

1 **Recovery of absorbance spectra of micrometer-sized biological and**
2 **inanimate particles**

3 R. Lukacs^{*1}, R. Blümel² B. Zimmerman¹, M. Bağcıoğlu¹, and A. Kohler¹

4

5 ¹Department of Mathematical Sciences and Technology, Faculty of Environmental Science
6 and Technology, Norwegian University of Life Sciences, 1432 Ås, Norway

7 ²Department of Physics, Wesleyan University, Middletown, Connecticut 06459-0155, USA

8

9 *Corresponding author:

10 Rozalia Lukacs,

11 Department of Mathematical Sciences and Technology (IMT),

12 Faculty of Environmental Science and Technology

13 Norwegian University of Life Sciences,

14 Drøbakveien 31, 1432 Ås, Norway

15 Tel: +47 64 96 54 29

16 Fax: +47 64 96 54 01

17 E-mail: rozalia.lukacs@nmbu.no

18

19

1

2 **Abstract**

3 In this paper, we first provide an overview of the Mie type scattering at absorbing materials and
4 existing correction methods, followed by a new method to obtain the pure absorbance spectra of
5 biological systems with spherical symmetry. This method is a further development of the recently
6 described iterative algorithm of van Dijk et al. [1]. The method is tested on FTIR synchrotron
7 spectra of *polymethyl methacrylate* (PMMA) microspheres and pollen grains with approximately
8 spherical shape. The imaginary part of the refractive index was successfully recovered for both
9 systems. Good agreement was obtained between the pure absorbance spectra obtained by this
10 method and the measured spectra.

11

12 Key Words: **Mie scattering, ripples, FTIR spectroscopy, recovery of pure absorbance,**

13 **PMMA**

14

15

1 **1. Introduction**

2 For several decades, infrared (IR) spectroscopy has been extensively applied within biological
3 sciences to investigate more or less intact biochemical structures [2-4]. The main advantage of
4 IR spectroscopy is that biological materials can be investigated without any extraction steps or
5 chemical sample pre-treatment. Biochemical information on biological samples is obtained
6 via interpretation of highly specific chemical absorption bands. Following the invention of
7 Fourier transform IR (FTIR) microscopes in the 1980s [5], infrared spectroscopy experienced
8 a further boost. As a consequence of this development, FTIR microspectroscopy of thin tissue
9 sections gained in popularity. Compared to traditional microscopy, the advantage of FTIR
10 microspectroscopy is that chemical information from intact tissue can be obtained without
11 using staining of the tissue sections [6, 7]. Thin tissue sections for FTIR microspectroscopy
12 have typical thickness of 6-10 μm , and can be obtained by cryo-sectioning. The thin sections
13 are placed on infrared-transparent material and measured in transmission by infrared
14 microspectroscopy. In addition, FTIR microspectroscopy has been used successfully for the
15 investigation of single cells such as plant and human cells [8, 9].

16 Recently, synchrotron light sources have attracted considerable attention. While conventional
17 light sources (black body radiators) have a higher total power in the infrared, synchrotron
18 light sources are more strongly collimated, resulting in a higher brilliance in the infrared. This
19 allows the use of aperture sizes down to a few microns.

20 Since tissue sections, cells and other biological materials vary strongly in size and
21 morphology, FTIR spectroscopy and microspectroscopy are strongly hampered by non-
22 chemical variations, such as scattering contributions, which often seriously distort pure
23 chemical absorption spectra. The strongest non-chemical variations in the FTIR spectroscopy
24 of thin dried films or tissue sections are due to differences in the sample thicknesses. When

1 the sample thickness changes the effective optical path length, a so-called multiplicative effect
2 is present, which can effectively be estimated and suppressed by extended multiplicative
3 signal correction (EMSC) [10]. Baseline variations, another non-chemical variation that are
4 typical for FTIR microspectroscopy, arise when the light intensity varies, e.g. during the time
5 interval between background and sample measurements. The resulting baseline variations are
6 constant baseline shifts over the whole spectral region, and can equally well be estimated and
7 suppressed by EMSC [10]. A third type of non-chemical variations, that are typical for the
8 FTIR microspectroscopy of cells, are Mie-type scattering variations [11]. Scattering effects
9 have been identified as a major obstacle for the reliable interpretation and further use of IR
10 spectra in biological and biomedical science, and therefore methods for the recovery of pure
11 absorbance spectra are needed.

12 Analytical solutions for the scattering of light at spheres are known and have for the first time
13 been provided by Gustav Mie [12]. Mie-type scattering at cells causes strong and broad
14 background oscillations, which can be effectively suppressed by EMSC, applying a meta-
15 model of the analytical Mie solutions [13]. Subsequently, it has been pointed out that the
16 broad oscillations caused by Mie scattering cannot be treated independently from chemical
17 absorption in the FTIR microspectroscopy of single cells [14]. The real and the imaginary
18 parts of the refractive index are determined by both optical and chemical properties of the
19 material. Since the real and the imaginary parts of the refractive index are further related by
20 the Kramers-Kronig relation [15, 16], they depend on each other and cannot be treated
21 independently. Excellent reviews of the field of scattering in IR spectroscopy are available
22 such as the recent review by Bhargava [17]. In this review the problem of the interdependence
23 of sample geometry, optical properties of the FTIR spectrometer, and scattering and
24 absorption is clearly outlined and the corresponding literature dealing with this complex issue
25 is reviewed. The two additional papers of the Bhargava group, that include the optical

1 properties of the spectrometer, provide a general framework for scattering from layered
2 samples [18] and spheres [19]. Therefore, while the general electromagnetic theory of the
3 apparatus-sample interaction and its influence on the complex refractive index of the sample
4 is known, the focus in this paper is on a concrete problem. Namely the question of whether
5 scattering (real part of the refractive index) and absorption (imaginary part of the refractive
6 index) may be treated independently in some geometries. The results of our study show that
7 state-of-the-art models describing Mie-type scattering and absorption deviate considerably
8 from the exact results.

9 Several approximate models have been established to explain Mie-type scattering and
10 absorption of FTIR spectroscopy of cells and tissues [14, 20-24]. As we will see, the problem
11 is complicated, and existing modelling algorithms and software are not yet completely up to
12 the task. Kohler et al. [13] and Bassan et al. [20-22] suggested an algorithm based on
13 multiplicative signal correction [25]. They used the van de Hulst approximation formula for
14 the calculation of the Mie scattering efficiency [26]. This equation is an approximation used
15 instead of the full Mie theory. While the approach is very efficient, it represents only a rough
16 estimate since it uses an approximation formula for the non-absorbing case and no numerical
17 aperture is considered.

18 Recently Van Dijck et al. [1] suggested a method for the recovery of the complex refractive
19 index of materials with spherical shape. The effects of the sample morphology on the
20 measured spectra can be removed, and using the imaginary part of the index, the shape-
21 independent IR absorption spectrum of the material is recovered. The authors have applied
22 their algorithm to polymethyl methacrylate (PMMA) spheres. The size of the PMMA spheres
23 is a required input parameter for their algorithm. In case the size of the spherical object is not
24 known *a priori*, the authors have suggested using the spectral region 2000-2600 cm^{-1} to obtain
25 an estimate of the size of the sphere. In the present paper we evaluate this strategy for

1 biological samples. We combine the method suggested by Van Dijck et al. with EMSC, and
2 develop an algorithm that allows recovering the pure absorbance spectra of biological systems
3 in the presence of Mie type scattering. Since the development of future inverse scattering
4 algorithms will benefit from any simplifying property of the scattering system, we first
5 investigate and answer the question to what extent scattering and absorption can be treated
6 independently. We then investigate the recovery of spectra for two separate systems, PMMA
7 spheres and pollen. Pollen grains are a very close approximation of biological Mie scatter
8 systems since they may have spheroidal shapes. By investigating this relatively simple system
9 we may, in the future, generalize the method for more complicated biological systems such as
10 cells.

11 Our paper is organized in the following way. In the section “Theory” we set the stage by
12 introducing our notation and theoretical background used and applied in our paper. In section
13 “Experiment” we provide the experimental details of the synchrotron FTIR measurements. In
14 section “Results and Discussion” we show that in the case of absorbing spheres the
15 assumption of independence of scattering and absorption is not justified. We find that current
16 state-of-the art models [see e.g. Kohler et al. [13] and Bassan et al. [20]] may predict spurious
17 line shifts of up to 12 cm^{-1} . Therefore the models that employ independence of scattering and
18 absorption need to be improved. We present our algorithm for the recovery of pure
19 absorbance spectra of two systems: 1) a model system (PMMA spheres), and 2) a biological
20 system (pollen grains with approximately spherical shapes). In some of our model
21 calculations we need to model chemical absorption lines, and this we do by using the Lorentz
22 model discussed in detail in Appendix A. There is an additional reason for presenting the
23 Lorentz model in Appendix A: There is tremendous, as yet untapped, potential in the Lorentz
24 model for extracting pure absorbance spectra. This is so, since (i) absorption bands, in
25 principle, can be traced back to their molecular origins and (ii) the absorption lines described

1 in the Lorentz model automatically satisfy the Kramers Kronig relation [15, 16]. This may be
2 put to use in the construction of forward models for scattering on isolated scatterers such as
3 cells or other micron-sized biological samples. In Appendix B we provide the analytical
4 solutions for Mie scattering that introduces the notation. It shows that the analytical scattering
5 and extinction formulas for dielectric spheres can be stated in compact form. In addition, the
6 formulas may be used as a basis for the extraction of pure absorbance spectra from
7 (approximately) spherical scatterers of a biological or inanimate nature.

8

9 **2. Theory**

10 The purpose of this section is to present the theoretical concepts and algorithms used, applied,
11 and referred to in our paper. In subsection ‘Basic definitions’ we start with some basic
12 definitions to set the stage and to introduce our notation. In subsection ‘Mie scattering’ we
13 discuss scattering at an absorbing sphere and we show a possible way to calculate the
14 absorbance when the numerical aperture of the optical system is included. In subsection
15 ‘Iterative algorithm’ we present the steps of the algorithm proposed by van Dijck et al.

16 **Basic definitions**

17 In the infrared spectroscopy of biological materials, measurements are usually performed in
18 forward direction. As illustrated in Fig. 1, infrared light impinges on a cell or a film of
19 biological material representing a scatterer. The incident intensity is denoted by I_0 ; it is
20 usually referred to as the reference intensity and it is characteristic for the light source. It is
21 experimentally obtained by moving the scatterer out of the light path. Part of the incident light
22 I_0 may be scattered, as illustrated by the blue arrows in Fig. 1; it may be absorbed by the
23 scatterer, as illustrated by the red area, which denotes a radiation sink; and it may be
24 transmitted, as illustrated by the purple arrows. The transmitted beam is recorded by the

1 detector as the sample intensity I . The intensity I is directly proportional to the power \wp
 2 measured by the detector. It is given by $\wp = IG$, where G is the area of the aperture in front of
 3 the detector (see Fig. 1). Along the same lines both the scattered power and the absorbed
 4 power can be expressed in the following way: The scattered power \wp_{sca} is given by $\wp_{sca} =$
 5 $I_0 \sigma_{sca}$ and the absorbed power is given by $\wp_{abs} = I_0 \sigma_{abs}$, hereby defining the cross sections
 6 σ_{sca} and σ_{abs} for scattering and absorption, respectively. Denoting by $\wp_0 = I_0 G$ the power
 7 recorded by the detector in the absence of the scatterer, the balance of power requires

$$8 \quad \wp_0 = \wp + \wp_{sca} + \wp_{abs} \quad (1)$$

9 With this equation, the transmission, defined as

$$10 \quad T = \frac{\wp}{\wp_0} \quad (2)$$

11 may also be written as

$$12 \quad T = 1 - \frac{\sigma_{ext}}{G}, \quad (3)$$

13 where $\sigma_{ext} = \sigma_{sca} + \sigma_{abs}$ is the extinction cross section. A commonly used quantity is the
 14 dimensionless extinction efficiency Q_{ext} , which is defined as

$$15 \quad Q_{ext} = \frac{\sigma_{ext}}{g} \quad (4)$$

16 where g is the geometrical cross section of the scatterer. Conventionally, in chemistry, a
 17 quantity called absorbance A is defined as

$$18 \quad A = -\log_{10}(T). \quad (5)$$

19 The following formulas are useful, since they allow us to go back and forth between A and
 20 Q_{ext} . For given A we obtain Q_{ext} according to

1
$$Q_{ext} = [1 - 10^{-A}] \frac{G}{g} \quad (6)$$

2 and for given Q_{ext} we obtain A according to

3
$$A = -\log_{10} \left[1 - \frac{g}{G} Q_{ext} \right] \quad (7)$$

4 Only in the case where σ_{sca} is very small compared to σ_{abs} , are A and T simply related to
5 σ_{abs} . In this case, with Eq. 3, we obtain

6
$$T \approx 1 - \frac{\sigma_{abs}}{G}. \quad (8)$$

7 As before, A is obtained according to Eq. 5. All these quantities are frequency dependent and
8 usually given as a function of the wavenumber $\tilde{\nu}$, which is the reciprocal of the wavelength λ .

9 Equation 3 is a ray-optical result and valid only if diffraction at the detector aperture is
10 negligible. This requires that both dimensions of G are much larger than λ . In addition, G has
11 to be large compared with σ_{ext} . This is readily apparent from Eq. 3, which predicts the
12 nonsensical result $T < 0$ for $G < \sigma_{ext}$. Here, we are using the correct value of σ_{ext} , determined
13 in an experiment with sufficiently large G . If, however, we measure σ_{ext} with a G that
14 approaches σ_{ext} in size, and since T cannot be negative, the measured σ_{ext} , in this case, will
15 differ from its asymptotic value obtained using a large G . This discussion acquires immediate
16 relevance in the context of synchrotron light scattering, where the width of the incident beam
17 may be of the order of, or narrower than, the size of the scattering particle, thus substantially
18 modifying σ_{ext} . This may occur, e.g., when strongly focused synchrotron light is used in the
19 study of single biological cells.

