Broadband Photon Time-of-Flight Spectroscopy on Pharmaceutical Tablets and Dairy Products

Sören Johansson
Supervisors: Dmitry Khoptyar, Stefan Andersson-Engels

Master’s Thesis
HT 2014

Lund University
Physics Department
Broadband Photon Time-of-Flight Spectroscopy on Pharmaceutical Tablets and Dairy Products

© 2015 Sören Johansson

Division of Atomic Physics
Department of Physics
Lund University
P.O. Box 118
SE-221 00 Lund
Sweden

http://www.atomic.physics.lu.se
http://www.atomic.physics.lu.se/biophotonics

Lund Reports on Atomic Physics, LRAP-497
In this thesis, Photon Time-of-Flight Spectroscopy (PTOFS) is presented as a method to measure absorption and scattering spectra of turbid media samples in the near infra red wavelength range. In particular, absorption and scattering spectra measured on pharmaceutical tablets and dairy products are used to analyse the composition of the samples.

In PTOFS, a temporally short pulse of a specific wavelength is sent through the sample. In the sample photons of the pulse are scattered and absorbed and thus the temporal shape of the pulse changes. This changed pulse shape is recorded in high temporal resolution and a model of light propagation in turbid media is fitted to the pulse shape and returns the absorption and reduced scattering coefficient. This is repeated for all wavelengths of interest in order to create a spectrum.

During the work on this thesis project, the experimental setup has been improved by completing an automated measurement software. A new and more powerful super continuum light source and a new spectral filter with superior spectral resolution have been implemented. It is shown, that the new experimental setup is beneficial for the quality of the measured data.

For the study on pharmaceutical tablets, the newly measured data of 70 tablets is compared to measurements on the same tablets with an older version of this setup and to data taken with Transmission Near Infra Red Spectroscopy. The results show, that the new data has less coupling between absorption and scattering. With these data it is therefore possible to create a more robust and precise compositional analysis with chemometric techniques.

The measured spectra of dairy products and dilutions of cream allow an approximate fat content prediction by a fit of the absorption spectra as a linear combination of the ingredients. It is also shown, that the fat content is related to the concentration of scatterers that has been retrieved from a fit as a power law of the scattering spectra. These fits also provide information about the particle size of the scatterers in the dairy products.
Popular science abstract

Optical ingredient analysis of tablets and dairy products

Pharmaceutical tablets and dairy products, such as milk or yoghurt have in common, that they have more than one chemical component. They also appear white and are turbid media. With the Photon Time-of-flight Spectroscopy (PTOFS) system at the Biophotonics group of Lund University, it is possible to get information about the composition of such turbid media and thus to recover the part of the ingredients of tablets and dairy products.

Scattering and absorption are physical effects that appear when light interacts with matter. Scattering randomly changes the direction of the light propagation and absorption diminishes the intensity of the light. Both of these effects are depending on the wavelength, the color of the light. Examples for scattering are the blue sky, clouds or foggy weather, while absorption can be observed in the different colors of e.g. red wine and apple juice. Absorption is connected to the chemical composition of the sample, while scattering is related to structural information.

In PTOFS a temporally very short pulse of light is sent through the sample. This pulse contains a high number of light particles (photons), that each are scattered and absorbed individually in the sample. As these processes occur by chance, some photons are taking a longer path than other and some photons get absorbed. This means, that the photons are taking different ways through the sample and thus also need different times to pass it. For that reason the temporal shape of the pulse is measured after leaving the sample. By fitting theoretical models to this shape, it is possible to retrieve absorption and scattering properties of the sample.

Pharmaceutical tablets are highly scattering and are usually made of more than one ingredient. But for optimal treatment and patient safety it is important to know the precise amount of drug inside every tablet. In this thesis work it is shown, that with PTOFS it is possible to predict the drug content of mixed tablets better than 1% of its total mass by using a suitable calibration. The advantage of PTOFS compared to other techniques is, that it is non-destructive and also independent on the particle sizes of the ingredients.

Yoghurt and Milk are available in different fat contents, but they all are produced from raw cow milk with 3%-5% fat content. With PTOFS it is possible to measure the fat content out of scattering and absorption properties. By taking both properties into account, it is possible to get more precise fat content predictions. So this technique might be interesting for quality control and customer protection.
ABBREVIATIONS

AOTF    Acousto-Optical Tunable Filter
APD    Avalanche Photo Diode
API    Active Pharmaceutical Ingredient
CFD    Constant Fraction Discriminator
FWHM    Full Width Half Maximum
IRF    Instrumental Response Function
LLTF    Laser Line Tunable Filter
LV    Latent Variables
MC    Monte Carlo
MCP    Multi Channel Plate
MSC    Multiplicative Scatter Correction
NIR    Near Infra Red
NIRS    Near Infra Red Spectroscopy
PCF    Photonic Crystal Fibre
PLS    Partial Least Squares
PMT    Photo Multiplier Tube
PTOFS    Photon Time-of-Flight Spectroscopy
RDP    Ratio of Deviation to Performance
RMSE(C/P)    Root Mean Square Error (of Calibration/Prediction)
RTE    Radiative Transport Equation
SD    Standard Deviation
SNR    Signal-to-Noise Ratio
TCSPC    Time-Correlated Single-Photon Counting
TI    Trans Impedance
TNIRS    Transmission Near Infra Red Spectroscopy
TR    Time Reference
UV-VIS    Ultra Violet - Visible Spectroscopy
WMC    White Monte Carlo
Popular science abstract

1 Introduction
1.1 Pharmaceutical tablets ........................................ 2
1.2 Dairy products .................................................. 2

2 Theory
2.1 Properties of light .............................................. 5
2.2 Light propagation in turbid media .............................. 5
2.2.1 Scattering ..................................................... 5
2.2.2 Absorption ................................................... 6
2.3 Radiative Transport Theory .................................... 7
2.3.1 Definition: Radiance ....................................... 7
2.3.2 Radiative Transport Equation ............................ 7
2.3.3 Monte Carlo simulations .................................. 8
2.3.4 Diffusion approximation ................................. 9

3 Material and Methods .......................................... 11
3.1 Principle of PTOFS ............................................ 11
3.2 Experimental setup ........................................... 11
3.2.1 Design ....................................................... 11
3.2.2 Super continuum source .................................. 12
3.2.3 Spectral filter ............................................... 12
3.2.4 Sample preparation ........................................ 13
3.3 Data collection ................................................ 13
3.3.1 TCSPC ........................................................ 13
3.4 Data evaluation ................................................ 13
3.5 Software ........................................................ 14
3.5.1 Lund TOFS .................................................. 14
3.5.2 TimeResolved_v3 .......................................... 15
3.6 Quality of PTOFS measured data ............................ 15
3.7 Issues ............................................................ 15
3.7.1 Finite spectral bandwidth filter effect .................. 15
3.7.2 Data evaluation ............................................. 15
3.7.3 Sufficient signal-to-noise vs. pulse broadening ....... 16
3.8 Material and Methods used for tablet study ............... 16
3.8.1 Tablet-set .................................................... 16
3.8.2 Experimental setup ....................................... 17
3.8.3 PTOFS data evaluation .................................. 17
3.8.4 Comparative data ......................................... 17
3.8.5 Chemometric evaluation ................................ 17
3.8.6 Quality of the evaluation ............................... 18
3.9 Material and Methods used for dairy product study .... 19
3.9.1 Dairy product properties .................................. 19
3.9.2 Experimental setup ....................................... 19
3.9.3 PTOFS data evaluation .................................. 19
3.9.4 Spectral analysis .......................................... 20
The interaction of electromagnetic radiation with atoms and molecules is the subject of optical spectroscopy. In turbid media, the main interactions of this radiation with the matter are scattering and absorption. Generally speaking, scattering is connected to structural properties of the media, while absorption is related to the chemical composition [1]. These two phenomena have similar impact on the light that is interacting with the matter. For that reason it is challenging to consider both, scattering and absorption.

Several spectroscopic techniques utilize the Near Infra Red (NIR) spectral region (750 nm - 2500 nm). In this spectral region, overtones and combination bands of primary absorption bands at higher wavelengths in the mid infra red are dominating [2]. Because these overtones and combinations are relatively weak, they are well suited to measure turbid media, in which the light has a relatively long path length. This is an advantage of NIR Spectroscopy (NIRS), that is commonly used for structural and chemical analysis of turbid media. NIRS has been used to analyse turbid media such as biological tissue [3], quality and production control of food products [4] and pharmaceutical products [5].

Still most of the techniques focus on measuring absorption (e.g. Transmission NIRS (TNIRS)) and try to minimize the influence of scattering with mathematical corrections. In Photon Time-Of-Flight Spectroscopy (PTOFS) it is possible to measure more or less independently both, scattering and absorption simultaneously. This has the advantage, that one gets structural and compositional informations in one measurement and in an ideal case neither of these informations affects the other.

In PTOFS, a temporally short pulse is sent through the sample, where the light gets scattered and absorbed, and is then recorded. The new, broader shape of this pulse then is used to extract scattering and absorption information from the sample by fitting a suitable theoretical model to the pulse shape. The Biophotonics group at Lund University developed a PTOFS system that is able to measure broadband spectra of absorption and scattering from 500 nm to 1400 nm [6]. Recently this system has been used to measure the change of oxygen saturation in arm tissue [7], human muscle tissue [8] and phantom samples [9].

The aim of this thesis work is to measure absorption and scattering spectra on pharmaceutical tablets and dairy products with an improved PTOFS system. It is shown, that the newly introduced experimental equipment improves the quality of structural and compositional analyses and the potential of this technique for these applications is discussed.
1.1 Pharmaceutical tablets

In the production process of pharmaceutical tablets it is of great importance to assure the precise quantity of drug in each tablet in order to fulfill the high quality requirements for patient safety and treatment efficiency. Often a tablet is, due to production circumstances, made of more components than the drug. It is then important to know the precise content of the drug in the tablet. In order to fulfill the quality regulations, it is necessary to test the composition of the mixed tablet to a precision better than 1%.

With NIRS it is possible to determine the composition of the absorption spectrum of a mixed tablet and thus it is a commonly used tool in the production process [5, 10–14]. Furthermore NIRS techniques are non-destructive, relatively fast and inexpensive and therefore together with chemometric techniques, NIRS is very interesting for pharmaceutical industries.

