5.5 Heat flowing into the cryostat through electrical wire

Suppose we use a commonly used Phosporus Bronze wire for the electrical connection from the helium cooling stage of the cryostat to the \(\mu\text{POT}\) (e.g. @ 20\(\text{mK}\)). The heat power to the \(\mu\text{POT}\) through the wire is calculated by the following formula:

\[
P = \frac{kA(T_1 - T_2)}{L},
\]  

(9)

with \(k\) the thermal conductivity \((1.6 \text{ Wm}^{-1}\text{K}^{-1}@4\text{K}, 0.22 \text{ Wm}^{-1}\text{K}^{-1}@1\text{K})\) \cite{4}, \(A\) the cross section of the wire \((3.2 \cdot 10^{-8}\text{m}^2)\) \cite{4}, with \(T_1\) the temperature of the helium cooling stage (4\(\text{K}\)) and \(T_2 = 20\text{mK}\) and \(L\) the length of the wire from the helium cooling stage of the cryostat to the \(\mu\text{POT}\) (we take 20\(\text{cm}\)). Filling in the values gives a heating via one electrical wire of 1\(\mu\text{W}\). For our \(\mu\text{POT}\) we will need four extra wires, so this makes 4\(\mu\text{W}\). The resistance of this wire from \(\mu\text{POT}\) to helium stage is negligible \((R = 0.7\Omega)\) \cite{4}, so the wire can be made longer to reduce the heating up. The costumer should use heat sinks along the different cooling stages in the cryostat.
6. Theoretical analysis: OptoGroove

6.1 General description

The idea behind this sensor is to engrave with a laser a pattern onto the sidewall of the PiezoKnob spindle. The dark regions reflect less light than the bright regions, due to the surface roughness induced by the engraving process. By counting bright spots (or intensity transitions) one can deduce the rotation of the spindle and from that the linear displacement of the stage. With a simple binary engraved pattern, no absolute positioning is possible: it is up to the electronics to keep the count, and infer the corresponding total linear translation of the stage. A possible improvement to reduce radiation in the cryostat is a collimating lens, mounted between the fiber facet and the PiezoKnob spindle.

A big advantage of this measurement device is that there is little temperature dependence due to the fiber [5]: it operates by discriminating between backreflected intensity levels, and not on interferometric effects that depend on the refractive index. We predict that this sensor will work both at room temperatures and at cryogenic temperatures.

6.2 Description

The sensing system as a whole is very simple, and can be realized with off-the-shelf components: pigtailed laser and corresponding photodiode, a fiber directional coupler (DC), a few meters of optical fiber, and an optical read-head mounted on the PiezoKnob fork. The active components and the DC are located outside the cryostat, which is an advantage of this design.

The optical read-head should be kept at a minimum operating distance \( L_0 \), to avoid accidental contact due to mechanical jitter of the PK spindle. Given this distance and the numerical aperture of the fiber NA, the spot size \( W \) can be approximated by

\[
W = 2L_0 \cdot \text{NA}. \tag{10}
\]

For a single-mode fiber without a lens, the numerical aperture is about 0.15. The distance \( L_0 \) between the fiber and the spindle should be at least 0.2 mm to account for the aforementioned
Figure 8. The set up of the optical measurement device. On the right is the cryostat with inside the encoder, on the left are the photo diodes (PD) sending and receiving the light and on the far left is the receiver front-end.

jitter. In this case the spot size $W = 60 \mu m$. If a collimating lens is used the numerical aperture could be reduced and the spot size can be reduced, thus allowing a finer encoder resolution and lower optical losses.

The encoder period $T$ (one dark spot and one bright spot) is limited by the resolution of the laser engraving. We use a value of $20 \mu m$, which can be achieved using common techniques. However, if the period of the laser engraving is much smaller than the spot size more power is needed to get a signal back that can be analyzed. This translates to power delivered to the cryostat, which should be as small as possible. Here, we take that the period $T$ of the laser engraving should be at least twice the spot size on the spindle.

The translational displacement of the stage after one turn of the spindle $L_{\text{turn}}$ is known for the PiezoKnob, namely $250 \mu m$. Given a period of the laser engraving and a spindle with diameter $D_s$ the resolution $R$ obtained by this measurement device is then given by

$$R = \frac{L_{\text{turn}}}{\pi \cdot D_s/(T/2)}.$$  \hspace{1cm} (11)

For the smallest possible engraving, a period of $20 \mu m$, and a spindle with a diameter of $18 \text{ mm}$, the resolution that can be obtained is $0.044 \mu m$, far less than required. For the required resolution of $1 \mu m$, the period can be much larger, namely $452 \mu m$. In the last case little power is needed to get back a useful signal. For the fiber described above, with a numerical aperture of 0.15, a period of $120 \mu m$ should be used and then the resolution is $0.26 \mu m$.

6.3 Read Out

The light source of the measurement device is outside the cryostat, see Fig. 8. Between the source and the cryostat there is a beam splitter. The light coming back from the encoder passes by the beam splitter and goes to a read-out. There are devices readily available for detecting and amplifying light signals, for example the Si Transimpedance Amplified Photodetector PDA36A by Thorlabs. The output of such a device is an analog electronic signal.

The noise equivalent power (NEP) of this device is $10^{-11} \text{W/} \sqrt{\text{Hz}}$ for a wavelength of about $800 \text{ nm}$, which is the light used in the optical encoder, the measuring device. For a bandwidth
of 1 kHz this means that the power of the noise is 0.3 nW. We want a signal to noise ratio of at least 100, which means that the incoming signal should have a power of at least 30 nW.

### 6.4 Optical power budget

If we know that a power $P_r$ of 30 nW should be delivered to the read-out we can calculate back what the power $P_t$ of the source has to be. The light passes through the beam splitter twice, one time going into the cryostat and one time when coming out of the cryostat. Every time the light passes through the beam splitter half of the power is lost (a possible improvement is to use an asymmetric coupler, e.g. a 10/90 configuration, which would allow a better efficiency of the readout path).

Most power is lost inside the cryostat because only a small fraction of the light reflected of the spindle is reflected back into the fiber. After reflection the area of the reflected light is twice the spot size, see Fig.9. The fraction of the light $\eta$ coming back into a fiber with diameter $D_f$ is then

$$\eta = \frac{\text{area of fiber}}{\text{area of reflected light}} = \frac{D_f^2}{(2 \cdot W)^2}. \quad (12)$$

The power $P_t$ then is give by

$$P_t = 4 \frac{P_r}{\eta}. \quad (13)$$

For a fiber with a diameter of 4 µm and a spot size of 60 µm the fraction of light getting back into the fiber is 0.0011. The power that has to come back into the fiber is 60 nW (twice the required power for the read-out). This means that the incoming power has to be at least 54 µW and that the source should deliver 108 µW.

In the above derivation it is assumed that the light is reflected by a surface perpendicular to it. However, the spindle on which the encoder is engraved jitters on its axis during rotation, which results in a periodic change in propagation distance and direction.
Figure 10. Spindle jitter.

In the simplest configuration, the optical fiber directly projects light onto the encoder, with no collimating lenses.

The lateral displacement $\delta$ concept and deviation in backreflected angle are detailed in Figure 10.

The maximum backreflection angle (with respect to the optical axis) can be computed as a function of the spindle diameter $D$ and of the maximum known lateral jitter as $\phi = \tan^{-1}(\frac{2\delta}{D})$.

Using a realistic value of $\delta = 100 \mu m$ and a spindle diameter $D = 18 mm$ we obtain a maximum angular deviation of 0.011 radians, or 0.63 degrees.

On the other hand, the incoupling angular condition for a fiber-air interface is $\theta_{max} = \sin^{-1}(NA)$, or about 8 degrees. This implies that, despite the mechanical jitter, incoupling is guaranteed.

However, any angular misalignment of propagation translates to a reduced power transmission efficiency:

$$\eta_{ang} = 1 - \frac{\theta n_0}{\pi \cdot NA}$$

where $n_0$ is the refractive index of the medium between the two facets (in the case of vacuum, $n_0 = 1$).

We can model the spindle jitter as a circular motion of the axis, centered at the nominal spindle position; at any instant, the spindle center can be written as $(x_c, y_c) = (\delta \cos(\omega_1 t), \delta \sin(\omega_1 t))$, where $\omega_1$ is the jitter angular frequency. This frequency is comparable to the spindle rotation frequency due to rotational drag, but we will maintain its functional dependence explicit for clarity. In the nominal configuration, the spindle lateral surface can be considered as a flat vertical mirror, perpendicular to the optical axis of the read-out system. The angular error (with respect to a nominal, vertical configuration of the reflector) due to spindle jitter, as a function of the jitter dynamical behavior, is therefore

$$\theta = \tan^{-1}\left(-\frac{\sqrt{R^2 - \delta^2 \cos^2(\omega_1 t)}}{\delta \cos(\omega_1 t)}\right) - \frac{\pi}{2}$$

which leads to a coupling efficiency loss plotted in Figure 15.

Therefore, a realistic lower bound for incoupling efficiency with a maximum jitter of 100 $\mu m$ is $\eta_{ang} = 97.5$. 
Figure 11. Backreflection efficiency loss due to spindle jitter.

The total power efficiency, considering angular and displacement mismatch, is therefore

$$\eta_{\text{tot}} = \eta \cdot \eta_{\text{ang}}$$

so that the required transmitted optical power (in order to guarantee a minimum SNR at the photodetector, as discussed above) becomes

$$P_t = 4 \frac{P_r}{\eta_{\text{tot}}} = 110.7 \mu W$$

Some of the light will be reflected back into the fiber already at the interface between the fiber and the air in the cryostat. The power of this reflected light is larger by a factor of 100 than the power of the light reflected from the spindle, which makes it impossible to detect the light coming from the spindle. This problem can be solved by using an anti reflection coating at the end of the fiber. This coatings reduces the reflection of light back into the fiber, making the power of this light comparable to the power of the light coming from the encoder.

### 6.5 Thermal power budget

Half the power of the source can be considered as the power delivered to the cryostat, because the fraction of light getting back into the fiber inside the cryostat is negligible. This is less power than the power delivered to the cryostat by moving the stage and acceptable.

There is also power delivered to cryostat by heat conductivity through optical fiber $P_{hc}$. It depends on the length $l$ and the cross section $A$ of the fiber, on the heat conductivity $\sigma$ of the fiber and the temperature difference between the outside and inside of the cryostat $\Delta T$,

$$P_{hc} = \frac{\Delta T \sigma A}{l}$$

The heat conductivity of a fiber is 0.05W/(mK). The fiber going into the cryostat has a length of 1 m and a diameter of 100 µm, the temperature difference between the outside and inside of the cryostat is 300 K. The power delivered to the cryostat is then 0.15 µW, which can be easily accounted for in realistic experimental protocols incorporating the PiezoKnob.
6.6 Bill of materials

Indicative unit prices (obtained from the Thorlabs 2012 catalogue) for possible components are given in the following Table.

<table>
<thead>
<tr>
<th>Notes</th>
<th>Part no.</th>
<th>Unit price (Euro)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si Photodiode (700-800 nm)</td>
<td>Resp. 0.4 [A/W], NEP 5 \cdot 10^{-14}</td>
<td>FD6010</td>
</tr>
<tr>
<td>Pigtailed laser diode</td>
<td>Non-flange type</td>
<td>T63-029</td>
</tr>
<tr>
<td>Optical fiber</td>
<td>5m FC-APC SMF</td>
<td>T55-517</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At the small optical powers required by this concept, no cooling of the laser is required, which translates to large savings in terms of system complexity and price. Please note that antireflection coating of the fiber end, optional collimating optics and extra manufacturing involved in the readhead mount were not considered in this calculation.

7. Conclusions and outlook

<table>
<thead>
<tr>
<th>Displacement sensor</th>
<th>Basic requirements</th>
<th>(\mu \text{POT})</th>
<th>OptoGroove</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental temperature</td>
<td>1 to 300 K</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Measuring range</td>
<td>5 mm</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Resolution</td>
<td>1 (\mu)m</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>2(\mu)m</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Absolute position accuracy</td>
<td>10 (\mu)m</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Measuring frequency</td>
<td>5 Hz</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Output</td>
<td>Digital (t.b.d.)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Max volume</td>
<td>10mm diameter, 30 mm</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Length</td>
<td>The smaller the better</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cost price</td>
<td>1000 Euro per axis</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Wires (feed-throughs)</td>
<td>(\leq 6)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Maximum dissipation sensor</td>
<td>(\leq 0.5)mW @4K</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>(\leq 1\mu)W @20mK</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

7.1 \(\mu\)POT

The \(\mu\)POT device can be used to measure displacements of the stage inside the cryostat. Advantages of this method is that there is absolute distance measure, the electronic wires needed into the cryostat are a reduced number. Furthermore, this method is easy to scale up to more dimensions without many complications.

One disadvantage we forsee is the need to engineer the resistance wire with the correct material in order to minimize the temperature dependence and the probe contact. Furthermore, due to the contact probe we are not able to determine the life–time of the device. Thus, we recommend special attention to these two points. A possible solution is to experimentally test the system.

7.2 OptoGroove

Advantages of this concept are its simplicity and thermal stability. For every actuation axis, a single optical fiber and simple collimation optics need to be inside the cryostat. Moreover, all
components needed are readily available off the shelf.

One major disadvantage in the current setup is that it cannot directly measure absolute displacements. The electronic backend or the software should keep track of the encoder every time the PiezoKnob is moved.

A possible solution to this problem might be to use a more complicated code engraved on the spindle. Instead of using only the sequence with one dark spot and then a bright spot, one could think of a higher cardinality encoding (grouping of engraved spots into "bytes") from which also the absolute position could be deduced.

Another possible drawback is the minimum bending radius of optical fiber, which for this application could be around 2 cm. However, detailed mechanical design of the read-head supporting structure can reduce the total volume occupation of the sensor.

8. References


[2] Contact mechanics


[4] Cryogenic accessories–wires 1, Cryogenic accessories–wires 2

MicroDish

Can physics tell the difference between a dead and living microorganism?

Evert Bosdriesz¹, Job Bouwman², Lisa Hartgring³, Yusuke Hashimoto⁴, Tjalle Hoekstra¹, Colin Ingham⁵, Karel van Laarhoven⁶, Wouter Roos¹, Willem van de Water⁶, Jan Weenink², Gijs Wuite¹

¹ VU University, the Netherlands
² Utrecht University, the Netherlands
³ FOM Nikhef, the Netherlands
⁴ Radboud University Nijmegen, the Netherlands
⁵ MicroDish, the Netherlands
⁶ Eindhoven University of Technology, the Netherlands

1. Abstract

In this work different techniques are explored to assess the viability of (bacterial) cells on the MicroDish culture chip. The culture chip is composed of thousands of miniature wells on a porous aluminium oxide layer. The viability of microorganisms is tested by looking at cell growth in the wells using electrical and optical techniques. For the electrical side, a simple setup was investigated where the filling of a well can be detected by attaching electrodes to the top of the well and measuring a change in resistance. Also, more sophisticated electrical techniques, such as finer nano-grids and impedance measurements of cells in suspension were explored. For the optical side, an overview was made of various microscopy techniques. A simple white light interferometer can in principle measure the change of the depth of wells on the MicroDish culture chip, thereby measuring growth of biomass. An experiment was conducted with a Mirau interferometer which showed that it is a potentially feasible method for the fast, cheap, and automated detection of bacterial cell growth. More sophisticated optical techniques may still be a possibility to detect the viability of cells.

2. Company profile

MicroDish is a biotechnology start-up (founded 2008 in Utrecht, 6 FTE) that grows microorganisms on porous aluminium oxide and in micron scale wells fabricated on top of this material (see figure). MicroDish sells miniaturized culture chips, performs contract research and is involved in a number of R&D and product development collaborations. Porous aluminium oxide (PAO) forms the basis of the MicroDish technology; it is a relatively inert brittle material with a low background for many detection processes (e.g. fluorescence). PAO is biocompatible and supports microbial growth. MicroDish has microbiology laboratories cleared for genetically modified organisms and in house production.
3. Problem description

Introduction
Despite many years of progress biologists do not know what life is. Even for simple organisms, such as bacteria (and despite the promises of synthetic biology to 'build life' or at least retroengineer it) this is a surprisingly poorly defined concept. For example, you may use microbial division and growth as an assay (a viable count, usually performed on nutrient agar plates). The viable count is slow (e.g. requiring weeks for tuberculosis), is poorly automated and of limited use in many cases as we still cannot grow most microorganisms on agar. Other methods (cell integrity, enzyme activity, measurements of energy metabolism) all have their flaws. There are indications that methods derived from disciplines outside of the biological sciences have a lot to offer. For example, electrical impedance is used both academically and commercially. Stable isotope probes are used to assess even very minimal levels of metabolism. Microcalorimetry may be promising and both Raman spectroscopy and mass spectroscopy (MALDI-TOF) are entering standard usage within microbial diagnostics. Given the threat of antibiotic resistance and the importance of industrial microorganisms progress in this area can be considered urgent.

The task
Develop a method using biophysical methods to tell us if bacteria within microcolonies or presented as individual cells are alive or dead. The method should work for organisms grown or at least inoculated on porous aluminium oxide (an inert surface) as this forms the basis of the MicroDish technology. It is likely that the students select a particular aspect of this broad topic. Examples might include how you can assess the growth of a single cell or how you can tell how quickly cells within a microcolony are killed by an antibiotic. The method devised may target a particular group of microorganisms or be more general. Support will be given by the supervisors in making this choice. Optical physics, biophysics, thermodynamics, nanophysics, modelling may all be relevant disciplines to a solution depending on the path chosen.

What is the commercial context?
The solution should help address one or more of the areas below giving an answer (live vs dead) in < 1 day; ideally faster:

1. Understanding what kills microorganisms (hygiene, diagnostics, antimicrobial screening).
2. Detecting living microorganisms that contaminate food, drink or pharmaceuticals.
3. Assessing the quality of microorganisms in industrial processes and foods (e.g. probiotics, starter cultures).
4. Figuring out how to interfere or modify the growth of microorganisms, for example in screening new therapeutics (e.g. antibiotics) or optimising industrial processes (e.g. in converting chemistry based processes to microbial fermentations as part of the biobased economy).

4. Problem solving strategy
The first problem in the MicroDish case is the definition of live cells. Live cells can be distinguished from dead cells in many ways. The conventional method to test the viability of cells is to do a so-called indirect viable cell count. This method uses cheap agar plates to grow cells. Single, live cells will grow into colonies, but they are only detectable after many cell divisions.

There are other ways to distinguish dead from living cell. Living cells take up nutrients, metabolize and secrete waste, e.g. gases, metabolites. However, different microorganisms
may take up different nutrients and secrete different metabolites and gases. In our case, we want to be able to detect the growth of a huge variety of microorganisms. Therefore we have discarded the possibility of adding labelled metabolites and the detection of secreted products because this is organism-specific. Another characteristic of living cells is that they produce heat. The heat production of a single cell is very small and therefore we chose not to pursue this method.

On the other hand it is possible to detect dead cells. Dead cells will have damaged membranes and some fluorescent-dye based techniques exist that make use of this property. These dyes are able to penetrate damaged membranes and thereby make cells detectable with fluorescence. A disadvantage of this method is the cost and availability of fluorescent dyes. In many cases, the addition of a dye or other chemical is lethal, and therefore only one time point for cell growth can be acquired.

We set out to design a method that is able to detect growth as fast as possible, without the need of adding a reagent or killing the cells during detection. In order to do so, we have defined alive as the ability to grow. Cells that do not grow are dead.

The detection of cell growth is done on the MicroDish culture chip. The culture chip consists of a porous aluminum oxide substrate. On top a grid is formed of many thousands of wells, which can vary in size. Typically the walls of wells are 10 µm high and the bottom of the well has a diameter ranging from 20 – 180 µm. Figure 1 shows the PAO layer and the wells.

![Figure 1: Images of the PAO (A), the wells at different resolution (B,C,D) and bacterial colonies stained with a fluorescent dye (E-H). Source: [Ingham 2007].](image)

The chip contains distinct wells into which a small amount of cells can be inoculated. Since nutrients are applied from the bottom through the pores of the porous aluminium oxide, the wells are only filled up to the height of the microcolony (Figure 2). The wells are small in size, diameters of 20-180 µm, and even a small amount of cell divisions leads to a significant increase in volume (see Figure 3).
Figure 2: Schematic cross section of the MicroDish culture chip. Nutrients are flowing through the PAO. A thin layer of the nutrient fluid covers the cells in the wells. When the viable cells duplicate, the biomass increases and the wells fill up.

Figure 3: Growth of microbe level in well for bacteria with a size of 0.5 \( \mu m^3 \) and a starting population of 1000 cells. For a well diameter of 20 \( \mu m \), a change in the well level of \( \sim 1 \mu m \) may be measurable after 1-2 doublings of the cell population.

We identified two possible pathways for growth detection in the MicroDish chip. The first option is to make use of electrical detection of the increase of cells per well. The second option we have explored is the use of microscopy techniques to detect the increase of cells. In the following sections a description is given of both approaches.

5. **Electrical detection of cell growth**

Measuring rise of the cell level with microelectrodes

The most basic way in which electrical signal could be used to detect cell growth is depicted in figure 4. The regular microdish chip is modified by building two electrodes into the sides of every well, which can be connected to a voltage source. While the well is empty, virtually no current will flow through the circuit due to the high resistance of air. However, when the well is filled with cells and medium, a measurable current flows through the culture. Provided the electrodes are implemented high enough, the well can be inoculated with a layer of cells that does not yet connect to the electrode. Growth of the cells could then be detected as soon as the volume increase fills the well up to electrode level.
Figure 4: Schematic representation of a basic set up that can detect cell growth.

Figure 5 depicts a basic wiring diagram of this setup. We assume that several centimeters of micrometer thin gold wire are needed to access a specific well on the MicroDish chip, and that these wires have a certain resistance $R_w$.

The current $I$ through the system can be obtained by measuring the voltage $V_m$ over an additional resistor $R_m$ included in the set up. The current through this setup is simply given by:

$$I = \frac{V}{R_w + R_m + R_{well}}$$

$V$ is the applied potential difference and $R_{well}$ the resistance between the electrodes in the well. The resistance of the well can then be deduced from the measured potential difference $V_m$ using:

$$R_{well} = R_m \left( \frac{V}{V_m} - 1 \right) - R_w$$

In order to see if this setup could yield a measurable signal, we may estimate the measured voltage $V_m = R_m * I$ with the some rough values of the relevant parameters. For a hypothetical
cuboid well, with electrodes fully covering the opposing walls the resistance of the well can be a priori estimated using:

\[ R_{\text{well}} \approx \frac{\rho d}{A} \]

where \( \rho \) is the resistivity of the medium between the electrodes (being either air or cell culture), \( d \) the distance between the electrodes and \( A \) the area of the part of the electrodes, which is in connection with the water.

