Mid-infrared laser measurements of aqueous glucose

Bujin Guo*
Yi Wang*
Yang Wang
Han Q. Le

University of Houston
Photonic Device and System Laboratory
Department of Electrical and Computer Engineering
and
Texas Center for Advanced Materials and Southwest Public Safety Technology Center
4800 Calhoun
Houston, Texas 77204
E-mail: hqle@uh.edu

Abstract. Mid-IR semiconductor lasers of two wavelength bands, 5.4 and 9.6 μm, are applied to measure aqueous glucose concentration ranging from 0 to 500 mg/dL with Intralipid® emulsion (0 to 8%) added as a fat simulator. The absorption coefficient $\mu_a$ is found linear with respect to glucose and Intralipid® concentrations, and their specific absorption coefficients are obtained via linear regression. These coefficients are subsequently used to infer the concentrations and compare with known values. The objective is to evaluate the method accuracy. Glucose concentration is determined within ±21 mg/dL with 90% confidence and ±32 mg/dL with 99% confidence, using <1-mJ laser energy. It is limited by the apparatus mechanical error and not the photometric system noise. The expected uncertainties due to photometric noise are ±6 and ±9 mg/dL with 90 and 99% confidence, respectively. The uncertainty is fully accounted for by the system intrinsic errors, allowing rigorous inference of the confidence level. Intralipid® is found to have no measurable effect on glucose determination. Further analysis suggests that a few mid-IR wavelengths may be sufficient, and that the laser technique offers advantages with regard to accuracy, speed, and sample volume, which can be small, $-0.4 \times 10^{-7}$ mL for applications such as microfluidic or microbioarray monitoring. © 2007 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2714283]

Keywords: glucose; laser glucose measurement; mid-infrared laser spectroscopy.

Paper 06129RR received May 19, 2006; revised manuscript received Jan. 4, 2007; accepted for publication Jan. 4, 2007; published online Apr. 13, 2007.

1 Introduction

There have been extensive efforts to develop optical techniques to measure blood glucose concentration for diabetes care.1–3 There is also the application of in situ real-time monitoring of glucose in cell culture, microarray arrays, and bioreactors. In both applications, accuracy, reliability, and speed are important. This work is about a laser-based photometric technique that can address these issues. This work aims primarily for the latter application, but the technique is potentially applicable for blood glucose.

A broad review of minimally or noninvasive blood glucose monitoring has been given by Klonoff.4 More recently, Khalil5 and Coté6 specifically reviewed optical spectroscopy, including absorption, polarimetry, Raman, and fluorescence. Absorption spectroscopy in particular, for both the near-IR (NIR) and mid-IR (MIR) spectral region, has been explored extensively by many research groups and commercial concerns over many years.7–24 For the MIR “fingerprint” region, Heise et al.7 in early work analyzed blood glucose in vitro using Fourier transform infrared (FTIR) spectroscopy. Bhandare et al.9 also used FTIR to investigate the spectral interference of other blood constituents on the glucose signature. In more recent works,10–13 various groups focused on using only a few key MIR wavelengths rather than the entire spectrum to determine blood glucose. Specifically, Shen et al.11 found that spectroscopy at just two wavenumbers, 1082 and 1093 cm$^{-1}$, was sufficient. Rhiel et al.14 employed FTIR to monitor not only glucose but also lactose and ammonia in cell culture in real time. Opposite to absorption, MIR radiation was also used to measure glucose based on Kirchoff’s law of thermal emission.15 Some other works are based on laser photoacoustic effects.25,26

There are generally three aspects in these studies: the spectroscopy aspect is concerned with spectral features and processing algorithms; the photometry aspect is concerned with optical system engineering and measurements; and the deployment aspect is concerned with application. Photometry is often undervalued and underdeveloped compared to spectroscopy but essential for real-life applications. The MIR works,7–15 for example, demonstrated the spectroscopic merit, but the photometry was simply lacking. The thermal-source FTIR instrument employed in these works is useful as a research instrument but not necessarily practical or cost effective. Thermal sources have low power spectral density, low radiance, high relative intensity noise (RIN), and large 1/f thermal noise. These impose limitations on dynamic range and speed for accuracy and real-time monitoring. In comparison, the NIR approach does not have any intrinsic spectro-
scopic advantage, but benefits greatly from more advanced photometric technologies.