However, as pharmaceutical tablets usually are not transparent, light is heavily scattering while passing through the tablet. This elastic scattering makes it challenging to measure the absorption spectrum, in a way such that it is not affected by scattering properties of the tablet. In ultra violet - visible (UV-VIS) spectroscopy this problem is solved by dissolving the tablet in solvents, such as e.g. methanol and water and filter out the scatterers [15]. The disadvantage of this technique is, that it is destructive, relatively slow and expensive, but it still is a good reference method for compositional analysis on mixed tablets.

PTOFS is able to separate between absorption and scattering during data evaluation. So with these PTOFS absorption spectra it should be possible to determine the tablet composition very precisely. It has been shown that it is possible to measure absorption and scattering spectra with PTOFS [16] and also that it is possible to retrieve the drug content of the tablet [17, 18]. Still the results of the PTOFS evaluated data are not superior compared to the conventional TNIRS. So in this thesis project it is examined, whether the drug content prediction from PTOFS evaluated data is enhanced, if the spectral resolution of the system is increased. This higher spectral resolution should minimize the coupling between measured absorption and scattering, especially at sharp absorption peaks, as described in section 3.7.

In this thesis, the potential of PTOFS in drug content evaluation of pharmaceutics is demonstrated. It is also shown, that the enhanced spectral resolution is reducing the coupling between evaluated absorption and scattering coefficients and that this improves the precision of drug content prediction.

1.2 Dairy products

Dairy products are important part of human nutrition, especially for children. Products, such as milk or yoghurt, available in grocery stores vary vastly in fat content. In order to achieve a dairy product with the desired fat content, the raw milk, that contains 3 to 5 % fat, is industrially processed. These processes have been done routinely for many years, but as in all industrial processes it is important to control the quality of the product. Another motivation for quality control could be consumer protection against products not meeting the specifications (e.g. diluted milk). But also during the production process information about structure and chemical properties might be interesting for optimization. Over time several tests and standards have been developed to meet these needs for quality control [19, 20].

Optical techniques are potentially interesting for this kind of monitoring as they are non-destructive, fast and potentially of low cost. Therefore several experimental optical techniques have been examined on milk[4]. It has been shown, that it is possible to measure the average particle size and monitor the gelation process of milk, using dynamic light scattering and diffusing wave spectroscopy [21]. Compositional analysis of milk has been done online with diffuse reflectance spectroscopy [22] and transmission NIRS has been used to
Introduction

examine protein content [23], the fat content in milk powders [24] and the composition of milk out of fat, protein and casein [25]. Other approaches compare reflectance and transmission NIRS techniques for measurements on raw milk for cow health monitoring [26]. Furthermore it has been shown, that also PTOFS can be successfully applied on dairy products [27] and fat particle concentration and sizes have been extracted out of PTOFS scattering spectra in the wavelength range between 500 and 1000 nm [28]. Little has been done so far on fat content analysis based on measurement of absorption peaks in the NIR. This might be due to the fact, that absorption peaks of fat and water almost overlap at this region, so the influence of scattering might destroy all sensitivity.
In this work, PTOFS absorption and scattering spectra are measured and used to evaluate the fat content of a dairy product in a wavelength range between 1000 nm and 1300 nm. This newly presented data could allow a fat content evaluation by using information from both, scattering and absorption. It is discussed if this additional information could lead to a precise fat content evaluation tool.

This thesis report is structured in the following way: First the basic theory of light-matter interaction is presented in chapter 2. This chapter should give the reader an overview about the used theory in PTOFS, but it is recommended to go into the references for a deeper understanding. Chapter 3 describes the methods and materials, used for this thesis work. This includes the experimental setup, with a focus on the improvements that have been done during this thesis work and the evaluation methods. The chapter closes with the very specific materials and methods that are used for the pharmaceutical tablets and dairy products. The results are presented in chapter 4, which is divided into two parts for tablets and dairy. The thesis report concludes in chapter 5, followed by closing remarks and reprints of papers, that have been published or submitted as part of this thesis work.
Photon time-of-flight spectroscopy (PTOFS) applies theory of light-matter interaction as well as radiation transport theory. These concepts are presented in this chapter.

2.1 Properties of light

In PTOFS several properties of light are used in order to measure optical properties of a turbid sample. For that reason this section should remind the reader of the two fundamental pictures of light. Light can be described by the Maxwell equations (2.1 - 2.4) and so it has all properties of an electromagnetic wave. The other fundamental picture of light describes light as particles (photons) that have a discrete energy (2.5).

2.2 Light propagation in turbid media

Turbid media is defined as media, in which light is highly scattered during propagation. Additionally, the light can be absorbed in the medium. The absorption spectrum of a material potentially contains information about the chemical composition. Therefore it is the goal of absorption spectroscopy to measure the absorption spectrum. As in turbid media both, scattering and absorption occur, both of these phenomena have to be taken into account for theoretical modelling. For that reason both phenomena are introduced here.

2.2.1 Scattering

Scattering in a material occurs on local changes of the index of refraction and it locally changes the direction of the propagating photon. If the energy of the incident and the scattered photon is conserved, the process is called elastic scattering. If the energy of the photons changes during the scattering it is called inelastic scattering (e.g. Raman scattering). Inelastic scattering can be used to extract information about the atomic or molecular energy levels, but is a minor process during light propagation [30]. Elastic scattering conserves the energy of the photon and does thus not contain any energetic information. It is by far the most common scattering type in media. For that reason PTOFS only considers elastic scattering, while the energetic information is extracted from the absorption spectrum. Elastic scattering can be described as Rayleigh and Mie scattering [31, 32].

Mie scattering expresses the scattering process with an exact solution of the Maxwell equations and is valid for all spherical shaped scatterers. It, however extensively applies mathematics. Rayleigh theory uses approximations for particles small compared to the wavelength of the scattered light. These approximations makes the theory mathematically easier. For that reason

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]

\[
\n\n\n\]
2.2.2 Absorption

Light propagating in a material can also be absorbed by the atoms and molecules in the material. If the energy of the incoming light matches one of the atomic or molecular transitions, the photons can be absorbed. This process transfers the absorbing molecule to a higher energetic state. In order to get back to the stable lower state, the absorbed energy can heat the material or emit another photon. If the re-emitted photon is of a lower energy than the absorbed photon, this process is called fluorescence. Any such light is neglected in PTOFS.

Absorption can be quantified with an absorption coefficient $\mu_a$. In a microscopic picture, the probability that a photon gets absorbed corresponds to a higher absorption coefficient. In a macroscopic picture, the absorption coefficient is defined as the inverse length after the incident light is attenuated to a fraction of $e^{-1}$. This is expressed in the Beer-Lambert law: [36]

$$I(L) = I_0 \cdot e^{-\mu_a L} \quad (2.11)$$

Rayleigh theory is commonly used.

### Scattering of a single photon

Consider a single photon with a direction of propagation $s$ reaches a particle, the photon gets scattered and so it will change its direction of propagation to $s'$ (see Fig. 2.1). The probability $p$ of the photon scattering in an angle $\theta$ to $s$ is given by the Henyey-Greenstein function [33]

$$p(\theta) = \frac{1}{4 \pi} \cdot \frac{1 - g^2}{(1 + g^2 - 2 \cdot g \cdot \cos(\theta))^{3/2}} \quad (2.6)$$

where $g$ stands for the anisotropy factor that is defined as the average of the cosine of the scattering angle $\theta$.

$$g = \langle \cos(\theta) \rangle \quad (2.7)$$

For isotropic scattering the factor $g$ vanishes and the probability $p$ gets the same for all possible angles $\theta$. For an increasing anisotropy factor $g$, the scattering becomes more anisotropic and on average the scattering is more directed forwardly.

### Scattering of multiple photons

For a large number of photons that are scattering in a material, the average behaviour of these photons is of interest. For this purpose a scattering coefficient $\mu_s$ is introduced as the inverse of the free mean path between two scattering events (2.8) [34]. The scattering coefficient thus gives a measure for the probability a photon gets scattered while passing a path in the material. Any anisotropy in the scattering is corrected in the reduced scattering coefficient $\mu'_s$, defined as

$$\mu'_s = \mu_s \cdot (1 - g) \quad (2.9)$$

In the data evaluation of this thesis, with the knowledge of previous studies it is assumed, that the scattering can be described as an approximation of the Mie theory as [35]

$$\mu'_s = A \cdot \left(\frac{\lambda}{\lambda_0}\right)^{-\beta} \quad (2.10)$$

Here $\lambda$ denotes the wavelength of the spectrum, $\lambda_0$ an arbitrary reference wavelength, $A$ the amplitude parameter, that is connected to the concentration of scatterers and $\beta$ a parameter, that is connected to the particle size of the scatterers. This equation is also used to fit the measured scattering spectra over the parameters $A$ and $\beta$.
Here $I_0$ represents the incident intensity and $L$ the length the light travels in the sample with the absorption coefficient $\mu_a$.

As the absorption spectrum of a material is connected to energetic transitions of the molecules, it contains information about the composition of the material. The absorption coefficient at the wavelength $\lambda$ of a sample containing $N$ ingredients can be calculated as a linear combination of the volume fractions $f_{V,i}$ of the component $i$ times the pure absorption coefficient $\mu_{a,i}$ [37]:

$$\mu_a(\lambda) = \sum_{i=1}^{N} f_{V,i} \cdot \mu_{a,i}(\lambda) \quad (2.12)$$

In the NIR spectral region, that is used in this work, the measured absorptions are overtones from vibrational transitions in the mid-infra-red region. For that reasons the absorption features in the NIR are relatively small. This results in rather long light path lengths in the sample. An overview of the absorption overtones in the NIR can be found in [38].

### 2.3 Radiative Transport Theory

The Radiative Transport Theory is an approach to mathematically describe the propagation of light in turbid media with known absorption and scattering properties. Because the data evaluation of the measured time-of-flight curves is based on this theory, the basic concepts are presented here.

#### 2.3.1 Definition: Radiance

In this part the radiance $L$ is a quantity that is mainly used to explain the radiative transport. The radiance can be understood as the product of the number of photons $N(r, \hat{s}, t)$ at a position $r$ with the direction $\hat{s}$ at time $t$ and the energy of the photons $h \cdot \nu$ times the speed of light $c$:

$$L(r, \hat{s}, t) = N(r, \hat{s}, t) \cdot h c \quad (2.13)$$

#### 2.3.2 Radiative Transport Equation

The Radiative Transport Equation (RTE) describes how the radiance $L$ is changed while propagating through a volume element of turbid material. A detailed derivation of the RTE can be found in [36], while here only the origin of the contributions are motivated.