20 A further important comment concerns the refractive index, which is an important quantity
21 when considering the scattering of light at biological materials. Since biological materials
22 absorb light, the refractive index, in general, has a non-zero imaginary part. We denote the

1 complex refractive index by $m(\tilde{\nu}) = n(\tilde{\nu}) + in'(\tilde{\nu})$, where $n(\tilde{\nu})$ is the real part of the
2 refractive index, describing the refractive properties of the material, and $n'(\tilde{\nu})$ is the
3 imaginary part of the refractive index, describing the absorptive properties of the material.

4 In order to model chemical absorption lines we assume a medium of absorbing dipoles
5 describing both absorption and scattering within the framework of the Lorentz model, briefly
6 reviewed in Appendix A. In the IR spectroscopy the dipoles are the absorbing functional
7 groups. The Lorentz model, providing the complex dielectric constant $\tilde{\epsilon}_r$ according to Eq.
8 A18, then allows the computation of the complex refractive index m via Eq. A20. In principle,
9 even complex biological materials containing many absorbing functional groups are exactly
10 described by Eq. A18. However, in practice a parameterization of the Lorentz model
11 according to Eq. A21 turns out to be very effective. The only input parameters to this
12 effective model are the positions, widths and strength of the absorption bands. Because of the
13 functional form of the effective Lorentz model, the real and the imaginary parts of the
14 complex refractive index m automatically fulfil the Kramers-Kronig relation [15, 16].

15 **Mie scattering**

16 *A rigorous Mie-type model for scattering of infrared light at cells*

17 The strong Mie-type scattering artifacts, that are often present in FTIR spectra of cells, have
18 so far been approximated using analytical expressions derived from the Mie theory [13, 14,
19 20]. The extinction cross section $\sigma_{ext}(\tilde{\nu})$ for the scattering of light at a spherical particle is
20 described by the Mie theory [26] and given by

$$21 \quad \sigma_{ext}(\tilde{\nu}) = \pi a^2 Q_{ext}(\tilde{\nu}), \quad (9)$$

22 where a is the radius of the scattering particle. Gustav Mie derived the exact solutions of this
23 problem, which are expressed in terms of spherical Bessel functions. The interested reader is
24 referred to the book by Van De Hulst [26]. It has been shown that for the case of $|m - 1| \ll$

1 1, where m is the complex refractive index, $Q_{ext}(\tilde{\nu})$ can be approximated by the following
 2 formula [26]

$$\begin{aligned}
 3 \quad Q_{ext}(\tilde{\nu}) \approx & 2 - 4e^{-\varrho \tan \beta} \frac{\cos \beta}{\rho} \sin(\rho - \beta) - 4e^{-\varrho \tan \beta} \left(\frac{\cos \beta}{\varrho}\right)^2 \cos(\varrho - 2\beta) + \\
 4 \quad & 4 \left(\frac{\cos \beta}{\varrho}\right)^2 \cos(2\beta) \quad (10)
 \end{aligned}$$

5 with

$$6 \quad \rho = 4\pi a \tilde{\nu} |m - 1| \quad \text{and} \quad \tan \beta = n' / (1 - n) \quad (11)$$

7 In the following study we will use the Van de Hulst approximation, since it provides a handy,
 8 analytical solution without noticeable loss of accuracy. All the previous and following
 9 considerations are valid by either applying the Van De Hulst approximation or the exact Mie
 10 solutions. With the help of either the Van De Hulst formula or the exact Mie solutions, the
 11 absorbance can be calculated via the extinction cross section $\sigma_{ext}(\tilde{\nu})$ according to Eqs. 3 and
 12 5. The exact Mie solutions and formula 10 (since they contain n and n') are generally valid
 13 even if the complex refractive index m has a non-zero imaginary part. In case of a real
 14 refractive index, i.e. when absorption is neglected, the Van De Hulst solutions can be
 15 simplified to

$$16 \quad Q_{ext}(\tilde{\nu}) \approx 2 - \frac{4}{\rho} \sin \rho + \left(\frac{4}{\varrho}\right)^2 (1 - \cos \varrho) \quad (12)$$

17 Several efforts have been made to treat the case of Mie scattering with absorption [14, 20, 21].
 18 Although Eq. 10 represents the approximation formula for the case of a complex refractive
 19 index, describing precisely the problem of absorption, it has not yet been applied in the
 20 literature in connection with the interpretation and extraction of information from measured
 21 data.

22 The extinction efficiency $Q_{ext}(\tilde{\nu})$, expressed by Eq. 12 with real index of refraction, is
 23 usually interpreted as the additive scatter contribution to the absorbance spectrum [11, 13,

20]. This is a very rough estimation as we will show in the following. When expanding the logarithm in Eq. 7 in powers of $Q_{ext}(\tilde{\nu})$, it can be seen that to first order in $Q_{ext}(\tilde{\nu})$ we obtain

$$A \approx \frac{\pi a^2}{G \ln(10)} Q_{ext}(\tilde{\nu}) \quad (13)$$

This shows that the absorbance and the extinction are only approximately proportional to each other. In order to investigate this further we model one absorption band at 1654 cm^{-1} applying the Lorentz model outlined in Appendix A. This band corresponds to the C=O stretching vibration of the peptide bond in proteins. We choose $\Lambda = 10^4 \text{ cm}^{-2}$, $\Gamma = 30 \text{ cm}^{-1}$, and $\bar{\epsilon}_r = 1.44$. Far away from the absorption band at 1654 cm^{-1} the real part of the refractive index is close to $\sqrt{\bar{\epsilon}_r} = 1.2$, and changes considerably in the vicinity of the band position. In Figs. 2a and b the approximated absorbance for a complex refractive index, calculated according to Eqs. 10 and 7 is compared with the approximated absorbance for a real refractive index obtained from Eq. 10 and 13, using $\frac{\pi a^2}{G} = 0.05$. The approximated absorbance for a complex refractive index is plotted in red, while the approximated absorbance for a real refractive index is plotted in blue. We see that the maximum value of the absorbance appears at the same shifted position $\tilde{\nu}_{max} = 1648 \text{ cm}^{-1}$. This shows that the approximation with the real refractive index still reveals the approximation with complex refractive index shifted position of the maximum. This is immediately clear, since the derivative of the approximation with complex refractive index formula and the derivative of the approximated formula with the real refractive index are both proportional to the derivative of $Q_{ext}(\tilde{\nu})$. Moreover, the maximum occurs at wavenumbers at which the derivative with respect to $\tilde{\nu}$ is zero. In the low frequency range (approx. below 2000 cm^{-1}) the two curves in Figs. 2 agree very well, while they differ significantly in the high frequency regime. This observation may be important for fitting spectra of single cells, where so far only the approximated formula with the real refractive index has been used [13, 14, 20, 21].

1 In addition, we point out that the formulas for $Q_{ext}(\tilde{\nu})$ in Eqs. 10 and 12 are applicable only if
2 the incident light is a plane wave, i.e. the size of the incident beam is infinite and the
3 incoming light rays are parallel. When beam sizes are of the same order as the particle sizes,
4 the computations of $Q_{ext}(\tilde{\nu})$ have to be modified and the expression for $Q_{ext}(\tilde{\nu})$ may change
5 significantly. This also shows that $Q_{ext}(\tilde{\nu})$ can be taken only as a rough approximation for the
6 scattering of infrared light at single cells when using highly focused infrared beams such as in
7 synchrotron infrared spectroscopy. A further investigation of this topic has to show how well
8 $Q_{ext}(\tilde{\nu})$ in Eq. 10 describes the experimental situation [27]. In addition, the incident beam in
9 synchrotron infrared spectroscopy has a high numerical aperture and the condition of parallel
10 light rays is not exactly fulfilled [27].

11 In order to take the numerical aperture into account when calculating the apparent
12 absorbance, we calculate the light intensity I from the exact Mie solutions. Since synchrotron
13 radiation is linearly polarized, the field vector \vec{E}_0 of the incident synchrotron light may be
14 decomposed into a component \vec{E}_{0r} perpendicular to the scattering plane and a component \vec{E}_{0l}
15 parallel to the scattering plane. Denoting by φ the (azimuthal) angle between \vec{E}_0 and the
16 scattering plane ([26], section 9.3), the magnitudes of the perpendicular and parallel
17 components of \vec{E}_0 are given by $E_{0r} = E_0 \sin(\varphi)$ and $E_{0l} = E_0 \cos(\varphi)$, respectively. Taking
18 the detector aperture G and the geometrical cross section g into account, the scattered
19 intensities I_r and I_l are given explicitly by ([26], section 4.4)

$$20 \quad I_r(\theta, \varphi) = \left(\frac{G}{g}\right) \frac{i_1(\theta)}{4\pi^2 \tilde{\nu}^2 r^2} |E_{0r}|^2 = \left(\frac{G}{g}\right) \frac{i_1(\theta)}{4\pi^2 \tilde{\nu}^2 r^2} I_0 \sin^2(\varphi) \quad (14)$$

21 and

$$22 \quad I_l(\theta, \varphi) = \left(\frac{G}{g}\right) \frac{i_2(\theta)}{4\pi^2 \tilde{\nu}^2 r^2} |E_{0l}|^2 = \left(\frac{G}{g}\right) \frac{i_2(\theta)}{4\pi^2 \tilde{\nu}^2 r^2} I_0 \cos^2(\varphi) \quad (15)$$

1 where $I_0 = |E_0|^2$, r is the distance between sample and detector, $i_{1,2}(\theta) = |S_{1,2}(\theta)|^2$, and
 2 $S_{1,2}(\theta)$ are the scattering amplitudes stated explicitly in Appendix B, Eqs. B1 and B2. In
 3 order to account for the numerical aperture, we integrate

$$4 \quad I(\theta, \varphi) = I_r(\theta, \varphi) + I_l(\theta, \varphi) \quad (16)$$

5 over the solid angle cone with opening angle θ_{NA} (the aperture angle) to obtain the intensity
 6 I_{NA} , corresponding to the given numerical aperture, according to

$$7 \quad I_{NA} = \int_0^{\theta_{NA}} \int_0^{2\pi} I(\theta, \varphi) \sin(\theta) d\theta d\varphi = \left(\frac{G}{g}\right) \frac{I_0}{4\pi\tilde{\nu}^2 r^2} \int_0^{\theta_{NA}} [i_1(\theta) + i_2(\theta)] \sin(\theta) d\theta. \quad (17)$$

8 In our simulations we used $r = 0.5$ m. This turned out to be the optimal value resulting from
 9 our calculations and is a reasonable value for our optical setup as well. The parameter G was
 10 estimated with the EMSC model. Inserting the transmission $T = I_{NA}/I_0$ into Eq. 5, we obtain
 11 the absorbance

$$12 \quad A = -\log_{10} \left\{ \left(\frac{G}{g}\right) \frac{1}{4\pi\tilde{\nu}^2 r^2} \int_0^{\theta_{NA}} [i_1(\theta) + i_2(\theta)] \sin(\theta) d\theta \right\} \quad (18)$$

13 For the calculation of the exact Mie solutions the algorithm proposed by Bohren and Huffman
 14 [28] was implemented in MATLAB.

15

16 *Iterative algorithm*

17 We will follow the iterative algorithm proposed by van Dijck et al. [1], in order to recover the
 18 pure absorbance spectrum A from measured and distorted absorbance spectra $A_{measured}$. The
 19 main steps of the algorithm are presented below.

20 Initialization: For $j = 1$, the complex refractive index $m_j = n_j + i n'_j$ is initialized, where j is
 21 the index of iteration. The real part of the refractive index, n_j , is initialized with an estimated
 22 constant n_0 . The imaginary part, n'_j , is initialized with zero.

23 Following initialization, the iteration proceeds as follows:

1 I. The formulas in appendix B together with Eq. 18 are used to predict the
2 absorbance spectrum $A^{(j)}$ from the complex refractive index m_j .

3 II. The difference $E^{(j)}$ between the measured spectrum $A_{measured}$ and the
4 predicted spectrum $A^{(j)}$ is calculated according to:

$$5 \quad E^{(j)}(\tilde{\nu}) = A_{measured}(\tilde{\nu}) - A^{(j)}(\tilde{\nu}) \quad (19)$$

6 III. From the difference and using Eq. C19, the next value for n' is calculated
7 according to:

$$8 \quad n'_{j+1}(\tilde{\nu}) = n'_j(\tilde{\nu}) + \frac{\ln(10)}{4\pi\tilde{\nu}d_{eff}} E^{(j)}(\tilde{\nu}), \quad (20)$$

9 where $d_{eff} = \frac{4a}{3}$ is the effective thickness of a sphere of radius a .

10 IV. The negative values of $n'(\tilde{\nu})$ are set to zero.

11 V. A new value for $n(\tilde{\nu})$ is predicted according to:

$$12 \quad n_{j+1}(\tilde{\nu}) = n_0 + Kramers\ Kronig[n'_{j+1}(\tilde{\nu})] \quad (21)$$

13 VI. The new complex refractive index is calculated according to:

$$14 \quad m_{j+1}(\tilde{\nu}) = n_{j+1}(\tilde{\nu}) + i n'_{j+1}(\tilde{\nu}) \quad (22)$$

15 The complex refractive index $m_{j+1}(\tilde{\nu})$ is updated and the next iteration
16 with $j + 1$ is started.

17

18 *Estimation of the radius and the constant part of the real part of the refractive index*

19 In order to estimate the constant part of the refractive index n_0 and the radius of the sphere a ,
20 intervals of n_0 and a may be considered. For every combination of these two parameters we
21 calculate the Q_{ext} with Eq. 12. The absorbance, A_{pred} , is predicted using Eq. 13. The parameter
22 G is estimated using EMSC model. For direct comparison of the predicted spectrum with the
23 spectrum measured at the discrete wavenumbers $\tilde{\nu}_i$, we suggest to calculate the root mean
24 square error (RMSE) according to

1
$$E^{(j)}(\tilde{\nu}_i) = A_{measured}(\tilde{\nu}_i) - A_{pred}(\tilde{\nu}_i) \quad (23)$$

2
$$RMSE^{(j)} = \sqrt{\frac{\sum_{i=1}^K [E^{(j)}(\tilde{\nu}_i)]^2}{K}} \quad (24)$$

3 where K is the number of measured discrete wavenumbers $\tilde{\nu}_i$. The iterative algorithm
 4 described in the previous section is very time-consuming when full Mie theory is applied and
 5 the integration over the numerical aperture is executed. Therefore, we suggest to use the van
 6 de Hulst approximation formula in Eq. 12 in order to obtain a first estimate of n_0 and a . This
 7 result can then be entered into the iterative algorithm using the full Mie theory and integrating
 8 over the numerical aperture. Since the van de Hulst approximation formula in Eq. 12 depends
 9 on the size factor

10
$$\alpha = a(n_0 - 1). \quad (25)$$

11 and not on n_0 and a , separately, it is sufficient to minimize the RMSE as a function of α .
 12 Thus, this first estimate results only in an estimate of α , but not in an estimate of n_0 and a ,
 13 separately. We emphasize that the size factor α needs to be carefully distinguished from the
 14 size parameter $x = k \cdot a$ and from the parameter $\rho = 2x|m - 1|$, defined in Eq. 11.