In reality, this simple model does not hold: the electrodes form only a small part of the opposing walls, the shape of the wells deviate from a cuboid (see Figure 1c) and the top surface of the cell culture is not necessarily flat. The combined effect of these more realistic conditions might be investigated by either numerical simulations or by experimental measurements. However, for now we will use the aforementioned equation, assuming that we can take a value for \( A \) that is somewhere in between the electrode surface and a coronal cross section of the specific well (\( A_{\text{average}} \)).

Other assumed values are listed in table 1. When the well is filled, this yields a signal of \( V_m \approx 3 \text{ mV} \), while virtually no signal can be detected when the cell is empty (\( V_m \approx 10^{-13} \mu \text{V} \)). That means this setup should be able to detect cell growth filling up the well.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Rough value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied voltage</td>
<td>( V )</td>
<td>1 V</td>
</tr>
<tr>
<td>Measuring resistance</td>
<td>( R_m )</td>
<td>1 k( \Omega )</td>
</tr>
<tr>
<td>Resistance of a few cm of gold wire</td>
<td>( R_{\text{gold}} )</td>
<td>3 k( \Omega )</td>
</tr>
<tr>
<td>Distance between electrodes</td>
<td>( d )</td>
<td>100 ( \mu \text{m} )</td>
</tr>
<tr>
<td>Area of electrodes</td>
<td>( A )</td>
<td>1 ( \mu \text{m}^2 )</td>
</tr>
<tr>
<td>Cross section of well</td>
<td>( A_{\text{well}} )</td>
<td>600 ( \mu \text{m}^2 )</td>
</tr>
<tr>
<td>Average Area</td>
<td>( A_{\text{average}} )</td>
<td>300 ( \mu \text{m}^2 )</td>
</tr>
<tr>
<td>Resistivity of air</td>
<td>( \rho_{\text{air}} )</td>
<td>( 10^{15} \Omega \text{m} )</td>
</tr>
<tr>
<td>Resistivity of cell culture (Sun, 2010)</td>
<td>( \rho_{\text{cel}} )</td>
<td>3 ( \Omega \text{cm} )</td>
</tr>
</tbody>
</table>

In principle such a simple set-up with two opposing electrodes gives a discrete outcome: the well is either full or not. A slightly more continuous outcome may be expected if the meniscus of the cell culture is concave.

A grid of wires needs to be embedded in the MicroDish chip in through which the electrodes of the different cells can be addressed. A possible layout of such a grid is depicted in figure 6.
Figure 6: A schematic top view of an array of wells with added electrodes and wiring. Wires A, B and C are separated from wires 1, 2, 3 by an insulating layer. The status of a specific cell can be obtained by measuring over the wires that connect to its two electrodes, e.g. the central cell can be read out by measuring over B2.

Where at the moment the MicroDish culture chip is constructed from only a layer of aluminum oxide and a grated polymer layer, adding the wires and electrodes will make the construction more complex. The wires plus electrodes need to be embedded in the structure in two layers, with insulation between them. Photolithographic and soft-lithographic techniques can be used to construct the kind of layered wire structure that would be needed for a setup like this (Black, 2000). For commercially available photolithographic production, the lower limit in structure length scales would be in the order of 250 nm (Xia, 1998). This means the construction of μm thick wiring and electrodes should be feasible, as long as wells on the chip are far enough apart to fit them in.

Because the proposed measurements are performed at DC or at very low frequency AC voltages, the induction effects in the grid will be negligible. Therefore, other than the resistance of the specific wires leading to the target well, the wiring will have very little effect on the signal in this case.

Possibilities for more sophisticated setups

Though the setup described in the previous paragraph seems easy and viable, the various electric properties of cells make more sophisticated options of detection through impedance possible. Firstly, the cell has detectable electric properties of its own. The cell membrane acts as an insulating shell with a conductivity of around $10^{-7}$ S/m and a capacitance of about $1\mu F/cm^2$, while the cell interior with a high conductivity of 1 S/m. Secondly, cells change the electric permittivity on the medium around them by generating small ions through metabolism and ion exchange.

Various impedance measurement methods and devices - exploiting one or more of these cell properties - have been developed in the academic environment, and even several commercial biosensors and analysis devices exist. These techniques could be interesting for MicroDish. We try to summarize the possibilities and their boundary conditions, especially with respect to the additional conditions imposed by the environment of a well in the MicroDish array.
In general one can distinguish three different approaches that are used for detection and quantification of cell growth or activity (for a nice overview see (Wang, Ye, Ying, 2012)). The first approach consists of measuring impedance changes of the medium. Both the changes induced by the cell metabolism and those caused by the increasing volume fraction of cell material need to be taken into account here. The second exploits the insulating properties of the cell membrane. When cells attach on the electrode surface they effectively reduce the electrode surface thereby increasing the impedance. A third method consists of first inducing lysis in cells and then measuring the subsequent change in impedance due to the release of the ionic cytoplasm. The last approach really doesn’t seem viable for MicroDish, which is looking for a non-destructive method.

A possible advantage of measuring changes in medium impedance could be that the method allows for real time monitoring of the cell activity. The reason is that both the cell growth and the ion production are continuous processes that induce gradual changes. This would for instance be valuable in investigating cell responses to changing conditions.

An important limitation in setups that have been developed so far is that a simple (DC) resistance measurement is not able to measure a change in medium composition. It has been empirically determined that the decrease in resistance of the fluid due to ion production is effectively countered by the increase in resistivity due to an increase in cell-volume fraction. The capacity of the medium, however, does change, as ion production increases the capacity of the ion double layer, and the presence of cells induces a capacity in the medium. The medium composition change can therefore be detected by measuring the frequency-dependent impedance. For MicroDish we call this a limitation because measurement of the frequency-dependent impedance will also pick up the inductive and capacitive impedance from the wiring between the wells on the chip. A much more complicated circuitry than the basic principle proposed in the previous paragraph would be necessary. Therefore, if frequency-spectra are to be measured, a specific design tuned to the MicroDish product will be necessary.

It should be noted that especially in the MicroDish setup, a DC solution could be possible. If the connection of the well to the culture-medium reservoir, through the PAO, would be sufficient to drain the well of the chemical waste produced by cell activity, the resistance of the medium would only be modified by the increase in cell volume. This would allow for a DC resistance measurement that would be more compatible with the chip wiring.

On the other hand, another difference between the classic medium-impedance cell and a setup in the MicroDish well would be that in the MicroDish case cells would not really be in suspension if they only lay on the bottom of the well with a water film over them. The volume fraction of cells would be very high, so trying to measure medium properties might not be sensible in the first place. If this is the case, then the second general strategy – measuring cell adherence – might be more interesting. A digitated array of electrodes on the bottom of the well might be used to detect the increased coverage of the well floor by the wet cells. If water would indeed largely be localized as a film at the cell surface, conductance across the setup would actually increase as cells grow between the electrodes. In a classic setup with a cell suspension, on the other hand, cell-electrode coverage would actually increase the overall impedance. Again, the situation in the MicroDish well might significantly alter the boundary conditions of the method. Effectively, this approach would become a more sensitive variation on the case described in the previous paragraph, detecting the gradual coverage of the well floor through the gradual increase in conductance. In the end, this might be the best option for real-time monitoring cells in a well. However, the method would only be sensible when wells are inoculated with a low number of cells that don’t cover the floor initially, and would only allow for growth monitoring up to the point where the floor is entirely covered by cells.
An example of a fluid cell impedance setup is described in (Yang, 2008). Here salmonella cell are suspended in deionized water (DI) or phosphate buffered saline (PBS) and the impedance response was studied over a large range of frequencies. It was found that bacterial cell suspensions in DI water with different cell concentrations result in different electrical impedance spectral responses. Other examples can be found in (Yang, Bashir, 2008).

Figure 7: A schematic diagram of the setup for electrical impedance measurements of bacterial cells suspensions in DI water or PBS. The chamber capacity is ~25 microliters. Taken from (Yang, 2008).

Figure 7 shows the experimental setup. The wells are composed of a ground glass substrate and mounted on silicone rubber walls. The finger electrodes are made of gold and are 15 by 15 micrometers while the well contains 25 micrometer of suspension. The response of the system can be electronically modeled by two capacitors in series separated by a resistor.

Figure 8: Impedance spectra of Salmonella suspensions in (A) DI water and (B) PBS. Taken from (Yang, 2008).

The measurements on the impedance of this system show that the low frequency domain is dominated by the double layer capacitance while the high frequency region is dominated by the resistance of the solution between the two electrodes (figure 8). The slope of the impedance curves depends on the concentration of the cells. In the case of PBS the concentrations shows no significant different at any frequency.

Another option is to use the insulating behavior of the cell membrane. This setup will use the growth of cells to reduce the surface of the electrodes. Again the growth of cells can be monitored in impedance measurements. A specific setup has been carried out (Wu, Ben, Chang, 2005). In this case diffusion of cells towards the electrodes is enhanced by the application of a vortex electric field.

**Literature**

6. Optical techniques to detect cell growth

Introduction

The most obvious way to observe cell growth is to look at the cells themselves. For a standard Petri dish with bacterial cells grown on agar, it is for example possible to see colonies of $10^8$ to $10^9$ cells with the bare eye. Of course, microscopy techniques provide a much better resolution, which makes it much easier to see cell growth.

Microscopy techniques have been used extensively in biological sciences. In fact, they form the basis of observations of biological processes. Over the years many different techniques have been developed. Depending on the situation (i.e. your sample, purpose, cost) one technique can be more suitable than the other.

In our case, the goal is to detect cell growth in the culture chip developed by MicroDish. The culture chip contains thousands of miniature Petri dishes on a porous aluminum oxide (PAO) substrate. The (bacterial) cells are trapped in these so-called wells (the miniature Petri dishes) and nutrients are constantly delivered to the cells via the PAO. One of the unique features of the culture cell is that there are many thousands of wells, which has many applications such as high-throughput screening.

Considering the observation of cell growth in the culture chip via optical methods, one can make a list of requirements. First of all, the technique is preferably non-invasive and non-destructive. Current techniques often use fluorescent dyes that may be toxic, destructive or non-specific. Secondly, the method should be fast, meaning it should not take too long to scan the micro-array. Taking for example a culture chip with several thousands of wells, it should not take more than a second to scan and analyze the signal of a single well. Moreover, in order to exploit the full potential of the culture chip, automation should be possible. Also, the cost should be suitable for the specific purpose and preferably the method is simple.

The desired resolution of the optical technique depends on the purpose of the experiment. If one is interested in seeing the properties of single cells, for example seeing changes in the shape of cells which could indicate cell duplication, or observing small populations of cells, the resolution should be at the sub-micron level.

On the other hand, if one is interested in changes of large populations of cells, only a micron resolution may be desirable. One could envision the following situation in the wells of a culture chip (see figure 2 and 3). A relatively large population is present in the wells, for example, when the cells together with the surrounding medium with nutrients partially fill the 10 micron-deep well. Then, if the cells are viable, the level in the well increases. An
optical technique can be used to measure the depth of the cells in the well with micron-accuracy. Even with a slow duplication rate, such a depth measurement may be a very simple and fast way to measure growth.

In this chapter we consider many different microscopy techniques that may be useful to detect cell growth in the MicroDish culture chip. We categorize the techniques by its systematics and purpose and we present a list of pros and cons for each technique. Finally, we critically assess each technique and present a list of optical techniques that may be explored further.

**Transmission**

In 1676 Antonie van Leeuwenhoek used a microscope to discover the bacteria. Some homemade lenses were enough to see single bacteria. He made use of transmission microscopy, in which the light source is on the opposite side of the sample with respect to the objective/detection.

In the MicroDish case, visualization of a single microorganism dividing would be the ultimate proof whether it is alive or not. In this section, we discuss multiple transmission microscopy methods.

**Direct visualization**

The culture chips have a thin layer (~60 µm) of PAO on which the cells grow. This layer has been shown to be translucent for visible light (Zhuo, Peng, Lin, Qu, & Lai, 2011). However, literature exists on the scattering of visible light by PAO (Woodman, 1972). The scattering makes direct visualization of cells or structures in the wells impossible. This conclusion is supported by experiments performed earlier by MicroDish, where it was not possible to directly visualize cells in the wells of the culture chip. For this reason, direct visualization of the cells in the wells is not pursued.

**Absorbance**

Cell cultures are often measured by determining the optical density at 600 nm (OD600). With the absorbance at this wavelength the concentration of cells can be calculated. Although the exact absorbance will depend on size of the cells - varying for different organisms - growth of cells can be measured by an increase in absorbance over time. This would be a very simple method to measure single cell growing. We were unable to determine whether enough sensitivity can be reached by the absorbance of a few cells. Moreover, the difficulty with this technique is the requirement of transmission of light through the PAO layer. The known issues with visible light passing through this layer and unknowns about sensitivity led to discarding absorbance of light of 600 nm as a possible method for measuring growth.

**IR**

Infrared light (IR) (3,000-15,000 nm/ 3500 - 700 cm-1) can be used to detect specific chemical bonds. Cells consist for a large part of proteins, DNA and lipids. These molecules contain bonds that absorb infrared light of the above wavelengths. The specificity of absorption is fairly high and can be used to determine the content of the sample. Also, different organisms have different compositions of these bonds. These differences can even be used to identify bacteria - or strains - by IR absorption (Lipkus, Chittur, Vesper, Robinson, & Pierce, 1990).

Infrared absorption is also used in the food industry to detect bacteria in food products (Ellis et al., 2002). However, a complication in these experiments is the presence of water. The absorbance spectrum for water over these wavenumbers is strong and close to that of the material a cell is composed of. The interference of water with the IR absorption was avoided in above studies by drying the cells before measuring them. In this case, we want to sample through the layer of nutrients. Therefore, the sensitivity of IR absorbance to detect growth of cells in the culture chip is questioned.
Besides the questions about sensitivity and the ability to pick up signal in the layout of the culture chip there is another aspect that makes IR absorption difficult. In all described cases, the measurement length was in the order of tens of seconds. This is not ideal for scanning thousands of wells in the culture chip. For these reasons, we discard IR absorption as a method for detecting living cells.

Because of the issues of above-mentioned techniques with the layout of the culture chip, we do not regard a transmission-based method as a possible solution for the observation of cell growth.

**Epi – microscopy**

In general, epi-microscopy techniques are those where the sample is lit and detected from the top. Simple light reflection off biological sample is typically poor. For example, cells are almost transparent and reflection of light from the substrate can overpower the signal from the cells. Therefore one often has to resort to more sophisticated techniques. Here we present some of these techniques.

**Interferometry**

Interferometry is based on the principle that the phase of light is changed when it propagates through, or is reflected by, a certain medium. In general, some reference light beam is compared to a beam that passes through a sample. When the phase of the sample beam is changed, an interference pattern appears. In this way it is possible to see structure (e.g. cells) that are not visible by conventional bright field microscopy.

The first microscope that used interferometry was the phase contrast microscope developed by Frits Zernike in the Netherlands. The discovery of phase contrast microscopy was awarded the Nobel Prize in Physics in 1953. A more modern and sophisticated version of the phase contrast microscope is the differential interference contrast (DIC) microscope.

**Differential Interference Contrast microscope**

The DIC microscope is very useful for seeing individual cells. The technique is relatively simple and readily available. Moreover, it is fast, as one can take real time pictures. The DIC is not very useful when the refractive index of the specimen does not differ much from the surrounding medium. In the case of bacterial cells on PAO, this appears to be the case. The cells are surrounded by nutrients in water, which reduces the contrast of the cells with respect to the fluid. Therefore it may be difficult to study single cells at the bottom of a well. Even so, we think it will be worthwhile to explore the visibility of single cells in the culture chip using DIC. Significant cell growth may still be seen, and it could be worthwhile to see how well the DIC microscope can image cells in the culture chip. On the other hand, there are most likely much simpler, faster and cheaper techniques to detect a significant growth of cells. Therefore we focused other these other techniques.

![Figure 9: A DIC image of single E.coli cells (Image from Athale lab, India).](image-url)
Optical Coherence Tomography (OCT)

Another technique that utilizes the phase properties of light is Optical Coherence Tomography. It uses a broadband source as input, an example being white light. It is somewhat analogous to ultrasound techniques that directly measure time delays for sound reflected by some tissue. With OCT the time delay is measured indirectly by looking at the optical path length of photons reflected by some tissue. It is possible to use OCT to look a millimeters into a sample with a resolution of a few micron. It is an established technique that is widely used in the medical sciences to look deep into tissue, for example the eye.

![An OCT image of the retina (Olson et al., Eye 2006).](image)

In the case of the culture chip, the OCT can in principle be used to get a 3D profile of the wells with bacterial cells inside. With 20-30 frames/s scanning of a complete culture cell can be performed in a short time. When the resolution of the OCT is a few micrometers, it may be possible to detect significant cell growth in the wells that are 10 micrometer deep. However, there are indications that achieving the micrometer resolution may be quite challenging. Therefore it is questionable whether or not this method may be used. Moreover, when it comes to just measuring the level of cells in a well, there are simpler methods to measure this.

Mirau Interferometer

An even simpler method to measure the distance to a certain surface is the Mirau interferometer (figure 11). It does not measure a depth profile like the OCT does, but it is a cheap and simple way to measure a certain profile. In fact, this technique was available to us and an experiment has been conducted as a proof of principle for depth measurements. In the following chapter the experiment is described.

![Mirau objective. Interference occurs when the sample is at the exact same distance as the reference mirror.](image)
Other

Two other epi-illumination methods were discussed: Raman spectroscopy and dark-field imaging.

Raman spectroscopy

In Raman spectroscopy molecules are excited by monochromatic light. The photons can excite the vibrational mode of the molecule to a higher virtual state. Relaxation of the molecule leads to the emission of a photon, which can be either of less or more energy than the initial excitation photon. Like infrared spectroscopy, detailed information on specific bonds can be obtained. For our purpose, Raman has an advantage over IR spectroscopy, because water does not interfere with the detection of cells and their contents. Many studies have used Raman scattering to identify bacteria (Notingher, 2007). It has even been used to distinguish between live and dead human lung cells (Notingher et al., 2003). In that study, near-infrared light (780 nm) was used to observe human cells over the time-course of hours.

The main disadvantage is that Raman scattering is in general very weak. Therefore, the measurement of an accurate spectrum requires fairly long observation times (1-30 seconds).

Hypothetically, Raman spectroscopy allows for observation of an increase of cells over time. It could be possible that shorter observation times will be sufficient to measure this change. For now, we do not reject Raman spectroscopy as a method for the detection of cell growth. It would, however, require sophisticated experiments and testing. Even if a change over time can be observed, the measurement of time might not allow for sampling thousands of wells within a reasonable time.

Dark Field Microscopy

Dark Field Microscopy is a technique that detects only the light scattered by a specimen. Light is often incident at an angle on the sample, and the microscope lens is directly above the sample. Therefore, reflected light is not detected in dark field microscopy. Objects appear therefore bright on a dark background.

For the case of bacterial cells on PAO, Dark Field Microscopy may greatly enhance the contrast of the cells with respect to the substrate. However, there are reasons to question this technique. Firstly, the sample must be illuminated quite intensely, as the image will be constructed by scattered light, which is much fainter than directly reflected light. This strong light source may harm the cells in the sample. Moreover, the substrate of PAO is not very smooth. We expect that also a lot of light will be scattered by the PAO substrate and the medium with nutrients. It remains therefore to be seen whether or not Dark Field Microscopy provides enough contrast to see the bacterial cells, and therefore also growth. Thus we recommend tests with a Dark Field microscope to test whether or not cells can be distinguished from the PAO and the nutrients medium.

Concluding, in general there are many optical techniques to possibly detect cell growth in the MicroDish culture chip. We have analyzed a number of these techniques and discussed the pros and cons. Table 2 summarizes the discussion. Considering simplicity
Table 2: Overview of considered optical techniques for detection of cell growth in MicroDish culture cell. TBT = To Be Tested.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Visibility</th>
<th>Speed</th>
<th>Cost</th>
<th>Result</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct visualization</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>X</td>
<td>Low transmittance through PAO</td>
</tr>
<tr>
<td>Fluorescence</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>X</td>
<td>Invasive, need dyes</td>
</tr>
<tr>
<td>IR Spectroscopy</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>Likely not enough contrast, too slow</td>
</tr>
<tr>
<td>Phase Contrast / DIC</td>
<td>++(?)</td>
<td>+</td>
<td>-</td>
<td>TBT</td>
<td>Interpretation difficult due to nutrients</td>
</tr>
<tr>
<td>Optic. Coherence Tomography</td>
<td>+/-</td>
<td>++</td>
<td>+</td>
<td>+/-</td>
<td>Challenging to see pm change in depth</td>
</tr>
<tr>
<td>Mirau interfer.</td>
<td>+/-</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>Cheap, simple, proof of principle for height measurement</td>
</tr>
<tr>
<td>Raman spectroce</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>TBT</td>
<td>Slow, but visibility may be good</td>
</tr>
<tr>
<td>DarkField Microscope</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>TBT</td>
<td>Possible scattering due to PAO</td>
</tr>
</tbody>
</table>

References


7. Pilot Study: Using Mirau microscopy to measure the well depth

We performed a pilot study to test the feasibility of using Mirau interference microscopy to measure the height distance between the top of the MicroDish culture chip and the bacterial cells.

Experimental Setup

We placed an agar containing Petri dish with a MicroDish culture chip inoculated with different densities of Escherichia Coli under a 20 times magnification interference objective. To avoid vibrations in the sample, some weights were placed on the edge of Petri dish. The
environment in which the experiment was performed was kept at a temperature of 20°C, which is well below the 37°C that is optimal for E. coli growth. Therefore, a small lamp was placed near the microscope as heat source. However, we were not able to record to what extent this heat source was able to raise the sample-temperature.

Reference measurements were taken by focusing the microscope on the top surface of the culture chip, where interference lines were clearly visible. Depth measurements were taken by manually moving the sample towards the microscope until the bottom of the wells came in focus. Again, interference lines were clearly visible. The depth was measured by recording how much sample had to be lifted to change the focus from the top to the bottom of the well.

**Results and interpretation**

As a null measurement, an area of the culture chip were no well were inoculated was measured. When focused on the top surface, interference fringes were clearly visible. Also, the wells were clearly visible as “holes” where fringes were absent (Figure 12). As the top surface was moved out of focus by “lifting” the sample, all interference fringes disappear. As the sample was further lifted, the bottom of the wells came in focus, and fridges were visible were the “holes” were previously. A well height of approximately 6 or 7 micron was measured, which is somewhat below the expected 10 micron. However, it should be noted that accurate measurements were hampered by “drift” in the focus.