The RTE is based on energy conservation, that in this case is equal to the conservation of the number of photons and hence to the conservation of radiance. Every change in radiance in a volume element $dV$ can be explained by physical processes. The radiance in direction $\hat{s}$ changes due to light coming in and leaving the volume (a). It also decreases due to scattering and absorption (b), summarised as extinction (with the extinction coefficient defined as $\mu_t \equiv \mu_s' + \mu_a$).

A source $S$ (d) increases the radiance as well as photons that get scattered from other directions to the examined direction $\hat{s}$ (c).

All these processes contribute to the RTE in the following way:

$$\frac{\partial L(r, \hat{s}, t)}{c \cdot \partial t} = -\hat{s} \nabla L(r, \hat{s}, t) - \mu_t L(r, \hat{s}, t) + \mu_a \int_{4\pi} L(r, \hat{s}', t) P(\hat{s}' \cdot \hat{s}) d\Omega' + S(r, \hat{s}, t) \quad (2.14)$$

The parts of the RTE are visualized in Fig. 2.2. This non-linear differential equation in general does not have an analytical solution. For that reason approximations and simulations are used to in order to enable finding solutions to the equations.
2.3.3 Monte Carlo simulations

Monte Carlo (MC) simulations use virtual particles that are behaving according to the equation that should be solved. In the case of the RTE the particles move through a material according to probability functions that contain the conditions from the specific RTE. For a sufficient high number of particles, the particle concentration gives an accurate solution for the radiance of the RTE. In this thesis MC is used for the evaluation of the PTOFS curves from dairy products. Therefore the basic concepts of MC are presented here. For a more detailed description of MC used for PTOFS evaluation it is referred to [34].

In the MC simulation process every virtual particle undergoes a loop of several steps: After the particle is initialized, it is moved according to random number generators and probability functions for the various mechanisms involved in the light transport, see e.g. (2.15). Then another function and a drawn random number decides if the particle survives (gets absorbed) or not. If it survives, the direction is changed according to the Henyey-Greenstein function (2.6) and another random number. These steps are then repeated.

The geometry of the sample plays an important role. The particle either leaves the sample or gets reflected at a boundary, according to the Fresnel equations (see [29, chapter 6]). It is important to implement the sample geometry very carefully in the MC algorithm.

For the PTOFS data evaluation in this work, a White Monte Carlo (WMC) algorithm was used. This algorithm assumes the material to have no absorption (to be "white") and so all photons survive at each step. This is one reason this algorithm is more efficient. In order to add the absorption to the simulated solution of the RTE, the radiance is scaled by using the Beer-Lambert law (2.11). WMC simulates for one scattering coefficient. If other scattering coefficients are of interest, this solution can be obtained by scaling the initially simulated result. This scaling is only possible in infinite homogeneous media.

MC simulations are very reliable and generally valid for any geometry (even though it might become more difficult to practically implement the geometry in the algorithm). For a high number of particles MC gives very precise solutions of the RTE and therefore MC (especially the MCML program [39]) became a gold standard for solving the RTE. The main drawback of this method is, that it is computationally very expansive and thus usually needs a lot of time.

\[ l = -\frac{\ln(\xi)}{\mu_x} \]  

Function to get a length \( l \) to move in a material with an absorption or scattering coefficient \( \mu_x \) with the random variable \( \xi \in (0,1] \) [34]
2.3.4 Diffusion approximation

The diffusion approximation can be used in certain cases to simplify the RTE to a diffusion equation that has known analytical solutions. The diffusion approximation was used in this work for the evaluation of the pharmaceutical tablets and for that reason the concepts of this method is presented here. A more detailed derivation can be found in [36] and also with a different approach in [34]. In order to bring the RTE to a diffusion equation form, approximations and simplifications are necessary. In principle the radiance $L$ is separated into an isotropic part $\Phi$ and a gradient part $J$. This radiance is then put into the RTE that then can be brought into the form of a Fick’s law

$$J(r, t) = -D \nabla \Phi(r, t)$$

(2.16)

with a diffusion constant $D$. Assuming $\mu'_s \gg \mu_a$ and a large distance between source and detector, the diffusion constant can be written as

$$D = \frac{1}{3 \cdot (\mu_a + \mu'_s)} .$$

(2.17)

When Fick’s law (2.16) is substituted to the RTE (2.14), it results in the diffusion equation.

$$\frac{\partial \Phi(r, t)}{\partial t} + \mu_a \Phi(r, t) - D \nabla^2 \Phi(r, t) = S(r, t)$$

(2.18)

For this diffusion equation solutions are known. An example for a solution for a temporal and spatial infinite narrow pulse can be written as

$$\Phi(r, t) = \frac{c}{4\pi Dct} \cdot e^{-\frac{r^2}{4Dct} - \mu_a ct}$$

(2.19)

As for the MC method the sample geometry plays an important role and therefore the boundaries have to be included into the model. Compared to MC this is not always possible for a diffusion equation as the boundaries must be integrated into an analytical expression. There are geometries (infinite, semi-infinite and slab-geometry) that can be used with the diffusion equation by using the concept of mirror sources, often used in electrodynamics [40]. For more sophisticated geometries it is, however, not possible to use the diffusion equation. Beside the drawback of the limited geometries, it is important to keep in mind, that the diffusion equation only works as long as approximations are valid. In general this is the case for a significantly higher scattering than absorption and a sufficiently large distance between source and detector. Otherwise the advantage of the diffusion method is its analytical solution that makes it computationally much faster than the MC method. Recently there has been work on extending the validity range of the diffusion method. Alerstam [41] introduced a diffusion theory for anisotropic materials and Liemert and Kienle [42] introduced a diffusion method for shorter distances.
In this chapter the methods of Photon Time-of-Flight Spectroscopy (PTFOS) is introduced. The method and setup presented here is used in all measurements performed during this thesis work. The last two sections of this chapter are specifically explaining the materials and methods for the two subjects that are examined in this thesis.

3.1 Principle of PTOFS

The PTOFS technique extracts the absorption and reduced scattering coefficient from the temporal broadening of a pulse after propagation through a sample. The principle idea is that light taking a long way through the sample needs a longer time until reaching the detector, than light taking a short path through the sample. This results in a temporal pulse broadening, as the speed of light is on average constant in one material. The length of the path the light is taking depends on the scattering and absorption properties of the material. For example, in a highly scattering material light will on average have a longer path through the sample than in a low scattering material (see Fig. 3.1). Also light propagating through a highly absorbing material tends to travel a shorter path than in a low absorbing material. This follows the Beer-Lambert law (2.11). Thus the temporal photon distribution contains information about the optical properties of the material.

In chapter 2 the Radiative Transport Equation (RTE) is discussed and in (2.14) it can be seen that it also describes the temporal distribution of the radiance. In Figs 3.2 & 3.3 PTOFS curves are shown that have been calculated with the diffusion model. In this figures it can be seen that the shape of the curve varies significantly with absorption and scattering properties. It is also important to note that the absolute photon count is unnecessary to discriminate between the curves.

3.2 Experimental setup

This section gives an overview of the instrumental equipment used in the Lund PTOFS setup. A detailed characterization of this setup can be found in Ref. [6].

3.2.1 Design

The schematic of the used setup is shown in Fig. 3.4. A Photonic Crystal Fiber (PCF) super continuum source (Model SC480-10 or Model SC500-6, Fianium Ltd, Southampton, UK) generates pulses with less than 100 ps Full-Width Half Maximum (FWHM) duration at a repetition rate of 80 MHz. This pulse then is spectrally filtered to select one wavelength by a Laser-Line Tunable
3.2.2 Super continuum source

During this thesis project work a new PCF super continuum source has been introduced to the system. The new source (Model SC480-10) has a higher optical power output of 10 W over all wavelengths (compared to 6 W of SC500-6) and its spectrum has a lower cut-off wavelength of 480 nm (compared to 500 nm of SC500-6).

3.2.3 Spectral filter

For spectral filtering a Bragg grating filter called LLTF and an AOTF are used. The LLTF is newly introduced to the setup and has a smaller spectral FWHM ($\Delta \lambda \approx 2 \text{ to } 5 \text{ nm}$) than the previously used AOTF ($\Delta \lambda \approx 6 \text{ to } 12 \text{ nm}$), meaning that the wavelength employed is better defined.

The LLTF is a tunable Bragg grating filter and its basic sketch can be found in Fig. 3.5. In the LLTF, the polychromatic incident beam is guided to an image of a grating, saved in a volume hologram. This hologram works as a diffraction filter (LLTF) (Model Contrast-SWIR, Photon etc., Montreal, Canada) or an Acousto-Optic Tunable Filter (AOTF) (Model AOTF-DUAL with crystals VIS, NIR1 or NIR2, Fianium Ltd, Southampton, UK) depending on the wavelength range and application. After the filter the light pulse is lead in custom made gradient index multimode fibers (400 \( \mu \text{m}/640 \mu \text{m} \) core/cladding diameter, Leoni Fiber Optics, Germany) to a glass plate that works as a beam splitter, where a 4% fraction is separated as a Time Reference pulse (TR). The main part of the pulse then is guided to and collected from the sample. Both pulses are then individually attenuated with a variable optical attenuator (OZ Optics Ltd, Ottawa, Canada) and recombined by another glass plate. The fiber lengths of both arms are chosen in a way that the TR pulse arrives around 2 ns earlier than the sample pulse. The recombined pulses are guided to a detector. Depending on the wavelength range either a Micro-Channel-Plate Photo-Multiplier Tube (MCP PMT) (R3809U-68 Hamamatsu Photonics, Hamamatsu, Japan), in the wavelength range of 950 to 1400 nm, or a silicon SPC avalanche photo diode (APD) (PD1CTC Micro Photon Devices, Bolzano, Italy), in the wavelength range of 400 to 1000 nm, is used. The detectors are protected by a shutter (Thorlabs, Newton, USA and OZ Optics Ltd, Ottawa, Canada, respectively). The signal of the PMT is amplified by a Trans-Impedance (TI) amplifier (HFAC-26, Becker & Hickl). The detector signal and a synchronization signal from the source are used by Time Correlated Single Photon Counting (TC-SPC) electronics. The TCSPC electronics is on a PC card (Model SPC-130 or SPC-130EM, Becker & Hickl, Berlin, Germany) and it records the PTOFS curve.