15

16 **3. Experiment**

17 Samples of pollen were collected at the Botanical Garden of the Faculty of Science of the
 18 University of Zagreb during the 2012 pollination season. The following samples, belonging to
 19 the *Cupressaceae* plant family, were measured: *Cunninghamia lanceolata*, *Juniperus*
 20 *chinensis*, *Juniperus communis*, and *Juniperus excels*. The pollen samples, of approximately
 21 spherical morphology, varied in diameter between 10 to 40 μm . The samples were collected
 22 directly from plants at flowering time by shaking mature male cones. The samples were kept

1 in paper bags at room temperature for 24 hours, and afterwards transferred to vials and stored
2 at -15 °C.

3
4 In addition to the pollen samples, polymethyl methacrylate (PMMA) microspheres of assorted
5 sizes were measured as well as a simple artificial system for modeling scattering from
6 biological materials. The PMMA samples were purchased from Microspheres-Nanospheres
7 (Corpuscular Inc, NY), and used without further modifications. The spheres had the following
8 diameters as stated by the vendor: 5.5, 10.8, 15.7, 20.0, 30.0 and 40.0 μm .

9
10 In order to obtain high-quality spectra of single particles, pollen and PMMA samples were
11 recorded by using synchrotron radiation at the SOLEIL synchrotron facility. The synchrotron
12 spectra were measured on the SMIS infrared beamline, details of which can be found
13 elsewhere [29]. The transmission spectra of all the samples were recorded with a resolution of
14 4 cm^{-1} by using the synchrotron radiation coupled to a Nicolet 5700 FTIR spectrometer with a
15 Nicolet Continuum XL IR microscope (Thermo Scientific, CA), equipped with a liquid
16 nitrogen cooled mercury cadmium telluride detector. The spectra were measured in the 8000-
17 650 cm^{-1} spectral range, with 128 scans each and using 15 \times and 32 \times objectives with different
18 aperture sizes, depending on the size of the sample (10 \times 10 μm , 15 \times 15 μm and 20 \times 20 μm). The
19 numerical aperture of the microscope was 0.65 [29]. For each measured sample an image was
20 recorded with an optical microscope.

21

22 **4. Results and Discussion**

23 *Comparison with existing Mie-type models for cells*

24 In the literature, the extinction $Q_{ext}(\tilde{\nu})$ in Eq. 12, i.e. the Van De Hulst approximation for a
25 real refractive index, has frequently been used as an approximation for the absorbance [13, 14,

20, 21]. In the paper by Bassan et al. [14], Mie scattering in the infrared spectroscopy of single cells is constructed theoretically, taking the chemical absorption of the scatterer into account. The case where absorption is considered was termed resonant Mie scattering. Starting from a practically scatter-free thin-film absorbance spectrum A , the imaginary part of the complex index of refraction, n' , is calculated, assuming proportionality between n' and A . Because of the explicit frequency dependence, as discussed in [1] (see Eq. 8), this is only approximately valid. After determining n' , a Kramers-Kronig transformation [15, 16] according to

$$n(\tilde{\nu}) = n_0 + \frac{2}{\pi} P \int_0^{\infty} \frac{s n'(s)}{s^2 - \tilde{\nu}^2} ds \quad (26)$$

is applied to obtain the real part of the index of refraction m . At this point the real part is inserted into Eq. 12 in order to obtain the extinction efficiency $Q_{ext}(\tilde{\nu})$. It is important to note that this involves neglecting the imaginary part of the refractive index m , since Eq. 10 is equivalent to Eq. 12 only if the index of refraction is real. The result obtained was called apparent absorbance [14]. The approximation involved in calculating the apparent absorbance is not negligible, as we will see in the following. In Figs. 3a and 3b, using the same model absorption line as in Fig. 2 at 1654 cm^{-1} , the apparent absorbance with real refractive index is plotted in blue, while the apparent absorbance result, obtained via Eqs. 3, 4 and 10, employing the complex refractive index, is plotted in red. It can be seen that the apparent absorbance is shifted to lower frequencies compared to the result with complex index of refraction. The apparent absorbance has a shifted maximum at $\tilde{\nu}_{max} = 1641 \text{ cm}^{-1}$, which corresponds to a low-wavenumber shift of 7 wavenumbers with respect to its expected location at $\tilde{\nu}_{max} = 1648 \text{ cm}^{-1}$ (see Fig. 2), a significant difference when interpreting spectral bands of biological materials. Concerning the system at hand, we have found that the shift is to the right (lower wavenumber region) if the band is located on the right wing of the associated Mie fringe, and the shift is to the left (higher wavenumber region) if the band is located on the left-hand side

1 of the associated Mie fringe. While this observation applies to our model and our current
2 choice of parameters, it is an open question whether shift directions and band locations are
3 correlated this way in general.

4 *A Mie-type model for a nucleus in a cell membrane*

5 It is important to note that the approximation expression with complex index of refraction for
6 the extinction efficiency $Q_{ext}(\tilde{\nu})$ (for the approximate version with real index of refraction
7 see Eq. 10) takes absorption and scattering into account and thus it is expected that this
8 expression can be used to obtain an expression for the absorbance of infrared light by a cell
9 nucleus. However, if the nucleus is embedded in a medium, the absorbance of the cell plasma
10 may add to the absorbance of the nucleus obtained via Eqs. 3, 4 and 10. In this case the
11 absorbance spectrum for the nucleus $A_{nucleus}$ can be obtained according to Eqs. 3, 4 and 10,
12 considering the nucleus as a scattering and absorbing sphere. The absorbance spectrum of the
13 plasma A_{plasma} may be a relatively undistorted spectrum, since the plasma can be considered
14 as a thin film, in which the nucleus is embedded. The measured spectrum $A_{measured}$ may
15 therefore be simply written as the sum of the two contributions according to

$$16 \quad A_{measured} = A_{nucleus} + A_{plasma} \quad (27)$$

17 The above findings may be used to improve the algorithm of Refs. [20, 21], which employs
18 the model constructed by Bassan et al. [14].

19 *Ripples*

20 The exact Mie extinction shows a rapidly fluctuating structure on top of the smooth wavelike
21 underlying structure. The sharp narrow structures are called ripples, while the broader,
22 smoother structures are called wiggles. The wiggles are always present in the extinction curve
23 and if we increase the size factor, ripples start to appear on top of wiggles. We modelled an
24 absorbance band at 1654 cm^{-1} using the Lorentz model presented in Appendix A. We
25 calculated the Mie extinction, Q_{ext} , and the scattering efficiency, Q_{sca} using the exact Mie

1 solutions (see Appendix B equations B3, B4, B7 and Eqs. 15-16) and the modelled complex
 2 index of refraction, according to the following equations:

$$3 \quad Q_{ext} = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \Re(a_n + b_n) \quad (28)$$

$$4 \quad Q_{sca} = \frac{1}{x^2} \int_0^{\theta_{NA}} [i_1(\theta) + i_2(\theta)] \sin \theta \, d\theta. \quad (29)$$

5 In Fig. 4 we plot the exact Mie extinction, Q_{ext} , and the scattering efficiency, Q_{sca} , for an
 6 absorbing sphere, for two cases, $n_0=1.14$, $a=10\mu\text{m}$ and $n_0=1.24$, $a=10\mu\text{m}$. In Fig. 4a, the case
 7 of $n_0=1.14$, $a=10\mu\text{m}$, we see only the large oscillations (wiggles) in Q_{ext} and Q_{sca} . In Fig. 4b,
 8 we see the ripples in Q_{ext} and Q_{sca} when $n_0=1.24$, $a=10\mu\text{m}$. The formal reason for the ripples
 9 in the Mie absorbance spectra is known: they correspond to partial-wave resonances in the
 10 Mie coefficients [30, 31]. However, the excitation mechanisms of these resonances and their
 11 consequences for the electromagnetic field distribution inside the scatterer have not been
 12 explored yet.

13 The synchrotron FTIR spectrum of a PMMA sphere with $10\mu\text{m}$ diameter is shown in Fig. 5a.
 14 In Figs. 5b and 5c, the spectrum has been divided into 1) the spectral region $7200\text{-}3600\text{ cm}^{-1}$
 15 and 2) the spectral region $3600\text{-}800\text{ cm}^{-1}$ respectively. While the spectrum in Fig. 5b shows
 16 only features caused by the Mie scattering of the sphere, the spectrum in Fig. 5c shows the
 17 additional features due to absorbance of the PMMA molecules. Zooming into the
 18 wavenumber region between about 6000 cm^{-1} and 7000 cm^{-1} (see Fig. 5d), we see sharp
 19 structures in the absorbance that may correspond to Mie ripples. To our knowledge, the
 20 observation of Mie ripples was not yet reported in the experimental FTIR literature. Thus, the
 21 features in Fig. 5d may constitute the first experimental observation of Mie ripples in an FTIR
 22 spectrum. Since the present paper does not focus on the ripple structure, but focuses instead
 23 on the extraction of pure absorbance spectra, we defer the discussion of ripples in absorbance
 24 spectra to a forthcoming paper.

25

1 *Recovery of pure absorbance spectra of PMMA spheres*

2 In order to test the algorithms, we used PMMA spheres as a model system. The value of the
3 constant part of refractive index, n_0 , for PMMA spheres is known approximately and thus a
4 good estimate for the radius of the spheres, a , can be obtained. We used the measured
5 spectrum from Fig. 5b to test the prediction method for the radius of the sphere and the
6 constant part of the refractive index as described in the theory section. We considered the
7 following intervals for n_0 and for a (radius of the sphere):

8
$$n_0 \in [1.1; 2.0] \quad \Delta n_0 = 0.05$$

9
$$a \in [1\mu m; 10\mu m] \quad \Delta a = 0.25\mu m$$

10 We calculated the RMSE as a function of α . The result is shown in Fig. 6, where RMSE is
11 plotted as a function of $\alpha \cdot 10^6$. We will refer to α as the size factor, as frequently done in the
12 literature. We find four major local minima in the RMSE function: at $\alpha=0.64\mu m$, $\alpha=1.1\mu m$,
13 $\alpha=2.1\mu m$ and $\alpha=3.1\mu m$ (see inset of Fig. 6). The global minimum is located at approximately
14 $\alpha=2.1\mu m$. Figure 7 shows the predicted absorbance, calculated with Eqs. 12-13, together with
15 the measured absorbance, for all four local minima. The best fit to the large oscillation in the
16 measured absorbance spectrum in Fig. 5b is found for $\alpha = 2.1\mu m$. We consider the value
17 $2.1\mu m$ a good solution. In the case of PMMA we know that the approximate value of n_0 is
18 1.48. Considering the minimum in the RMSE function around $\alpha=2.1\mu m$, we obtain the radius
19
$$a = \frac{\alpha}{n_0 - 1} = 4.4 \mu m.$$

20 In our case we know the radius of the PMMA sphere and this helps to choose the correct n_0 .
21 But if we do not have information about the radius or the refractive index, we can choose a
22 and n_0 values which satisfy the $\alpha=2.1\mu m$ condition. We will use these values (n_0 and a) as
23 initial parameters in the van Dijck iterative algorithm. If we can recover the pure absorbance
24 spectra and if we have a good predicted absorbance, then the initial parameters are a good

1 choice. If we cannot recover the imaginary part of the refractive index, or the predicted
2 absorbance is not good, we discard the n_0 and a combination.

3 Consider the measured spectrum in Fig. 5c. To test our initial parameters, $n_0=1.48$ and
4 $a=4.4\mu\text{m}$, we selected three different combinations of n_0 and a as inputs for the the van Dijck
5 iterative algorithm, i.e. (i) $n_0=1.7$ and $a=3\mu\text{m}$, (ii) $n_0=1.48$ and $a=4.4\mu\text{m}$ and (iii) $n_0=1.3$ and
6 $a=7\mu\text{m}$. Following the choice of n_0 , all three values of a were obtained via $a = \frac{2.1 \mu\text{m}}{n_0-1}$. The n_0
7 values were chosen such that the resulting a values box in and thereby test our initial
8 parameter set $n_0=1.48$ and $a=4.4\mu\text{m}$. Running the van Dijck iterative algorithm for the three
9 selected values of n_0 and a , we predicted our apparent absorbance using the exact Mie
10 solutions from Eq. 18. This way we take into account the numerical aperture. After 12
11 iteration steps the algorithm converged and the imaginary part, n' , of the refractive index and
12 the predicted absorbance was obtained (Figures 8-9 respectively). As can be seen in Fig. 9,
13 the predicted absorbance, for the values $n_0=1.7$ and $a=3\mu\text{m}$ does not provide a good fit. The
14 solution $n_0=1.3$ and $a=7\mu\text{m}$ provides a better fit both for the imaginary part and the predicted
15 absorbance. The best combination is $n_0=1.48$ and $a=4.4\mu\text{m}$, as both the imaginary part of the
16 refractive index and the predicted absorbance agree best. With this obtained n' we calculate
17 the pure absorbance spectrum of a PMMA thin film and consider the thickness to be $d_{eff} =$
18 $\frac{4a}{3} = 5.9\mu\text{m}$:

$$19 \quad A_{pure} = \frac{4\pi n' d_{eff} \tilde{\nu}}{\ln(10)} \quad (30)$$

20 Figure 10 shows this calculated pure absorbance spectrum of a PMMA thin film computed
21 with Eq. 30. The main absorbance peaks in the calculated spectrum are compared with other
22 experimental measurements. The following bands were found in the calculated absorbance
23 spectrum: 2987 cm^{-1} , 2937 cm^{-1} , 2839 cm^{-1} , 1763 cm^{-1} , 1724 cm^{-1} , 1497 cm^{-1} , 1466 cm^{-1} , 1441
24 cm^{-1} , 1394 cm^{-1} , 1333 cm^{-1} , 1261 cm^{-1} , 1225 cm^{-1} , 1182 cm^{-1} , 1047 cm^{-1} and 951 cm^{-1} . The

1 FTIR spectra of PMMA powder have been measured in the range of 4000 – 400 cm^{-1} by Haris
2 et al. The following characteristic absorption bands were found: 3000 cm^{-1} , 2953 cm^{-1} , 2840
3 cm^{-1} , 1727 cm^{-1} , 1483 cm^{-1} , 1447 cm^{-1} , 1437 cm^{-1} , 1398 cm^{-1} , 1367 cm^{-1} , 1267 cm^{-1} , 1239 cm^{-1} ,
4 cm^{-1} , 1197 cm^{-1} , 1147 cm^{-1} , 1050 cm^{-1} , 990 cm^{-1} , 967 cm^{-1} , 913 cm^{-1} , 840 cm^{-1} , 807 cm^{-1} and 750
5 cm^{-1} [32]. Our calculated absorbance is in good agreement with the experimental data. With
6 this method it was possible to get the pure absorbance spectra and to give a realistic estimate
7 of n_0 and a . The method will now be tested on biological systems.