![Figure 12: Images of the MicroDish culture chip with a Mirau objective. First the top of the well is in focus causing interference to be around the well. When the focus is shifted (~10 micron) to the bottom of the well, interference is seen in the well.](image)

To get an estimate of the resolution of the measurement, the bottom of the well was brought into focus. By either moving the focal plane 1 micron up or down, all nearly all interference fringes disappeared, indicating that micron-resolution is definitely feasible.

In order to measure cell growth, the microscope was put over a part of the culture chip which was thought to be inoculated. A null measurement indicated a cell height of approximately 5 or 6 micron. After approximately 2 hours, no significant difference in height was observed. This can either mean that no cell growth had taken place due to non-optimal growth conditions. Or that the water/cell surface is not sufficiently reflective, and one is actually measuring the distance to the PAO well bottom and not to the cells.

Therefore, a positive control experiment was conducted by inoculating a chip at a very high density. It was expected that the wells were completely filled with cells. The resulting image (figure 13) clearly differs from that of an empty well. There is no abrupt end in interference fringes at the edge of the wells anymore. Rather, concentric circle towards the center of the well were observed. It appeared as though there was a smooth transition in height of the
medium/cell surface from the sides of the wells to a depth of roughly 7 micron at the center of the well.

![Figure 13: In a full well, no clear change between the well and the surroundings could be detected.](image)

**Outlook/Discussion**

There are a number of reasons why Mirau interferometry is a potentially feasible method to determine cell growth in MicroDish wells. At a price of approximately $2000 for a Mirau objective, the method is relatively cheap. There appears to be a clear difference between an empty and a full well. The interference fringes are probably easily detectable by imaging software, which facilitates automation. In principle, a fairly large number of wells can be observed simultaneously, making quick scanning feasible. A problem might be that one needs to search for the depth at which the bottom comes in focus. However, this problem can be solved by taking “snapshots” at 1 or 2 micron depth intervals. If the well depth is lower, the number of snapshots that are needed to scan the entire well is reduced. An alternative, even simpler strategy that can be considered is to keep the focus on the top surface and wait until the “hole” in interference pattern disappears because the cells grow out of the well. A potential problem with this strategy is that the chip surface needs to be in a very flat plane. Since the chip is typically placed on an agar plate, it does not seem feasible to keep the entire chip surface within a 1 micron plane.

**8. Conclusions and outlook**

The most general indication of bacterial viability is cell division. Therefore we have decided to focus on cell growth. In the environment of the MicroDish culture cell even a small increase in biomass fills a relatively large fraction of your measuring volume. We exploit this with two different techniques.

The electrical technique is based on detecting current through a cell culture that becomes possible as soon as the increased volume of cells connects two measuring electrodes. A crude setup consisting of two short electrodes above the well floor is estimated to already yield a usable signal. A finer electrode setup would allow for continuous monitoring of growth but would be more complex to fabricate and it would limit the range of volume increase that could be measured.

The other method is based on measuring the height of bacterial cells in the wells of the MicroDish culture chip using optical techniques. One of the simplest ways to do this is by a
white light interferometer. A proof of principle experiment has been conducted and a height measurement seems feasible. The technique is fast and in principle allows for automation. More sophisticated optical techniques such as Raman spectroscopy may still be feasible.

9. Acknowledgements

We would like to thank Dr. Martin van Exter and Dr. Wolfgang Löffler for the use of and help with the Mirau interferometer.
NXP Semiconductors

**Electrical sensing and actuating of LED wavelength**

Michele Buscema¹, Toon Coenen², Matthijs Cox³, Lourens van Dijk⁵, Femius Koenderink², Lucas Kunneman¹, Abbas Mohtashami², Radu Surdeanu⁴, Barbara Witek¹, Haifeng Yuan⁶

¹ Delft University of Technology, the Netherlands
² FOM Institute AMOLF, the Netherlands
³ Eindhoven University of Technology, the Netherlands
⁴ NXP Semiconductors, Belgium
⁵ Utrecht University, the Netherlands
⁶ Leiden University, the Netherlands

1. **Abstract**

The light output of light emitting diodes (LEDs) in terms of flux and wavelength varies because of the fabrication process, which is undesirable for most applications. Currently, the LEDs are binned into different wavelength categories prior to being sold. Firstly, this is an expensive and logistically complicated procedure. Secondly, the peak wavelength of the LEDs is influenced by temperature, operating conditions, and aging making binning alone insufficient. It would be useful to have an automated CMOS-integrated process which identifies the optical properties in terms of flux and wavelength of the LEDs. In that case the LED driver can adjust the driving conditions to shift wavelengths to desired values, or even give active feedback on the LED to maintain the desired performance.

Here we discuss how to implement such a wavelength and flux sensing device on the electronic driver chip of the LED without any prior knowledge about its exact optical properties. In particular, we present two possible routes that might be promising for implementation. In the first method the signal is detected optically, which can be precise enough but it requires part of the LED light to fall onto the sensor. In the second method, all the sensing is performed electrically, which is appealing because it always works regardless of the sensor’s environment, i.e. its relative position with respect to the LED.

In addition to providing two working solutions, we quantitatively show that the small package available for a device in a CMOS chip precludes many optical filtering solutions.

2. **Company profile**

NXP Semiconductors is a company of 27,000 employees which is headquartered in Eindhoven, and has R&D and manufacturing facilities in 25 countries. NXP’s solutions drive innovations in the fields of RF, analog, power management, interface, security and digital processing electronics. Key application areas of these products are automotive, secure identification, satellite TV, computing, mobile, lighting, industrial electronics, healthcare and
wireless infrastructure. It is NXP’s aim to develop electronic circuits that improve the energy efficiency, connectivity and security of future products.

3. Problem description

Solid State Lighting (SSL), i.e., lighting using (Organic) Light Emitting Diodes, (O)LEDs, has become in the past years a fast growing research and development topic, as well as a large emerging market. By offering many benefits, such as reduced energy consumption, better quality of light, reliability and long life, SSL has entered a wide range of application domains, from general illumination, displays to automotive. The workhorse material system for SSL is the III-nitride system, in which, e.g. for blue and green LEDs, epitaxially grown GaN on sapphire is a base substrate, where LEDs are defined through multiple quantum wells. Color is tuned from blue to green by introducing Indium. However, LED systems have several fundamental challenges related to the way the LEDs are manufactured. For the same epitaxial growth conditions, the produced LEDs have a large distribution of central output wavelengths, light flux outputs and forward voltages. These variations have several causes:

- Due to the lattice mismatch between the layers of the double heterostructure (AlInGaN for Green and Blue or AlInGaP for Red) and the substrate (at this moment sapphire) there is intrinsic stress in the structure. Hence large amounts (~10^{10} - 10^{12} cm^{-2}) of defects are present (stochastic defects, such as: stacking faults, threading dislocations), and interface states are present at the interfaces between the contacts and the LED structure and the GaN buffer layer and the Multiple Quantum Wells (MQW).
- Material segregation: highly efficient light emitting quantum dots of various dimensions are present in the MQW due to local variations in Indium concentration.
- Via the piezoelectric effect, in GaN structures, unavoidable strain in the epitaxial layers leads to high internal electric fields, leading to significant Stark shifts in the wavelength of the emitted light. The Stark effect hence causes a spread in wavelengths. As the LED is driven, the wavelength will vary with driving current, as the Stark effect is partially compensated by the driving field.
- Process variations across the wafer.

All these fundamental issues will stay for years to come and are affecting several aspects of LED systems:

- The logistics of produced LEDs is cumbersome. The LED manufacturers bin the LEDs into various wavelength, flux output or forward voltage bins, in order to address their customers’ needs. This binning process is expensive and time consuming, which reflects in the end price of the LED products. Even after binning, LEDs in the same bin still have significant spreads in light output and wavelength, which is detrimental to the quality of the final products.
- Driving electronics: for optimal control over output and management of aging effects, the driving electronics needs to know the spread in forward voltages of the LEDs in the system.
- Color point control: wavelength value is necessary information, i.e. initial or continuous calibration of the LED wavelength is needed to maintain a unique and optimum color point. Whether from datasheet or read out at luminaire assembly time, this information should be written in the control electronics. This is a problem both for white light LED systems assembled from red, green, and blue diodes, and for white light sources based on blue LEDs pumping yellow phosphors.
NXP as IC manufacturer is looking for solutions to overcome the binning problem at the IC level, by sensing electrically the wavelength, and changing/actuating the wavelength to the required value for the envisaged application.

To study this technology further within NXP, a novel method has to be developed which enables electrical sensing and actuating of the LED wavelength. The problem formulation is two-fold:

i. Which electrical measurements can be performed to sense (first step) and actuate (second step) the wavelength of the LEDs, independent on the color or manufacturer?

ii. Can a method be developed that is easily (negotiable what easy is…) integratable in an IC?

4. Background information

4.1 Introduction

Although the goal is to measure and actuate the wavelengths emitted by LEDs, NXP cannot change the technology inside the LEDs, and instead seeks a technology that is ‘LED-agnostic’, i.e. applicable to any LED as packaged with its secondary optics, remote phosphors, ceramic packaging, etc. Therefore, the solution must be implemented on the driver, which is a small integrated circuit that supplies the current-driven LED with power and provides feedback on its performance. New features require a smart driver, which is able to measure voltages, do timing, calculations, provide variable current output, and has memory storage. It is the complementary metal-oxide-semiconductor (CMOS) technology that permits all these functionalities in a small package of about 1 mm². An ideal LED driver would keep track of the temperature, the flux, and center wavelength of the emitted light. Already, the accurate temperature measurement and tracking is implemented (as described in US8,278,831 patent), in purely electrical fashion. The next step is for the driver to measure the wavelength, either electrically or optically, in a LED-agnostic fashion, i.e. applicable to any LED.

4.2 Demands/limitations

There are several demands and limitations that have to be taken into account for any sensing geometry. The human eye can detect a 1.25 nm spectral shift and a 2.5% variation in flux for two LEDs in close proximity. The LED driver chip therefore should be sensitive enough to detect flux and wavelength variations in this order.

There are limitations on how much space the wavelength/flux sensor can occupy on the driver chip. Laterally there is at most 250 x 250 µm² area available. Also the height of the device is limited to a maximum of 5 µm.

Furthermore all the materials and fabrication steps that are required for the sensor have to be compatible with CMOS processing. This means that we can use several types of oxides, nitrides, silicon, and some metals like aluminium and copper. Using organic materials, gasses, liquids, and other metals is not preferable.

4.3 Use cases

We discern three different ‘use cases’, which correspond to solving a different family of problems. The first bifurcation is between an absolute determination of the peak wavelength (λₚ), or to obscure the difference in λₚ between neighboring LEDs. We see two approaches for measuring λₚ. The first is to use optical sensors to determine the wavelength emitted by the
LED. The second option is to extend the Pulse Width Modulation (PWM) driving to include frequency dependent capacitances, and to convert such information to $\lambda_p$. With these different approaches in mind, we discern the following three use cases:

1. “Once upon a time in the factory”, which entails optically measuring $\lambda_p$ once under controlled conditions, in order to calibrate the driver. This can be done by incorporating sensors on the driver, so minimal external hardware is necessary for calibration. We envision that the driver chip would then first be held head on in front of the LED, so sufficient light is available for calibration. The $\lambda_p$ value is stored directly in the driver, replacing the need for external measuring apparatuses, such as spectrometers and computers interfacing with the LED driver for wavelength storage. After storing the calibration data, the geometry of driver and LED is unrestricted.

2. “Wavelength fuzzing”, shifting around $\lambda_p$ by modulation of the driving current. Doing this for multiple LEDs would result in uniformity without precise control over $\lambda_p$. No sensor is needed.

3. “Measuring all the way”, i.e. $\lambda_p$ determination throughout the service life of the LED. This can be either electrical or optical.

![Figure 1. Two proposed ways of sensing the wavelength: (a) Electrically, and (b) Optically.](image)

The following paragraphs will go into detail about the concepts and measurements of the electrical and optical determination of $\lambda_p$, and on the effect of the current on the LED wavelength and flux. On the basis of these analyses we will describe the feasibility of the technical implementation of the three use cases. In section 5 we discuss optical sensing, which could be applied to use case 1 in particular. In section 7 we discuss electrical sensing, which could be applied to use cases 1 and 3. In section 9 we discuss the feasibility of wavelength fuzzing (use case 2).

### 5. Possible device geometries using optical sensing

#### 5.1 Optical sensing with narrowband color filtering geometries

##### 5.1.1 Introduction

A sensor for optical measurements contains an element that converts photons into either current or voltage. To determine the peak wavelength, it is necessary to obtain information over different color selective channels. To distinguish color, light of different wavelength can
be spatially separated before it reaches the photodetection pixels, or the pixels themselves must be sensitive to different wavelengths. The concepts that we discuss for spatially separating the light include gratings and photonic crystals, whereas the possible filters mounted on photodiodes include Bragg mirrors, Fabry-Pérot interferometers (MEMS) and Mie scatters. Finally, a detailed discussion is made on differential wavelength sensing using multiple photodiodes with different response curves. A main issue in several concepts is that a spread in incident angle prohibits a decomposition of the spectral components of the incident light. This is a severe problem, as CMOS integration entails thin, planar structures, which will receive light from a large solid angle. Only in the factory (use case 1) controlled lighting situations can be realized, and as a result, sensing geometries that are insensitive to the incident angle are the most promising solution.

5.1.2 Bragg reflectors and Fabry-Pérot resonators

We now turn to the method of using multiple detectors, each covered with a wavelength selective filter. Bragg stacks consist of a periodic stack of alternating layers with different indices of refraction. If the thickness of the layers is tuned properly, it is possible through interference to achieve strong transmission of light for a specific wavelength, so it can act as an effective bandpass color filter. The bandwidth of transmission depends on the index contrast between the layers and the number of layers. On the driver chip, the number of layers is limited by the height budget. Because these structures rely on interference, they are not only very sensitive to wavelength but also to incidence angle. To study the angular dependence of the transmission, we performed calculations using the open source program ‘Open Filters’, which is an implementation of the transfer matrix method, taking Fresnel reflection coefficients as input. We calculate the transmission for a layer stack consisting of 9 silicon nitride layers of 63 nm and 8 silicon dioxide layers (7 with a thickness of 84 nm and one central layer of 168 nm), resulting in a total thickness of 1.3 µm. In Fig. 2 we show the calculated transmission for 0, 5, 15, and 30° incidence angle. The FWHM of the peaks is ~4 nm so the transmission curve is sufficiently sharp to resolve small changes in LED wavelength. Multilayer optics is hence very well suited to the limited height available. However, the transmission peak position very strongly depends on angle (here the transmission peak shifts ~30nm for a 30° change in angle). Because it is not possible to control the incidence angle of the LED light onto the filter, this type of filter is not suitable for precise wavelength detection.

![Figure 2. Transmission of a 1.3 µm Bragg stack as function of wavelength for different incidence angles](image-url)
We concluded that any geometry using a Fabry-Pérot effect has problems that are similar to the Bragg stack as it also relies on interference, and has the same functional dependence of peak wavelength on angle.

5.1.3 Nanoantennas as narrow bandpass filters

Metallic and high index dielectric/semiconductor nanoantennas can strongly scatter light and have a wavelength dependent response because of Mie type resonances that are supported in the antenna. If placed in a randomized array, they are less sensitive to variations in the angle of incidence. For metallic antennas aluminum is preferable as a material because of its CMOS compatibility and relatively low absorption losses on the blue side of the spectrum. Metallic nanoparticles and holes have been well studied, and it is known that for efficient light scattering antennas the quality factors of the resonances are low (Q~5). This response is too broad to make a decent bandpass filter. One can increase the quality factor by arranging the nanoantennas in periodic arrays at the expense of a strong increase in the angular sensitivity which is also undesirable (see sections 5.1.2 and 5.1.4).

In semiconductor nanoparticles like silicon particles for instance, the quality factor of the resonances can be higher because the amount of absorption in the particle is lower. Possibly the resonances are sharp enough to be a good bandpass filter. The peak position of the resonances can be tuned by changing the diameter. Figure 3a shows resonant spectra for silicon spheres in vacuum calculated using a Mie calculation method. A more detailed map of the resonant response as function of wavelength is shown in Figure 3b. In this case, absorption within the nanoparticle is not taken into account so in reality the peaks will be broader. The presence of a substrate will redshift the resonances but the quality factor remains roughly equal if the index of the substrate is low enough (silica would be fine). Although the peak wavelength is nicely tunable by changing the diameter, it is very sensitive to small changes in diameter. In particular, a 1 nm change in diameter leads to a ~4 nm change in peak position. This puts unrealistically stringent demands on the fabrication tolerances so we do not recommend this geometry as a narrowband filter.

![Figure 3. Scattering cross section for silicon nanoparticles with different diameters.](image)

5.1.4 Gratings

A grating utilizes periodicity on the interface between two media with different refractive index, to redirect light of different colors in different directions. In principle, this can conveniently be implemented on CMOS, as the surface corrugation is easily fabricated above a row of photodiodes. The physical principle is conservation of the \( k \)-vector component parallel to the surface of the interface:
\( k_{\text{out}} = k_{\text{in}} + G = k_{\text{in}} + m \frac{2\pi}{d} \) (1)

where \( G \) specifies the periodicity of the grating, and \( k_\theta = (\omega/c) n \sin(\theta) \). Solving this equation for \( \theta_{\text{out}} \) yields:

\[
\theta_{\text{out}} = \sin^{-1} \left( \frac{m\lambda + n_1 d \sin(\theta_{\text{in}})}{n_2 d} \right)
\]

(2)

where \( n_1 \) and \( n_2 \) are the refractive indices of the two media, \( m \) is the order of the diffracted mode, \( d \) is the grating pitch and \( \lambda \) the wavelength of the light in vacuum. Not all combinations of parameters yield a solution.

The CMOS implementation of the grating plus an array of photodiodes on which the spatially separated spectrum is detected allows for a depth of about 5 \( \mu \)m between the grating on top of the wafer and the different pixels below. Figure 4 shows the geometry with a ray that is incident normal to the grating containing light of 400 nm (blue arrow) and 460 nm (red arrow), which are spatially separated by 1.2 \( \mu \)m on the photodiodes below.

**Figure 4.** For normal incidence, a separation of 1.2 \( \mu \)m is realized between the 400 nm (blue arrow) and 450 nm (red arrow) rays (1st order). Grating pitch is 400 nm, and the refractive index \( n_2 \) is 1.5.

In CMOS technology the minimum pitch between photodiodes in the photodiode array is well below 1 \( \mu \)m, so resolving a distance of 1.2 \( \mu \)m is feasible. However, note that in this example the wavelength difference is still 50 nm. If we go to a wavelength difference of 1 nm at 450 nm central wavelength the lateral displacement is only 29 nm which is very difficult to resolve.

The main complicating factors are the angular spread and spot size of the incident light. Figure 5 shows how both complications restrict the use of a grating for detectors implemented on CMOS. Typically, in order to maintain high enough flux to do a measurement, the spot size accepted by the detector should preferably be of the order of 5 \( \mu \)m or above, i.e., at least of the same order as the device height. When a grating deflects the incoming light by about 50°, the spot size as projected on the detector array will prevent resolving the colors. In general, this can be overcome by increasing the height of the device to beyond ~10x the spot size, but in a small CMOS package, this is not possible. Figure 5 shows that gratings are likewise detrimentally sensitive to angular spread. An angular spread of only 10 degrees prohibits precise determination of the wavelength. Without imaging optics between the LED and driver, the light cannot be collimated. To conclude, the grating approach suffers from three serious problems (poor color separation, large smearing by spot size, angular sensitivity) prohibiting application in the driver.
5.1.5 Photonic crystals

Similar to a 1D grating, photonic crystals also have the ability to diffract light into different angles depending on the wavelength. It has been proposed\(^1\) that photonic crystal cavities could act as narrow bandpass filters because very high quality factors can be attained. Spatially separated cavities over distinct diode areas could result in wavelength determination. For this application we are mainly interested in blue light which means that the dimensions of the cavity have to be much smaller than for infrared light (the spectral region where photonic crystal cavities are usually used). As an estimate, the typical pitch of a photonic crystal in a best-case \( n = 3.5 \) material would be approximately \(<\lambda/3\), which amounts to a pitch of \(<150\) nm. A fabrication accuracy of better than 1% in feature size is required for reasonable \( Q \) photonic crystal cavities. This is probably beyond the resolution that is attainable with CMOS processing. Furthermore the material in which the cavity is patterned has to be transparent at 450 nm, yet CMOS compatible. Silicon nitride and titanium dioxide are examples of such a material but with indices \(<2.5\), greatly reducing the quality factor compared to the best case scenario (Si scaled down from telecom range). Gallium nitride would be an alternative material with higher index, but it would only logically be incorporated in the LED substrate, not in a CMOS driver.

5.2 Differential wavelength and flux sensing with \( n \) photodiodes

5.2.1 Introduction

Rather than using narrowband spectral filtering, one can also envision using two photodetectors with different broadband responses. Similar to the functioning of the human eye, the wavelength and flux can be determined by using just two or more photodetectors with different broad spectral responses, if sufficient signal to noise ratio is available. The spectral response can be altered by changing the electrical pn-junction characteristics, changing the applied voltage on the photodiode, or applying a color filter. The advantage over narrowband filtering is that broadband responses can be obtained in an almost angle-independent manner. The color filtering can be achieved by one of the methods described above, or by taking advantage of the fact that the absorption length in silicon strongly depends on wavelength. In that case one can engineer the spectral response of the photodiode by changing the junction depth in the wafer. This method also allows stacking of junctions. Note that the key property used is that the path length of the light to reach a certain depth depends on wavelength.

**Figure 5.** As Figure 4, but with two extra rays. One ray (dotted), displaced ~6 μm, visualizes the effect of finite beam diameter, while the other ray (dashed) shows how a different angle of incidence mixes the colors on the sensor again.
To understand all the important factors in this detection scheme we have pursued two routes. We have employed a simple quasi-analytical code to assess the performance of schemes using just two diodes. In this route, we have assessed firstly a scheme based on junction responsivity curves alone, and secondly a scheme based on Gaussian color filters. For these schemes we have examined performance as measured against noise figures. As a second route we have developed a single analytical model that explains how measurements on \( n \) detectors should generally be treated to obtain the best spectral reconstruction. In this second route we have not examined noise figures, but do find recommendations on how to define detector functions.