![Figure 3.3](image-url)  
*Figure 3.3. Calculated PTOFS curves for fixed absorption and varying scattering coefficients. The calculation is using the Diffusion approximation.*

![Figure 3.4](image-url)  
*Figure 3.4. Schematic of the used setup*
grating. One wavelength is diffracted, all other wavelengths are transmitted. The diffracted wavelength is selected by the incident angle on the grating. [43]. The diffracted light is then back-reflected by a corner cube reflector to again being diffracted by the Bragg grating, because the angle is still the same. This doubly diffracted beam is then lead to the filter output. It is possible to change the filtered wavelength by tuning the angle of the Bragg grating and the corner cube reflector, as this changes the incident angle of the light and therefore the diffracted wavelength. [44, 45]

3.2.4 Sample preparation

The distance the light pulse travels inside the sample plays an important role in the data evaluation. For that reason it is important to keep close control of the distance between light injection and collection points in the experimental setup. For samples that are measured in slab geometry, a sample holder has been built (Thorlabs, parts SM1L10 and S120-FC) that brings the fiber tips to direct contact with the sample. For solid samples (tablets) it is though possible to measure the thickness of the sample to get the distance between the fibers. The measurements on samples in infinite or semi-infinite geometry are performed with needles containing the fibers. These needles are positioned and fixed and the distance can be measured between the needle tips. Another possible method is to place an object of known thickness between the needles during positioning.

3.3 Data collection

The aim is to measure the shape of the broadened light pulse with high temporal resolution. With available photon detectors this is not possible, because the temporal resolution of the conventional detectors is not sufficient. In order to reach the aimed temporal resolution for example Streak Cameras [46] or the TCSPC technique can be used.

3.3.1 TCSPC

As the TCSPC technique is used in this work, it is briefly presented here. A detailed description about this technique can be found in Ref. [47]. As the name suggests, only a single photon out of every light pulse is measured in a TCSPC measurement. The time for the photon detection is measured relative to a synchronization signal from the source. The time difference can be measured very precisely by employing Constant Fraction Discriminators (CFD). The time measurement is repeated for many pulses. As maximum one photon out of every pulse is making it to the detector, the pulse shape can be statistically restored by doing many measurements and making a histogram of detected arrival times. As the temporal resolution of the single photon detection with CFDs is considerably better than the temporal resolution of the photon detector itself, a better resolution of the pulse shape is possible using TCSPC.

3.4 Data evaluation

The purpose of the data evaluation is to retrieve the absorption and the reduced scattering coefficients ($\mu_a$ and $\mu'_s$) from each measured PTOFS curve. For this the solutions of the RTE (see Sec. 2.3) are fitted to the measured time curves. As discussed in the previous chapter, either the Diffusion equation or the WMC simulation is used as a fitting model. For the fitting process, each PTOFS curve of a sample is assigned with a Instrumental Response Function (IRF) curve, measured for the same wavelength. The IRF is measured by replacing the sample with a both side double black printed paper. A typical IRF has a temporal FWHM $\sim 50$ ps and an example can be seen in Fig. 3.6. For each pair of sample-IRF curves the temporal shift of the
TR pulse is calculated and corrected in the IRF curve. Then the theoretical curve for initially guessed $\mu_a$ and $\mu_s'$ are calculated with the chosen method. In order to take the system response into account, the theoretical curve is convoluted with the experimental IRF. This is repeated in the Levenberg-Marquardt algorithm until the best fitting $\mu_a$ and $\mu_s'$ are found. An example of the fit is shown in Fig. 3.6.

3.5 Software

The software used for data collection and data evaluation were both developed in the Lund Biophotonics group and part of this project work has been to improve this software. For that reason the main purpose and functions are presented in this report.

3.5.1 Lund TOFS

The Lund TOFS software has been developed in collaboration with a computer science group of Danmarks Tekniske Universitet and is written in C#. The program is used to collect and save PTOFS curves. As it is necessary to scan all wavelengths in order to create a broadband spectrum, the software provides an automatic measurement routine to measure a whole spectrum on a sample or IRF. For that purpose the program controls the optical attenuators, the spectral tunable filter and it collects and saves the data from the TCSPC card.

Contributions during thesis project work

As Lund TOFS has been written for a PTOFS system using another AOTF (SuperK Select, NKT Photonics A/S, Birkerød, Denmark), it was necessary to implement the driver of the used AOTF and LLTF in order to perform automated measurements with the available experimental equipment. The new filters that should be controlled by the software come with driver Dynamic Link Libraries and with a documentation that includes basic examples in C++. In order to implement the driver functionality to the Lund TOFS, C++ command line based programmes are written for each filter, that are then called as external programmes within the C# programme.

![Figure 3.6. Example of a PTOFS curve measured on a tablet (blue), its diffusion model fit (red) and the instrumental response (black); all measured at 1168 nm. The fit results in an absorption coefficient $\mu_a = 0.47$ cm$^{-1}$ and a reduced scattering coefficient $\mu_s' = 232$ cm$^{-1}$](image)
3.5.2 TimeResolved_v3

The software TimeResolved_v3 is used for evaluation of the data, recorded with Lund TOFS and is written in MATLAB. It implements the routines described in 3.4 and returns the absorption and reduced scattering spectra. In TimeResolved_v3 it is possible to prepare the data, adjust the fitting method for specific geometry and then fit the model to the data. For each step of the data evaluation procedure, the PTOFS curves and fits are shown in order to control the quality of the evaluated data.

3.6 Quality of PTOFS measured data

The stability of the performance of the PTOFS system has been examined in several studies and the system has been developed continuously for a better data quality [9, 16]. For example, it has been an issue, that the laser was temporally shifting. This caused less precision for the measurements, as the sample and IRF measurements are done at different times. The shift has been thus mis-interpreted as a shift due to the optical properties of the sample. With the introduced TR pulse, the precision has been improved significantly. Another improvement in measurement quality was done during this thesis work by introducing an enhanced spectral filter.

The performance of the PTOFS system is also frequently controlled by reproducing absorption and scattering spectra of home made and commercially available phantoms. [9, 16]

3.7 Issues

The PTOFS system is able to take high quality measurements, but still there are some issues to work on. Some of these issues are discussed below.

3.7.1 Finite spectral bandwidth filter effect

If a light source with a finite spectral bandpass is used to record PTOFS, light at wavelengths with the lowest absorption within the spectral width will dominate the time-resolved curve [48]. The AOTF and LLTF filters used in this PTOFS setup of course have a finite spectral resolution and thus this effect also occurs in the PTOFS measured data. In Fig. 3.7 this effect can be seen for both filters, measured on a 3.64 mm thick BG36 phantom made of a crushed coloured glass filter (BG36, Schott glass, Germany, [49]), titanium dioxide particles (Sigma-Aldrich, USA) and epoxy resin. In the marked regions of the scattering spectra we notice both, increase and decrease in spectral regions with absorption peaks. Also the amplitude of the absorption peaks is lower than it should be.

In Fig 3.7 it can be also seen, that this effect occurs significantly less in the spectra recorded with the LLTF. This is believed to be due to the higher spectral resolution of this filter reducing the impact of this effect.

3.7.2 Data evaluation

As described in 3.4, the accuracy of the measured spectra relies strongly on a suitable light propagation model for fitting the PTOFS curves. In the TimeResolved_v3 data evaluation software only a limited number of models and measurement geometries are implemented. These models usually are suitable for most of the measured samples, but in some cases it would be helpful to work on the implementation of new models or geometries.
3.7.3 Sufficient signal-to-noise vs. pulse broadening

In order to achieve a good quality fit of the PTOFS curves, a good Signal-to-Noise Ratio (SNR) is desired. Noise is introduced by the detector, electronics, photon statistics, reflections on optical components and due to ambient light. A desired high SNR can thus be achieved by increasing the signal and measure with a suitable photon count rate. For high absorption coefficients the optical power of the source might be not sufficient. The count rate is though higher for a shorter distance between source and detector fibres, so a sufficient SNR might be achieved by measuring with a short fibre distance.

On the other hand, it is important for data evaluation, that the PTOFS curve is sufficiently broadened by the propagation through the medium. If the measured pulse width after leaving the sample is not considerably broader than the IRF, it will not be possible to retrieve absorption and scattering from this pulse. The pulse will be broader, if it travels a longer distance in the sample. In principle one can determine absorption and scattering properties more accurately with increasing temporal broadening and thus longer inter-fibre distance.

In order to achieve a sufficient SNR and a sufficient pulse broadening it is thus important to choose a suitable fibre distance. It is, however, not always possible to find such a distance, as the SNR could be very low while the pulse is not broadened at all. This usually occurs for samples and wavelengths with high absorption. This problem limits the range of samples that can be measured with this PTOFS system. This limitation could partly be overcome by using a light source with more optical power.

3.8 Material and Methods used for tablet study

3.8.1 Tablet-set

The tablet-set used in this study contains 70 tablets that were all manufactured by AstraZeneca R&D Mölndal, Sweden, for a study with Raman spectroscopy [50]. All tablets contain three main ingredients: Ibuprofen as the Active Pharmaceutical Ingredient (API), Mannitol as filler material and Magnesium stearate as lubricant. The tablets vary in the designed drug content (16, 18, 20, 22, 24 %w/w (weight percentage)), the average particle diameter of the API (71, 95, 154 µm) the average particle diameter of the filler (91, 211, 154 µm) and the compression force during manufacturing process (8, 12, 16 kN). With the variation in the particle sizes and compression force the scattering properties of the

Figure 3.7. Effect of a finite spectral bandwidth filter (marked) in scattering and absorption spectra of measurements on the BG36 phantom with AOTF and LLTF.
Material and Methods

3.8.2 Experimental setup

The pharmaceutical tablets are measured in the spectral region of 950 to 1350 nm with a spectral resolution of 4 nm. For these measurements the MCP-PMT is used as detector and the LLTF as spectral filter. The tablet is placed in a tablet holder and the signal is recorded in transmission geometry. The PTOFS curves were captured over an integration time of 25 s with a photon count rate of approximately $1 \times 10^5$ photons per second.

3.8.3 PTOFS data evaluation

As the measurement geometry demands a model for slab geometry and the requirements for the Diffusion model are fulfilled, a slab geometry Diffusion model from Contini et al. [40] was chosen for fitting the data. An example of a fitted PTOFS curve measured on a tablet can be found in Fig. 3.6.

3.8.4 Comparative data

The absorption spectra taken in this study with PTOFS and LLTF as spectral filter are compared to absorption spectra of the same tablets, measured with other experimental techniques. The used tablet-set has been part of a previous study, measuring the absorption spectra with a similar PTOFS setup, using an AOTF as spectral filter [18] and TNIRS spectra for the tablets were captured by AstraZeneca. The TNIRS technique measures extinction spectra that can be processed by various techniques to effective absorption spectra, that are used in this study. In order to have a fair comparison of all techniques, exactly the same tablets were selected for data evaluation for each technique.