8

9 *Recovery of pure absorbance spectra of Pollen*

10

11 Pollen grains are an ideal real-world model system for characterization of scattering
12 phenomena of biological samples. As opposed to the vast majority of cells and tissues that are
13 easily deformed, pollen grains have stable and reproducible morphology because of their thick
14 and shape-persistent grain walls. Moreover, the desiccated nature of the grains provides
15 relatively stable biochemical composition, and thus enables simple manipulation and
16 measurement. Finally, the diversity of pollen morphologies, with a variety of shapes, textures,
17 and sizes (ranging from less than 5 μm to over 200 μm), enable a wide range of experimental
18 conditions for the measurement of scattering phenomena.

19 Pollen grains belonging to the *Cupressaceae* plant family were chosen due to their
20 approximately spherical shape and appropriate range of radius sizes. The synchrotron
21 spectrum of *Juniperus chinensis* pollen grain, with 27 μm diameter, is shown in Fig. 11a. In
22 Fig. 11a we can distinguish three main parts of the spectrum. The first part, between 6300-
23 3600 cm^{-1} , is shown in Fig. 11b. The second part, between 3600-1000 cm^{-1} , is plotted in Fig.
24 11c. In Fig. 11b we can see the large oscillations due to Mie scattering. In this interval Mie
25 scattering is the main physical origin of the spectrum. Figure 11c shows the absorption part of

1 the spectrum. In this area the molecular absorption is very strong and together with the
2 scattering is causing an FTIR spectrum. If we zoom into Fig. 11c, we see that for
3 wavenumbers between 2771-1880 cm^{-1} , absorption is not present and Mie scattering is again
4 the guiding phenomenon. Figure 11d shows this zoomed-in spectrum. Within this region the
5 chemical constituents of pollen do not show any absorbance. Thus, in this region n' is zero.
6 In all the spectral regions shown in Fig. 11, we can see the ripples caused by Mie scattering.
7 To the best of our knowledge, this is the first time that Mie scattering ripples were observed in
8 IR spectra of a biological system. As stated before, we do not focus on them in this paper. Our
9 main purpose here is to recover the pure absorbance spectra of the pollen and make a good
10 guess concerning the refractive index and the radius of the pollen grains. In order to achieve
11 this, we use the scenario developed for the PMMA spheres. We will consider the spectrum
12 from Fig. 11b as a measured absorbance and we will try to predict the radius of the pollen
13 grain, a , and the constant part of the refractive index, n_0 . The following intervals were chosen
14 for n_0 and for a :

$$15 \quad n_0 \in [1.1; 2.0] \quad \Delta n_0 = 0.05$$

$$16 \quad a \in [5\mu\text{m}; 15\mu\text{m}] \quad \Delta a = 0.25\mu\text{m}$$

17 For every combination of these two parameters we predicted the absorbance spectrum with
18 the van de Hulst approximation formula from Eq. 12. We scaled the prediction with EMSC
19 and compared with the measured spectrum. We calculated the root mean square error (RMSE)
20 function according to Eqs. 23-24. The same investigation is followed as in the case of
21 PMMA: In Fig. 12 the resulting RSME function is plotted as a function of the size factor,
22 clearly revealing four distinct minima at $\alpha=1\mu\text{m}$, $\alpha=1.3\mu\text{m}$, $\alpha=2.2\mu\text{m}$ and $\alpha=3.2\mu\text{m}$ (see inset
23 of Fig. 12). Figure 13 shows the predicted absorbance for some selected values of the size
24 factor. All combinations of n_0 and a around the parameter value $\alpha = 2.2\mu\text{m}$ give a good
25 approximation to the large oscillation in the measured absorbance spectrum, and therefore we

1 consider the value 2.2 as the best solution. We took this value as the guiding value. From the
2 optical image of the pollen grain we estimated the radius of the pollen grain, $a=13.5\mu\text{m}$.
3 Choosing the size factor 2.2 for the constant part of the refractive index, we get $n_0=1.16$.
4 Considering that all the biological materials inside the pollen grain (cellulose, sporopollenin,
5 etc.) have a constant part of the index of refraction of about 1.5 in the visible spectrum of
6 light, this value is quite small. However, taking the porous structure of the outer part of the
7 pollen grain into account, this value makes sense. Therefore, from biological point of view the
8 smaller value of n_0 is the better choice.

9 We fed different values of n_0 and a as initial parameters into the van Dijck iterative algorithm.
10 If the pure absorbance spectrum is recovered, and if we get a good predicted absorbance, then
11 the initial parameters were a good choice. If the imaginary part of the refractive index is not
12 recovered, or the predicted absorbance is not good, the corresponding combination of n_0 and a
13 was discarded.

14 Before running the iterative algorithm we tested the prediction method suggested by
15 van Dijck et al.. To predict n_0 , these authors propose to use the region from the measured
16 spectrum between 2600 cm^{-1} and 2100 cm^{-1} , as here most organic materials do not show
17 absorption. In the case of *Juniperus chinensis* pollen grains we used the region of the
18 spectrum between 2771 cm^{-1} and 1880 cm^{-1} , shown in Fig. 13c, to predict a and n_0 . If we use
19 the same method presented above, with the same interval and resolution of n_0 and a , we
20 obtain for the RMSE function the values plotted in Fig. 14. Employing our prediction method
21 only in this region results a size factor around 3.2. We conclude that in our case the region
22 recommended by van Dijck et al. is not enough to predict n_0 and a .

23 We continue with our investigation to obtain the pure absorbance spectrum of the
24 pollen grain in the following way. We considered the measured spectrum shown in Fig. 11c.
25 Then, we ran the van Dijck iterative algorithm for two values of n_0 and a ($n_0=1.16$,

1 $a=13.5\mu\text{m}$; $n_0=1.5$, $a=13.5\mu\text{m}$). Using the iterative process, we predicted the apparent
2 absorbance using the exact Mie solutions according to Eq. 18 taking into account the
3 numerical aperture. In the region from 2771 cm^{-1} to 1880 cm^{-1} n' is kept zero. The imaginary
4 part, n'' , of the refractive index and the predicted absorbance converged after 20 iteration steps
5 in each case. Figures 15-16 show the imaginary part of the refractive index and the predicted
6 absorbance, respectively. For $n_0=1.16$, $a=13.5\mu\text{m}$, the imaginary part of the refractive index
7 shows vibrational bands at the expected positions. Yet, the obtained imaginary part of the
8 refractive index does not resemble a scatter-free pollen spectrum. Especially the region from
9 1800 cm^{-1} to 1000 cm^{-1} shows an unexpected signature. In case of $n_0=1.5$, $a=13.5\mu\text{m}$ no
10 meaningful peaks in imaginary part could be reconstructed as can be seen in Fig. 15.
11 The predicted spectrum obtained for the case $n_0=1.16$, $a=13.5\mu\text{m}$ represents a meaningful
12 prediction. The regions with absorption bands were nicely predicted. The case $n_0=1.5$,
13 $a=13.5\mu\text{m}$ does not reveal a meaningful prediction.

14

15 **5. Conclusions**

16 In recent years, infrared spectroscopy of biological materials has been challenged by samples
17 of increasing morphological complexity, with the consequence that infrared spectra are
18 strongly distorted by scattering. Several efforts have been made to explain the observed
19 scattering phenomena, and approximate models for the different situations have been
20 presented. All the models presented so far are ad hoc models. In this paper we have presented
21 an exact description of the absorbance spectrum for the scattering and absorption of infrared
22 light at spheres of absorbing materials.

23 When dealing with Mie scattering, absorption and scattering cannot be treated as independent.

24 In this paper we have shown that current models that treat absorption and scattering as
25 dependent yield approximate absorption bands that are considerably shifted, as opposed to the

1 exact models presented in this paper. We therefore propose to implement the exact models in
2 the existing EMSC algorithms for reconstructing absorbance spectra. Concerning the Mie
3 formalism, both the approximate, analytical or the exact Mie formulas [26] may be used.
4 Our method for the estimation of n_0 and a works well in the case of homogeneous spheres, but
5 may be too simple in the case of pollen. It may be necessary to take the layered structure of
6 the pollen into account.
7 The pure absorbance spectra were successfully recovered for PMMA spheres. For biological
8 systems no satisfactory recovery of the pure absorbance spectrum could be obtained. As we
9 demonstrated in this paper, good starting values for n_0 may often be obtained on the basis of
10 physical and biological considerations. Using these values of n_0 as starting values, pure
11 absorbance spectra can be obtained with a high level of confidence. Motivated by the need of
12 accurate n_0 values, we are currently developing new methods of extracting n_0 from the
13 scattering data themselves. Combining these new methods for determining n_0 with the
14 methods outlined and demonstrated above brings out the full power of the new techniques
15 advanced in this paper and contributes decisively to the solution of extracting pure absorbance
16 spectra from measured FTIR spectra.

17 **Appendix A: Lorentz Model**

18 The Lorentz model is a classical model for the computation of the dielectric constant of a material in
19 the presence of absorption resonances. Although based on classical electrodynamics, including
20 quantum effects only phenomenologically via the (measured) spectrum of discrete molecular
21 frequencies, it is surprisingly effective in explaining the frequency dependence of the complex index
22 of refraction m . In particular, it correctly predicts the Lorentz-type shape of the imaginary part of m ,
23 which closely resembles the line shape of an absorption resonance. For the derivation of the Lorentz
24 model, we follow the excellent presentations by Griffiths [33] and Parson [34]. The Lorentz model in
25 its simplest form assumes that an electron is bound to an atom or molecule with a harmonic binding
26 force

$$F_b = -M\omega_0^2 x \quad (\text{A1})$$

where M is the mass of the electron, ω_0 is the natural oscillation frequency of the electron in the harmonic binding potential and x is the amplitude of vibration of the electron in the direction of the external electric field \vec{E} which, in our case, is the infrared light field. It is not necessary to restrict ourselves to electrons. Any charged particle or active group of charged particles that may execute a vibration, such as, e.g., O-H or C=O stretches, are successfully described by the Lorentz model. Therefore, from now on, we imagine a "particle" with effective mass M and charge q subject to the binding force in Eq. A1, and substitute "electron", or "chemically active group" for "particle", as the case may be. For electrons, for instance, $q = -e$, where $e = 1.602 \times 10^{-19}$ C is the elementary charge. For chemically active groups, q is substituted with the polarization charge δq of the polar ends of the group. When the particle vibrates, it loses energy, for instance by electromagnetic dipole radiation or by energy transfer to the backbone molecule or the medium, via the long-range Coulomb force. We model this energy loss with a damping force

$$F_\gamma = -M\gamma \frac{dx}{dt} \quad (\text{A2})$$

which depends linearly on the speed $\frac{dx}{dt}$ of the vibrating particle, and γ is the damping constant, which depends on the details of the energy dissipation processes. In addition to the binding force F_b and the damping force F_γ , the particle also experiences the driving force

$$F_d = qE(t) + \frac{qP(t)}{3\epsilon_0} \quad (\text{A3})$$

where

$$E(t) = E_0 \cos \omega t \quad (\text{A4})$$

is the field strength of the infrared light as a function of time, ω is its frequency and

$$P(t) = Nqx(t) \quad (\text{A5})$$

is the polarization induced by the external field and N is the number of particles per unit volume. We assume here a linear dielectric, in which the polarization of the medium is directly proportional to the radiation field. Newton's equation of motion for the particle,

$$M \frac{d^2x}{dt^2} = F_b + F_\gamma + F_d \quad (\text{A6})$$

1 leads to the differential equation

$$2 \quad M \frac{d^2 x}{dt^2} + M\gamma \frac{dx}{dt} + M\omega_0^2 x = qE_0 \cos \omega t + \frac{qP(t)}{3\varepsilon_0} \quad (\text{A7})$$

3 for the position of the particle. The solution of Eq. A7 is greatly simplified if we consider Eq. A7 as

4 the real part of the complex equation

$$5 \quad \frac{d^2 \tilde{x}}{dt^2} + \gamma \frac{d\tilde{x}}{dt} + \omega_0^2 \tilde{x} = \frac{q}{M} \tilde{E} + \frac{q}{3\varepsilon_0 M} \tilde{P} \quad (\text{A8})$$

6 where x is the real part of the complex quantity \tilde{x} ,

$$7 \quad \tilde{E} = E_0 e^{-i\omega t} \quad (\text{A9})$$

8 is the complex electric field and

$$9 \quad \tilde{P} = Nq\tilde{x} \quad (\text{A10})$$

10 is the complex polarization. In infrared spectroscopy we are not interested in the transient solutions of