5.2.2 Route 1: Two-detector schemes to determine wavelength
Diode type response functions.

The wavelength and flux of the LED can be determined by using two or more photodiodes (PD) with a different spectral response. The spectral response of the photodiode can be altered in the following ways:

1. changing the electrical pn-junction characteristics,
2. changing the applied voltage on the photodiode or
3. applying a color filter.

1. **pn-junction characteristics.** An example of a CMOS photodiode is shown in Figure 6. By changing the junction depth the absorption changes for each wavelength in a different manner. In this way the shape of the spectral response of each PD can be tuned. The maximum of the photo response will shift towards longer wavelengths when the depth of the junction increases.

2. **Applied voltage.** When the voltage over the junction is altered the collection efficiency of charge carriers will change. The higher the current \( I \) becomes the more charge carriers will be collected. A high spectral resolution can be obtained when this change is wavelength dependent.

![Figure 6. Geometrical representation of a photodiode.](image)

3. **Color filtering.** Color filtering before the light reaches the diodes can be achieved by one of the methods described in section 5.1.3. As opposed to sections 5.1.2-5.1.5, here we propose that very broadband, and hence angle independent filters, might suffice.
To understand the important factors in this detection scheme we developed a simple analytical model to determine the wavelength $\lambda$ and flux $\phi$ of the LED with sufficient accuracy. For this we start with the spectral response $S(\lambda)$ of a number of photodiodes.

Let us assume that we designed photodiodes with known spectral response $S(\lambda)$. For the rest of this report we assume $S(\lambda)$ to be known and we neglect any errors in this function. For the industrial application it is important that $S(\lambda)$ is well reproducible, and we assume no significant difference between the PDs on different drivers.

5.2.3 First rough estimation of the wavelength accuracy

The measured spectrum of an actual LED, as measured by a simple commercial USB spectrometer can be seen in Figure 7. It contains a narrow peak around 460 nm due to the blue LED, and a broad phosphor band around the blue and red. We focus on the LED spectrum, and propose that it can be approximated by a Gaussian distribution like:

$$\phi_{\lambda_{\text{peak}}} (\lambda) \sim \text{Exp} \left[ - \left( \frac{\lambda - \lambda_{\text{peak}}}{\sigma} \right)^2 \right].$$

(3)

where the spectral width of a single LED is set by a full-width half-maximum (FWHM or $\sigma$) of roughly 30 nm. The peak wavelength is itself distributed, with typical values for the distributions of blue LEDs taken as $\lambda_{\text{peak,average}} \sim 460$ nm and with a spread $\sigma \lambda_{\text{peak}} \sim 15$ nm. This represents the case in which no binning of the LEDs is done. An example of the response of a PD can be seen in Figure 8. The FWHM of the PD is several hundred nanometer. The LED thus has a narrow spectral band width compared to that of the photodiode (PD). Therefore we can for a rough estimate assume that the response of the PD is linear in LED peak wavelength between a wavelength of 440 and 480 nm.

![Figure 7. A measured spectrum of a blue LED with phosphors on top.](image)
Figure 8. A sketched example of a photodiode response curve.

5.2.4 Wavelength determination with two diodes

The following linear electrical signals are generated by the two photodiodes:

\[
s_1(\lambda_{peak}) = a_1 \cdot \phi(\lambda_{peak}) + b_1,
\]

\[
s_2(\lambda_{peak}) = a_2 \cdot \phi(\lambda_{peak}) + b_2
\]

(4) (5)

There are two PD slope combinations we distinguish:

1. The slope of both photodiodes are both positive (Figure 9a)
2. The slope of both photodiodes are opposite: negative and positive (Figure 9b)

Figure 9. Two possible sets of photodiode response curves: (a) Diodes with similar slope and (b) diodes with opposite slopes. This configuration is difficult to get in the blue without a color filter.

The second case is expected to result in the highest accuracy. However it might be difficult to make a PD that has a negative slope around 460 nm. Short wavelengths (~300 nm) are absorbed much faster than long wavelengths (~700 nm). The spectral response can thus be tuned by adapting the absorption depth which is correlated to the junction depth. So a negative slope @460nm implies a very shallow junction depth (less than 10 um). Such shallow junction depths would result in so called outdiffusion of the dopants and thus the PD doesn’t function as intended.
By combining the output signal of both PDs the wavelength $\lambda_{\text{peak}}$ can be determined. To obtain a high accuracy it is important that this output function has a steep slope. Two output signals are considered for this paper. One can consider the ratio between both spectral responses:

$$\text{Ratio}(\lambda_{\text{peak}}) = R(\lambda_{\text{peak}}) = \frac{s_1(\lambda_{\text{peak}})}{s_2(\lambda_{\text{peak}})}.$$  \hspace{1cm} (6)

The advantage of using the ratio is that the result is flux independent. The wavelength $\lambda_{\text{peak}}$ can now be derived from the known inverse function of $R(\lambda)$. This is expected to be the most useful approach when the slopes are opposite. The ratio function is less suitable for two PD having both a positive slope because the slope of this output signal is relatively shallow. As an alternative one can use the difference-to-sum signal $F(\lambda)$ defined as:

$$F(\lambda_{\text{peak}}) = \frac{s_1(\lambda_{\text{peak}}) - s_2(\lambda_{\text{peak}})}{s_1(\lambda_{\text{peak}}) + s_2(\lambda_{\text{peak}})}.$$  \hspace{1cm} (7)

For the error propagation $\sigma V$ of an output signal $V$ we use the following formula:

$$\sigma V(s_1(\lambda), s_2(\lambda)) = \sqrt{\left(\frac{\partial V(\lambda)}{\partial s_1(\lambda)}\right)^2 (\sigma s_1)^2 + \left(\frac{\partial V(\lambda)}{\partial s_2(\lambda)}\right)^2 (\sigma s_2)^2}.$$  \hspace{1cm} (8)

In the explicit scenarios we walk through below, this is evaluated for $\lambda_{\text{peak}}$ varying around 460 nm to get a figure of merit for the obtainable accuracy. It is obvious that a large $\frac{\partial V}{\partial \lambda}$ is desirable for a small $\sigma \lambda$.

### 5.2.5 Detailed calculation

Here we consider a more detailed calculation for sensing by two diodes. We envision using junctions with different response curves that are similar, but shifted in wavelength. In other words, one deals with response functions that have a slope of identical sign. We assume that we prepare two photodiode junctions of different depths, and model detector response functions as

$$D_n(\lambda) = \text{erfc}\left(\frac{\lambda - \lambda_n}{w_\lambda}\right)$$  \hspace{1cm} (9)

(where erfc stands for the error function defined by integrating a Gauss from minus infinity to $x$). These detector functions (example plots in Figure 10) approximate datasheet response functions of photodiodes (e.g. Figure 8), by having a steep turn on in the blue controlled by a steepness parameter $w_\lambda$, a peak at around $\lambda_n$, and a gradual decrease towards the red. We consider two detectors with equal steepness in the turn on, yielding a rise from 0 to full response in about 50 nm. As optimization parameters, we optimize the shift of one diode to the blue of the wavelength of interest (460 nm), and the shift of the other diode response function to the red of the target wavelength. Figure 10a shows the response functions for a relatively advantageous case, where one diode has its peak response at 410 nm, and the other at around 485 nm.

We assume a Gaussian LED spectrum $\phi_{\text{peak}}(\lambda)$ of about 30 nm FWHM, and evaluate the variation of the detected signal $s_n$ defined as the overlap integral of spectrum and detector function

$$s_n(\lambda_{\text{peak}}) = \int \phi_{\text{peak}}(\lambda) D_n(\lambda) \, d\lambda$$  \hspace{1cm} (10)

with varying central wavelength $\lambda_{\text{peak}}$ of the LED around 460 nm. As output parameter from which we can extract the wavelength, we use the difference-to-sum-ratio $F$ (see Equation 7). This dimensionless ratio is immune to flux variations, and is always between 0 and 1 for the
given family of detector functions. This difference-to-sum ratio shows a clear rollover around the central wavelength of 460 nm. At the target wavelength $\lambda_{\text{peak}} = 460$ nm, the relative change in $F$ when the input wavelength of the LED is varied by 1 nm is around 2%. We have calculated how noise in the detectors propagates through into $F$. Suppose that both detectors have a signal-independent noise level of $d s$ compared to the full range signal. Error propagation through Equation 7, yields noise $d F$ in $F$, that varies linearly with $d s$, with a wavelength dependent noise amplification factor. Figure 10c shows the relative change $\Delta F/F$ for a 1 nm wavelength change, as well as the relative noise levels $d F/F$ for different assumed detector noise levels. For detector noise levels $d s < 1.5\%$, the relative change $\Delta F/F$ for a 1 nm wavelength shift is above the noise, implying that measurement of wavelength to 1 nm accuracy is possible. For detector noise levels $d s = 2\%$ or higher, the noise in $d F$ prevents measuring wavelength with 1 nm resolution. Higher resolution at equal noise levels requires steeper response curves. We did not optimize steepness further, as we believe that higher steepness is not manufacturable.

Figure 10. (a) Modeled response curves $D_{1,2}(\lambda)$ of two photodiodes with different peak response wavelengths. (b) We calculated diode response signals $s_1$ and $s_2$ by overlap integrals of the response $D_{1,2}(\lambda)$ with Gaussian LED spectra centered at different wavelengths $\lambda_{\text{peak}}$. We plot the differential response of two photodiodes ($F = (s_1 - s_2)/(s_1 + s_2)$) as a function of the LED wavelength $\lambda_{\text{peak}}$, assuming a LED bandwidth of 30nm. (c) Relative change in the differential response function $\Delta F/F$ for a wavelength change of 1nm in terms of the LED peak wavelength $\lambda_{\text{peak}}$.
(blue curve). Cyan, green and red curves represent the accuracy of the differential response of the two photodiodes at given noise values of each photodiode at 0.1%, 1% and 2% respectively. (d) Optimization of the accuracy by shifting the response peaks of photodiodes away from the central wavelength of 460 nm.

In Figure 10d we show the result of a parameter sweep. The reported quantity is \( \frac{\Delta F}{F} \) normalized to the noise \( dF \) (at \( ds=1\% \)) as a function of the wavelength shifts of either diode away from the central wavelength. The reported quantity is hence a ‘visibility’ of wavelength shifts in LED peak wavelength \( \lambda_{\text{peak}} \) given a 1% noise level in both detector signals \( s_1 \) and \( s_2 \). The optimization clearly favors the ‘blue detector’ to be shifted well away from 460 nm by 50 nm or more to the blue. A clear problem is that in actual devices the feasible blueshift is limited, while red shifts are easily obtained by having a deeper junction. Finally we note that as the signal is really generated by the difference between \( s_1 \) and \( s_2 \), this scheme only operates in a narrow 100 nm range. Resolving a full spectrum would require more (redshifted) diodes.

5.2.6 Gaussian type response functions.

In a second scheme we consider the use of filtered detectors, with Gaussian detector functions \( D_{1,2}(\lambda) \). We assume that broad Gaussian filters could be made. On basis of the Mie calculations on the Si nanoscatterers reported in section 5.1.3, and based on data on extraordinary transmission through plasmonic bull’s-eyes, we assume that bandwidths of 70 or 100 nm could be achievable in a manner insensitive to fabrication inaccuracies, using polydisperse random assemblies of scatterers, or holes in metal films.

We consider two detectors of equal bandwidth, and once again optimize their position relative to the target detection wavelength of 460 nm. As LED spectrum, we assume once again a Gaussian spectrum of about 30 nm FWHM, and evaluate the variation of the detected signal \( s_n \) defined as the overlap integral of spectrum and detector function, with varying central wavelength of the LED around 460 nm. The optimization clearly favors two detectors that are symmetrically shifted away from the target wavelength, by approximately their bandwidth. In this case the parameter \( F \) has a zero crossing around the target wavelength, and varies from -1 to 1. At the target wavelength, the absolute change \( \Delta F \) around the zero crossing is around 0.05 (about 3% of the full range that \( F \) sweeps) per nanometer variation in LED wavelength. The derivative in \( F \) shows a clear optimum around 460, evidencing that the two-detector spectrometer is a sensitive wavelength probe only over a 40 nm bandwidth. Further extension of the bandwidth would require more detectors to be included.

We have again calculated how noise \( ds \) in the detectors propagates through \( F \). We assume a noise \( dF \) that is independent of the signal level. Error propagation through Equation 8, yields noise \( dF \) in \( F \), that varies linearly with \( ds \). Figure 11c shows the absolute change \( \Delta F \) for a 1 nm wavelength change, as well as the absolute noise levels \( dF \) for different assumed detector noise levels \( ds \). Evidently the Gaussian filter solution is significantly more robust to noise than the shifted-diode methodology, and detector noise levels of below 5% would be acceptable. The anticipated fluctuation in size due to lithography disorder would not deteriorate, but rather aid the detector, since a broad response is desired. Artificially introduced disorder in size would be recommended.
Figure 11. (a) Response curves $D_{1,2}(\lambda)$ of two photodiodes modeled as two gaussian functions. (b) Differential response of two photodiodes ($F = (s_1-s_2)/(s_1+s_2)$) as a function of the peak LED wavelength $\lambda_{\text{peak}}$ with 30 nm bandwidth. (c) Absolute change in the differential response function ($\Delta F$) for a wavelength change of 1 nm in terms of the LED peak wavelength (blue curve). Cyan, green and red curves represent the accuracy of the differential response of the two photodiodes at given noise values of each photodiode at 0.1%, 1% and 2% respectively. (d) Optimization of the accuracy by shifting the response peaks of photodiodes away from the central wavelength.

5.2.7 Second route: Decomposing signal using the linear algebra of Gramm-Schmidt decomposition

It is evident that two detectors, whether chosen as two junctions with shifted response, or as two filtered diodes, only provide a unique and sensitive measurement of wavelength in at maximum a 100 nm bandwidth, i.e., in the band between the two detector slopes/peaks. If a broader bandwidth is to be sensed, multiple detectors are required. Suppose you have $n$ detectors, each with a different detector function $D_n(\lambda)$. Suppose the LED supplies an input spectrum $\phi_{\lambda_{\text{peak}}}(\lambda)$. The $n$ measurement values one has to derive wavelength from are:

$$s_n = \int \phi_{\lambda_{\text{peak}}}(\lambda) D_n(\lambda) \, d\lambda = \left\langle \phi_{\lambda_{\text{peak}}} \right| D_n \right\rangle$$

(11)
Where \(< | >\) denotes inner product. A significant problem for reconstructing the spectrum, is that the \(n\) detector response functions are not orthogonal functions on the wavelength axes. Suppose that we define a new set of basis functions via Gramm-Schmidt orthogonalization:

\[
\tilde{S}_1 = \frac{D_1}{\langle D_1 | D_1 \rangle} \\
\tilde{S}_2 = D_2 - \langle D_2 | \tilde{S}_1 \rangle \tilde{S}_1 \text{ and normalize} \\
\tilde{S}_3 = D_3 - \langle D_3 | \tilde{S}_1 \rangle \tilde{S}_1 - \langle D_3 | \tilde{S}_2 \rangle \tilde{S}_2 \text{ and normalize}
\]

Note the following: (1) The overlap integrals determine the degree to which detectors give linearly independent results, and (2) the order in which you normalize could ultimately be relevant to minimize noise. We hence anticipate that there must be a cleverer way than the Gramm-Schmidt algorithm to derive linearly independent detector functions. Note that you can write the Gramm-Schmidt result as a linear transformation matrix \(C\) as

\[
\begin{bmatrix}
\tilde{S}_1 \\
\tilde{S}_2 \\
\tilde{S}_3 \\
\quad \cdot \\
\end{bmatrix} = C \begin{bmatrix} D_1 \\
D_2 \\
D_3 \\
\quad \cdot \\
\end{bmatrix}
\]

with \(C = \begin{bmatrix}
1/\langle D_1 | D_1 \rangle & 0 & 0 & \cdots \\
-\langle D_2 | \tilde{S}_1 \rangle/\langle D_2 - \langle D_2 | \tilde{S}_1 \rangle \tilde{S}_1 \rangle & 1/\langle D_2 - \langle D_2 | \tilde{S}_1 \rangle \tilde{S}_1 \rangle & 0 & \cdots \\
\quad \cdots & \cdots & \cdots & \cdots
\end{bmatrix}\) (13)

The problems of detector response function choice are embedded in this matrix. For instance, the matrix will be singular, if you have two almost equal detectors, which is the worst choice for a linear decomposition case. Note furthermore that you can program the matrix \(C\) into the chip if you know the response functions. Now we can use the machinery of linear algebra to say that the best approximation to the spectrum, given the orthonormalized detector functions at our disposal, is:

\[
\phi_{\lambda\text{ peak}}(\lambda) = \sum_{n=1} S_n \tilde{S}_n \text{ with } \tilde{S}_n = \left( \phi_{\lambda\text{ peak}}(\lambda) \right) \tilde{S}_n
\]

Expressed in the measured values \(s_n\), you have (normalization not precise here):

\[
\tilde{s}_1 = \frac{s_1}{\langle S_1 | S_1 \rangle}, \quad \tilde{s}_2 = s_2 - \frac{\langle S_2 | \tilde{s}_1 \rangle \tilde{s}_1}{\langle S_2 - \langle S_2 | \tilde{s}_1 \rangle \tilde{s}_1 \rangle} \text{ etc.}
\]

In short:

\[
\tilde{s} = Cs
\]

Note how this transformation of measurement values has a very clear interpretation. Suppose you have two detectors, each with a narrowband almost delta-peak non-overlapping response. The independent pieces of linearly information are simply \(s_1\) and \(s_2\). Suppose you have two detectors each with a wideband response, say, shifted by a small amount. In this case the two independent pieces of information are essentially the information in \(s_1\) (total flux as measured by the broadband detector) and the narrowband information in \(s_1 - s_2\), i.e., in the minute difference between \(s_1\) and \(s_2\). In that sense the case two broadband shifted responses is formally equivalent to taking one broadband diode to simply measure flux, and one narrowband diode to measure wavelength, given that you know the total flux.. Noise considerations might favor one realization (narrowband plus broadband) over the other (two shifted broad responses).

You now have the best reconstruction of the spectrum through the coefficients \(\tilde{s}_n\). You now need to derive the quantities you want, which would depend on the usage scenario. If we think of the definition of peak wavelength as a mean
\[ \bar{\lambda} = \frac{\int d\lambda \phi_{peak}(\lambda)}{\int d\lambda \phi'_{peak}(\lambda)} \quad (17) \]

then
\[ \bar{\lambda} = \frac{\sum s_n \int d\lambda S_n(\lambda)}{\sum s_n \int d\lambda S_n(\lambda)} \quad (18) \]

The evaluated integrals can be programmed as numerical coefficients in the chip and do not involve any measurement. The only operation the driver is required to perform is taking the ratio of two weighted sums over the measurement values. Generally, to simultaneously measure more characteristics of an LED system, it would make more sense to not use Eq. (17, 18), but to adapt the retrieval algorithm to the number of wavelengths you wish to sense.

6. Recommendations for the optical approaches

- Narrowband filtering either via spatially dispersive or stacked filter systems is not recommended. Angular tolerances and the height budget are insufficient, even in the simple head-on scenario. We also do not recommend nanoantenna geometries because the spectral sensitivity is not high enough.

- Differential absorption using just a few diodes with distinct spectral responses will provide sufficient wavelength resolution, at noise levels of a few percent in detection.

Points of attention:
- to convert the proposed junction responses to implantation parameters.
- Two diodes optimize for one wavelength range. We propose that 3 or 4 diodes could span the entire visible range.
- Enhanced flux accuracy. For the determination of the flux the slope of the PD negatively affects the accuracy of the flux determination:

\[ \Delta \varphi \propto a_1 * \Delta \lambda. \]

It might be beneficial to add a third diode with the slope \( a_1 \) as low as possible and a large \( b \). This diode than is only used for the flux determination.
- Steep slopes. For the determination of \( \Delta \lambda \) it was found that a large slope leads the the highest \( \lambda \) sensitivity.
- More than two PDs. If the use of 2 PD does not result in sufficient accuracy additional PD can be added. However this can increase the accuracy only several times.
- Angular selectivity. By placing a open cylinder on top of the PD angular selection of incident light at the expense of flux can be obtained, to improve angle-independence of the final measurement. The diameter can be tuned such that a proper numerical aperture (or solid angle) is obtained. This limits the incoming flux however, so a one should balance the several constrains.

There are several recommendations to obtain a good signal to noise ratio, s/n in this geometry:
- Area. The signal to noise ratio scales roughly with the area. A large area is beneficial for the final accuracy.
- #of diodes. The signal of several equal PD can be used of which the signal is averaged.
7. Electrical Sensing

7.1 Introduction

In the introduction, the main part of the problem was explained: the binning of the LEDs in different wavelengths bins is a tedious, labor-intensive and costly job. Therefore, a possible solution would be to design a smart driver that can provide the appropriate driving conditions for any LED. Of course, the solution should be CMOS-fabrication compatible and provide information without deteriorating the optical properties of the LED, especially the flux. In this respect, electrical techniques could offer a viable solution as they rely on components that are readily integratable in CMOS technology and do not interfere with the regular functioning of the LED.

After an elaborate literature survey, we decided to pick two techniques that could deliver valuable information on the band structure of a Multiple Quantum Well (MQW) LED: Deep Level Transient Spectroscopy (DLTS) and Admittance Spectroscopy (AS). These techniques can deliver crucial information on trap states in a device. Trap states are in the bandgap of the device, generally caused by defects. The traps basically capture electrons and holes, which then recombine non-radiatively (leading to lower flux) or radiatively (leading to a broadening in the linewidth of the emission spectrum of the LED).

In the next two sections we will first describe the DLTS and AS technique. In section 7.4 we discuss the result of our preliminary AS experiments on different LEDs.

7.2 Deep Level Transient Spectroscopy

7.2.1. Idea of measurement scheme

To perform deep level transient spectroscopy (DLTS) measurements, a voltage pulse is sent to the device and then the electrical response of the device (such as the capacitance) is recorded over time. The difference between the capacitance at a short time after the pulse and the capacitance at a long time after the pulse is measured. The time difference between the two measurement instants is called measurement window (see Figure 12, top part). The signal sign delivers information over the trap concentration while the sign discriminates between minority carriers (positive) and majority carrier traps (negative). The decay rate of the LED capacitance is equal to the occupation/decay rate of the traps. As these trap states are thermally activated, a temperature dependent study must be performed to gather the exact energy position of the levels, with respect to the either of the band edges.
Both the pulse width, pulse voltage and the measurement window can be adjusted to deliver more information on the energy position of the traps. The pulse width (i.e. the voltage pulse duration in time) has an influence on the quantity of trap states that are populated: a short pulse might not populate all the trap states that are allowed by the magnitude of its voltage. The pulse voltage influences the energy level of the traps that can be probed/filled. The higher the pulse voltage, the higher the energy of the trap states that are being probed. A variation in the measurement window ($t_1 - t_2$, see Figure 12 (top part)) will modify the amplitude of the signal, but not its sign.