The evaluated drug contents from Raman spectroscopy measurements [50] are used as reference values. Despite that Raman spectroscopy is not the gold standard for the analysis of the chemical composition of a mixed tablet normally, it is here used as the benchmarking reference technique. Otherwise the typical reference method is UV-VIS spectroscopy. This technique needs though to dissolve the tablet in liquids and is thus destructive. For that reason there is no UV-VIS spectroscopy data available for the measured tablets. It is though possible to estimate an error of Raman compared to UV-VIS spectroscopy from the data set of tablets measured first with Raman and then with UV-VIS spectroscopy. The root mean square error of Raman spectroscopy on tablets from this tablet-set is 0.5%. All these measurements are conducted by people at AstraZeneca.

3.8.5 Chemometric evaluation

The evaluated absorption spectra were used to perform a Partial Least Squares (PLS) regression analysis. This chemometric technique uses the eigenvectors obtained from a training set of absorption spectra and corresponding reference API contents in order to create a calibration. This calibration then can be used to evaluate absorption spectra with unknown API contents. The used eigenvectors are called Latent Variables (LV) and should represent a chemical component of the sample. [51]

In this study, the MATLAB function plsregress was used to build the PLS calibration. The absorption spectra data of each measurement method were pre-treated in a way it returned the best result: All spectra are mean-centred by wavelength; the mean absorption value of all tablets for each wavelength is calculated and subtracted. The TNIRS spectra also are corrected for scattering influences with the Multiplicative Scatter Correction (MSC) [52]. The spectra from both PTOFS techniques also are mean-centred for each tablet; the mean absorption of each tablet is calculated and subtracted. The data pretreatment
3.8.6 Quality of the evaluation

It is of great importance to determine the quality of the evaluated data and for that purpose some statistical tools are used. The meaning of all these tools is not always obvious and for that reason the main ideas behind them are explained in this section in general. The main idea of all used statistical tools is, to compare the evaluated data with reliable reference data and quantify the differences. The Root Mean Square Error (RMSE) is calculating the mean of the squared differences between reference and evaluation and it is defined as [53]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_{i,\text{eval}} - y_{i,\text{ref}})^2}$$

(3.1)

Here $N$ denotes the number of evaluated samples, $y_{i,\text{eval}}$ the evaluated, predicted value and $y_{i,\text{ref}}$ the reference value of the $i$-th sample. This RMSE can be calculated for the same data set that was used to build the calibration. In that case the RMSE for calibration (RMSEC) gives a measure about the goodness of the calibration set. The RMSE can also be calculated for predicted values (RMSEP) for a test set different to the calibration set. Then the RMSEP quantifies the quality of the prediction with the calibration set on the test set.

Another measure for the quality of the correlation is the explained variation $R^2$ and is calculated as [54]

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_{i,\text{eval}} - y_{i,\text{ref}})^2}{\sum_{i=1}^{N} (y_{i,\text{eval}} - \bar{y})^2}$$

(3.2)

where $\bar{y}$ denotes the mean value of the evaluated values. Again $R^2$ can be calculated for the calibration and predicted data.

Another measure that is popular to use is the Ratio of Performance to Deviation (RPD). This ratio should classify if a model is good enough for screening (RDP${}>3$), for quality control (RDP${}>5$) or for any application (RDP${}>8$). But after all RDP is just another way of expressing the information contained in the commonly used $R^2$ value [55]

$$RDP = (1 - R^2)^{-0.5}$$

(3.3)

It is also important to know the robustness of a method. For that purpose, a lot of calibrations should be done and the quality should be checked with the quantities mentioned above. In this study, the number of tablets is limited, but the robustness is checked by randomly taking out some tablets from the calibration set and so calculating the calibration many times. In that way the average of the RMSE and $R^2$ values should be more reliable than for only one calibration. The robustness of the calibrations can be seen in the variation of these values. All values are assumed to be normal distributed and so the variation is quantified as the standard deviation $\sigma$. In order to test, if the number of calibration sets is sufficient, a Student’s t-test is done on all RMSE and $R^2$ distributions and a confidence interval for the mean value is calculated with the MATLAB routine ttest [56]. This confidence interval gives a measure how much the mean value could change, if the calculations of RMSE or $R^2$ was repeated endlessly [57].
3.9 Material and Methods used for dairy product study

3.9.1 Dairy product properties

The dairy products used for this study are milk, cream and yoghurt with different fat contents. All products are bought at a common grocery store and the product properties from the package labels are listed in Tab. 3.1. According to these values, the main difference in the product composition is in the fat content. In [26] it is suggested, that lactose and proteins mostly absorb at longer wavelengths. For that reason, only water and fat are considered as main absorbers in the measured absorption spectra. Literature values of water and fat absorption spectra can be seen in Fig. 3.8 [58]. These data are also used for fitting the measured absorption spectra of the dairy products.

The scattering of these products though is considered to come from all other ingredients than water. But because the lactose and protein content is approximately the same for all products, the main variation in scattering is expected to come from the different fat contents. The difference between milk and yoghurt products is, that in yoghurt the proteins are forming gel. This is done by adding a yoghurt culture to the milk. Another way of producing a protein gel, is to add other enzymes (e.g. from rennet) and/or apply heat. This is how usually cheese is produced. [19]

3.9.2 Experimental setup

The dairy products are measured in the spectral region of 1000 to 1300 nm with a spectral resolution of 4 nm. For these measurements the MCP-PMT is used as detector and the LLTF as spectral filter. As all measured dairy products are liquid, it is possible to measure in infinite geometry. For practical reasons the fibre ends were positioned facing each other. The inter-fibre distance was varying between 5 and 13 mm, depending on the measured sample. In Fig. 3.9 the experimental conditions can be seen. A silicone baking form filled with the liquid studied is punctuated for needles containing the fibres from the source and to the detector.

The PTOFS curves are captured over an integration time of 25 s with a photon count rate of approximately $5 \cdot 10^4$ photons per second.

3.9.3 PTOFS data evaluation

From previous studies [28] it is known, that especially for low fat content dairy products the Diffusion approximation does not hold for the expected absorption and scattering properties. For that reason it is decided to use a WMC model, that has been created for Intralipid samples in infinite geometry [59]. Only for the measurements on 15.0 % cream the Diffusion approximation holds. It turns out that the cream data are fitted better with the Diffusion model and thus for the cream spectra the Diffusion evaluated data are presented.

<table>
<thead>
<tr>
<th>Product name</th>
<th>Distributor</th>
<th>Product type</th>
<th>Fat</th>
<th>Lactose</th>
<th>Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mellanmjölk</td>
<td>Coop Sverige</td>
<td>Milk</td>
<td>1.5%</td>
<td>4.9%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Mjölk</td>
<td>Coop Sverige</td>
<td>Milk</td>
<td>3.0%</td>
<td>4.8%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Matlagningsgrädde</td>
<td>Coop Sverige</td>
<td>Cream</td>
<td>15.0%</td>
<td>4.2%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Mild Lättyoghurt</td>
<td>Arla Foods</td>
<td>Yoghurt</td>
<td>0.5%</td>
<td>4.0%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Mild Naturell Yoghurt KRAV</td>
<td>Skånemejerier</td>
<td>Yoghurt</td>
<td>1.7%</td>
<td>4.2%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Naturell Yoghurt</td>
<td>Skånemejerier</td>
<td>Yoghurt</td>
<td>3.0%</td>
<td>3.6%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Mild Yoghurt Grekisk</td>
<td>Arla Foods</td>
<td>Yoghurt</td>
<td>6.0%</td>
<td>3.6%</td>
<td>3.3%</td>
</tr>
</tbody>
</table>

Table 3.1: Properties of the measured dairy products according to the package labels.
3.9.4 Spectral analysis

Absorption spectra

According to theory in Sec. 2.2.2, the absorption spectrum of a mixed sample should be a linear combination of the ingredients’ spectra (see (2.12)). In case of these dairy products, the ingredients absorbing in the studied spectral region are water and fat. The absorption spectra of water and fat are known from literature (see Fig. 3.8) and so it is possible to get the volume fractions \( f_{V,i} \) for water and fat.

In order to get the mass concentrations \( c_i \) out of these volume fractions, it is necessary to normalize the volume fractions weighted by the mass densities \( \rho_i \).

\[
c_i = \frac{f_{V,i} \cdot \rho_i}{\sum_j f_{V,j} \cdot \rho_j}
\]  

(3.4)

Scattering spectra

As explained before, it is possible to fit the scattering spectrum as a power law (see (2.10)). This formula is used to perform a fit on the scattering spectra. As reference wavelength \( \lambda_0 = 1000 \text{ nm} \) is chosen. The fit parameter \( A \) is indicating the particle concentration, while the other fit parameter \( \beta \) relates to the particle size. As in the measured products only the fat particles vary significantly in concentration, parameter \( A \) should be connected to the fat contents of the product.
Chapter 4

Results

In this chapter, the results of the measurements on tablets and dairy are presented. According to the two kind of studied samples, this chapter is split into two parts.

4.1 Pharmaceutical Tablets

4.1.1 Absorption and reduced scattering spectra

In Fig. 4.1, the absorption and reduced scattering coefficient spectra for all measured tablets are shown together with the spectra of pure component samples. In the absorption spectra of the mixed tablets it can be seen that the spectrum of the filler dominates, while the influence of the API is visible at approximately 1150 nm. The influence of the lubricant is however not visible. The lubricant is also known to have the smallest content in the mixed tablets, and its main absorption peak is also at around 1200 nm, where all ingredients exhibit high absorption. In the spectra of the mixed tablet a slight variation between the individual tablets can be observed.

The scattering spectra show a huge variation in the amplitude and slope of the mixed tablets. This is expected due to the designed variation in particle sizes and concentrations and allows to examine the influence of absorption in the compositional analysis.

In Fig. 4.2, absorption and scattering spectra are shown for one tablet measured with TNIRS and PTOFS using an AOTF and LLTF as spectral filter. The shape of the absorption spectra of all three techniques is similar, but the TNIRS spectrum has a significant different amplitude. The TNIRS data does not show pure absorption, rather extinction. In this plot, the TNIRS also is scaled by a constant factor of 0.83. The absorption spectrum of the AOTF and LLTF data are almost the same, except of the magnitude of the absorption peaks at 1200 nm and 1270 nm.