11 Eq. A8, i.e. solutions which are generated by switch-on and switch-off of the infrared light. These

12 solutions quickly decay exponentially in time. Once the transient solutions have decayed, the system

13 settles into the steady-state solution

$$14 \quad \tilde{x} = \tilde{x}_0 e^{-i\omega t} \quad (\text{A11})$$

15 Inserting this into Eq. A8 and using Eq. A10, we obtain

$$16 \quad \tilde{x}_0 = \frac{q/M}{\omega_0^2 - \omega^2 - \frac{q^2 N}{3\varepsilon_0 M} - i\gamma\omega} E_0 \quad (\text{A12})$$

17 The complex dipole moment of the particle is

$$18 \quad \tilde{p} = q\tilde{x} = \frac{q/M}{\omega_0^2 - \omega^2 - \frac{q^2 N}{3\varepsilon_0 M} - i\gamma\omega} \tilde{E} \quad (\text{A13})$$

19 We now assume that we have N_m active molecules in our sample and each molecule consists of f_s

20 particles with masses M_s , charges q_s , resonance frequencies ω_s and damping constants γ_s . We define

$$21 \quad \Omega_s^2 = (\omega_s)^2 - \left(\frac{q_s^2 N_m}{3\varepsilon_0 M_s} \right) f_s \quad (\text{A14})$$

22 The polarization \tilde{P} , i.e. the dipole moment per unit volume, is given by

$$23 \quad \tilde{P} = \varepsilon_0 \tilde{\chi} \tilde{E} \quad (\text{A15})$$

24 where ε_0 is the permittivity of the vacuum,

$$25 \quad \tilde{\chi} = N_m \sum_s \frac{q_s^2 f_s / M_s}{\Omega_s^2 - \omega^2 - i\gamma_s \omega} \quad (\text{A16})$$

1 is the susceptibility and \tilde{E} is the complex electric field defined in Eq. A9. The dielectric constant is

$$2 \quad \tilde{\epsilon} = \epsilon_0 \tilde{\epsilon}_r \quad (\text{A17})$$

3 where the relative dielectric constant $\tilde{\epsilon}_r$ depends on $\tilde{\chi}$ according to

$$4 \quad \tilde{\epsilon}_r = 1 + \tilde{\chi} \quad (\text{A18})$$

5 In Eq. A16 we have to sum over all resonances over the entire electromagnetic spectrum including the
6 radio frequency region below the infrared frequency range, the infrared frequency region, and the
7 spectral region above the infrared. We are interested in the infrared frequency region. Therefore, for
8 frequencies larger than infrared frequencies, e.g. in the visible and UV, we may neglect $\gamma_s \omega$ with
9 respect to ω^2 and Ω_s^2 , and expand Eq. A16 to first order in $\left(\frac{\omega}{\Omega_s}\right)^2$. As a consequence, the summation in

10 this frequency range contributes approximately a real term

$$11 \quad \alpha(\omega) = N_m \sum_k \frac{q_k^2 f_k / M_k}{\Omega_k^2} \left[1 + \left(\frac{\omega}{\Omega_k}\right)^2 \right] \quad (\text{A19})$$

12 to the susceptibility in Eq. A16, where the sum over k in Eq. A19 extends over all resonances with
13 frequencies above the infrared range. In the case of the resonances below the infrared frequency range,
14 i.e. the far infrared region and radio frequency region, we may neglect Ω_s^2 and $\gamma_s \omega$ with respect to ω^2 .

15 Thus, this frequency range, approximately, contributes the frequency dependent term

$$16 \quad \beta(\omega) = -\frac{N_m}{\omega^2} \sum_l \frac{q_l^2 f_l}{M_l} \quad (\text{A20})$$

17 to the susceptibility in Eq. A16, where the sum over l in Eq. A20 is over all the resonances below the
18 infrared frequency range. Thus, all together, we now obtain

$$19 \quad \tilde{\epsilon}_r = \bar{\epsilon}_r + N_m \sum_{s \in IR} \frac{q_s^2 f_s / M_s}{\Omega_s^2 - \omega^2 - i\gamma_s \omega} \quad (\text{A21})$$

20 where the sum in Eq. A21 extends only over the infrared (IR) resonances and

$$21 \quad \bar{\epsilon}_r = 1 + \alpha(\omega) + \beta(\omega) \quad (\text{A22})$$

22 is the frequency dependent effective relative dielectric constant of the medium, i.e. the background
23 dielectric constant, on which the infrared resonances are built. The complex index of refraction is now
24 given by

$$25 \quad m = \sqrt{\tilde{\epsilon}_r} \quad (\text{A23})$$

1 At this point two important comments are in order: (1) In optics we are familiar with the phenomenon
 2 of dispersion, i.e. the change of the index of refraction with increasing wavelength. For glass, for
 3 instance, we know that the index of refraction decreases with increasing wavelength, which gives rise
 4 to the familiar observation of the splitting of white light into its constituent colours with the help of a
 5 prism. This decrease in the index of refraction is now easily explained. According to Eqs. 19 and 20,
 6 both $\alpha(\omega)$ and $\beta(\omega)$ cause a decrease in $\bar{\epsilon}_r$, and therefore, according to Eq. A23, they also cause a
 7 decrease in m , when the wavelength increases. Therefore the Lorentz model explains this basic
 8 observation. (2) Since the quantities N_m , q_s , f_s , and γ_s are usually not readily available, we write

$$9 \quad \tilde{\epsilon}_r = \bar{\epsilon}_r + \sum_s \epsilon_{IR} \frac{\Lambda_s}{\tilde{\nu}_s^2 - \tilde{\nu}^2 - i\Gamma_s \tilde{\nu}} \quad (\text{A24})$$

10 where $\bar{\epsilon}_r$, $\tilde{\nu}_s$, Λ_s , and Γ_s are adjustable parameters and

$$11 \quad \tilde{\nu} = \frac{1}{\lambda} = \frac{\omega}{2\pi c} \quad (\text{A25})$$

12 where λ is the wavelength and c is the vacuum speed of light. This is the microscopic basis for the
 13 usual practice in spectroscopy of fitting Lorentzian lines to resonance structures in the index of
 14 refraction.

15 Instead of using quantum mechanics to solve for the quantized excitations of the molecule in
 16 the presence of the infrared radiation field, the Lorentz model uses classical mechanics to
 17 solve the forced, damped harmonic oscillator equation (A8). This, apparently, introduces two
 18 errors, (i) the oscillator (A8) is not quantized and (ii) neither is the radiation field, i.e. it is not
 19 treated as consisting of photons. The question is: how serious are these approximations? The
 20 answer is the following. (i) Quantum mechanics has been partially included by providing the
 21 Lorentz model with the discrete set of molecular frequencies Ω_s , a direct result of the
 22 quantization of the molecule via the many-body Schrödinger equation. (ii) Although the
 23 radiation field consists of photons, the light intensities in infrared spectroscopy are so high
 24 that we can safely neglect the quantization of the radiation field. Of course there remains the
 25 question of the quantization of the Ω_s modes, whose amplitudes are treated as a classical,
 26 continuous variable, although, according to quantum mechanics, they should be quantized.

1 This, however, is not serious. As soon as the Ω_s oscillators are appreciably excited,
 2 corresponding to a few absorbed photons, the classical approximation is practically
 3 indistinguishable from the exact quantum treatment, which is due to the Bosonic nature of the
 4 oscillator excitations. Thus, because of the relatively large intensities of the infrared light
 5 field, the classical approximation of both the molecular oscillators and the radiation field is
 6 justified.

7 **Appendix B: Mie Formulas**

8 The Mie scattering amplitudes are defined as:

$$9 \quad S_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{a_n \pi_n(\cos\theta) + b_n \tau_n(\cos\theta)\} \quad (\text{B1})$$

$$10 \quad S_2(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{b_n \pi_n(\cos\theta) + a_n \tau_n(\cos\theta)\} \quad (\text{B2})$$

$$11 \quad a_n = \frac{\psi'_n(y)\psi_n(x) - m\psi_n(y)\psi'_n(x)}{\psi'_n(y)\zeta_n(x) - m\psi_n(y)\zeta'_n(x)} \quad (\text{B3})$$

$$12 \quad b_n = \frac{m\psi'_n(y)\psi_n(x) - \psi_n(y)\psi'_n(x)}{m\psi'_n(y)\zeta_n(x) - \psi_n(y)\zeta'_n(x)} \quad (\text{B4})$$

$$13 \quad \psi_n(z) = \sqrt{\frac{\pi z}{2}} J_{n+\frac{1}{2}}(z) \quad (\text{B5})$$

$$14 \quad \zeta_n(z) = \sqrt{\frac{\pi z}{2}} H_{n+\frac{1}{2}}^{(2)}(z) \quad (\text{B6})$$

$$15 \quad x = 2\pi a \tilde{\nu} \quad (\text{B7})$$

$$16 \quad y = mx \quad (\text{B8})$$

$$17 \quad m = n + in' \quad (\text{B9})$$

$$18 \quad \pi_n(\cos\theta) = \frac{1}{\sin\theta} P_n^1(\cos\theta) \quad (\text{B10})$$

1
$$\tau_n = \frac{d}{d\theta} P_n^1(\cos\theta) \quad (\text{B11})$$

2 where, m (complex in general) is the refractive index of the homogenous sphere, a is the
3 radius of the sphere, J is the Bessel function of the 1st kind and $H^{(2)}$ denotes the Hankel
4 functions. The argument z in Eqs. B5 and B6 is an arbitrary complex number; it may be equal
5 to x or y . P^l denotes the first order associated Legendre polynomial.

6 The purpose of listing these equations is twofold: It establishes our notation and shows that all
7 aspects of Mie scattering may indeed be written down analytically. In this context we mention
8 that m in this paper is defined according to B9, with negative n' for positive absorbance. This
9 is the opposite sign convention from the one used in the standard reference book of Van De
10 Hulst [26].

11 **Acknowledgments**

12 The research was supported by the Norwegian Research Council (ISP project N°. 216687), by
13 the SOLEIL, French national synchrotron facility (project N°. 20120345), and by the
14 European Commission through the Seventh Framework Programme (FP7-PEOPLE-2012-IEF
15 project N°. 328289). We thank Paul Dumas and Christophe Sandt from the SMIS beamline at
16 SOLEIL.

1 References

- 2 [1] T. van Dijk, D. Mayerich, P. S. Carney and R. Bhargava, *Appl Spectrosc* **67** (5), 546-552
3 (2013).
- 4 [2] M. Jackson and H. H. Mantsch, *TrAC, Trends Anal. Chem.* **11** (6), 206-210 (1992).
- 5 [3] J. D. Krusejarres and G. Janatsch, *Clinical Chemistry* **33** (6), 961-961 (1987).
- 6 [4] D. Naumann, D. Helm and H. Labischinski, *Nature* **351** (6321), 81-82 (1991).
- 7 [5] D. L. Wetzel and J. A. Reffner, *Cereal Food World* **38**, 9-20 (1993).
- 8 [6] J. Pijanka, G. D. Sockalingum, A. Kohler, Y. Yang, F. Draux, G. Parkes, K. P. Lam, D.
9 Collins, P. Dumas, C. Sandt, D. G. van Pittius, G. Douce, M. Manfait, V. Untereiner and J. Sule-
10 Suso, *Laboratory Investigation* **90** (5), 797-807 (2010).
- 11 [7] J. K. Pijanka, N. Stone, G. Cinque, Y. Yang, A. Kohler, K. Wehbe, M. Frogley, G.
12 Parkes, J. Parkes, P. Dumas, C. Sandt, D. G. van Pittius, G. Douce, G. D. Sockalingum and J.
13 Sule-Suso, *Spectroscopy-an International Journal* **24** (1-2), 73-78 (2010).
- 14 [8] P. Lasch, A. Pacifico and M. Diem, *Biopolymers* **67** (4-5), 335-338 (2002).
- 15 [9] M. C. McCann, M. Hammouri, R. Wilson, P. Belton and K. Roberts, *Plant Physiol.* **100**
16 (4), 1940-1947 (1992).
- 17 [10] A. Kohler, C. Kirschner, A. Oust and H. Martens, *Appl. Spectrosc.* **59** (6), 707--716
18 (2005).
- 19 [11] B. Mohlenhoff, M. Romeo, B. R. Wood and M. Diem, *Biophys. J.* **88** (5), 3635–3640
20 (2005).
- 21 [12] G. Mie, *Annalen der Physik* **25** (1908).
- 22 [13] A. Kohler, J. Sule-Suso, G. D. Sockalingum, M. Tobin, F. Bahrami, Y. Yang, J. Pijanka,
23 P. Dumas, M. Cotte, D. G. van Pittius, G. Parkes and H. Martens, *Appl Spectrosc* **62** (3), 259-266
24 (2008).
- 25 [14] P. Bassan, H. J. Byrne, F. Bonnier, J. Lee, P. Dumas and P. Gardner, *Analyst* **134** (8),
26 1586-1593 (2009).
- 27 [15] R. D. L. Kronig, *J. Opt. Soc. Am.* **12** (6), 547-556 (1926).
- 28 [16] H. A. Kramers, (presented at the Atti Cong. Intern. Fisica, (Transactions of Volta
29 Centenary Congress) Como (unpublished), 1927).
- 30 [17] R. Bhargava, *Appl Spectrosc* **66** (10), 1091-1120 (2012).
- 31 [18] B. J. Davis, P. S. Carney and R. R. Bhargava, *Analytical Chemistry* **82**, 3474–3486
32 (2010).
- 33 [19] B. J. Davis, P. S. Carney and R. R. Bhargava, *Analytical Chemistry* **82**, 3487–3499
34 (2010).
- 35 [20] P. Bassan, A. Kohler, H. Martens, J. Lee, H. J. Byrne, P. Dumas, E. Gazi, M. Brown, N.
36 Clarke and P. Gardner, *Analyst* **135** (2), 268-277 (2010).
- 37 [21] P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, P. Dumas, M. Brown,
38 N. Clarke and P. Gardner, *J Biophotonics* **3** (8-9), 609-620 (2010).
- 39 [22] P. Bassan, A. Sachdeva, A. Kohler, C. Hughes, A. Henderson, J. Boyle, J. H. Shanks, M.
40 Brown, N. W. Clarke and P. Gardner, *Analyst* **137** (6), 1370-1377 (2012).
- 41 [23] B. Bird, M. Miljkovic and M. Diem, *Journal of Biophotonics* **3** (8-9), 597-608 (2010).
- 42 [24] M. Miljkovic, B. Bird and M. Diem, *Analyst* **137** (17), 3954-3964 (2012).
- 43 [25] J. L. Ilari, M. Martens and T. Isaksson, *Appl Spectrosc* **42** (5), 722-728 (1988).
- 44 [26] H. C. v. d. Hulst, *Light scattering by small particles.* (Wiley, New York., 1957).
- 45 [27] G. L. Carr, L. M. Miller and P. Dumas, in *Biomedical Applications of Synchrotron*
46 *Infrared Microspectroscopy*, edited by D. Moss (Royal Society of Chemistry, 2011).