If the temperature is swept, while keeping all the other parameters constant, then the signal will present a peak at the temperature at which a certain transition is activated. For the LEDs, this can be done by initially giving a well-adjusted high current pulse to the LED and the temperature can be measured very accurately using the sensorless sensing method described in US8,278,831 patent. Therefore, performing a temperature sweep will deliver information about the position in energy of the trap states, at least relative to the edge of one of the bands. This information can be possibly retrieved by performing measurements while the LED is cooling down after having been heated up with a long driving pulse in the PWM driving scheme applied to drive LEDs. The sensitivity of this type of measurements should be in order of ~10 meV. The sensitivity in the energy position is most likely determined by the accuracy in the temperature measurements. Note that at higher temperature, the thermal occupation is already broadened by $K_BT$, where $K_B$ is the Boltzmann constant and $T$ is the operating temperature ($K_BT \sim 23$ meV at room temperature).

7.2.2 Implementation possibilities
Measuring the transient response to a pulsed excitation requires a large bandwidth of the circuitry around the device being probed. If the device (in the present study a multiple quantum well – MQW - LED) is already enclosed in its final packaging, the extra wiring
might lead to a reduction in the bandwidth and sensitivity of the measurements. Hence, there might be a systematic error induced by transient (or in general high frequency methods) on a packaged device. On the other hand, pulsed measurements should be easily and readily integrated in the LED driver. This stems from the fact that LEDs are driven with a Pulsed Width Modulation (PWM) circuitry and, therefore, trains of pulses are already present. Therefore, including an impedance measurement unit on the driver board should allow for ready integration of this technique.

7.2.3 Fundamental question
To gain the direct conversion between the energy value and the emission wavelength, the technique should allow for measuring the whole bandgap, not only the relative distance of the traps from either the conduction band or the valence band. From the literature, it is not clear if this requirement can be met with DLTS. Further information beyond standard DLTS may be accessible by performing DLTS in a different fashion: instead of sweeping the measurement window (see above), it is possible to measure more curves at the same temperature with different sampling rates. Then, performing a Laplace transformation, it is possible to get the same information. This technique is called Laplace DLTS or LDLTS; performing a literature survey, we found that LDLTS is considered to be more accurate than simple DLTS. Since the deep traps are mostly non-radiative, the literature suggests that by using DLTS one can measure the non-radiative density of states.

7.3 Admittance spectroscopy
A second technique, related to DLTS is called admittance spectroscopy (AS) – see Figure 12 (bottom part). This technique can deliver both information on the deep traps (like DLTS) and on the shallow dopants that determine the full band gap of the MQW, as reported in Ref.5. To perform admittance spectroscopy, an oscillating driving signal is fed into the LED, with or without a dc offset, and the response of the system is measured. The impedance (or admittance) can then be determined. Impedance can be converted into capacitance and conductance. Conductance goes through a peak and the capacitance experience a sharp decrease when the frequency hits a resonance with the trap level. Therefore, it is possible to estimate the rate of occupation/deoccupation of the trap states from the peak frequency. In this case, the MQW needs to be modeled as a capacitor in parallel with a resistor (for the unbiased, un-depleted part of the MQW) and just a capacitor for the biased/saturated part of the MQW. With this schematic in mind, it is possible to calculate the relevant parameters namely the global capacitance $C(\omega)$ and the conductance $G(\omega)$ as a function of frequency $\omega$. The same thermal method as with DLTS can be applied to obtain precise information on the energy position of the trap states. In principle, the frequency response can be converted to the corresponding time response and vice versa with Fourier transformations.

Both DLTS and AS can, in principle, provide information about the geometrical location of the trap states in the junction. This is determined by performing measurements at different voltage bias, and then by running a deep analysis of the curves with a model on the junction geometry. This determination is highly demanding and it requires making certain assumptions regarding the band structure of the MQW. Therefore, the result of the analysis will always be dependent on the assumptions made and is, therefore, not general. For practical purposes in LED device control, such a full understanding may not be required.

By combining the two methods, DLTS and AS, with the temperature sweep given by well-controlled pulses from the driving circuitry and measuring accurately the temperature of the LED junction with NXP’s sensorless sensing concept, we can extract the emission
wavelength of the LED, given by extracting the accurate position of the radiative traps in the band gap.

Figure 13. Admittance Spectroscopy results from commercially available LEDs. (a) Parallel capacitance vs excitation frequency for green (green dots) and blue (blue squares) LEDs. (b) Conductance over angular frequency vs driving frequency with the same color/shape coding as in panel (a). Inset: zoom-in of the low frequency response for green LEDs.

7.4 Experiments

We performed admittance spectroscopy measurements using a HP4191A impedance analyzer on two types of LED, green and blue LED from LUMILEDS (blue: LXX2-PB12-K00, green: LXX2-PM12-R00). The results are shown in Figure 13. Figure 13(a) shows the capacitance response against the driving frequency. Starting from 1MHz, the green diode capacitance (green dots) is declining towards a more stable value around 20MHZ before experiencing a drastic drop at around 100MHz. The blue diode capacitance (blue squares) is constant at low and high frequencies and shows again a drop at frequencies around 100 MHz. Figure 13(b) shows the conductance versus the driving angular frequency ($\omega = 2\pi f$) for both the blue (blue squares) and the green diodes (green dots). The inset shows a zoom in of the low frequency (up to 10MHz) response of the green diode. By comparing the low frequency behavior of the green diode both in Figure 13(a) and 13(b, inset) with literature, we might assign this type of behavior to the presence of trap states in the MQW region of the LED. This difference is qualitatively reproduced over a series of measurements conducted over 3 blue LEDs and 1 green LED. Moreover, the difference in magnitude could be used to distinguish a blue LED from a green LED. This measurement can be easily integrated in the pulse width modulation: the conductance difference at a certain driving frequency could be used to distinguish the emission wavelength of the LED. To quantitatively correlate the difference in conductance to the emission wavelength, a more complete study should be performed.

As the frequency rises, the behavior of the tested LED devices changes profoundly: there is a sudden drop in the capacitance and in the conductance. This is associated with a transition from purely capacitive behavior (low frequency) to a purely inductive behavior (high frequency); this is a well-known phenomenon for any kind of capacitor-like electronic component. The position of the transition in frequency is influenced by, for instance, the cabling/package of the device and the electrostatic environment in which the measurements are performed. Therefore, we do not ascribe this behavior to any relevant physical properties of the LED.
Figure 14. Temperature activated admittance spectroscopy of commercially available LEDs. Conductance over angular frequency vs driving frequency for green (a) and blue (b) LED. The green (a) and blue (b) dots correspond to room temperature measurements, while the red dots correspond to higher temperature. The inset in (a) is a zoom in of the low frequency response of the green diode.

It is known that the frequency peaks due to traps states change their position, in frequency, as a function of the temperature of the system. Therefore, to confirm that the peak in conductance in the inset of Figure 13(b) is related to trap states we perform another set of measurements at higher temperature the result of which are shown in Figure 14. In panel (a) we show the conductance frequency response of the green LED at room temperature (green dots) and at higher temperature (red dots); the inset shows a zoom in of the low frequency region. It can be seen that increasing the temperature has a noticeable effect only in the low frequency part of the spectrum. The peak shift to higher frequency as the temperature increases. This shift towards higher frequencies after heating is also in agreement with what has been reported in literature and can be understood as follows: the higher the temperature, the faster the population of the trap states can change, hence the higher frequency for the peak. This strengthens our hypothesis that this conductance peak around 2~3 MHz should be related to trap states in the green LED.

There are also visible changes in the high frequency region for both LEDs. The conductance increases and there is a slight peak shift to slower frequencies as the temperature is increased. This is in agreement neither with the literature nor the physical intuition, as presented above and should be further investigated.

We would like to acknowledge the friendly support of the electronics workshop at the University of Leiden for the possibility of performing the presented measurements, especially in the persons of Bert Crama and René Overgaauw.

8. Design recommendations for electrical sensing

- **More elaborate experimental studies** on the electrical characteristics of LEDs should be performed. This implies measuring either the time or frequency response over a wide range of applied voltages, temperatures and LEDs from various manufacturers. From our preliminary experiments, we see the most promising results in the 1-10 MHz region, and so we recommend to focus on the features in this regime.
- We are confident that not only the spectroscopy itself can be implemented in the CMOS electronics, but also that the required temperature dependence can be
integrated: during operation the LED temperature will raise, changing the impedance characteristics. The temperature can be accurately monitored with the technique already patented by NXP, therefore this can yield very precise data on for the determination of the position (in energy) of the trap states.

- **The preferred method** for implementation on a chip seems to be a combination of the two methods, DLTS and AS, with the temperature sweep given by well-controlled pulses from the driving circuitry and measuring accurately the temperature of the LED junction with NXP’s sensorless sensing concept. The wavelength is given by extracting the accurate position of the radiative traps in the band gap. The transient technique relatively easily integratable in state-of-the-art drivers, which are already capable of producing almost perfect current or voltage block pulses. It is then necessary to time-resolve the electrical response of an LED to a driving pulse in order to perform DLTS measurements. A frequency method could possibly require somewhat more complex electronics to be implemented, but it might be more desirable for certain applications. For the solution to be integrated in actual application, the time required to perform a measurement should be known as this determines the slowest link in a hypothetical feedback loop. A full frequency sweep can be in the order of milliseconds, which is fast enough for video rate applications.

- **Absolute wavelength measuring** is desired and requires the above methods to probe the depth of the quantum wells in the LEDs. In the literature we found that it is possible to probe a single quantum well as one giant trap, but whether this is possible in an LED structure is still unknown. The results of the electrical study should be compared to the actual emission spectrum of each studied LED to find out whether absolute wavelength probing is possible.

9. Tuning the wavelength and flux of the LED

Blurring the LED spectrum (fuzzing) without sensing the LED

For certain applications it might be enough to simply ensure the colour uniformity among different LEDs rather than setting them all to a predefined wavelength. Two monochromatic light sources emitting light that differs only by 1.2 nm in wavelength will be already seen as different colour shades by a human eye. As an example the two diodes with emission peaks at 460 and 462 nm are presented in Figure 15 (solid lines) which will appear different to the human eye. The idea is to blur the peaks and in this way make them relatively more similar (Fig 15, dotted lines).
Proceedings Physics with Industry 2012

Figure 15. Two blue LEDs emitting at different wavelengths (solid lines) can be made more uniform to the human eye by smearing out the peaks (dashed lines).

If the diode peak-wavelength is shifted back and forth at a sufficiently fast rate (>60Hz), then the human eye will see it as a static color. We assess if we can use the Stark effect for shifting the wavelength. The LED would be driven by a constant, operating driving current and a fast varying AC component that will be responsible for the varying Stark shift. Although the reported wavelength shift due to the Stark effect for blue LEDs is up to 6 nm (NXP data), around the operating current (~360 mA) the wavelength is not subject to steep changes. An AC current amplitude of 100 mA will only shift the wavelength by ~1 nm. We hence conclude that fuzzing LED output by a fast varying AC component to the current will impair the flux, but not smear the wavelength.

10. Conclusions

In conclusion we have studied two ways of measuring the emission wavelength of LED. For optical detection we have concluded that using narrow bandpass filters is not practical except when the lighting conditions are very well-controlled in terms of incidence angle. This may be possible in use case scenario 1, i.e., calibration under controlled conditions in the factory, but even then many LED packages exclude controlled conditions. Using differential wavelength sensing with two photodiodes with a broad response is more viable, as it does not involve a strong angle dependent dispersion. Still, it requires light to reach the diodes which is not possible in every LED application, thereby possibly precluding use case scenario 3 (Measuring continuously).

We have also proposed an electrical method “admittance spectroscopy” in time or frequency domain to sense the wavelength, which only makes use of the electrical contacts that are already present on the LED. If this method is sensitive enough it could possibly be used for all LED geometries. Precise calculations and experiments should be done for both the optical and electrical sensing techniques to determine their potential for application.

Finally, we have explored the concept of wavelength fuzzing by varying Stark shifts through the drive current. A limiting factor in this approach is a very small wavelength tuning range that Stark effect offers (~1nm in the operating current range).
This also presents an important constraint for LED wavelength actuation. Using the Stark effect induced only by the drive current is not sufficient to overlap the emission of different LEDs, even if the wavelength spread is minimal (5 nm). Indeed, it suggests that even if the problem of wavelength measurement will be solved on a driver without the need of modifying the LED itself, compensating for different LED wavelength might be difficult to solve in such a simple fashion. In order to get a sufficient tuning range for the LED wavelength, one should have direct access to the LED to mount a piezo for strain control, or to mount an electrode for external Stark effect control. If this is not possible, binning remains unavoidable. A precise measurement device on the CMOS driver means that LEDs could be binned more accurately and efficiently, eliminating the binning step from the LED distribution process. This would, however, shift the binning to a different part of the value chain, i.e., away from the LED distributor, to a solution at the luminaire assembler.

11. References


[3] Debbar, N., Biswas, D. & Bhattacharya, P. Conduction-band offsets in pseudomorphic In_{x}Ga_{1-x}As/Al_{0.2}Ga_{0.8}As quantum wells (0.07≤x≤0.18) measured by deep-level transient spectroscopy. Physical Review B 40, 1058-1063 (1989).


Abstract

Tissue samples that are taken during a biopsy need to be snap-frozen in order to preserve their properties and use the tissue for contemporary molecular biology technologies that may improve the treatment of the patient. There is currently a lack of (safe) methodologies or devices for snap-freezing tissue. Furthermore, there is a lack of knowledge on the optimal cooling rate, which depends on the type of tissue and is important to know in order to avoid damage to the cells.

This report comments on the biological background of the acceptable cooling rates and also describes a design for a new biopsy snap-freezing device. The suggested device fulfills the requirements for use inside a hospital environment. The device consists of a cooling unit and a base station. The copper cooling unit can be pre-cooled on the base station until used. After biopsy, the tissue sample inside a cryovial can be deposited into the cooling unit and is then cooled down at rates between 1-10 K/sec, which is within the biologically safe range for several tissue types. The cooling unit may then be transported for several hours while keeping the tissue sample below 193 K.
1. Company profile

PamGene was founded in December 1999 and is a company engaged in the research, development, manufacturing and commercialisation of life science applications based on its proprietary micro array technology. PamGene is a biomarker company, focusing on opening up new opportunities for the development of Personalized Medicine (pharmaceutical discovery, translational and clinical). Most of PamGene’s projects are in oncology and many of these involve medicines that inhibit or modulate cellular kinases and kinase pathways as well as nuclear receptors and their signaling mechanisms. Our collaborators apply our technologies for biomarker research and to support drug discovery in areas including tumor tissue and compound profiling in diseases concerning kinases in Oncology and also in several other diseases and fields of expertise such as the Central Nervous System, Immunology and Obesity.

2. Problem description

Access to frozen biopsies is of tremendous importance to improve the treatment choices for patients. To enable contemporary molecular biology technologies, such as kinase activity testing, fresh/frozen tissue is a prerequisite. The problem is that frozen patient material (biopsy) is not seen as a routine source for diagnostic testing in clinical practice. Current practice in a hospital does not include a simple and reliable solution for this yet. When cancer is suspected, a variety of biopsy techniques can be applied. One type of surgery, called a core needle biopsy, uses a large, fitted needle to extract a sample of tissue about the size of a piece of pencil lead. A core needle biopsy can take place in a clinic or hospital and it can be performed by e.g. an internist, radiologist, or surgeon. Current practice in a hospital uses chemical fixation (formalin) directly after tissue acquisition to preserve tissue from degradation, and to maintain the structure of the cells. The latter is needed as a definite diagnosis of cancer is almost always based on the histological examination of tissue samples.

To enable contemporary molecular technologies fresh/frozen tissue is a prerequisite. Fresh tissue is obtained by means of snap freezing for which solid carbon dioxide, liquid nitrogen ($\text{N}_2\text{\textsubscript{lq}}$) or isopentane cooled with liquid nitrogen are used as coolants to keep good morphology of the tissue and to keep the molecular activity intact. Snap freezing tissues for diagnostic and research purposes are therefore often time consuming, laborious, even hazardous and not user friendly and are therefore not applied at the location of biopsy acquisition.

A simple device is foreseen which facilitates easy and reliable logistics of frozen patient material from the location where the biopsy is removed to the pathology laboratory where the biopsy is examined and archived. This device is composed out of two parts:

- a special designed micro-tube facilitating easy removal of the tissue from the sample notch of the biopsy needle (the PamTainer)
- a cryogenic device enabling temporary freezing of the tissue collected in the micro-tube (the PamFreezer).

In this workshop we focus on the PamFreezer design. The PamFreezer should be able to replace the currently used coolants while snap freezing the tissue directly after its acquisition, to allow temporary storage and to bridge the transportation time of the frozen tissue to the pathology laboratory.
Main characteristics of the device include:

- Fast cooling ratio (cooling well below 273 K in seconds)
  Note: details of the cooling rate as well as the required lowest cool temperature are not
  known though literature shows sufficient results at cooling times of 10 - 30 seconds
  either at N\textsubscript{2}\textsubscript{liq} or dry ice. Cooling in a 193 K fridge however, is not allowed.
- The sample should remain below 193 K for at least one hour, during transport from the
  place of biopsy taking to the storage/analysis location somewhere else in the hospital.
- Preferably only electricity is allowed as utility, alternatively CO\textsubscript{2} cartridges may be used
  (no safety vessel).
- Low weight, user friendly, limited servicing required, low noise, etc.
- Autoclavable.

2.1 Geometry

The biopsy tissue sample is taken with a 8G needle (largest case), which has an inner diameter
of 3 mm. The length of the notch in this needle is 1 cm (average case); its volume is then 70
mm\textsuperscript{3}. A cylindrical geometry of the sample is assumed. The type of tissue, and therefore also
the (thermal) properties, may vary, however that variation is typically within a few percent.
The cryovial in which the sample is placed has typical dimensions of 10 mm outer diameter, 5
cm in length and a wall thickness of 1.5 mm, and is made of polypropylene.

2.2 Energy estimates

For the “back of an envelope” estimate of the demanded cooling power, the following
assumptions are used:

- Tissue sample is initially at body temperature (310 K).
- Vial is initially at room temperature (293 K).
- The vial and sample have to be cooled down to 77 K within 10 seconds.
- The specific heat of the materials is temperature dependent.
- As the specific heat decreases and is in literature often only specified at specific
  temperatures, linear interpolation is used as the trend is linear for most materials in first
  order approximation.

Using these assumptions and the material properties listed in table 1, an estimate can be made
for the amount of heat to be extracted.

The heat to be extracted from the tissue sample is given by:

\[
E_{\text{tissue}} (T = 310 \text{ K} \rightarrow 273 \text{ K}) = \rho V_{c} c_{p} \Delta T = 1 \cdot 0.07 \cdot 4 \cdot 37 = 10.4 \text{ J}
\]
\[
E_{\text{phase transition}} (T = 273 \text{ K}) = \rho V L = 1 \cdot 0.07 \cdot 334 = 23.4 \text{ J}
\]
\[
E_{\text{tissue}} (T = 273 \text{ K} \rightarrow 77 \text{ K}) = \rho V_{c} c_{p} \Delta T = 1 \cdot 0.07 \cdot 2 \cdot 193 = 27.0 \text{ J}
\]

Total amount of heat from the tissue: 60.8 J
### Table 1. Material properties of the biopsy tissue sample, the cryovial and copper

<table>
<thead>
<tr>
<th>Material</th>
<th>Density ($\rho$)</th>
<th>Heat capacity ($c_p$) T $&gt; 273$ K</th>
<th>Heat capacity ($c_p$) T $&lt; 273$ K</th>
<th>Thermal conductivity ($k$) T = 293 K</th>
<th>Heat capacity ($c_p$) T $&lt; 77$ K</th>
<th>Thermal conductivity ($k$) T $= 77$ K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological tissue [1]</td>
<td>1 g/cm$^3$</td>
<td>4 J/g K</td>
<td>2 J/g K</td>
<td>0.5 W/m K</td>
<td>2 J/g K</td>
<td>0.16 W/m K</td>
</tr>
<tr>
<td>Cryovial (polypropylene) [2]</td>
<td>0.85-0.95 g/cm$^3$</td>
<td>1.5 J/g K</td>
<td>0.5 J/g K</td>
<td>1.6 W/m K</td>
<td>0.2 J/g K</td>
<td>0.16 W/m K</td>
</tr>
<tr>
<td>Copper</td>
<td>8.96 g/cm$^3$</td>
<td>0.4 J/g K</td>
<td>0.2 J/g K</td>
<td>857 W/m K</td>
<td>0.2 J/g K</td>
<td>857 W/m K</td>
</tr>
</tbody>
</table>

with $c_p$ the temperature dependent specific heat, $\Delta T$ the change in temperature, $\rho$ the density and $V$ the volume of the sample. The latent heat $L$ is required for the phase-transition from liquid to solid around the freezing point of water (second line of calculation).

The heat to be extracted from the vial is calculated in a similar way by:

$$E_{\text{vial}}(T = 310 \text{ K} \rightarrow 77 \text{ K}) = m_{\text{polypropylene}}c_p\Delta T = 1 \cdot 1 \cdot 233 = 233.0 \text{ J}$$

$$E_{\text{air in vial}}(T = 310 \text{ K} \rightarrow 77 \text{ K}) = m_{\text{air}}c_p\Delta T = 0.001 \cdot 1 \cdot 230 = 0.2 \text{ J}$$

Total amount of heat from the vial: 233.2 J

Total amount of heat from the vial and the tissue sample: $\approx$ 294 J

In order to cool the vial and sample to the desired temperature of 77 K within 10 seconds, a cooling power of 29 W is needed:

$$P_{\text{desired}} = \frac{\text{total energy}}{\text{cooling time}} = \frac{294 \text{ J}}{10 \text{ s}} \approx 29 \text{ J/s} = 29 \text{ W}.$$ 

Note that the amount of heat to be withdrawn and hence the cooling power, increases linearly with the volume of the biological material and/or the size of the vial. As small and low-noise cryogenic coolers (suitable for application in a hospital environment e.g. pulse-tube or Stirling) have a typical cooling power of 10 W, the demanded cooling rate of the sample cannot be achieved directly. As a solution, we suggest to use a “cold reservoir” that is pre-cooled to a low temperature (e.g. 77 K), in which the sample is inserted and hence cooled down rapidly.