The scattering spectrum can only be measured by the PTOFS technique. These spectra are fitted with equation 2.10 using the parameters $A$ and $\beta$ for fitting and the result and the residuals are shown in Fig. 4.2. The residuals of the LLTF data is much less than for the AOTF data. This can be explained with the enhanced spectral resolution of the LLTF which affects the absorption and scattering at spectral regions with high slopes in the absorption curve (see Sec. 3.7). The coupling between evaluated absorption and scattering coefficient is less for the LLTF data. This better fit would be beneficial in the chemometric analysis. In the spectral region $\lambda > 1320$ nm scattering decreases significantly for both PTOFS techniques. For that reason this part of the spectrum is not considered for further evaluations.

Figure 4.1. Measured spectra of all examined tablets and their pure components.
4.1.2 Evaluated drug content

The measured spectra are pretreated and a PLS calibration is conducted as described in Sec. 3.8.5. In Fig. 4.3 the first three Latent Variables (LV) for all spectra of each technique can be seen. From this plot it is suggested to consider only two LV for all techniques, as the third LV is noisy and close to zero. In order to see the influence of an additional LV, the PLS prediction results for calibrations with three LV are also shown in one example.

Randomly chosen calibration and evaluation set

For the first attempt to compare PLS evaluated drug contents of the tablets, the whole tablet set is randomly split into two parts of 50% each (see Fig. 4.4a). One part is used for building a PLS calibration, the other one is evaluated and the quantities described in Sec. 3.8.6 are calculated. This is repeated 100 times with new random selections each time. The results can be found in Tab. 4.1. It is aimed to reach a $R^2$ value close to 1 and a low RMSEP value.

The TNIRS data show the least RMSEP for this selection. In the PTOFS techniques the LLTF is significantly better than the AOTF for calibration and prediction. The big discrepancies between calibration and prediction values can be explained by a poor choice of calibration. As the calibration is randomly selected, it is possible to choose calibrations that only contain tablets with similar API content. From literature it is known that a good calibration should span a huge variation in the samples [51]. This poor calibration also explains the relatively high Standard Deviation (SD) in the RMSEP.

A more sophisticated way to build a calibration is to choose a calibration set that spans all available API contents. This is done in a second attempt, where first all tablets are separated by the designed API content and then half of each group is selected as calibration and the rest as validation set (see Fig. 4.4b). This again is repeated 100 times. The prediction error values in this evaluation are lower for all measurement methods as it can be seen in Tab. 4.2. Again the TNIRS predictions are better than PTOFS predictions, but the PTOFS data show a better calibration. It also again shows, that the LLTF produces a better

---

**Figure 4.2.** Comparison of absorption and scattering spectra of the same tablet ($N2ln2C$) with different techniques. The scattering is fit as a power law (2.10) and for this fit the parameters are $A = 282.8\ cm^{-1}$, $\beta = 0.99$ for AOTF and $A = 269.9\ cm^{-1}$, $\beta = 0.91$ for LLTF with a reference wavelength arbitrarily set to $\lambda_0 = 1000\ nm$.

**Figure 4.3.** Latent variables (LV) of all TNIRS (a), AOTF (b) and LLTF (c) spectra in case of a randomly selected calibration set.
calibration and prediction than the AOTF. But the better predictions and lower SD for all techniques indicate, that this way of randomly selecting tablets for calibration is better. For the next evaluations this random method is again used. For this random selection also results from a prediction using a calibration model with three LV are presented. The RMSEP is significantly lower for TNIRS and AOTF data, while the RMSEP for LLTF data is approximately the same as for a calibration model using two LV. The RMSEC are lower for all data for a model built on three LV.

**Calibration and evaluation set chosen by scattering properties**

Another aspect of the quality of the evaluated data is, how much the evaluation is influenced by the scattering properties of the tablets. For this purpose, the calibration and validation sets are chosen by the scattering properties of the tablets. This evaluation method should show if the absorption spectra are influenced by the scattering properties of the tablets.

In Fig. 4.5 the correlation of evaluated API contents with the reference data is shown. For this evaluation, all tablets with high scattering properties are used for building the PLS calibration, while all tablets with low scattering are evaluated. In the correlation plot it can be seen that the evaluated API content correlates with the reference data for all techniques, while the TNIRS data spreads more.

In this evaluation of all tablets with one scattering property it is not possible to see how much the calibration relies on every single tablet. In the next step, the calibration set is thus again altered randomly. This time all tablets with high scattering are put into groups depending on their designed API content. Out of each group one tablet is randomly taken out. Still only tablets with high scattering are left in the tablet set, but it is slightly varied (see Fig. 4.4c). In Tab. 4.3 the statistical quantities describing the quality of 100 of these evaluations are shown. Tab. 4.4 shows results of the same method, but with tablets with low scattering properties in the calibration set (see Fig. 4.4d).
For these evaluations the LLTF and TNIRS data show good prediction and calibration results. The AOTF data show a larger error in prediction and the TNIRS data show differences in prediction for smaller and bigger scattering particles in the calibration tablets. For all evaluations the RMSEC is lower than the RMSEP.

Table 4.1: API content results of PLS regression using randomly selected 50% of all measured tablets as calibration set and the other 50% as evaluation set. The evaluation is repeated with 100 different randomly selected sets. LV = number of used Latent Variables, $R^2 = \text{explained variation}$, RDP Ratio of Performance to Deviation, RMSE = Root Mean Square Error of Calibration and Prediction and given in %, SD = standard deviation in %. All confidence intervals are calculated from a Student’s t-test. The quantities are explained in Sec. 3.8.6

<table>
<thead>
<tr>
<th>Method</th>
<th>LV</th>
<th>$R^2$ cal.</th>
<th>$R^2$ pred.</th>
<th>RPD</th>
<th>RMSEC</th>
<th>RMSEP</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNIRS</td>
<td>2</td>
<td>0.895±0.005</td>
<td>0.79±0.03</td>
<td>2.4±0.1</td>
<td>0.81±0.02</td>
<td>1.06±0.05</td>
<td>0.26</td>
</tr>
<tr>
<td>AOTF</td>
<td>2</td>
<td>0.810±0.009</td>
<td>0.68±0.03</td>
<td>1.9±0.1</td>
<td>1.05±0.02</td>
<td>1.30±0.05</td>
<td>0.23</td>
</tr>
<tr>
<td>LLTF</td>
<td>2</td>
<td>0.866±0.007</td>
<td>0.77±0.03</td>
<td>2.3±0.1</td>
<td>0.88±0.02</td>
<td>1.14±0.05</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 4.2: API content results of PLS regression using randomly selected 50% of all measured tablets as calibration set and the other 50% as evaluation set, but with the restriction that each API content of the calibration set is represented equally in the calibration set. The evaluation is repeated with 100 different randomly selected sets.

<table>
<thead>
<tr>
<th>Method</th>
<th>LV</th>
<th>$R^2$ cal.</th>
<th>$R^2$ pred.</th>
<th>RPD</th>
<th>RMSEC</th>
<th>RMSEP</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNIRS</td>
<td>2</td>
<td>0.899±0.005</td>
<td>0.86±0.01</td>
<td>2.8±0.1</td>
<td>0.80±0.02</td>
<td>0.92±0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>AOTF</td>
<td>2</td>
<td>0.817±0.009</td>
<td>0.75±0.01</td>
<td>2.0±0.1</td>
<td>1.03±0.02</td>
<td>1.20±0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>LLTF</td>
<td>2</td>
<td>0.873±0.005</td>
<td>0.83±0.01</td>
<td>2.7±0.1</td>
<td>0.88±0.02</td>
<td>1.01±0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>TNIRS</td>
<td>3</td>
<td>0.921±0.003</td>
<td>0.90±0.01</td>
<td>3.2±0.1</td>
<td>0.71±0.01</td>
<td>0.83±0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>AOTF</td>
<td>3</td>
<td>0.913±0.004</td>
<td>0.80±0.02</td>
<td>2.3±0.1</td>
<td>0.74±0.02</td>
<td>1.09±0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>LLTF</td>
<td>3</td>
<td>0.934±0.003</td>
<td>0.85±0.01</td>
<td>2.5±0.1</td>
<td>0.66±0.01</td>
<td>0.97±0.02</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Table 4.3: API content results of PLS regression using tablets with high scattering for calibration and evaluates tablets with low scattering. In the calibration set one tablet of each nominal API content is randomly taken out of this calibration set. The evaluation is repeated with 100 different sets.

<table>
<thead>
<tr>
<th>Method</th>
<th>LV</th>
<th>$R^2$ cal.</th>
<th>$R^2$ pred.</th>
<th>RPD</th>
<th>RMSEC</th>
<th>RMSEP</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNIRS</td>
<td>2</td>
<td>0.927±0.003</td>
<td>0.902±0.003</td>
<td>3.21±0.03</td>
<td>0.69±0.01</td>
<td>0.96±0.01</td>
<td>0.06</td>
</tr>
<tr>
<td>AOTF</td>
<td>2</td>
<td>0.910±0.006</td>
<td>0.748±0.011</td>
<td>2.03±0.05</td>
<td>0.76±0.02</td>
<td>1.13±0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>LLTF</td>
<td>2</td>
<td>0.959±0.002</td>
<td>0.845±0.006</td>
<td>2.57±0.05</td>
<td>0.53±0.01</td>
<td>0.87±0.01</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 4.4: API content results of PLS regression using tablets with low scattering for calibration and evaluates tablets with high scattering. In the calibration set one tablet of each nominal API content is randomly taken out of this calibration set. The evaluation is repeated with 100 different sets.

<table>
<thead>
<tr>
<th>Method</th>
<th>LV</th>
<th>$R^2$ cal.</th>
<th>$R^2$ pred.</th>
<th>RPD</th>
<th>RMSEC</th>
<th>RMSEP</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNIRS</td>
<td>2</td>
<td>0.911±0.003</td>
<td>0.867±0.004</td>
<td>2.77±0.04</td>
<td>0.79±0.01</td>
<td>0.82±0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>AOTF</td>
<td>2</td>
<td>0.936±0.001</td>
<td>0.871±0.008</td>
<td>2.87±0.08</td>
<td>0.66±0.03</td>
<td>1.07±0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>LLTF</td>
<td>2</td>
<td>0.950±0.002</td>
<td>0.926±0.002</td>
<td>3.71±0.05</td>
<td>0.61±0.01</td>
<td>0.86±0.02</td>
<td>0.08</td>
</tr>
</tbody>
</table>
4.2 Dairy products

The aim of the measurements on the dairy product is similar to the aim of the tablets. In both cases, the content of one ingredient should be predicted. However, the number of samples in this study is less. For that reason chemometric techniques such as PLS are less reliable and thus the evaluation methods are different in this section.