- 1 [28] C. F. Bohren and D. R. Huffman, in Absorption and Scattering of Light by Small Particles
2 (Wiley-VCH Verlag GmbH, 2007), pp. 475-476.
- 3 [29] P. Dumas, F. Polack, B. Lagarde, O. Chubar, J. L. Giorgetta and S. Lefrancois, Infrared
4 Phys Techn **49** (1-2), 152-160 (2006).
- 5 [30] P. Chýlek, J. Opt. Soc. Am. **66** (3), 285-287 (1976).
- 6 [31] P. Chýlek, J. T. Kiehl and M. K. W. Ko, Appl Optics **17** (19), 3019-3021 (1978).
- 7 [32] S. K. Mas Rosemal H. Mas Haris, S. Mohanc, Der Pharma Chemica **2** (4), 316-323
8 (2010).
- 9 [33] D. J. Griffiths, Introduction to Electrodynamics, 3 ed. (Prentice-Hall, Upper Saddle River,
10 1999).
- 11 [34] W. W. Parson, Modern Optical Spectroscopy, Student Edition ed. (Springer, Heidelberg,
12 2009).

13

1 **Figure captions:**

2 **Figure 1** Illustration of the scattering of incident light with intensity I_0 at an arbitrary-shaped absorbing
3 scatterer. A plane wave is incident from the left. In general, part of the incident light is scattered into
4 different directions, part of the light is chemically absorbed by the scatterer, and part of the incident light
5 is transmitted to the detector. The part of the incident infrared light chemically absorbed by the scatterer is
6 indicated by the red area, representing a radiation sink.

7 **Figure 2** (a) Apparent absorbance spectrum (red curve) for a sphere assuming a single absorption band
8 located at 1654 cm^{-1} , corresponding to the C=O stretching vibration of the peptide bond in proteins. As
9 parameters for the calculation of the refractive index, according to the Lorentz model (see Appendix A),
10 $\Lambda = 10^4\text{ cm}^{-2}$, $\Gamma = 30\text{ cm}^{-1}$ and $\bar{\epsilon}_r = 1.44$ were chosen (corresponding to a background refractive index
11 of 1.2). The apparent absorbance spectrum with complex index of refraction is compared to an
12 approximation of the absorbance with real index of refraction (blue curve) often found in the literature,
13 assuming that absorbance and extinction are proportional. (b) Enlarged subfigure of (a) in the spectral
14 range of the absorption band.

15 **Figure 3** Apparent absorbance spectrum of the absorbing sphere of Fig. 2 (red) is compared to the
16 *apparent absorbance* (blue) according to Bassan et al. [14]. Compared with the result considering
17 complex index of refraction, the *apparent absorbance* considering real index of refraction is shifted to a
18 lower frequency.

19 **Figure 4** Extinction efficiency in the forward direction and scattering efficiency, including a numerical
20 aperture, for a $10\mu\text{m}$ radius sphere assuming a single absorption band located at 1654 cm^{-1} for two cases
21 $\Lambda = 10^4\text{ cm}^{-2}$, $\Gamma = 30\text{ cm}^{-1}$ and $\bar{\epsilon}_r = 1.3$ (corresponding to an $n_0=1.14$) (a) and $\Lambda = 10^4\text{ cm}^{-2}$, $\Gamma =$
22 30 cm^{-1} and $\bar{\epsilon}_r = 1.54$ (corresponding to an $n_0=1.24$) (b).

23 **Figure 5** Synchrotron spectrum of a PMMA sphere with radius $10\mu\text{m}$ (a), in the wavenumber range 3600-
24 7200 cm^{-1} (b); in the wavenumber range $800\text{-}3600\text{ cm}^{-1}$ (c) and zoomed in the wavenumber range 5850-
25 7000 cm^{-1} (d).

1 **Figure 6** Root mean square error (RMSE) as a function of size factor α for the spectrum of a PMMA
2 sphere shown in Fig. 5b. Inset: RMSE function in the vicinity of the four main local minima.

3 **Figure 7** Measured absorbance and predicted absorbance for $\alpha = 0.64\mu\text{m}$ (a), $\alpha = 1.1\mu\text{m}$ (b), $\alpha =$
4 $2.1\mu\text{m}$ (c) and $\alpha = 3.1\mu\text{m}$ (d) .

5 **Figure 8** Imaginary part of the complex refractive index for three different values of n_0 and a after 12
6 iterations.

7 **Figure 9** Predicted absorbance for three different values of n_0 and a after 12 iterations.

8 **Figure 10** Calculated pure absorbance spectrum of a PMMA thin film with thickness $d_{\text{eff}}=5.9\mu\text{m}$ and the
9 imaginary part of refractive index obtained with the van Dijck iterative algorithm.

10 **Figure 11** *Juniperus chinensis* pollen synchrotron FTIR spectrum (a). Synchrotron FTIR spectra of
11 *Juniperus chinensis* pollen ranging from 3600 cm^{-1} to 6300 cm^{-1} (b); from 1000 cm^{-1} to 3600 cm^{-1} (c) and
12 from 1880 cm^{-1} to 2771 cm^{-1} (d).

13 **Figure 12** Root mean square error (RMSE) as a function of α for the spectrum of a *Juniperus chinensis*
14 pollen grain shown in Fig. 11b. The inset shows the RMSE function in the vicinity of the four main local
15 minima.

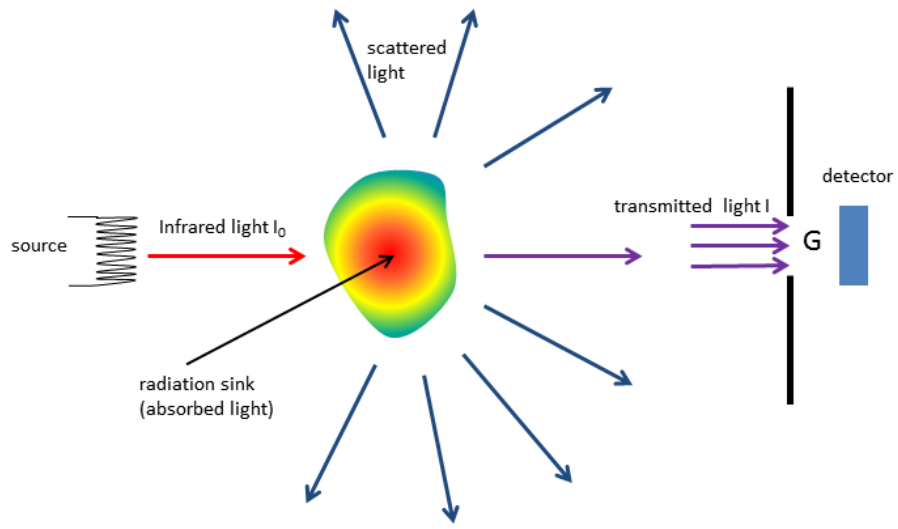
16 **Figure 13** Measured absorbance and predicted absorbance for $\alpha = 1.0\mu\text{m}$ (a), $\alpha = 1.3\mu\text{m}$ (b), $\alpha =$
17 $2.2\mu\text{m}$ (c) and $\alpha = 3.2\mu\text{m}$ (d) .

18 **Figure 14** Root mean square error as a function of size parameter for a *Juniperus chinensis* pollen grain,
19 considering the measured spectrum in the range between 1880 cm^{-1} and 2771 cm^{-1} .

20 **Figure 15** Imaginary part of the complex refractive index for two different values of n_0 and a after 20
21 iterations.

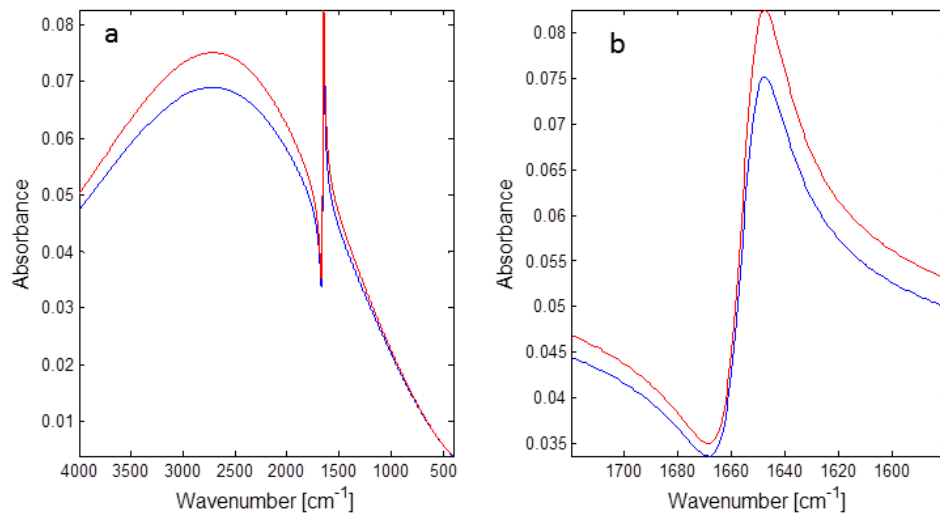
22 **Figure 16** Predicted absorbances for two different values of n_0 and a after 20 iterations.