As the cold reservoir absorbs the heat from the sample and vial, the temperature of this reservoir will increase inevitably. To calculate the temperature increase of such a reservoir made out of copper, we use the following assumptions:
• The cold reservoir is at cryogenic temperature (77 K).
• The cold reservoir is perfectly insulated from the environment, i.e. heat flux to surrounding is zero.
• The heat from the tissue sample and vial is fully transferred to the cold reservoir.
• The specific heat of the cold reservoir is temperature dependent.
• As the specific heat decreases and is in literature often only specified at specific temperature, linear interpolation is used as the trend is linear for most materials in first order approximation.
• The mass of the copper block is 232 grams.

The increase in temperature of the cold reservoir upon taking all the heat from the sample and vial is given by:

\[
\Delta T = \frac{E_{uptaken}}{c_p m} = \frac{291}{0.3 \cdot 232} = 4 \, K
\]

with \(\Delta T\) the change in temperature, \(E_{uptaken}\) the energy to be withdrawn from the sample and the vial, \(c_p\) the temperature dependent specific heat of the cold reservoir and \(m\) the mass of the cold reservoir. So a block of 232 grams of copper cooled down to 77 K will heat up to 81 K when the vial and the sample are inserted and have obtained a thermal equilibrium with the cold reservoir.

3. Benchmark: biopsy in cryovial in \(N_2^{liq}\)

The current procedure for biopsy snap-freezing (submersion of the cryovial in liquid nitrogen \([N_2^{liq}]\)) was evaluated as benchmark case. No literature on the temperature evolution of the tissue sample during this procedure is known. On the other hand, protocols are not standardized across hospitals or operators. Commonly used protocols include submersion of the tissue in \(N_2^{liq}\) for a varying number of seconds. A Leidenfrost effect (thermal shielding by vaporization of \(N_2^{liq}\) near the surface) may occur, which may be related to the ‘hissing’ sounds that are reported. To prevent this, cycles of submersion into and extraction from \(N_2^{liq}\) are also reported to be used, which may reduce the cooling rate significantly. However, details on the temperatures involved are missing.

We have investigated numerically the temperature evolution of a piece of tissue inside a vial, which is immersed in a liquid nitrogen bath at 77 K.

3.1 Numerical model

A one-dimensional numerical model was constructed, which assumed a piece of tissue \((T_t = 310 \, K, \text{ thickness of } 1.5 \, \text{mm})\) resting against a polypropylene wall \((T_w = 298 \, K, \text{ thickness of } 1.5 \, \text{mm})\), with liquid nitrogen \((T_n = 77 \, K)\) on the other side of the wall.

Inside the tissue, the wall and the liquid nitrogen, the one-dimensional heat (diffusion) equation is solved:

\[
k \frac{\partial^2 T}{\partial x^2} = \frac{\partial T}{\partial t}
\]

where \(x\) is the distance from the center of the tissue outward.
At the boundary between the vial wall and the liquid nitrogen, heat transfer is modeled:

\[ Q = -hA \frac{\partial T}{\partial x} \]  

(2)

The heat transfer coefficient \( h \) is not known exactly for these two materials, but for liquid nitrogen with steel this number is around 250 W/m\(^2\)-K [3] and we assume it to be 100 times smaller for plastic (which is based on the ratio of heat conduction of plastics and steel). When there is a thin layer of vaporized nitrogen near the wall (Leidenfrost effect), \( h \) is reduced by another factor of 6 (which is based on the ratio of heat conduction of nitrogen in liquid and vapor phase). All thermal properties of nitrogen in liquid and gas phase are obtained from the NIST database [4]. The thermal properties of the tissue and the vial are listed in table 1. At large \( x \) the temperature of the liquid nitrogen is assumed constant at 77 K. At \( x = 0 \) there is a symmetry boundary condition.

\[ \text{Figure 1. Temperature profiles at 6 selected time steps, showing the cooling of a tissue sample} \]
\[ \text{\( x = 0 - 1.5 \text{ mm} \) next to a wall \( x = 1.5 - 3 \text{ mm} \) immersed in liquid nitrogen \( x = 3 - 30 \text{ mm} \).} \]

\[ \text{Figure 2. Temperature evolution of the center of the tissue sample as a function of time. A nitrogen vapor layer formed on the vial surface leads to significant thermal shielding and therefore a reduced cooling rate.} \]
Equations 1 and 2 are solved numerically using explicit finite-difference discretization, with spatial steps of 0.1 mm and temporal steps of $10^{-5}$ s. The phase change of the tissue when freezing near 273 K is incorporated by setting the heat capacity to 334 W/K when the temperature of the tissue is between 272.5 and 273.5 K.

3.2 Results

Figures 1 and 2 show that the temperature of the tissue sample decreases rapidly towards zero, where there is a phase change, before the temperature decreases rapidly towards 193 K and below. A nitrogen vapor layer near the vial wall leads to significant thermal shielding that slows down the cooling of the tissue sample by a factor of 6. The tissue sample reaches a temperature of 193 K within 5 seconds (including $\text{N}_2^{\text{liq}}$ vaporization), with cooling rates around 40 K/sec. This cooling rate is likely to be overestimated for several reasons, e.g. the heat transfer across the several interfaces may not be so perfect as assumed here. Nevertheless, the model gives an order-of-magnitude estimation of the cooling of the tissue sample and of the important factors for this cooling.

4. Cooling rate limits

It is generally thought that snap-freezing should involve cooling rates as high as possible, in order to preserve the tissue as good as possible. Determining the exact limits for the cooling rates for the new snap freezing device require a more careful look at the underlying biological phenomena, which will be done in this section.

4.1 Biological aspects influencing the cooling rate

Liquid water is essential in both function and structure of living cells. Freezing water can be lethal for the cells and paradoxically can also preserve cells for long periods of time in a viable state, and it permits the long-term storage of tissues and organs. Depending on the rate of freezing or thawing, the sub-cellular constituents and the details of cell structure can be preserved or become disrupted. Thus the rate of freezing and thawing of the cells is very

\[ \text{Figure 3. Schematic presentation of crystal formation in cells.} \]
critical. As demonstrated cells are subjected to chemical, thermal and mechanical forces during cryopreservation, which can profoundly affect their biological function [5]. It has been observed that during freezing the intracellular ice formation induces significant damage to the cells. Two main damage mechanisms have been proposed during freezing. At slow freezing, the chemical potential difference across the cell membrane (as a result of extracellular ice formation) may lead to cell dehydration. The dehydration can cause significant cell damage [6]. Additionally extracellular ice formation can cause significant thermo-mechanical stress leading to mechanical cell damage (figure 3).

The cooling rate will determine the size of the crystals, and depending on the size of the crystals the damage to the cells will be in a higher or lower degree. Therefore, the speed of cooling down the tissue plays an important role in cryopreservation. To design our cooling device, it is essential to take the optimum rate of cooling/freezing into account.

4.2 Method for estimating the optimal cooling rate

The rate of intracellular ice nucleation depends on the type of cells. The key parameters are permeability of the cell membrane to water ($L_{pg}$), apparent activation energy ($E_{LP}$), and the ratio of the available surface area for water transport to the initial volume of intracellular water ($SA/WV$).

The optimized cooling rate $B_{opt}$ (K/minute) can be determined using the following equation [7]:

$$B_{opt} = 1009.5e^{-0.0546E_{LP}L_{pg}}\frac{SA}{WV}$$  \hspace{1cm} (3)

The $B_{opt}$ values for several types of cells have been calculated, see table 2. These values show that the cooling rate for different types of cells including tumor cells can be different. To have the maximum survival of cells we need to define the optimum cooling rate range (figure 4). Depending on the cell type, this optimal cooling rate range can be different.

Toner and coworkers developed a model from which they calculated the intracellular nucleation rate [8]. Later Karlsson et al. improved Toner’s model [5]:

$$J(T) = \Omega e^{-\kappa T^3\Delta T^-2}$$  \hspace{1cm} (4)

Here $J(T)$ is the nucleation rate of the intracellular ice formation in a given cell population undergoing cryopreservation, $\Omega$ is the kinetic coefficient, $\kappa$ is the thermodynamic coefficient, and $\Delta T$ is the supercooling. The $\Omega$ and $\kappa$ coefficients are different for each cell type. Toner et al. used mouse oocyte. Based on the $\Omega$ and $\kappa$ coefficients of the mouse oocyte they suggested an optimum cooling range of 60 to 600 K/minute (or 1 to 10 K/sec). We used this range further to design our apparatus.

5. PamFreezer design

Here we present a design idea (figure 5) for a PamFreezer cooling unit to snap freeze the biopsy tissue sample, which also includes a PamFreezer base station. The design is based on

<table>
<thead>
<tr>
<th>cell type</th>
<th>$L_{pg}$</th>
<th>$E_{LP}$</th>
<th>$SA/WV$</th>
<th>% error</th>
<th>$B_{opt}$ (K/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lymphocyte</td>
<td>0.10</td>
<td>15.5</td>
<td>0.74</td>
<td>5.3</td>
<td>0.5</td>
</tr>
<tr>
<td>tumor cells</td>
<td>2.71</td>
<td>55.4</td>
<td>0.3</td>
<td>31</td>
<td>0.7</td>
</tr>
<tr>
<td>spermotosoma</td>
<td>0.01</td>
<td>29.2</td>
<td>11.1</td>
<td>10</td>
<td>0.3</td>
</tr>
<tr>
<td>oocyte</td>
<td>0.04</td>
<td>13.3</td>
<td>0.10</td>
<td>5</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 2. Calculated values of $B_{opt}$ for several types of cells using the values report in [7].
the cooling rates and thermal conditions stated above.

5.1 Cooling unit

The central piece of the device is a cylindrical copper block of 232 grams weight, 1.75 cm radius and 3.25 cm height. It has a chamber in the centre to place the cryovial with the tissue sample inside, and an extra piece of copper on the top that acts as a cover. The cooling unit is pre-cooled in the base station (described below) until it needs to be used. Copper was chosen because of its good thermal conductivity and high heat capacity. Furthermore, it is cheap (ca. € 6,-/kg). The total mass of copper was calculated to fulfill its two main functions: as a “cooler” and as a “cold maintainer” inside the cryovial. For the first function, the block is estimated to be pre-cooled to 77 K in about 30 minutes. The second function is to absorb external heat once the block is separated from the cooler, thereby

Figure 4. Schematic presentation of maximum cell survival depending on the cooling rate.

Figure 5. Sketch of the PamFreezer design.
cushioning the temperature increase of the cryovial. The cooling unit is light in order to be transported through the hospital for at least one hour to the storage/analysis location.

5.2 Cryovial chamber

The cryovial and the cryovial chamber both have a conical shape, which allows for pushing the cryovial into the chamber to ensure maximum contact. A soft plastic for the cryovial material may enhance the fitting of the vial in the chamber. Additionally (or alternatively), the wall of the chamber may be covered with a thermal conduction material that enhances the contact between wall and vial, such as Velbond [11].

5.3 Insulating cover

In order to ensure that the temperature of the biological sample stays below 193 K for 1 hour without active cooling, the insulation of the cold reservoir and its mass i.e. heat capacity have to be taken into account. As the mass of the block is limited for practical reasons to lower than 0.5 kg, the insulation that is used is critical. First choice (table 3) could be low-density polystyrene as it is low cost and easy to manufacture/shape. However, as the device has to be heat-sterilized (typical temperatures 393 - 473 K), polystyrene cannot be used (melting point of 373 K). A better candidate would be a ridged container made from e.g. fiber glass (already widely used in non-magnetic cryostats) filled with e.g. glass wool or Perlite. A recently developed material, aerogel has proven to be a very good insulator with the extra benefit of having an extremely low density. It is thus an ultralight material suitable for insulating, due mainly to its low thermal conductivity (30 mW/m·K), which can be even further enhanced by introducing a vacuum (or low pressure). Furthermore, the material is relatively cheap (ca. € 20,- for one cooling unit). In the PamFreezer, the function of an aerogel cover is to insulate the copper block, in order to prevent heat transfer from the ambient. Except the part of the bottom wall that is in contact with the cooler, all the surface of the copper block should be covered by at least 5 cm of aerogel. In addition, the cold part of the cooler is also covered with aerogel to reduce heat gains from the surroundings. An extra piece of aerogel with the same thickness is reserved to replace the cryocooler once the device is ready for transport.

5.4 Base station: cooler selection

The base station is a fixed device that can cool down the cooling unit from room temperature to 77 K, and compensates the heat absorbed from the surroundings until the cooling unit needs to be used. The main component of the base station is a cryocooler. There are some requirements for the cooler selection that should be satisfied, like the cooling power and the minimum temperature attainable. The copper is cooled down from room temperature to 77 K in 30 minutes, so the required cooling power is 5 W. Moreover, since the
device is intended for use in hospitals, it should meet certain requirements for facility of use and safety reasons. To satisfy the requirements of the biopsy taking workplace, the device should not have vibrating parts to avoid excessive noise.

Initially, thermoelectric (Peltier) devices were considered as it is easy to control, cheap and has a low level of noise. However, the temperature limit they can reach is 193 K, which is not enough for our purpose.

The second option was the Stirling cooler, capable to reach cryogenic temperatures with powers up to 15 W. Although they work with helium as the working fluid, they satisfy the safety requirements since they are sealed systems and are therefore without any possible contact of the operator with hazardous fluids. Reasons against this cooler are its weight, price and noise. However, the cooling device is intended to remain in the surgery room, hence the weight is not a significant problem.

The last option is pulse tube cryocoolers, that also operate with helium and are capable to reach cryogenic temperatures. In contrast with Stirling, this cryocooler can be operated without moving parts (less noise) in the low temperature side of the device, making the cooler suitable for our application. Pulse tube cryocoolers are available from Qdrive (USA), Sunpower (USA), Thales Cryogenics (Europe), AIM (Europe), Honeywell Hymatic (UK) and Ricor (Israel) at prices in the k€ range.

5.5 Usage protocol

The PamFreezer has a cooling down time of approximately 30 minutes and may be (auto-)switched on during the night so that it is ready for operation at 8 AM. The biopsy is removed from the biopsy needle and placed inside the cryovial, preferably making as much contact with the cryovial as possible. The cryovial is then gently pushed inside the cryovial chamber of the PamFreezer cooling unit. The insulation cap is then added to the PamFreezer cooling unit. After one minute, the PamFreezer cooling unit may be undocked from the PamFreezer base station, and the bottom insulation should be added to the PamFreezer cooling unit. Once the PamFreezer cooling unit is undocked, the cooling unit starts to warm up and there is a time limit of 1 hour before the temperature of the tissue sample has increased to 193 K.

After returning to the location where the biopsies are taken, the PamFreezer cooling unit can be re-docked on the PamFreezer base station after removing the bottom insulation and replacing the cooling cap (which has a cryovial chamber notch that prevents condensation of air on the chamber surface). The PamFreezer cooling unit is then cooled down (from room temperature) within 30 minutes.

5.6 Additional device considerations

Biopsy sample inside the cryovial

The heat transport between the biopsy tissue sample and the cryovial is affected by the positioning of the sample inside the vial. The heat transport encounters a resistance at each interface, however this is very much material-dependent and is not easily obtained from the literature. The heat transport increases linearly with contact area (equation 2). Ideally, the sample would fill the entire lowest part of the conical cryovial.

Auto-switch on/off

The operator of the PamFreezer may be assisted by including an auto-switch in the base station design, that detects the docking of the PamFreezer cooling unit (using an electrical contact or pressure sensor) and switches the base station off once the cooling unit is undocked.
5.7 Budget solution

As currently available cryocoolers are expensive (k€ range), we envision a solution that is based on using liquid nitrogen. Instead of using it close to the patient (not desired), it may be used in a central facility in a hospital. By carefully choosing the mass of the copper block (1 kg) and the insulation (5 cm thick vacuum aerogel; thermal conductivity 0.001 W/m·K, heat capacity 0.84 J/g·K and density 1 kg/m\(^3\)), the temperature of the cold reservoir increases only from 77 K to 100 K within 10 hours. After cooling down this cooling unit in the central facility using N\(_2\)\(_{liq}\), it can be transported to and stored in the area where it is used for snap-freezing. After snap-freezing, the tissue sample can stay in the device for another few hours as long as the device with sample is transported to the final destination within 10 hours after pre-cooling. Although it still requires on N\(_2\)\(_{liq}\), this solution can be used in a pilot study for the applicability of cooling biological tissues using a solid material (e.g. copper).

6. Simulations of the PamFreezer cooling device

The performance of the proposed PamFreezer design is tested using finite element analysis executed with COMSOL Multiphysics software. A two-dimensional axisymmetric model is employed to calculate the heat transfer by conduction during the different stages of operation of the PamFreezer. Radiative and convectional heat transfer are assumed to play no significant role. The design has three important requirements: the pre-cooling time, snap-freezing speed, and the warm up time of the biopsy tissue without cooling. Before we address these properties, we discuss the details of the finite element analysis.

6.1 Numerical setup

Figure 7 shows the model used in the finite element analysis. The model has a cylindrical symmetry and the symmetry axis is shown in red. We have used two configurations to calculate both the pre-cooling time and snap-freezing properties. Firstly, for the pre-cooling study we remove insulation underneath the copper and add a heat sink on the bottom boundary of the copper. And secondly, we use the configuration as shown in figure 7 for the snap freezing. COMSOL calculates automatically an appropriate number of cells, with higher mesh densities near corners and small spaces.

We have three different boundaries within the model. Firstly, we have a symmetry boundary on the axis of symmetry. Secondly, we have the outer boundaries of the isolation which are set to room temperature. And thirdly, we have a boundary between a solid and a liquid or gas, where the condition is such that the temperature is continuous across the boundary.
The temperature of the insulation is always the room temperature at the start of the calculations. This is the worst case scenario for the snap-freezing and the warm up time of the biopsy tissue without cooling. In reality, the warm-up time will be larger than our result and the equilibrium temperature reached after snap-freezing will be lower than our result.

6.2 Pre-cooling

The pre-cooling of the copper was calculated using the cooling power characteristics of three types of commercially available cryocoolers. The three coolers have a cooling power of 2.5 W, 5 W and 10 W respectively at a temperature of 77 K. Figure 8 shows the temperature measured in the center of the copper lid, starting from the moment that the (insulated) copper is placed on the base station and the cooler is powered. The copper and insulation is assumed to be at room temperature (293 K) before cooling. As we observe, both the 5 W and 10 W cryocooler are able to cool the copper cold reservoir to 77 K within 20 minutes. The cooling rate of the 2.5 W cryocooler is significantly slower because the amount of heat leaking through the insulation is almost equivalent to the cooling power.

6.3 Snap-freezing

The snap-freezing of the tissue is modeled by inserting a conically shaped polypropylene vial with thickness of 1 mm and content volume of 3 mL in the pre-cooled copper. The vial contains a 0.07 mL piece of tissue that is assumed to consist purely out of water. To account for the influence of surface roughness of both the polypropylene vial and the copper cold reservoir, the vial is assumed to be surrounded by 0.1 mm of air. We also take the latent heat (23.4 J) of freezing water into account in order to model the phase transition correctly. Figure 9 shows the temperature of the center of the biopsy tissue after insertion of the cryovial into the cooling unit. The center of the tissue starts to cool down after ca. 10 seconds, as first the vial and the outer parts of the tissue have to cool down. In figure 9a we also observe that a larger temperature gradient between the copper and tissue results in a shorter snap-freeze.
time. We find a snap-freeze time of 20 s (for 77 K), 11 s (for 150 K), and 8 s (for 193 K). The cooling rates are determined at the freezing point of water and are approximately equal to, -8 K/s (for 77 K), -5 K/s (for 150 K), and -3 K/s (for 193 K), which is within the acceptable ranges as stated in section 4.

Another important observation of figure 9a is the temperature of the tissue after it reached its equilibrium temperature. We observe the equilibrium temperature to be higher than the starting temperature of the pre-cooled copper. If the pre-cooled copper is 77 K, we find an equilibrium temperature which is 23 K higher. For the other pre-cooled temperatures, 150 K and 193 K, we find that the equilibrium temperature is respectively 13 K and 10 K higher. We

Figure 9. (a) Temperature as a function of time measured in the center of the biopsy tissue when inserted in a pre-cooled copper. The different colors represent the different temperatures of the pre-cooled copper: 193 K (red), 150 K (blue), or 77 K (green). (b) Temperature as a function of time while snap-freezing different tissues with the same volume but different contact areas. The temperature of the pre-cooled copper is 77 K. The contact area increases from 6.6% (olive green), to 13% (red), to 21% (blue), and to 27% (green).
attribute this difference to the heat transferred from the tissue and vial to the copper block and to the heat transferred from the insulation, which is assumed to be at room temperature, into the copper, vial, and tissue.

We assume the contact area of the biopsy tissue with the vial to be of significant importance, because heat conduction through air is worse than through the vial-tissue contact. Figure 9b shows the snap-freeze curves for different contact area. We express the contact area as a percentage of the total surface area of the biopsy tissue, which is a cylinder with constant volume of 0.07 mL. The contact area increases from 6.6% (olive green), to 13% (red), to 21% (blue), and to 27% (green). We observe that a larger contact area does not result in a faster snap-freeze time, but the snap-freezing does happen earlier in time. This indicates that the snap-freezing is limited by the thermal conduction of the tissue material and especially the latent heat necessary to freeze the tissue.

6.4 Warm up time of the biopsy tissue during transport

The last important issue to address is the time the sample needs to warm up to 193 K as this is the temperature at which the sample preservation is lost. For this end, we calculated the temperature within the sample an hour after snap-freezing. Figure 10 shows a typical warm up curve of a tissue snap-freezed in a pre-cooled copper block of 77 K. The temperature after one hour is 183 K, which is still below the critical 193 K. It is important to stress that there are two effects which influence this behavior. Firstly, the insulation starts at room temperature during snap freezing, thereby emptying the cool copper bath. In reality the insulation is also cooled, therefor the warm up time will be larger than an hour in reality. And secondly, the actual device will likely have a heat leak between the insulation lid and the copper, thereby possibly increasing the heat leak to the environment, this would reduce the warm up time. However, we expect that one hour warm up time is feasible within the current model.

![Temperature as a function of time](image)

Figure 10. Temperature as a function of time in the biopsy tissue, which was snap-freezed in a pre-cooled copper block of 77 K.
7. Discussion & Outlook

The simulations show that the proposed PamFreezer design should be able to cool the tissue sample in a sufficiently fast way, while also being able to retain at low temperatures for at least one hour during transport. Furthermore, the materials for the cooling unit are cheap, however the cryocooler itself is expensive but may be replaced (initially) by liquid nitrogen.