4.2.1 Dilutions

First, the sensitivity of the measured spectra to the fat content is tested in a dilution experiment. Cream with 15% fat content is measured and then diluted with water in different steps. The spectra measured on these dilutions are shown in Fig. 4.6. The absorption spectra are similar for all products in the wavelength range between 1000 nm and 1150 nm. The magnitude of the spectra at 1200 nm varies between the spectra. The spectrum of the 2.5% dilution is limited to the region below 1220 nm due to experimental issues during the measurements. The spectra can be used to find a correlation with fat contents of the products. A first attempt to see the sensitivity of the absorption spectra to the fat content of the dilutions is done by creating a PLS calibration. The first five LV are shown in Fig. 4.7. The two first LV that could physically correspond to the water and fat components, while the remaining mostly contain noise. As it shows, this calibration cannot be used for prediction, as the number of the samples is not sufficient for a reliable prediction.

The scattering spectra all follow a power law, as it is expected from theory (see Sec. 2.2.1). The amplitudes of all scattering spectra are increasing for higher fat content. This could also provide a possibility to predict the scatterer concentration and thus the fat content from the scattering by fitting a power law to the spectra.

Figure 4.6. Absorption and reduced scattering spectra of cream (15% fat) and dilutions of this cream with different steps of lower fat content.

Figure 4.7. LV of a PLS regression on the dilutions.
4.2.2 Absorption and reduced scattering spectra of dairy products

The absorption and reduced scattering spectra of the dairy products described in Tab. 3.1, can be found in Fig. 4.8. As for the dilutions, there is a variation at the 1200 nm peak in the absorption spectra. The absorption spectrum of 1.5% milk seems also to be different below 1030 nm, but this can be identified as artefacts either from the measurement or the PTOFS curve fitting. For that reason this part of data is ignored in all further analysis. In the scattering spectra more variations are visible than for the dilutions. Again, there is a connection between fat content and amplitude of the spectrum, but also different slopes are visible. This could be due to structural differences of different products as milk and yoghurt. Again the spectrum of 1.5% milk is different below 1030 nm due to the artefact and this data is thus ignored.

4.2.3 Absorption as linear combination of water and fat

In this evaluation part, all dairy products are considered to have only water and fat as ingredients that contribute to the absorption spectrum in the measured region. With this assumption it is possible to fit the measured absorption spectra as a linear combination of the pure spectra (see Fig. 3.8). The fit result of the absorption spectrum of yoghurt with 3% fat content is shown in Fig. 4.9. For the best fit results, the spectra of the ingredients are shifted by 6 nm to the longer wavelengths. This correction could be necessary due to different wavelength calibrations of PTOFS and the setup from literature. The figure shows, that the fit basically corresponds to the measured spectrum, but at the valley and the peak the residuals are higher than on average. These fits are done on all measured dairy products and dilutions. The fit coefficients are converted to mass coefficients according to equation (3.4). These evaluated mass contents are correlated to the fat contents from the package or the dilution in Fig. 4.10. These evaluated fat contents show a RMSE of 1.59%
4.2.4 Scattering as power law

In the scattering spectra of the various dairy products (Figs. 4.6 & 4.8) a variation depending on the fat content is obvious. Also previous studies show that it is possible to correlate scattering properties and fat content well [28]. The scattering spectra were fitted for all measured samples. An example of such a fit using equation (2.10) is shown in Fig. 4.11 for 3% yoghurt. Again, the highest residuals can be seen at 1150 nm due to the finite spectral resolution effect (see Sec. 3.7).

In Fig. 4.12 the fit parameters $A$ and $\beta$ are shown for all measured dairy samples. The meaning for these parameters is discussed in the discussion chapter 5. The $A$ parameter is linearly increasing with fat content for all samples, while for the different type of samples there are different slopes and offsets in the linear behaviour. The $\beta$ parameter is decreasing for milk and cream with increasing fat content, slightly decreasing for yoghurt and approximately constant for all dilutions except of the 2.5% dilution.

and $R^2$ of 0.862. In the correlation plot the fat content is more likely overestimated, as the majority of points lies above the correlation line.

4.2.4 Scattering as power law

Figure 4.10. Correlation of the evaluated fat content from the fit of the absorption spectra as linear combinations of the ingredient’s spectra and uncertainties from the fit.

Figure 4.11. Fit of the scattering spectra of 3.0% fat yoghurt as a power law (2.10).

Figure 4.12. Evaluated $A$ and $\beta$ coefficients for the fat content of each product. The errors from the fit are relatively small compared to possible systematic errors introduced from the measurement. For that reason no error bars are given here. It is referred to the discussion in section 5.2.
In this chapter the results from chapter 4 are discussed separately for the pharmaceutical tablets and dairy. The chapter ends with a conclusion of this thesis work.

5.1 Pharmaceutical tablets

5.1.1 Quality of absorption and scattering spectra

In the results in Sec. 4.1.1, the spectra of one pharmaceutical tablet taken with different experimental techniques are presented (Fig. 4.2). In these results it can be seen, that only the PTOFS techniques can measure the scattering spectrum. So PTOFS can provide additional information for the tablet in the same measurement. This scattering spectrum could be used for structural analysis of the tablet (particle size and concentration, see [18]) and also to check the quality of the measurement. From theory (Sec. 2.2.1) it is known that scattering should decrease as a power law with increasing wavelength. If the scattering locally behaves differently, it indicates a problem with the measurement, such as the finite spectral resolution effect (see Sec. 3.7). Compared to the LLTF scattering spectrum, the AOTF scattering spectrum shows this finite spectral resolution effect much more clearly. This is indicating a better data quality of the spectrum obtained with the LLTF, as absorption and scattering are less coupled.

The finite resolution effect can also be clearly seen in the absorption spectra. The peak magnitude is higher for the LLTF than for the AOTF spectrum. According to the scattering spectra, the LLTF spectrum is more precise and also indicates a better data quality of the absorption data. The TNIRS data shows the same shape as the PTOFS data, but at a different intensity level. The measured absorption coefficient is generally higher and some absorption peaks are relatively higher than for PTOFS data. The differences can be explained by the fact, that TNIRS does not separate the scattering effects from the absorption. For this reason the accuracy is low for TNIRS, while the PTOFS data is assumed to provide more accurate results.

The measurement on tablets has been performed several times before on previous versions of the used PTOFS setup. For that reason the control of the experimental conditions is very good and only few errors are believed to have been introduced, assuring that the PTOFS LLTF tablet spectra are of high quality.

5.1.2 Comparison of the quality of PLS calibration drug content prediction

In the drug content evaluation in Sec. 4.1.2 it is shown, that it is possible to build working PLS calibrations for all techniques and also to predict drug contents with these calibrations employing the obtained absorption data. The results further show a possibility to build a PLS calibration out of the TNIRS spectra that is
working well to predict the drug content from recordings. In order to create these calibrations, it is however necessary to know the scattering properties of the selection in order to choose a suitable number of LV for the calibration. That means, that these calibrations are not universally usable, drawback of the TNIRS technique. With the PTOFS LLTF technique it is possible to build calibrations, useful for any scattering properties in the calibration set. The preferable number of LV in all examined calibration sets is the same.

In the PTOFS techniques the prediction with the LLTF data is better than the AOTF data. This can be seen in the lower RMSE and $R^2$ values in the tables in section 4.1.2 as well as in the lower number of considerable LV used for the LLTF data. The precision of the prediction also does not alter considerably for the LLTF data, depending on selection of calibration and evaluation set. The prediction errors for a random selection (Tab. 4.2) are for all techniques higher than for a selection by scattering (Tabs. 4.3 & 4.4). This is not expected, as the scattering selected calibration the model is challenged and a worse error would be expected in that case. As this is not yet fully understood, further analysis is required and under way. This is planned especially, as my group would like to publish the data in a peer-reviewed scientific paper. For this a deep understanding of these observations are required.

### 5.1.3 Number of considerable LV

From the tablet design it is known that all tablets contain three ingredients (filler, drug, lubricant), as it is described in Sec. 3.8.1. Considering the very low and constant designed content of the lubricant and that the absorption spectrum of the lubricant overlaps with all other ingredients (see Fig. 4.1) it is expected to only measure two components of the tablets.

In Fig. 4.3 it can be seen how many LV contribute to a PLS calibration. For the TNIRS and AOTF technique it shows, that either two or three LV are useful. The third LV is only useful if the scattering of the tablets is varying in the calibration set. This third LV is representing the scattering influence to the absorption spectra. This influence, that only in some situations appears, makes it difficult to use the TNIRS or AOTF technique to build a reliable calibration for drug content prediction.

The number of considerable LV for the PTOFS LLTF technique stays constant at two LV. The AOTF data also shows good results for a calibration using two LV, but for a random selection for calibration a third LV improves the RMSEP. As discussed before, the AOTF absorption spectra are in some spectral regions influenced by the finite spectral resolution effect. This influence can be represented in the PLS calibration as an additional LV.

In Tab. 4.2 it also can be observed, that the difference between RMSEC and RMSEP increases if three LV are used for calibration. This indicates that the calibration is over-fitted and too many LV have been used for this calibration.

### 5.2 Dairy products

#### 5.2.1 Quality of absorption and scattering spectra

The absorption spectra of the dairy products and the dilutions (Figs. 4.6 & 4.8) show the characteristic absorption peaks of water and fat (see Fig. 3.8). The quality of the absorption spectra is good, but around the absorption peak at 1200 nm some noise is visible on the spectra. This noise is connected to a low photon count rate due to the high absorption and the limited laser power. The scattering spectra also show noise again due to few detected photons. For some samples also the slope of the scattering spectrum seems to change at some point around 1200 nm. This is not according to theory and might be an artefact from the measurement.

The measurements on dairy products in this broad spectral range is experimentally challenging and some systematic errors might have been introduced. As
described in section 3.9.2 the products are measured in a silicone container with integrated fibres. Due to the flexibility of the silicone it is possible, that the position of the fibres are slightly changed before or during the measurement. This would cause an error of the distance between source and detector of great importance for the PTOFS curve fitting. Other errors may have been introduced due to temperature change of the samples. All samples are measured directly coming from the fridge, but during the experiment the temperature is not well controlled. It is known, that temperature has a some influence to the optical properties of dairy products, especially the water absorption, but this has been neglected in this work.