23



- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Fig. 1



1

2

3

4

5

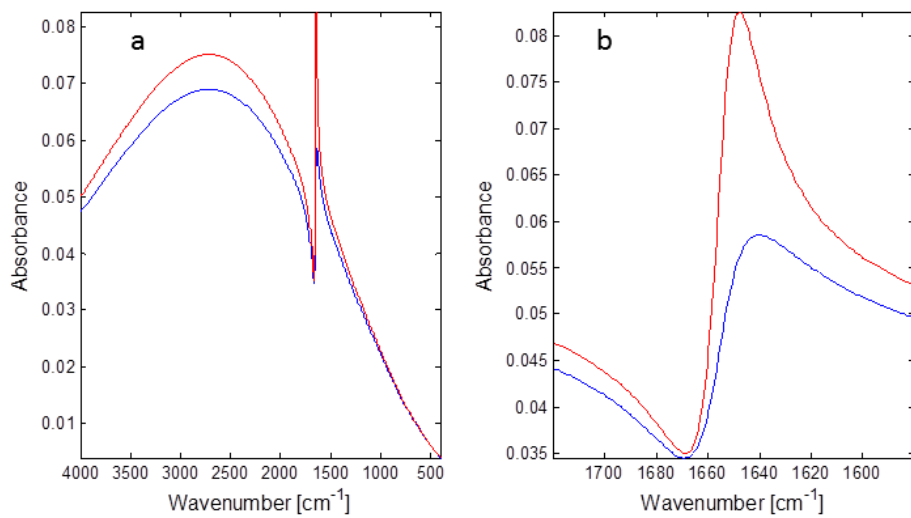
6

7 Fig. 2

8

9

10



1

2

3

4

5

6

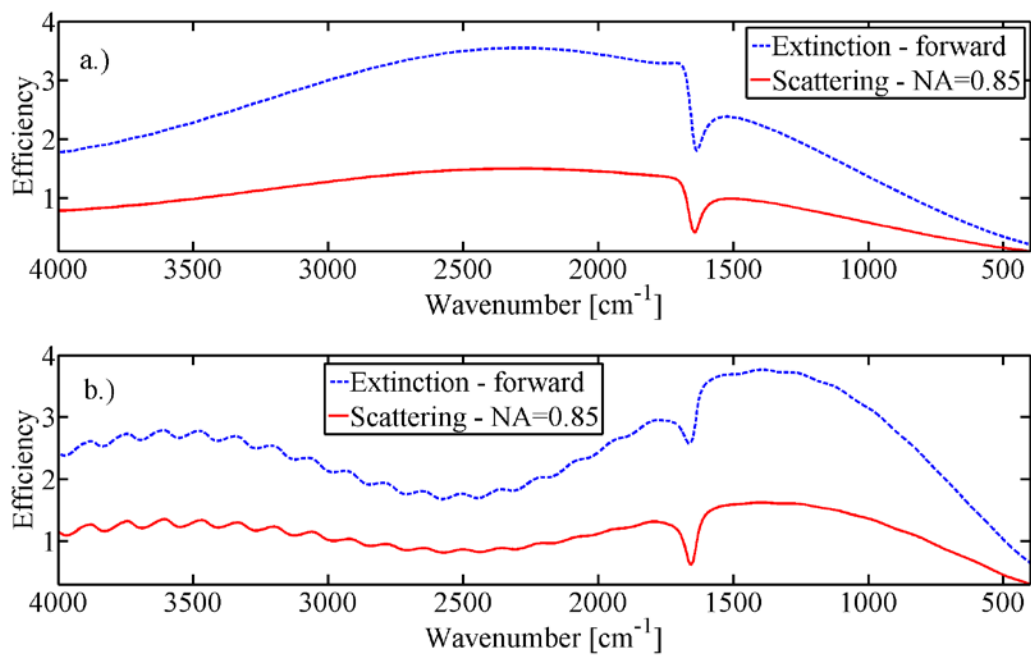
7 Fig. 3

8

9

10

11



1

2

3

4

5

6

7

8 Fig. 4

9

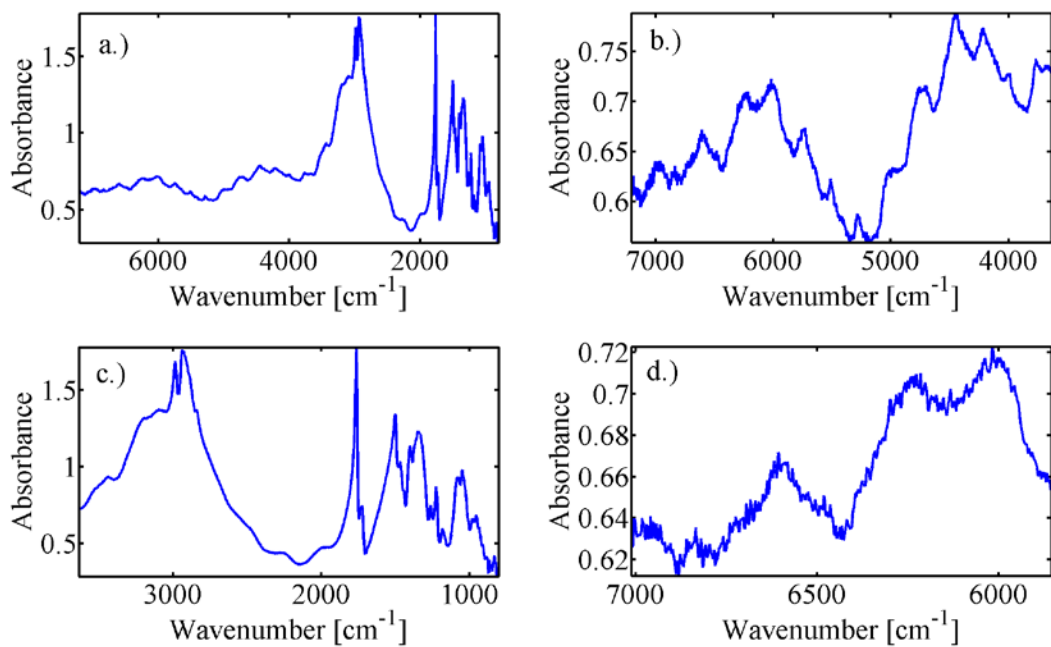
10

11

12

13

14



1

2

3

4

5

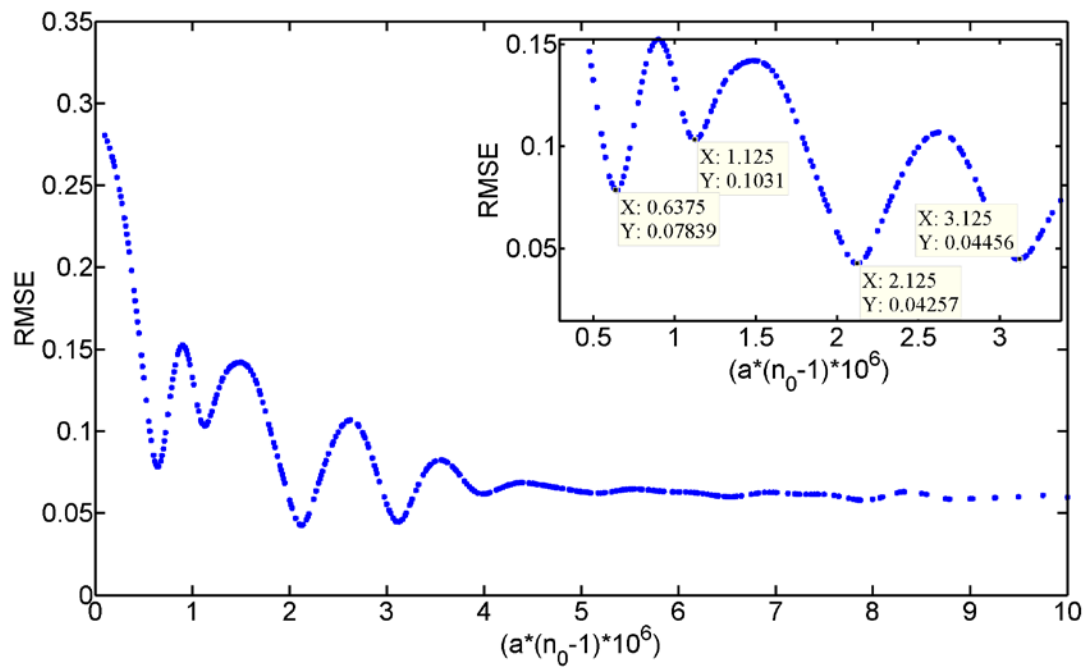
6

7

8 Fig. 5

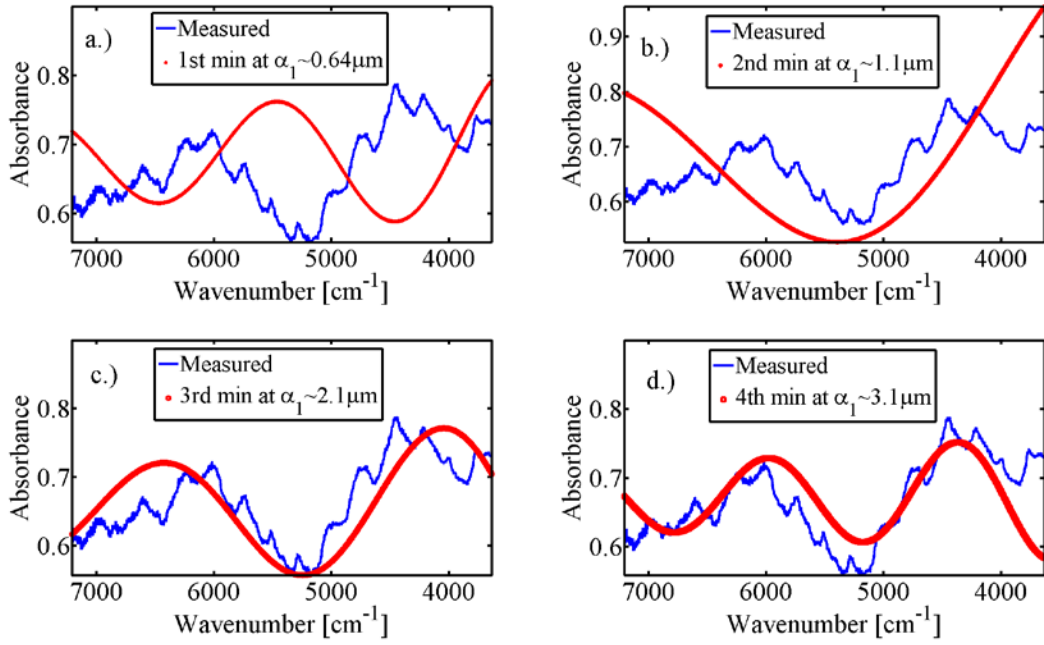
9

10



1
2
3
4
5
6
7
8
9
10
11
12

Fig. 6



1

2

3

4

5

6

7 Fig. 7

8

9

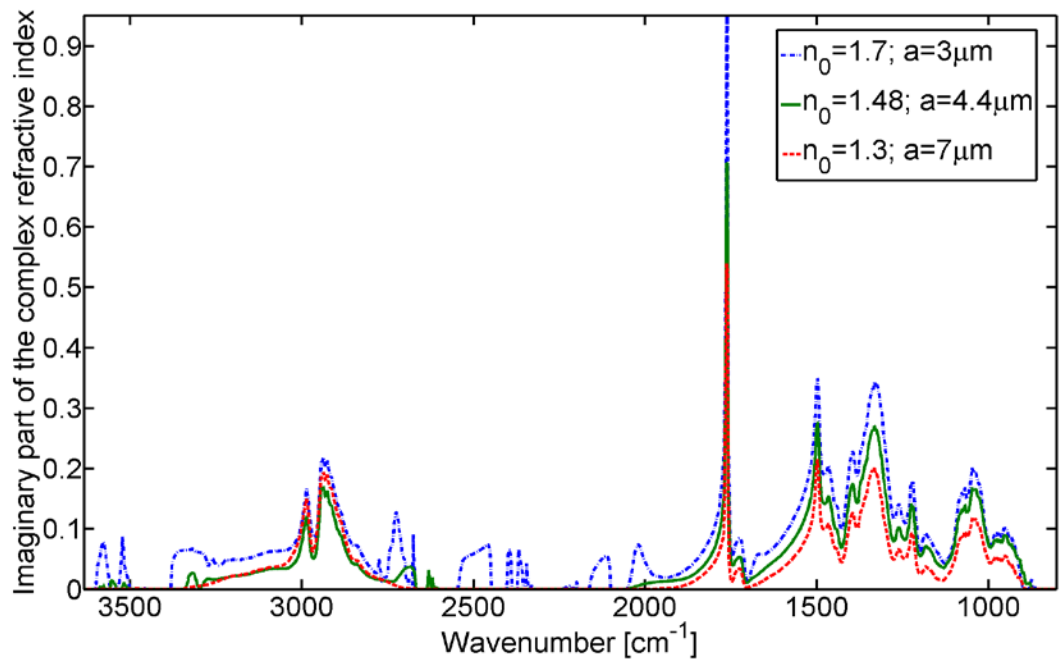
10

11

12

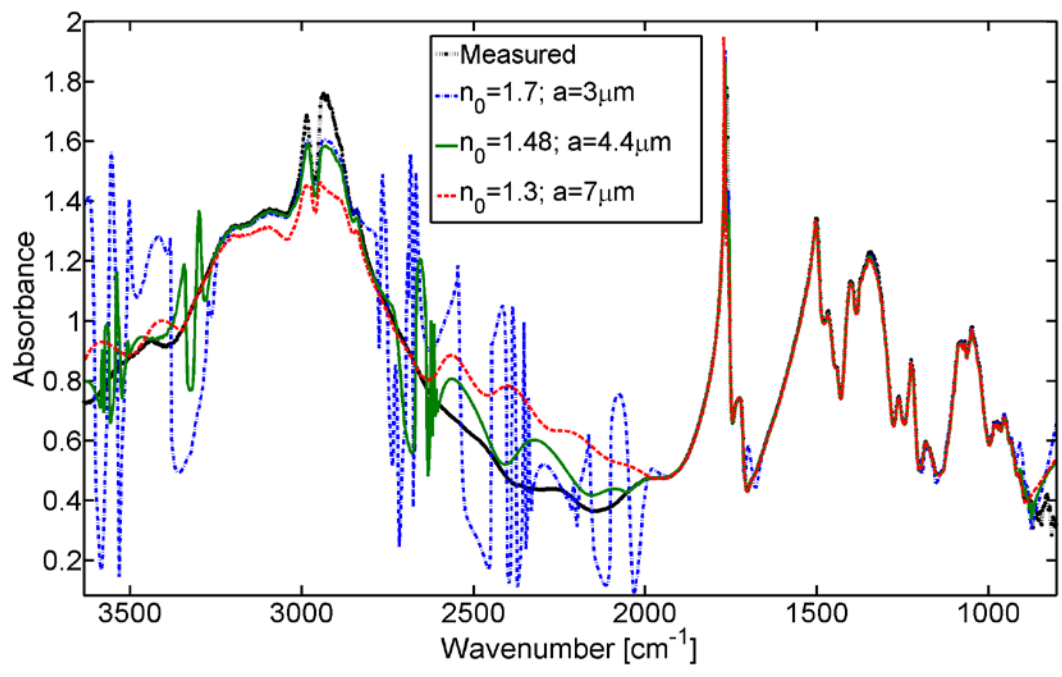
13

14



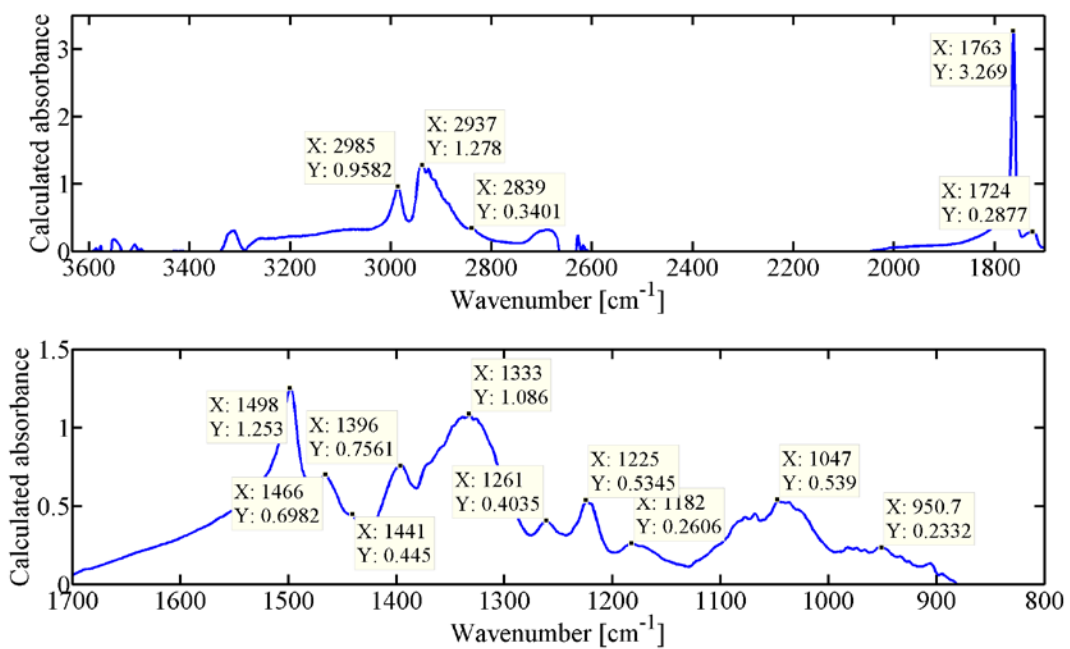