Important next steps are the (thermal) validation of both simulations and a more thorough investigation into the optimal cooling rates.

7.1 Model validation

The numerical models need to be verified before their predictions can be accepted.

Numerical validation

The numerical models for snap freezing in liquid nitrogen and in the new cooling device can be validated by comparing their typical cooling rates and characteristics. The COMSOL model furthermore needs internal verified with a grid-independency check.

The validity of the COMSOL model for the new cooling device is supported by the numerically calculated increase of the Copper block, which matches the estimation made in the Problem statement section.

Experimental validation

The two numerical models have to be validated experimentally, by time-resolved measuring the temperature of a tissue sample during snap-freezing. No literature is known on the cooling rate of a tissue sample snap-frozen in liquid nitrogen.

Temperature measurement of a tissue sample can be done in several ways [12, 13]:

- **Thermocouple**
  A junction of two different metals may generate a voltage when heat is applied. This phenomena is used in thermocouples, which can be as small as 0.1 mm in diameter. Such a thermocouple could be placed inside the tissue sample for temperature measurements. Heat exchange between the tissue sample and the metals should be fast in order to assure that the thermocouple reads accurately.

- **Thermistor**
  The resistance of a resistor or a wire depends on its temperature, and this phenomena may be used to measure the temperature of a sample in which a wire or resistor is placed. Also for this method the wire/resistor needs to be small.

- **Thermal noise measurement**
  The amount of thermal noise on an electrical wire or component is a measure of its temperature. Proper calibration is required (also for the two methods above).

- **Thermal camera**
  Thermal cameras can record the IR radiation of objects and in that way measure the temperature of such an object. Cooled thermal cameras typically operate in the 60-100 K range; uncooled thermal cameras operate around room temperatures. The temperature resolution is generally around 2 K and these cameras are usually expensive. Furthermore, they only measure the temperature of the outer surface; the temperature of the core can then only be estimated using assumptions for the material properties.
7.2 Cooling rate

• A systematic study of optimum cooling range for different types of cells is required.

• Employment of cryomicroscopy for identification and study of ice formation (darkening of intracellular space) is beneficial to determine the amount of damage occurred in the cells during the freezing process [9, 10].

8. Summary

We have proposed and simulated a design for a PamFreezer biopsy snap-freezing device, which fulfills the requirements for use inside a hospital environment. The device consists of a cooling unit and a base station. The copper cooling unit can be pre-cooled on the base station until used. After biopsy, the tissue sample inside a cryovial can be deposited into the cooling unit and is then cooled down at rates between 1-10 K/sec, which is within the biologically safe range for several tissue types. The cooling unit may then be transported for several hours while keeping the tissue sample below 193 K.

(Thermal) validation of this design is an important next step.

9. References


The physics of water and wax in the pores of a working Gas-to-Liquids catalyst

M. Dukalski¹, J. Harting², H. Huinink², F. Kako Joibari¹, W. Onderwaater³, E. van der Bijl⁴, B. van Oort⁵, G. Wijts⁶, S. Zhang²

¹ Kavli Institute of Nanoscience, TU Delft, The Netherlands
² Department of Applied Physics, TU Eindhoven, The Netherlands
³ Universiteit Leiden, ESRF, The Netherlands
⁴ Universiteit Utrecht, The Netherlands
⁵ Vrije Universiteit Amsterdam, The Netherlands
⁶ Institute of Physics, Universiteit Leiden, The Netherlands

Problem definition by Shell - focal points: R. Bos, H. Geerlings, H. Kuipers and S. van Bavel

1. Abstract

The so-called Fischer-Tropsch catalysis allows to convert natural gas into liquid products and is the underlying mechanism of commercially used “Gas-to-Liquids” plants. The actual reaction takes place in millimetre sized porous pellets in which active metallic particles are dispersed as catalysts. Due to the reaction the pores of the pellets will become filled with the reaction products (“wax” and water), but it is uncertain if the fluid in the pores can be understood as a single liquid phase, a liquid-gas mixture, or multiple continuous phases. The answer to this question is important for a thorough understanding of the transport processes inside the reactor and can be utilized to improve its efficiency. In this project, a theoretical analysis of the behaviour inside the pores is performed. It is concluded that a liquid water phase might well exist next to the wax phase. However, the analysis is based on very limited experimental data of unknown quality. Therefore, we propose a number of possible experiments to validate the theoretical concepts.

2. Company profile

Shell consists of a global group of energy and petrochemicals companies with around 90,000 employees in more than 80 countries and territories. Shell is currently the largest investor in research and development among the major oil firms. In 2010 they spent around 1 billion dollars in research and development for technologies that will be needed to produce cleaner energy and more efficient fuels.

3. Problem description

In the “Gas-to-Liquids” (GTL) process natural gas – methane – is converted first into so-called “syngas” and then into liquid products – e.g. diesel, kerosine or heavier hydrocarbons – that
for example can be used as transportation fuels. This process is performed commercially on a very large scale by both Shell and Sasol. These plants of Hyde Park size cost 10 – 20 billion dollar and produce up to 140.000 barrels per day. But the core of the chemical conversion process takes place on the sub nanometer scale; important physical processes take place at the scale of nano to micrometers. It is that latter scale that this problem focuses on. The chemical reaction at hand starts from the “syngas” mixture of carbon monoxide and hydrogen, which on the “active sites” of a metal catalyst reacts or polymerizes to long –linear– alkanes and water according to

$$n \text{CO} + 2n \text{H}_2 \rightarrow (\text{CH}_2)_n + n \text{H}_2\text{O} \; ; (\text{CH}_2)_n \text{ actually being CH}_3 - (\text{CH}_2)_{n-2} - \text{CH}_3.$$  

The conditions are $T = 200 – 250^\circ\text{C}, \; P = 20 – 80\text{bar}$. In practice $n$ will be any number and hence a whole range of products are made, but for simplicity we can here assume $n$ to be a single and high number, e.g. $n = 60$, i.e. the product is a liquid under reaction conditions. We call this product “wax”. Note that for each gram of wax we approximately make 18/14 grams of water.

The wax molecules are formed on the active metal surface (the “catalyst”). In order to have a very high metal surface area the nano scale metal particles are dispersed within a porous “carrier” material. This carrier is typically in the shape of spherical, cylindrical or more complex shaped particles with a diameter of 0.1 – 3 millimeter, with an interconnected pore structure where the pore sizes are typically in the range of 5 – 100 nm.

Due to the reaction that produces both water and wax these pores will become “liquid filled”. This might be a single phase of wax in which water is dissolved. In that case the gaseous components CO and H$_2$ flowing on the outside of the particles first have to dissolve into the liquid and then diffuse into the internal pores to reach the active sites where the reaction actually take place. Vice versa, the produced wax and water detach from the metal surface and then have to transport through the pores to the external surface of the particles. Since the reaction is fast relative to the rate of transport there will be concentration gradients of the reactants: the CO and H$_2$ concentration will be lower in the centre of the particle compared to the outer part. In a simple “one phase system” this also implies that the product H$_2$O has an inverse type of profile, i.e. the H$_2$O concentration – which may be zero at the outer surface – will be higher in the core and transport of H$_2$O primarily takes place via diffusion and a bit via convection (along with the wax that is simply being pushed out of the pores). However, this “one phase system” assumption may be completely wrong!

The key question of the project is: will there be a “simple” single wax phase inside the pores – with H$_2$O dissolved in it – or do we have a more complex situation, for example, a liquid-gas mixture, a micro-emulsion of water rich droplets in a wax continuous phase, vice versa wax in water, two or more different continuous phases, or something even different? Moreover, how can this be validated experimentally? And how does this affect the transport mechanisms of the reactants CO and H$_2$ and particularly also the product H$_2$O into and respectively out of the particle?

This is of imminent importance because:

- CO, H$_2$ and H$_2$O and their concentration profiles – from a metal surface perspective – will be quite different dependent on the type and nature of the phases.

- these concentrations – or chemical potentials – strongly determine the efficiency of the chemical reaction and hence the overall carbon and energy efficiency of the GTL process.

4. Problem solving strategy

The key question of this project can be rephrased by developing different “phase scenarios”: simplified models of the spatial distribution of liquid water and wax in the nanopores (1). CO and H$_2$ are not included in these scenarios. This is reasonable, because it seems likely that no water vapor is present (see below), and that the diffusion rates of CO and H$_2$ through liquid
water and wax are similar. Therefore the concentration profiles of CO and H$_2$ in the pore are expected not to vary much between the phase scenarios.

Scenario a) is the most simple scenario, in which all water mixes on the molecular level with the liquid wax (it dissolves in the wax), creating a single continuous phase. In the other scenarios phase separation occurs between wax and liquid water. This could result in a continuous phase of wax covering the pore walls, enveloping a continuous water phase (scenario b), or vice versa (scenario c). Alternatively, there may be “alternating phases”, i.e. alternating pore sections filled with water and with wax (scenario d)). Finally, a (micro)emulsion of water droplets in wax (scenario e)) or vice versa (scenario f)) could occur. In scenarios e) and f), the droplets may or may not be interconnected.

Mixtures of the scenarios may occur, and the scenario may depend on e.g. pore diameter and distance to the surface of the pellet. Nonetheless, these scenarios offer a lead for further research, and can be tested numerically and experimentally. For example, scenario c) (water on pore walls) seems unlikely, because it would lead to much faster catalyst degradation than observed in situ in reactors.

Following these arguments, the project is split into two sub-tasks:

- Perform a theoretical analysis based on literature data and concepts known from classical physics to answer the following questions:
  - What is the phase of pure water in GTL reaction conditions?
  - Does all water dissolve in wax?
  - Are water bubbles/droplets stable within the wax?
  - Are water bubbles/droplets energetically more favourable than a supersaturated mixture?

- Design experiments to validate the theoretical analysis.
5. Theoretical analysis

5.1 The phase of pure water in GTL reaction conditions

For our calculations below, we will assume that in the GTL plant, a 50 bar mixture of CO (33 bar) and $H_2$ (17 bar) is pumped into the top of the reactor vessels. Once the gas mixture has reached the bottom of the reactor, up to 70% of it would be converted to the reaction products $H_2O$ and $C_{60}H_{122}$. This means that the partial pressure of water in the lower part of the reactor would be around 12 bar. The temperature in the reactor is between 475 and 525K. At these conditions, bulk water exists only in the vapor phase [1].

However, due to the presence of the nanoporous pellets, we should also consider whether capillary condensation of water in nanoscale cavities is possible. The Kelvin equation [2] predicts the equilibrium radius of a condensed droplet for a fixed temperature $T$:

$$k_B T \ln \frac{p}{p_{sat}} = \frac{2 \gamma \nu}{R_K}$$  \hspace{1cm} (1)

Here, $p$ is the vapor pressure, $p_{sat}$ is the saturation vapor pressure, $\gamma$ is surface tension, $\nu$ is molecular volume, and $R_K$ is stable droplet radius (actually the radius of curvature of the water interface). The surface tension of water [3] is 36.6 mN/m at 478K. The saturation vapor pressure at that temperature [4] is $p_{sat} = 19.3$ bar. For a vapor pressure of $p = 12$ bar, the equilibrium radius of a condensed droplet is -0.5 nm. Since this corresponds to only 3 atomic distances, we can safely assume that, in the absence of wax, water will not condense even in the smallest cavities present.

5.2 Does water dissolve in wax?

In order to study the behaviour of water first it has to be understood whether or not the water is soluble in the wax. Therefore the questions we will try to answer is whether it is possible for the concentration of water in the wax to exceed the solubility limit during the Fischer-Tropsch (FT) process? In this section we will calculate the potential water concentration profiles as a function of system parameters and with that try to estimate whether or not water will ever exceed the solubility limit under the reactor conditions.

In order to get a qualitative feeling for the concentration of the FT reactants and products, we will model the reaction taking place in pores of millimetre size pellets using a one dimensional reaction-diffusion model. This model is just an extension to the Fick’s second law of diffusion with an additional rate of reaction, which is given by

$$\partial_t c_i = D_i \partial_{xx} c_i + R_i,$$

where $c_i$ stands for the concentration in units mol dm$^{-3}$, $D_i$ is the diffusion constant, and $R_i$ is the rate of production or consumption of species $i$.

From empirical evidence as well as mathematical modelling [9], it is known that the diffusion coefficient of carbon monoxide CO, is a factor of three smaller than that of molecular hydrogen $H_2$. The rate of the reaction equation is approximately first order in the concentration of hydrogen and zeroth order in the concentration of carbon monoxide (for sufficiently high concentrations of both reactants)

$$rate_{CO \text{ loss}} = k C_{H_2}.$$  \hspace{1cm} (2)

This rate equation simply expresses the fact that the binding of carbon monoxide to the catalytic sites proceeds at a much faster rate than the reaction rate between the bound carbon monoxide and the molecular hydrogen. For low enough concentrations, however, the reaction rate takes the form

$$rate_{CO \text{ loss}} = \frac{kaC_{H_2}C_{CO}}{1 + aC_{CO}},$$
which shows that the depletion of carbon monoxide inside the pore terminates the progress of the reaction, however for large carbon monoxide concentrations we get the behaviour as given by 2. Here \( n \) is simply a fitting parameter.

The equation above is rather difficult to implement in the reaction-diffusion equation due to highly sensitive behaviour of the rate function, which easily leads to physically unrealistic results. For this reason we are assuming that the molar ratio of the gas present in the reactor is adjusted such that the carbon monoxide reaches further into the pellet pore, and that the concentration of molecular hydrogen is still the rate determining factor. This model is mainly meant to present an order of magnitude analysis of the concentrations, and since the diffusion rates all the gases differ by less than one order of magnitude, our assumption should be a safe one to make.

In this way, following the stoichiometric equation for alkane chains of about 60 carbon atoms in steady state \( \partial_t c_i = 0 \), we get

\[
D_{H_2} \partial_{xx} c_{H_2} - 2k c_{H_2} = 0, \\
D_{CO} \partial_{xx} c_{CO} - k c_{H_2} = 0, \\
D_{H_2O} \partial_{xx} c_{H_2O} + k c_{H_2} = 0,
\]

which with the boundary conditions of \( c_i(x=0) = c_{0,i} \) and \( (\partial_x c_i)_L = 0 \) \( (L \) being the depth of the pore) for \( i = CO, H_2O \) and \( H_2 \), give the following solutions

\[
c_{CO} = c_{0,CO} + c_{0,H_2} \frac{D_{H_2}}{2D_{CO}} \left( \text{sech} \left( \frac{L}{\xi} \right) \cosh \left( \frac{L - x}{\xi} \right) - 1 \right), \\
c_{H_2} = c_{0,H_2} \text{sech} \left( \frac{L}{\xi} \right) \cosh \left( \frac{L - x}{\xi} \right), \\
c_{H_2O} = c_{0,H_2O} - c_{0,H_2} \frac{D_{H_2}}{2D_{H_2O}} \left( \text{sech} \left( \frac{L}{\xi} \right) \cosh \left( \frac{L - x}{\xi} \right) - 1 \right), \tag{3}
\]

where the natural length scale is \( \xi = \sqrt{\frac{D_{H_2}}{2k}} \). Here everything is working under the assumption that the reaction is taking place with the same rate constant homogeneously throughout the pore, i.e. \( \partial_x k = 0 \). The profiles above are shown in Figure 2.

From this model we can calculate the maximum water concentration inside the pore \( (x = L) \), giving

\[
c_{\text{max}, H_2O} = c_{0,H_2O} + c_{0,H_2} \frac{D_{H_2}}{2D_{H_2O}} \left( 1 - \text{sech} \left( \frac{L}{\xi} \right) \right),
\]

where we will check whether the concentration of water is greater than the critical concentration of water that can dissolve in wax under the reactor conditions.

We estimate the carbon monoxide and hydrogen concentration at the entrance to the pore by calculating the concentration of wax present there. Since the density of wax is approximately 850 g dm\(^{-3}\), and the molar mass of a hexacontane is about 840 g mol\(^{-1}\), then the concentration of wax is about 1mol dm\(^{-3}\). Additionally, at the reactor conditions (assumed to be 500 K and 50 bar) the hydrogen and carbon monoxide molar fractions are about 0.10 [9], then the concentration of molecular hydrogen at the entrance to the pore is of the order 0.1mol dm\(^{-3}\).

From the wax production data received with the assignment \( (100-1000 \text{ g of wax per decimeter-cubed of catalyst per hour (g dm}^{-3}\text{ cat hr}^{-1}) \), equivalent to \( \approx 1.6 \times 10^{-4} \text{mol (dm}^{-3}\text{ s}^{-1}) \), we can get that the rate of water production (equal to the rate of carbon monoxide consumption), which is of the order 10\(^{-2}\)mol (dm\(^{-3}\) s\(^{-1}\). From this we obtain the rate constant \( k = \frac{\text{rate of water production}}{c_{\text{H}_2O}} = 10^{-1}\text{s}^{-1} \), giving the characteristic length of \( \xi \approx 400\mu\text{m} \). For a pellet of 1 mm size, the possibly deepest pore is of the order of hundreds of
Figure 2: The profiles (solutions to 3) of carbon monoxide, water and hydrogen concentrations inside a pore.

micrometers, putting us in a regime of $L/\xi \approx 1$, with the maximum water concentration of

$$c_{\text{max}, \text{H}_2\text{O}} = c_{0, \text{H}_2\text{O}} + c_{0, \text{H}_2}(1 - \text{sech}1) = c_{0, \text{H}_2\text{O}} + 3.5 \times 10^{-2}\text{mol dm}^{-3}.$$

A typical value for the critical solubility of water in hydrocarbons of about sixty atoms (a boiling point of about 600 K) is of the order of $\approx 60 \text{ppm g} / \text{g} \approx 10^{-3}\text{mol dm}^{-3}$. This an order of magnitude lower than the water concentration present inside the pellet.

5.3 The stability of water vapor bubbles in the wax

Here we try to answer the question whether water vapor bubbles can exist within the liquid wax. The pressure drop $\Delta p$ that is associated with the crossing of the interface between water vapor and wax is given by the Young-Laplace equation:

$$\Delta p = -\gamma \hat{V} \cdot \hat{n}. \quad (4)$$

The tendency to minimize this wall tension leads to bubbles of spherical shape (Laplace’s law). Then, the above equation simplifies to

$$\Delta p = -\gamma \frac{2}{R}. \quad (5)$$

For a water vapor bubble inside hexacontane at 475K and $p_o = 50\text{bar}$, we estimated a surface tension $\gamma_{\text{hexacontane}} = 20\text{mN m}^{-1}$. The pressure inside the bubble is then

$p_i = p_o + \gamma_{\text{hexacontane}}(2/R)$, which for a radius of $R = 25\text{nm}$ corresponds to $p_i = 66\text{bar}$. At these conditions, bulk water will be in the liquid phase. We conclude that water vapor bubbles are unstable inside the wax in the nanopores.

5.4 The nucleation of liquid water droplets

We have concluded in section 5.2 that a mixture of water and hexacontane inside the pore would have a supersaturation factor of $W/W_{\text{max}} > 10$. We have also shown that the formation of water vapor bubbles inside the wax will be highly suppressed, since the water vapor phase is not stable at the reaction conditions. Now, we consider the nucleation of liquid water droplets inside the supersaturated state. We ask ourselves whether the formation of droplets
with a radius that is smaller than the nanopore radius is energetically favourable over the
supersaturated mixture state. For the sake of simplicity, we consider the classical theory of homogeneous nucleation [6],
which gives the free energy $F$ of a liquid nucleus inside a vapor:

$$F = \gamma A - \rho_l * V * k_B T \ln \frac{W}{W_{max}}$$

(6)

Here, $\gamma$ is the surface tension of the nucleus, $A$ is its surface area, $\rho_l$ is the molecular density
inside the droplet, $V$ is the volume, $W$ is the molar concentration in the vapor, and $W_{max}$ is the
molar saturation concentration. Basically, this formula tells us that the formation of a droplet
costs an energy $\gamma A$ to create the interface between the liquid and the vapor phase, but because
of the increase of entropy, we also gain a free energy that scales with the volume of the droplet.
The free energy has a maximum at the critical radius $R^*$:

$$R^* = \frac{2\gamma}{\rho_l k_B T \ln \left( \frac{W}{W_{max}} \right)}$$

(7)

Droplets that occur in the mixture are stable, and continue to grow, if they have a radius
$R > R^*$. Droplets with smaller radii will not be stable and dissolve.

In order to estimate the stability of water droplets in hexacontane under GTL reactor
conditions, we assume the water that is solved in the hexacontane to be a vapor. Furthermore,
we disregard the free energy we have to pay in order to remove a volume $V$ of hexacontane at
the position where a droplet is formed, since the molar concentration of hexacontane is 50
times smaller than that of water. It should be noted that in this assumption, we also disregard
the higher entropy per molecule of the hexacontane with respect to the water. For the interface
tension between water and hexacontane we estimated a value of $\gamma_{wax-water} = 50 \text{ mN/m}$, based
on values for shorter alkanes [7].

![Figure 3](image)

*Figure 3*: Free energy of nucleation versus nucleus radius for water in hexacontane at GTL
reactor conditions. Note that the functions decrease quickly towards negative values for radii
$R > R^*$. Because of the logarithmic scaling on the vertical axis, these negative values are not
plotted.

In Fig. 3 we have plotted the free energy versus radius of a water nucleus inside the
hexacontane, for several values of the supersaturation factor. For the lower estimation of
$\frac{W}{W_{max}} = 10$, we obtain a critical radius of 2pm, which is much lower than the atomic spacing
between water molecules. Since the supersaturation in reality is even larger, we conclude that the water droplets can easily be formed.

The mass production rate of water and wax is almost equal, and these two have nearly the same mass density. Therefore, during the process, we have approximately equal volume production of wax and water in the pore. A possible scenario for the demixing inside the pores is that we have alternating layers of the wax and water, perpendicular to the pore wall, with almost equal volume, which move towards the exit point of the pore. This scenario is illustrated in Fig. 4 for two cases: 1) water is wetting on the Alumina pore wall, or 2) water is non-wetting on the pore wall.

![Figure 4](image)

**Figure 4**: Two possible scenarios for the demixed water-hexacontane structure inside the alumina nanopores: Water is non-wetting on the alumina (top), and water is wetting on the alumina (bottom).

Observations on alumina that is used inside the GTL reactor are reported to show that the surface tends to become non-wetting for water during the reaction. This could be due to the adsorption of hydroxides on the pore walls. Non-wetting of water would lead to a layer of hexacontane on the pore walls, which is in agreement with the observation that the catalyst particles, in the steady state of the process, are not oxidizing fastly as they seem to do when in contact with bulk liquid water.