5.2.2 Fat content predictions

The content prediction of yoghurt and milk by fitting the scattering as a power law (see Fig. 4.12a) has been done before in a different wavelength range with a PTOFS-AOTF system [28]. The fitted $A$ coefficients in this thesis are reproducing the shape of those previous data. Interestingly, also the newly presented dairy products data are linearly related to the designed fat content. With a sufficient number of evaluated samples, it would thus be possible to create a linear calibration useful for fat content predictions. A separate calibration for each sort of dairy might however be necessary, as shown in Fig. 4.12. The dilutions have a lower $A$ coefficient than the milk products and the yoghurt. This can be explained by the meaning of the $A$ coefficient, directly related to the scatterer concentration. This is directly related to the fat content and the mass percentage of fat. By using fat particles of different sizes it is possible to have a variation in the concentration of scatterers, even if the mass fat content is the same. On the other side it is possible to have a range of different fat contents in the samples, but the same particle concentrations with different particle sizes. Additionally, some products as yoghurt also contain other scattering particles than milk, like for example protein gels (see Tab. 3.1). So this is a limitation for the fat content prediction by analysing the scattering spectra.

This problem does not show up if the absorption spectrum is used for fat content prediction, because the absorption spectrum is in principle not affected by structural properties. In Fig. 4.10 the fat content, evaluated from a fit of the absorption spectrum as a linear combination of the absorption spectra of the ingredients, is linear to the designed fat content. For the low fat samples there is a higher variation from the linear trend. According to theory the evaluated fat content should correlate to the designed fat content. In Fig. 4.10 almost all evaluated values are higher than the design. This indicates a systematic error either in the measurement or in the data evaluation. As discussed in section 5.2.1 there are some uncertainties introduced from the measurement. But also the data evaluation might have some problems. The absorption spectra of the ingredients and of the samples have been taken by different experimental techniques. So each technique might have introduced a systematic error, that is influencing the fitting of the spectra. Otherwise it could be possible that other ingredients of the samples also contribute to the absorption in that region.

5.2.3 Structure analysis

As described before, the scattering spectra provide structural information about the samples. In the previous section the fitting parameter $A$ of a power law fitting is related to the concentration of scatterers. The other fitting parameter $\beta$ is connected to the particle size of the scatterers. In Fig. 4.12 both parameters are shown for all samples. The scatterer concentration and its ability to predict the fat content has been discussed in the previous section. For the dilutions it is expected to have always the same particle size and a linear increase of the scatterer concentration. These expectations are met for all dilutions except of the particle size of the 2.5% dilution. This outlier can be explained by the smaller wavelength range this sample is measured and also by the low absolute values
of scattering. For lower absolute values, errors have a relatively higher impact. For milk and yoghurt the $\beta$ coefficient is higher for the low fat samples, what is related to a smaller particle size. As explained before, a smaller fat particle size allows a higher number of fat particles for the same mass fat content. So in the production process it might be desired to have more fat particles in order to have a similar product consistency (yoghurt) or a similar appearance for all available fat contents.

5.3 Conclusion

It has been established in this thesis report, that PTOFS is a powerful tool for compositional and structural analysis of pharmaceutical tablets and dairy products. In this technique the advantages of NIRS such as non-destructibility and reliability are extended by the ability to separate between absorption and scattering properties of the samples. During the last year the automation of the measurement procedure has been completed, what clearly simplifies the measurement process. Also the newly introduced LLTF spectral filter significantly reduces the impact of the finite spectral resolution effect and allows measurements of absorption and scattering spectra in a high quality. The possibility of measuring both, absorption and scattering in one measurement empowers to get compositional and structural information of the sample at the same time. It is also possible to control the quality of the measured absorption spectrum by analysing the scattering spectrum. If this is according to theory this indicates a good quality of the measurement.

The success of the improved experimental system is evident from the results of the measured pharmaceutical tablets. It is shown, that the new LLTF filter allows to build better PLS calibrations and also reliable drug content predictions. Compared to TNIRS and PTOFS with an AOTF spectral filter, the newly presented data allow calibrations and predictions that are more robust to changes in the scattering properties of the samples. This could be desirable in industrial applications, when the scattering properties of the samples are not always known. Still the errors in the prediction compared to Raman evaluated tablets (see section 3.8.4) are relatively high, so a comparison to the gold standard method UV-VIS would be interesting, even though it would destroy the samples.

The measurements with PTOFS on dairy products achieved similar results to recently published data, but in a different wavelength range and with different experimental equipment. In this work not only scattering but also absorption spectra are used to analyse the fat content of the samples. It shows that it is beneficial for the evaluation also to use the absorption spectra for this evaluation, as this evaluation method is not dependent on the type of dairy product. Still the evaluated fat content from the absorption spectra is in this analysis systematically higher than the designed fat content. Finally, it has been shown, that considering both absorption and scattering spectra for fat content prediction is advantageous suggesting PTOFS should be considered as a promising tool for fat content analysis in dairy products, even though it is experimental challenging.
OUTLOOK

Many things have been pursued recently to improve the PTOFS system at the Biophotonics group at Lund University, while still some improvements are suggested. During this thesis work, the automated measurement process software has been completed. The next task should be to optimize the measurement sequences in order to achieve a faster data collection. Currently a spectrum of 100 wavelengths needs about one hour to be measured and the same time for IRF collection. This could clearly be too long for some applications. The main time during measurements is used for data collection. In order to decrease this time, it is important to know how the quality of the spectra changes for a decreased collection time. On samples with high absorption it is also necessary to integrate even longer in order to achieve a suitable SNR. This problem could be solved with a light source with more optical power in the NIR region. By introducing the LLTF to the system the influence of the finite spectral resolution effect has been reduced. Still it would be interesting to investigate, if it is possible to minimize this effect mathematically by processing the PTOFS curves after the measurement. As it is described in the methods, it is only possible to evaluate PTOFS curves of samples described by one of the available model of light propagation. There will be samples that demand for a measurement geometry that is not available as a fitting model. It is thus important to continuously update these and also implement new models.

It would be also interesting to increase the measurement speed for the measurements on the tablets. For that purpose it should be investigated which wavelength range is most important for evaluation and if the spectral resolution of the spectra are taken in could be decreased. For the data evaluation it would be interesting to work together with an experienced chemometric expert, who has more experience about data pretreatment and PLS. This would also come along using a more sophisticated PLS software than the Matlab function used in this thesis work. Probably such a cooperation would also help to understand the physical meaning of the additional LV in the PLS calibrations of PTOFS data. Finally it would be interesting to compare the evaluated data not only to Raman evaluated values, but to the UV-VIS gold standard values, even though this would mean to destroy the tablets.

The study on dairy products in this thesis work is based on a limited number of measured samples and it would thus be necessary to confirm the results presented in this report with more experimental data. This additional data could also be used to work with chemometric techniques such as the PLS that has been used for the tablets. This would further mean, that no literature data for fitting the absorption spectra is needed, what has been discussed as a possible source of errors in the analysis. Finally the scattering spectra measured on dairy should be further investigated, as they show at some spectral regions dips, that cannot be explained by the finite spectral resolution effect. Still these dips cannot be explained by the presently employed theory and it thus seems to be an artefact that should be understood in order to ensure the quality of the measurements.
Acknowledgements

I would like to thank everyone who supported and encouraged me during the last year of my thesis work. First of all I want to thank my supervisors Dmitry Khoptyar and Stefan Andersson-Engels. Your doors were always open for me and I enjoyed the fruitful and interesting discussions with you. Especially I would like to thank for encouraging me to visit the summer school in Jena, presenting my work on a poster in London and bringing my results to publications. These experiences are very valuable for me and went far beyond my expectations of this thesis project.

I also want to thank Staffan Strömblad and Karolina Dorozynska for their pleasant company in the lab. We had very interesting discussions reasoning about malfunctions of the lab equipment and software problems and it always turned out, that we understand the experiment much better after these discussions. The Biophotonics group at Atomic Physics division also created a friendly environment to work in. For that I specially would like to acknowledge Nina Reistad, Haiyan Xie, Therese Alburg, Arefeh Monirehalsadat Mousavi, David Kraus, Hugo Söderlund and all other members of the group.

I gratefully thank Jonas Johansson and Anders Sparén for the fruitful discussions about the chemometric evaluations. Next, I acknowledge AstraZeneca for providing the tablet set, tablet characterisations and TNIRS spectra and Faisal Kamran for his foregoing work on this tablet set.

Finally, I thank my family and friends who supported me during my studies abroad for their encouragement and keeping me in touch with home.

This work was kindly supported by Lunds Tekniska Högskola (LTH), the Swedish Research Council (VR) and AstraZeneca.


List of publications

Broadband photon time-of-flight spectroscopy as a prospective tool in biomedicine and industrial process and quality control
Spectroscopy Europe 26, 3 (2014).
In this paper, possible applications of PTOFS on samples from every day life are presented and possible applications are suggested.
I participated in performing the measurements, creating the figures and in discussions about the text.

Broadband photon time-of-flight spectroscopy for drug content analysis of pharmaceutical tablets
In this article, we present absorption and reduced scattering coefficient spectra measured on pharmaceutical tablets with Photon Time of Flight Spectroscopy (PTOFS). We introduce a new spectral filter technology with better spectral resolution to the experimental setup, that significantly decreased the coupling between the measured absorption and scattering properties at spectral regions of absorption features. Additionally we evaluate the measured spectra with a Partial Least Squares regression method to evaluate the drug content of each tablet. These evaluated drug contents then are compared to drug contents evaluated of comparative data (previous PTOFS measurements with a broader spectral filter and compared to Transmission Near Infra Red Spectroscopy), as well as to reference values. The results show a more accurate and robust drug content prediction with the improved setup data compared to the comparative data.

Broadband photon time-of-flight spectroscopy for fat content analysis of dairy products
In this article, we present newly measured absorption and reduced scattering coefficient spectra, measured on milk, yoghurt and cream products and dilutions of cream. The spectra are measured with a Photon Time Of Flight Spectroscopy setup, that is able to measure broadband spectra. Recently the spectral resolution of the system has been improved. The absorption spectra are fitted as a linear combination of pure water and fat absorption spectra to evaluate the fat content. These results are slightly overestimating the fat content, compared to the labelled fat content. The scattering spectra are fitted as a power law and the fit parameters reveal the concentrations and sizes of the scattering fat particles. So the fat content can be estimated in two ways out of the data of one measurement.