Recent progress can help exploit the MIR region. This work describes the use of advanced MIR semiconductor lasers for in vitro aqueous glucose measurements. Lasers have well-established photometric advantages including high spectral intensity, brightness, high modulation bandwidth, low RIN, and 1/f noise. They are also well suited for applications such as microfluidic or microarray bioassay, as the beam can be focused to a small volume. For practicality, they are robust, compact, and potentially low cost. Advanced MIR semiconductor lasers have been demonstrated for tissue imaging, application for conjunctiva glucose measurement is being developed. A potential issue with MIR lasers is the lack of large spectral coverage, but some studies have shown that only a few wavelengths may be needed. Also, recent progress in tunable MIR lasers promises broad wavelength coverage.

Photometry is crucial to the issue of accuracy. There is often some ambiguity on what contributes to glucose measurement errors: how much it is due to spectroscopic problems (interference, distorted spectra, etc.), and how much is due to photometric errors. Glucose errors were usually determined by comparing with wet chemistry results, but not analytically linked to instrument performance. As a basic principle of system engineering, a system should fully account for all errors, and establish an intrinsic measurement confidence level based on the causal and deterministic relation with the instrument noises and errors. This is a focus of this work.

The experiments were performed for glucose in water and Intralipid® (IL) solution. It is chemically much simpler than blood, since the objective is to establish the baseline photometric merit without the issue of chemical complexity. It is hoped that it may eventually motivate blood applications with many-wavelength systems to handle the chemical complexity.

The rest of the work is organized as follows. Section 2 describes the experimental approach. Results and discussion are presented in Sec. 3; and Sec. 4 presents a summary and conclusion.

2 Experimental approach

2.1 Overall Concept

2.1.1 Regression Approach

Given a glucose solution with concentration \( C_{\text{glu}} \), which also contains an interfering substance such as IL of concentration \( C_{\text{IL}} \), a system can determine these concentrations via the equation:

\[
\begin{pmatrix}
\tilde{C}_{\text{glu}} \\
\tilde{C}_{\text{IL}}
\end{pmatrix} =
\begin{pmatrix}
q_{\text{glu,1}} & q_{\text{glu,2}} & \cdots & q_{\text{glu,n}} \\
q_{\text{IL,1}} & q_{\text{IL,2}} & \cdots & q_{\text{IL,n}}
\end{pmatrix}
\begin{pmatrix}
\Delta\mu_{\text{a,glu}}(\lambda_1) \\
\Delta\mu_{\text{a,glu}}(\lambda_2) \\
\vdots \\
\Delta\mu_{\text{a,glu}}(\lambda_n)
\end{pmatrix}
\]

+ higher order terms,

(1)

where \( \Delta\mu_{\text{a}}(\lambda_n) \) is the difference between the absorption coefficient of the solution and that of pure water \( \Delta\mu_{\text{a}}(\lambda) = \mu_{\text{a,solution}}(\lambda) - \mu_{\text{a,water}}(\lambda) \), and coefficients \( q \) are quantities to be obtained from calibration. For glucose and IL solution, it is expected that \( \mu_{\text{a,glu}} \gg \mu_{\text{a,IL}} \) in the MIR owing to Rayleigh’s \( \lambda^{-4} \) scattering law and the latter can be neglected. In Eq. (1), it is seen that higher-order terms are negligible; and it is convenient to use the matrix format:

\[
\tilde{C} = Q \cdot \Delta\mu_{\text{a}},
\]

(2)

where the notation \( \tilde{C} \) indicates that it is only an inferred estimate of the true value \( C \). In more complex solutions with \( k \) substances, \( \tilde{C} \) would represent \( k \) values of concentration.

The calibration for \( Q \) is obtained from \( \Delta\mu_{\text{a}}(\lambda) \) measured for solutions of known \( C_{\text{cal,glu}} \) and \( C_{\text{cal,IL}} \) based on the relation:

\[
\Delta\mu_{\text{a}}(\lambda) = \mu_{\text{a}}(\lambda; C_{\text{cal,glu}}, C_{\text{cal,IL}}) - \mu_{\text{a}}(\lambda; 0,0) = \sigma_{a,\text{glu}}(\lambda)C_{\text{glu}} + \sigma_{a,\text{IL}}(\lambda)C_{\text{IL}} + \text{higher-order terms}(C_{\text{glu}}^{\text{cal}}C_{\text{IL}}^{\text{cal}}, C_{\text{glu}}^{\text{cal}}C_{\text{IL}}^{\text{cal}} - C_{\text{glu}}^{\text{cal}} - C_{\text{IL}}^{\text{cal}}),
\]