1
2
3
4
5
6
7
8
9
10
11
12

Fig. 8



1
2
3
4
5
6
7
8
9
10
11
12

Fig. 9



1

2

3

4

5

6

7

8 Fig. 10

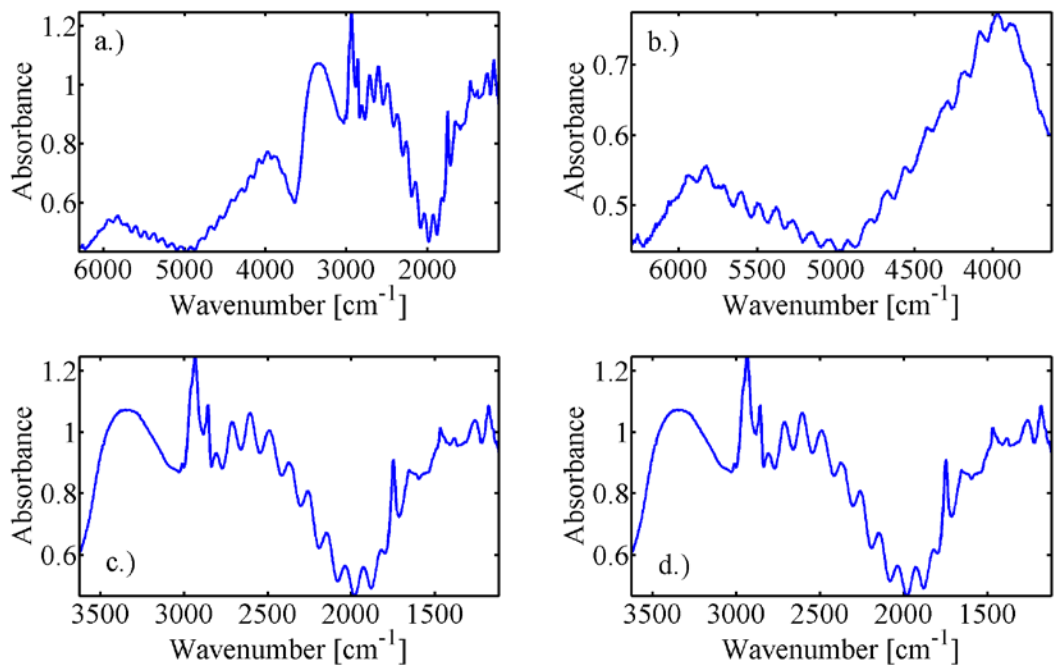
9

10

11

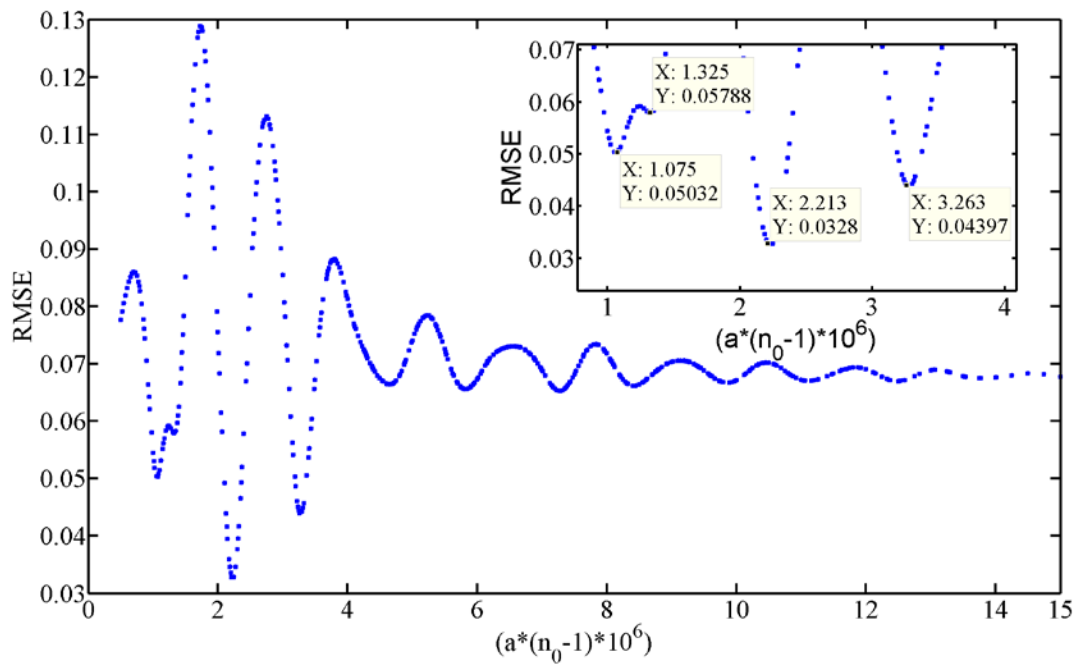
12

13



1
2
3
4
5
6
7
8
9

Fig. 11



1

2

3

4

5

6

7

8 Fig. 12

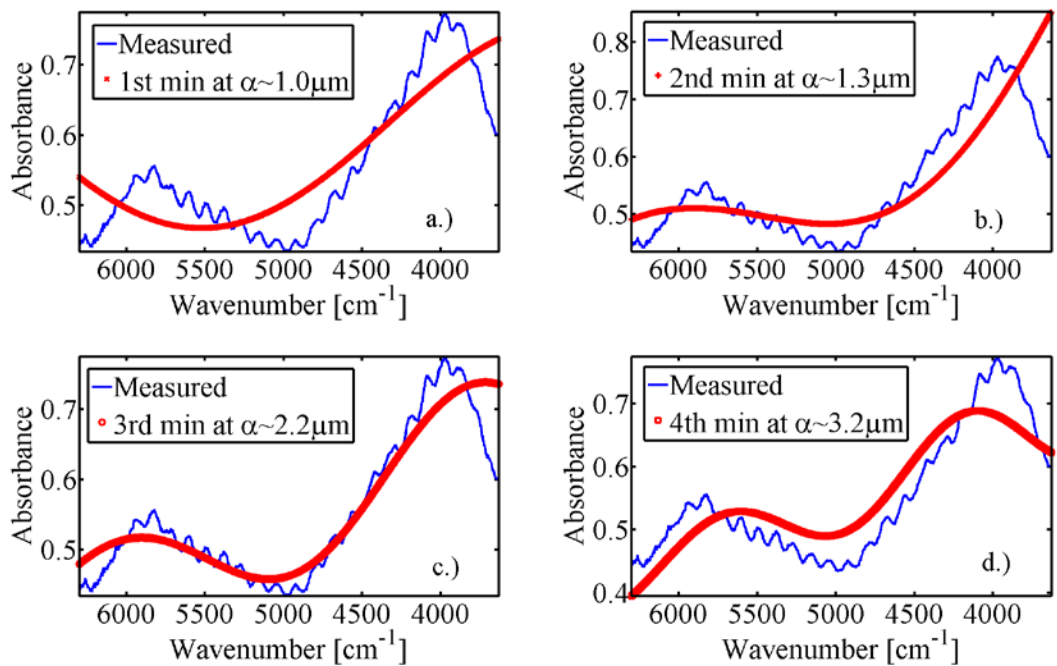
9

10

11

12

13



1

2

3

4

5

6

7

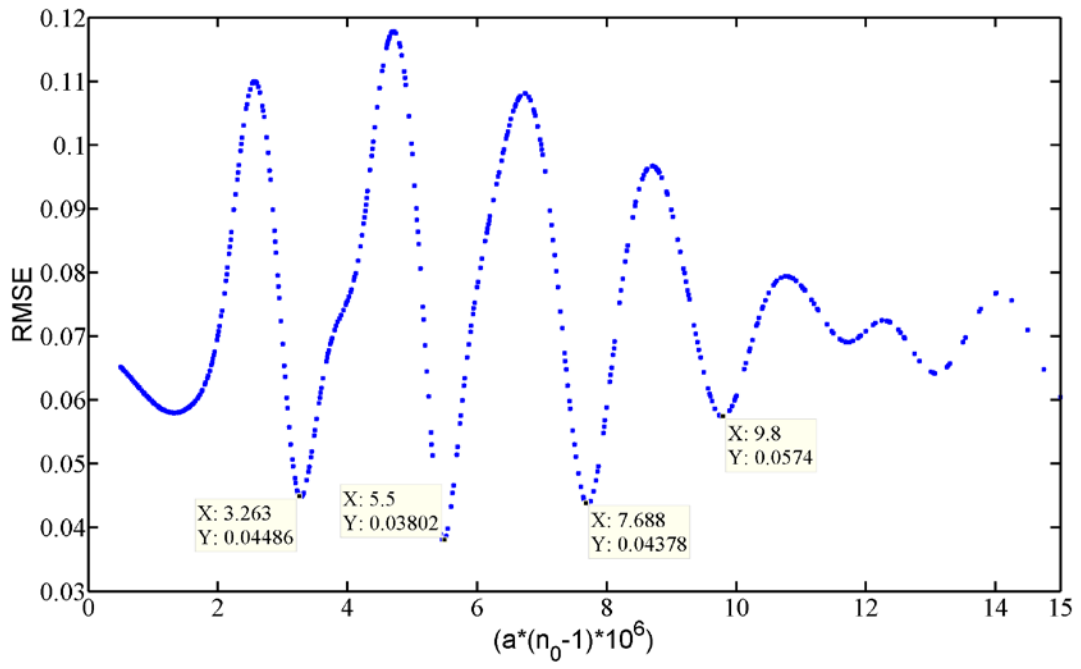
8 Fig. 13

9

10

11

12



1

2

3

4

5

6

7

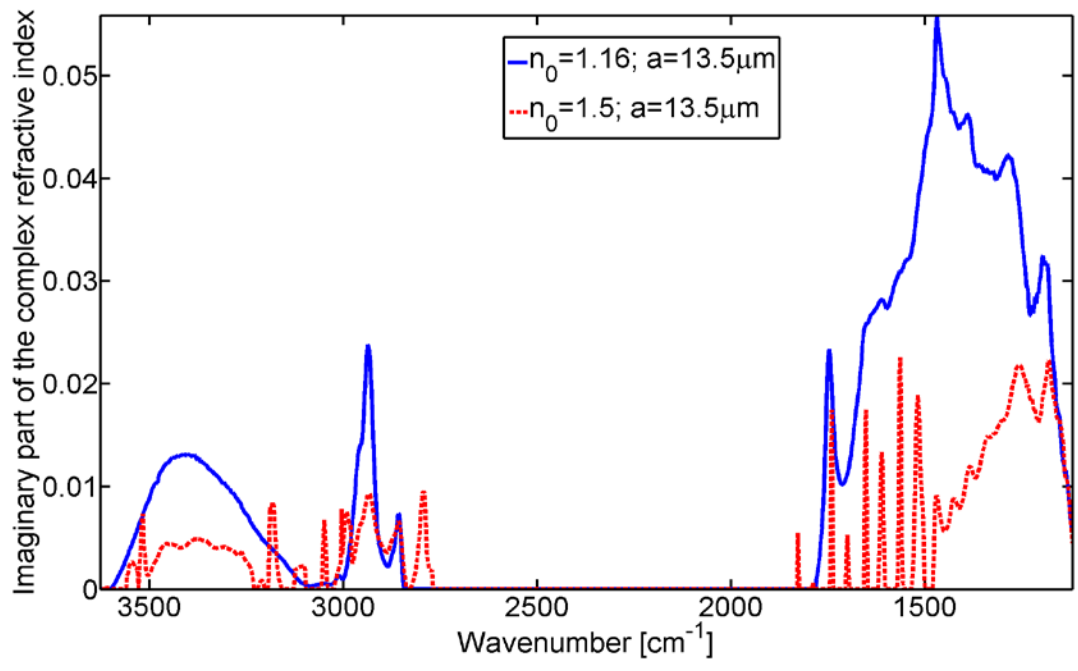
8

9 Fig. 14

10

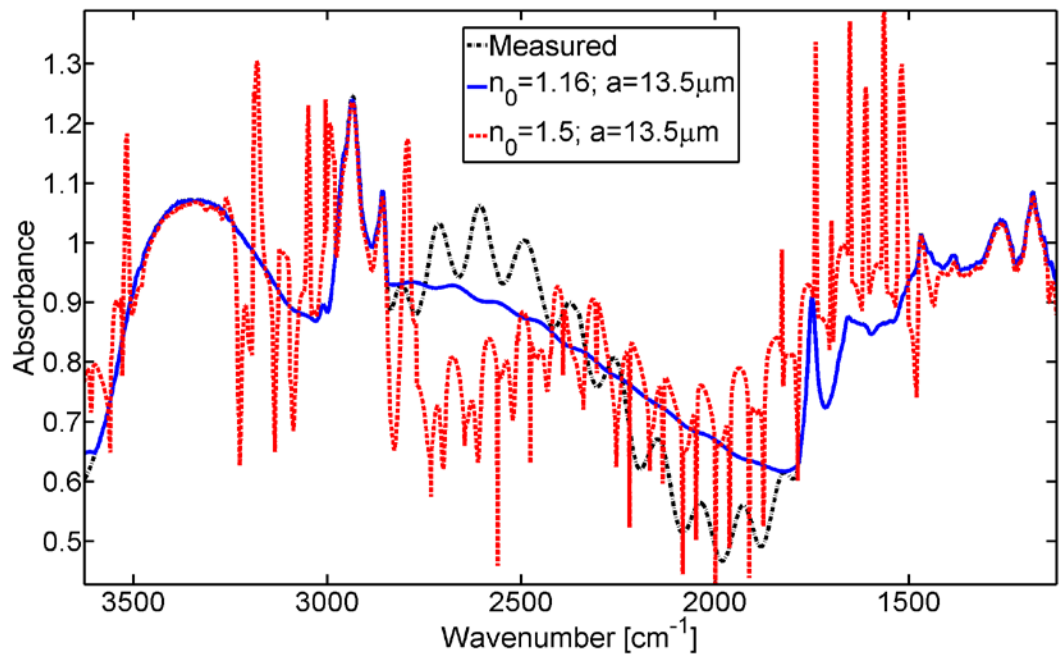
11

12



1
2
3
4
5
6
7
8
9
10
11
12

Fig. 15



1
2
3
4
5
6
7
8
9
10
11
12
13

Fig. 16