### 6. Proposed experiments

This section is devoted to experimental methods to probe the physical situation within the porous medium where the Fischer-Tropsch (FT) reaction takes place (see the phase scenarios in Figure 1). In the preceding sections it was clear that some bulk physical parameters are necessary to do modelling of the system in order to discriminate between various possible scenarios. Additionally, some model systems are proposed that mimic the inside of the pellet but offer more control of experimental conditions, giving a view on the non-equilibrium steady state present near catalyst sites.

This section is organised as follows. We first list the crucial bulk parameters needed to get a basic model of the system. Then we describe different (model-) systems for later reference when we discuss various experimental techniques.
Proceedings Physics with Industry 2012

Figure 5: (Model) systems. (a) pressure cell, (b) array of nanochannels, (c) single pellet in sample cell, (d) reactor system: a tube filled with pellets.

Bulk parameters are associated with the equilibrium properties of the system. In the non-equilibrium system (active reactor), these parameters play a crucial role in the understanding of and discrimination between different non-equilibrium processes. The parameters can be measured in absence of the FT reaction (but at reactor pressure and temperature) and without the complications of measuring inside the porous medium. For simplicity we refer to the porous medium as Al₂O₃.

- **Surface tensions** between H₂O/Wax and Al₂O₃/H₂O and Wax/Al₂O₃. Surface tensions should be measured both for liquid water and for water vapour. These surface tensions determine the nucleation radius of water droplets within the wax and determine the contact angles and wetting phenomena that determine if the Al₂O₃ surface is covered with water or wax. The Al₂O₃ may (in time) become covered with a layer of alcohols and carboxy-acid FT by-products, which will strongly affect the surface tensions, which should therefore also be measured in this Al₂O₃ state.

- The **Solubility** of H₂O in Wax, and for detailed analysis also CO and H₂ in liquid H₂O and Wax.

- **Diffusion Coefficients** of dissolved H₂O, CO and H₂ in wax, and CO and H₂ in liquid H₂O.

We propose several (model) systems for the experiments in the next section, see Fig. 5.

- Bulk wax and water, e.g. in a pressure cell, to measure the equilibrium properties described above under reactor conditions (Fig. 5a).

- Nano-channels: an assembly of parallel nanochannels, e.g. by etching channels in a glass plate, coating it with a thin layer of Al₂O₃ and covering the glass plate with a second plate (Fig. 5b).

- Single pellet in sample (flow) cell under reactor conditions (Fig. 5c).

- Reactor system: a tube filled with pellets, under reaction conditions (Fig. 5d). This system is the most heterogeneous, and therefore least simple to experiment on. However, it may yield the most representative information for what happens in situ in a reactor. Moreover, this system contains a large amount of all material, which may be crucial for experiments with weak signals, such as electron paramagnetic resonance (see below).

### 6.1 Methods

In this section we list possible approaches to probe the relevant physics inside the pellets. Note that we focused on providing a range of creative solutions, that have to be worked out in future work.
Figure 6: (a) Proposed pressure-jump setup, to study non-equilibrium processes, such as bubble/droplet formation. (a) Wax becomes oversaturated with H$_2$O by moving the piston up (releasing vapour pressure). (b) Injection of H$_2$O in the setup mimics the continuous generation of H$_2$O by the FT reaction.

Pressure jump experiments

Two methods of creating and studying water bubbles/droplets in liquid wax in a controlled way is illustrated in Fig. 6a. A pressurized cylinder filled with liquid wax is in equilibrium with the water vapor above it, at pressure $p_i$, where the amount of water dissolved in the wax is determined by the Henry constant. Releasing the piston (or a valve) to reduce the pressure to $p_f$ ($p_f < p_i$), will lead to over saturation of water in the wax. Consequently water will move out of the wax, through diffusion, bubble formation and/or droplet formation. The water movement could be measured with a range of methods (imaging, scattering, spectroscopy), provided the cylinder is transparent. This would yield information on several important bubble/droplet properties: the location and number of nucleation sites, the rate of nucleation, speed of movement along the surface/through bulk wax, number of bubbles/droplets and their size distribution. This measurement system enables studying the individual effect of several “actors” which all act simultaneously in the pellets, such as the extent of oversaturation, the nature of the pore surface ($\text{Al}_2\text{O}_3$, or $\text{Al}_2\text{O}_3$ covered with FT (by)products), and the presence of syngas at different concentrations. An extension of the system includes a nozzle in the cylinder wall (Fig. 6b) to inject water vapor at a controlled rate (mimicking water production at the catalytic sites). This may offer a way of measuring a critical injection rate leading to liquid water phase in the wax. Adding catalysts is another way of producing water in the cylinder, but this has the disadvantage that the effect of CO and H$_2$ concentrations cannot be varied independently from the water production rate.

$^1$H-NMR diffusion measurements

The $T_2$ relaxation time of the protons of water in small droplets is determined mainly by the average time for a water molecule to diffuse to the interface with the wax. Thus $T_2$ measurements can be used to estimate droplet radii, provided the diffusion coefficient of water
is known. The droplet size inside nanopores may be too small (too short $T_2$) to be detectable by standard NMR $T_2$ methods [12]. Before experimenting, it is therefore recommended to estimate the minimal detectable droplet radius, which depends on the time-resolution of the NMR apparatus. Dedicated pulse sequences may increase the radius/time resolution [13]. Detection of bubbles of water vapor is probably much more difficult, due to the low proton density and the higher diffusion coefficient than liquid water, which increases the minimal detectable radius. The ferromagnetic cobalt catalyst will cause inhomogeneities of the magnetic field, leading to fast $T_2$ relaxation. For low catalyst densities in the pellets this may not be a problem; it will render the water molecules in close proximity of the catalysts undetectable, without effecting the $T_2$ of the water at longer distances. For high catalyst densities the cobalt catalyst has to be replaced by a paramagnetic ruthenium catalyst, although this may render the experiment less representative for what happens in an active reactor.

Dielectric constant

Dielectric spectroscopy measures the complex relative permittivity for a wide range of frequencies from Hz - THz. The complex relative permittivity is given as $\varepsilon = \varepsilon' - i\kappa/\varepsilon_0\omega$ where $\varepsilon'$ is the relative permittivity, $\kappa$ is the conductivity, $\varepsilon_0$ is the vacuum permittivity and $\omega$ the angular frequency. The real part of the complex relative permittivity $\varepsilon'$ will show up as a phase shift in dielectric spectroscopy and the imaginary part is associated to attenuation of the signal. The dielectric constant of single-phase systems depends on its molecular properties. However, for an emulsion consisting of water and oil with complex dielectric constants $\varepsilon_w$ and $\varepsilon_o$, there will be Maxwell-Wagner-Sillars polarization at the boundary between the two media. This opens the possibility to study the properties of the emulsion [10]. The catalyst particles will probably interfere strongly with the dielectric measurements. But one could study a single pellet or the nano-channel setup by appropriate subtraction of the background signal.

Markers

The presence of liquid water at the pore surface could be probed by (sparsely) attaching a “marker” on the pore surfaces. When the marker is in contact with liquid water it should give a lasting response, but not in the presence of water vapor or dissolved water. An advantage of this approach is that the response can be measured offline, i.e. post-FT-reaction, outside the reactor and under ambient conditions. This greatly increases the number of techniques to probe the response, e.g. examination of pellet slices with electron microscopy. It may be necessary to use secondary markers such as such OsO$_4$, to selectively label the response area. We propose two categories of markers:

- **Conditionally mobile markers** are only mobile in the presence of liquid water. Suggested markers are (i) small water soluble salt crystals, (ii) compounds that strongly binds to the pore wall via multiple hydrogen bonds (H$_2$O-sensitive covalent bonds with the walls are unsuited, because they are probably also sensitive to dissolved H$_2$O), and (iii) cobalt particles that are too small to be active catalysts (such particles may even form fresh active catalytic sites via Ostwald ripening, providing a means of replenishing degraded catalyst particles). The presence of liquid water would lead to detectable spatial redistribution of the markers, starting from either a spatially homogeneous distribution of markers, or from an inhomogeneous distribution.

- **Immobile markers** lead to a local (and lasting) response in the presence of liquid water, e.g. chemical/physical changes of the pore wall or of the marker itself.

This quality of the marker makes or breaks this type of experiments. A good marker has to meet several requirements:
• **Stable** under the reactor conditions

• **Non-interfering** with the FT reaction and movement of reactants and reaction products. Interference could be chemical but also physical (e.g. steric hindrance). This requirement probably implies using low labelling densities. In view of the offline measurement of the response this seems to pose little limitation. Interference should be checked by comparing the FT reaction rate in presence and absence of the markers.

• **High selectivity.** The markers should have high contrast ratios for response to liquid water and the other phases/compounds present in the pores. Selectivity could be checked on model systems, e.g. Fig. 5a,b or on a flat surface of Al$_2$O$_3$ covered with markers.

• **Suitable sensitivity.** The sensitivity should be such that the inverse response rate is of similar magnitude as the typical exposure time of the marker to liquid water.

The markers described above have the advantage of measuring offline. Ideally a response would be measured in situ, in an active reactor. This limits the measurements techniques, and therefore the marker types. One measurement that may work is the detection of spin labels detected by EPR (electron paramagnetic resonance). The line shape of the EPR spectrum is sensitive to the mobility of the label, which depends on its local environment. It may therefore be used to distinguish between liquid water/liquid wax/water vapor. EPR spectroscopy has been demonstrated on substrates bound to porous Al$_2$O$_3$, albeit at ambient temperature [14]. In the ideal case spin probes could be attached to the pore surfaces via different lengths of stiff linkers, and consequently probe the local environment at different distances from the pore surface (distinguishing e.g. between Figure 1b and c). The EPR marker-response experiment is particularly challenging, because:

• **Unpaired electrons** of the spin labels, may be chemically active under reactor conditions.

• **Labelling density** required for sufficiently strong EPR signal may be prohibitively high and interfering with the FT process.

• **Cobalt catalyst** may interfere with EPR experiments, as also expected for the NMR experiments, in which case it should be replaced by ruthenium.

**Isotopes**

The average time H$_2$O molecules spend within the pellets depends on their mobility. This mobility will differ significantly for H$_2$O dissolved in wax, a water-in-oil emulsion and a percolated network of water. In order to measure the average dwelling time of H$_2$O molecules within a pellet we propose the following methodology. For water dissolved in wax and percolated water networks the mobility will be determined by the diffusion constants of water-in-wax and water-in-water respectively. For the emulsion the droplet size will play a role since larger droplets will be less mobile than smaller droplets via Stokes’ law. We propose the following methodology. Inject syngas with one isotope taking the place of hydrogen, oxygen or carbon. The syngas will diffuse into the pellet and the FT-reaction will form H$_2$O and Wax containing the isotope we injected. For simplicity we assume the FT reaction is insensitive to the isotopes. After the reaction all carbon isotopes will end up in the wax, all oxygen isotopes will end up in the water and deuterium the isotope of hydrogen will end up 50% in wax and 50% in water. In a time-resolved measurement of the concentration of isotopes ejected from the reactor different timescales will show up, see Fig. 7. There will be a timescale associated to the diffusion of syngas through the reactor. This will cause the pulse of isotope concentration released to be detected slightly later and a broadened with respect to the injected pulse of isotope concentration, even in the absence of the FT
Figure 7: Sketch of the isotope concentration as a function of time. At $t = t_0$ a concentration of an isotope is injected in the reactor. After drift and diffusion through the reactor the isotope appears at the other side of the reactor. The pulse has slightly broadened due to diffusion (top left figure). The other subfigures show the decline of the various isotope concentrations ejected from the reactor and the associated timescales due to different mobilities associated to the reaction products that the isotopes become.

reaction. The lower mobility of the wax will result in a longer timescale for the isotopically marked carbon atoms than the timescale for marked oxygen which will end up in the water. Using various isotopes to mark the carbon, hydrogen and oxygen one would in principle be able to disentangle different diffusion coefficients for the syngas entering the pellet and may even detect more timescales associated to the ejection of water from the pellets which would hint at combined diffusion by some dissolved single molecules and droplets of liquid water.

**Opto-Acoustic**

We propose to use a pulsed laser beam to heat the liquid water generated in the porous structure. The proposed methodology is used in medical diagnostics and is capable of determining microemulsions in biological tissue[15]. The water will show a thermoelastic expansion and pressure waves will be emitted giving an acoustic signal. The relevant physics is summarized in the equation

$$\left( v^2 - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \right) p(x, t) = -\beta \frac{\partial H}{C_p \partial t},$$

where $p(x, t)$ is the pressure field, $H$ is the heat deposited in the medium per unit volume and time, $v$ is the speed of sound in the medium, $\beta$ is the thermal expansion coefficient and $C_p$ is the heat capacity at constant pressure. The propagation of acoustic waves depend on droplet size and bulk properties of the wax and water[11] and can be detected via ultrasound transducers. We think this approach is feasible due to the large absorption of IR radiation by liquid water at a wavelength 2.9µm. The wax does not absorb much radiation in the infra-red and also the $\text{Al}_2\text{O}_3$-pellet scatters the light but does not absorbs the radiation so heating of the $\text{Al}_2\text{O}_3$ is limited.

Moreover, even if all components in the pellets would heat due to the applied laser power, the thermoelastic response of liquid water droplets probably is much stronger than the expansion
of the wax. Providing us with a clear signal.

**Diffusion of inert Gas**

The diffusion rate of an inert gas through a pore filled with the FT mixture offers a way to probe the presence of liquid water and water bubbles, by measuring the transition time of a brief pulse of inert gas from one end of the pore to another. For an inert gas with identical diffusion coefficients in liquid water and in wax, the diffusion time through the pore depends on the fraction of water bubbles. In the absence of water vapor bubbles, an inert gas should be used with different diffusion coefficients in liquid water and liquid wax. The diffusion time of such an inert gas through the pore will depend on the fraction of liquid water. For both types of inert gas, the (distribution of) transition times depend not only on the volume fraction of bubbles/droplets, but also on their mesoscopic “structure” and size: An interconnected network of bubbles/droplets may offer “highways” for the inert gas movement, resulting in short transition times. A large number of small bubbles/droplets may result in different diffusion times than the same volume fraction of larger bubbles/droplets. The net results of these effects unknown a priori, but can be assessed through simple 1 or 2 dimensional diffusion models.

**Bubble exit from pores**

![Figure 8: Scenario for a bubble/droplet (blue) exiting a pore by breaking through the film of wax that covers the pellet (yellow).](image)

The presence of water bubbles/droplets may be assessed from their exit from the pores. In a reactor the outer surface of each pellet is covered by downward flowing wax. Water exiting a pore has to travel through this file of wax. This can occur by (i) diffusion through the wax, (ii) a vapor bubble erupting at the surface, and (iii) a liquid droplet arriving at the film, where it evaporates due to the low water vapor pressure outside the pellet. Diffusion will leave the surface shape of the film unaffected. By contrast, bubbles and droplets will lead to a deformation of the film, followed by rupture (“popping”) of the film (Fig. 8). Several experiments can distinguish between these two cases.

- **Shape distortion** measured by changes in light scattering properties of the film or by changed reflection/deflection of laser beam by the wax film.
- **“Popping of bubbles”** measured from the resulting acoustic wave.

These types of experiments rely on detecting single droplet/bubble exit events (detecting multiple events simultaneously results in a smeared out signal, which is expected be very similar to that of respect to exit via diffusion). Single exit event detection requires a
time-resolution better than the average time between two events. It may be necessary to reduce the rate of exit events by measuring on single (or small numbers of) pellets, or by reducing the number of pores per pellet. Optical methods probably yield higher time-resolution than acoustic methods, because acoustic waves may experience more dispersion.

**Quench**

In the preceding sections we showed that we expect liquid water and wax to fill the pores of the pellets. Nucleated water droplets would coalesce and could form percolation networks of liquid water inside the pores. If this theory is valid, one way to observe these networks or droplets would be by quenching the reaction in such a way that the water or wax cannot redistribute during and after the quench. We thought of two possible quenching techniques which we list below. After quenching the structure of the percolated network of wax would be visible using for example electron microscopy of a slice of the pellet. The quenching possibilities are

1. *Freeze* the pellet by cooling it fast to a temperature at which the wax solidifies. A problem with this approach is the liquid water in the pores would expand and the structure after the quench would not be representative for the structure under operating conditions.

2. *Irradiate* the sample such that the alkanes in the wax cross-link and form a strong network that fixes the wax. As illustrated in Fig. 9.

6.2 Suitability of experiments for model systems

To summarize we made a table showing the viability of the sketched methods for the different model systems. The Pressure jump experiments need to be done in the cylinder-piston geometry. For the other methods we refer to table 1.

7. Conclusions

The aim of this study was to understand the physical state of water in porous catalyst particles during a Fischer-Tropsch reaction. Does water mix with the wax produced in this reaction or
Table 1: Feasability of the methods in the different model systems.

<table>
<thead>
<tr>
<th>experiment</th>
<th>nano-channels</th>
<th>single pellet</th>
<th>minireactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1H-NMR diffusion measurements</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Markers</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Diffusion of inert gas</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Bubble exit from pores</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Dielectric constant</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>Isotopes</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Thermo-Acoustic</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>Quench</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
</tr>
</tbody>
</table>

does a second phase develop? And if a second phase develops, what is the nature of this phase: a vapor bubble or a water droplet. By analysing the 1D reaction-diffusion equations it was shown that in the operation window of the reactor water always exceeds the solubility limit. So, the formation of a second water rich phase seems plausible. At this point it has to be remarked that there is little quantitative data on the mixing behaviour of water and wax at the pressures and temperatures of interest. On the basis of thermodynamic analysis it was proven that water will not form gas bubbles, but preferentially nucleates in droplets. At the given process conditions nucleation of these droplets is possible within the porous matrix of the catalyst. In order to come up with more reliable prediction of droplet formation there should be more quantitative data on the surface tension of the wax/water interface at the processing conditions. Further, it seems that the formed droplets are more stable in the core of the pellets than at close to the external surface of the pellet. However, the stability of a water droplet in relation to its location in the pellet deserves more attention. The same is true for the motion and aggregation of the droplets inside the pore system. The conclusion that water does not mix with wax and forms a separate (liquid) phase seems to be contradictory with experiences with respect to the life-time of the Co-catalyst. In general Co-particles quickly deteriorate in water at the given temperatures and pressures, but this does not happen in the so fast in the reactor. Possible explanation could be that the wax that grows directly on top of a Co-particle actually protect the particle from attacks by liquid water.

An important conclusion of the presented study is also that the theoretical analysis has to be built on a small amount of experimental data of unknown quality. The problem is that there is little information on the water/wax phase behaviour and the surface properties of the water/wax interface at high temperatures and pressure. Therefore, it is concluded that proper validation of the theoretical concepts discussed in this paper is only possible when more experimental data is collected. High quality experimental data of the phase behaviour of water/wax are the key to breakthrough in our understanding of the wax/water distribution in the porous catalyst. To validate the idea of droplet formation an experiment is proposed to monitor the birth and growth of droplets at the catalyst particles during the reaction. Further various techniques are proposed that are capable of identifying liquid water in situ in a Fischer-Tropsch reactor.

In summary, it has to be concluded that the presented analysis points to the presence of a co-continuous network of wax and liquid water in the porous pellet. Key to progress is high quality data on the wax/water phase behaviour.

8. Outlook

As an outlook we would like to propose PhD projects that could be carried out by a prospective student at a physics or chemistry department. This research projects were thought up based on the conclusion of the research conducted during the workshop week.
One of the possible PhD project could involve building a more accurate mathematical model which would attempt to describe the local dynamics of composition fluctuations inside the pore. On the scale of tenths or hundreds of nanometers, one could study the creation and dynamics of two phases (water and wax, perhaps even with its composition modelled by the Anderson-Schulz-Flory distribution). This model could analyse whether one would indeed see phase separation, with nucleation of water droplets. One could also include in said model spacial differences in catalytic synthesis conditions which give rise to different product formation.

First this will be studied on pore-to-pore basis. Secondly, the outcomes will be scaled-up to the pellet and reactor level to be able to calculate experimentally as well as industrially measurable effects on the FT synthesis.

A second possible PhD project involves to perform some of the experiments which were proposed during the workshop. This is will allow to collect high quality data on the surface tension between wax and water, the solubility of water in wax, or the corresponding diffusion coefficients. These findings could then be used by the first PhD to improve the theoretical models.

The third PhD project we propose involves computer simulations. By using molecular dynamics together with simplified models for the chemical reaction itself, the transport of reaction products and their phase behaviour can be studied on the single pore level. In order to upscale the obtained results, we propose to combine the microscopic simulations with continuum multiphase fluid dynamics simulations to understand the transport of water and wax on the pellet- or reactor level.

9. References


List of participants

Mireia Alos palop TUD  Xander Janssen TUD
Jildou Baarsma UU  Fateme Kako Joibari TUD
Marcel Bartels FOM  Akashdeep Kamra TUD
Sander van Bavel Shell  Femius Koenderink AMOLF
Felipe Bernal Arango AMOLF  Herman Kuipers Shell
Erik van der Bijl UU  Lucas Kunneman TUD
René Bos Shell  Karel van Laarhoven TU/e
Evert Bosdriesz VU  Kees Landheer UU
Job Bouwman UU  Theo van der Leij PamGene
Marcel ter Brake UT  Roy Leyte-Gonzalez TU/e
Michele Buscema TUD  Abbas Mohtashami AMOLF
Victor Calado TUD  Willem Onderwaater LEI
Toon Coenen AMOLF  Bart van Oort VU
Bram Colijn UT  Gonzalo Rios Cruellas UT
Matthijs Cox TU/e  Wouter Roos VU
Lourens van Dijk UU  Upasana Roy LEI
Michiel de Dood LEI  Radu Surdeanu NXP
Marcin Dukalski TUD  Srinivas Vanapalli UT
Matthieu Fleys Shell  Pim Veldhuizen UU
Marjan Fretz STW  Gerard Verbiest LEI
Hans Geerlings TUD  Bram Verhaagen UT
Lisa Hartgring Nikhef  Willem van de Water TUD
Jens Harting TU/e  Jan Weenink UU
Maryam Hashemi-Shabestari LEI  Geert Wijts LEI
Yusuke Hashimoto RU  Barbara Witek TUD
Tjalle Hoekstra VU  Gijs Wuite VU
Henk Huinink TU/e  Hai Feng Yuan LEI
Colin Ingham Microdish  Shi Qi Zhang TU/e
Martijn de Jager FOM  Marco Zocca TUD
Huub Janssen JPE