(3)

where superscript “cal” indicates that these are calibrated concentrations, and \( \sigma_{a,\text{glu}}(\lambda) \), \( \sigma_{a,\text{IL}}(\lambda) \) are the specific absorption coefficients. As it turns out, all higher-order nonlinear terms were found insignificant and the relation becomes:

\[
\Delta\mu_{\text{a}}^{\text{cal}} = \Sigma \cdot C^{\text{cal}},
\]

(4)

where:

\[
\Delta\mu_{\text{a}}^{\text{cal}} = \begin{pmatrix}
\Delta\mu_{\text{a,glu}}(\lambda_1) & \Delta\mu_{\text{a,glu}}(\lambda_2) & \cdots & \Delta\mu_{\text{a,glu}}(\lambda_n) \\
\Delta\mu_{\text{a,IL}}(\lambda_1) & \Delta\mu_{\text{a,IL}}(\lambda_2) & \cdots & \Delta\mu_{\text{a,IL}}(\lambda_n)
\end{pmatrix}, \quad \Sigma = \begin{pmatrix}
\sigma_{a,\text{glu}}(\lambda_1) & \sigma_{a,\text{IL}}(\lambda_1) \\
\sigma_{a,\text{glu}}(\lambda_2) & \sigma_{a,\text{IL}}(\lambda_2) \\
\vdots & \vdots \\
\sigma_{a,\text{glu}}(\lambda_n) & \sigma_{a,\text{IL}}(\lambda_n)
\end{pmatrix},
\]

\[
C^{\text{cal}} = \begin{pmatrix}
C_{\text{glu}}^{\text{cal}} & C_{\text{IL}}^{\text{cal}} & \cdots & C_p^{\text{cal}}
\end{pmatrix}^T,
\]

(5)

where \( p \) is the number of calibrated samples. With linear regression, the estimate for \( \Sigma \) is:

\[
\tilde{\Sigma} = \Delta\mu_{\text{a}}^{\text{cal}}(C^{\text{cal}}C^{\text{cal}})^{-1}.
\]

(6)

The relation between \( Q \) and \( \Sigma \) is simply: \( Q = (\Sigma^T\Sigma)^{-1}\Sigma^T \) and Eq. (2) becomes:

\[
\tilde{C} = (\Sigma^T\Sigma)^{-1}\Sigma^T \cdot \Delta\mu_{\text{a}},
\]

(7)

2.1.2 Uncertainty and confidence level analysis

The critical issue and an objective of this work is the confidence level of \( \tilde{C} \). Mathematically, it is a function of two independent random variables \( \Delta\mu_{\text{a}} \) and \( \Sigma \). The error of \( \Delta\mu_{\text{a}} \) is
based on the photometric system: laser, detector, electronic noises, and other measurement apparatus errors. Σ is a random variable not in the sense that it varies from one measurement to the next, but that it has an intrinsic uncertainty associated from imperfect calibration.

Ideally, if Σ is very accurately calibrated, the uncertainty of Ĉ was caused only by photometric measurement error. In this case, Δµu, error can be characterized by a normal distribution N(0, δ). Then, C is also a normal distribution function with an estimated variance:

\[ \text{variance}(C) = (\Sigma^T \Sigma)^{-1} \delta^2; \quad \text{or variance}(C_i) = (\Sigma^T_i \Sigma_i)^{-1} \delta_i^2. \] (9)

Realistically however, there is no perfect calibration and this is a key issue for optical glucose technologies. Calibration may have to be done frequently, as it can be unique to an individual or apparatus and can change over time. The expected variance of Σ from Eq. (7) is:

\[ \text{variance}(\Sigma) = (C^{cal} C^{cal^T})^{-1} \text{variance}(\Delta \mu_u^{cal}), \] (10)

which also depends on Δµu0 photometric error, though it is not necessarily the same as that in Eq. (9). Photometry thus has a two-fold impact on error, and the quality of calibration is critical.

With Σ uncertainty taken into account, the uncertainty of C can then be complex, and variance can be a useless concept. For example, consider the simplest case when Σ is a normal-distribution scalar with mean δ and standard of deviation u; the probability density function (PDF) of C of Eq. (8) is then:

\[
\text{PDF}(C) = \frac{u}{\pi(\delta^2 + C^2u^2)} \exp \left[-\frac{1}{2} \left(\frac{\Delta \mu_u^2}{\delta^2} + \frac{C^2}{u^2}\right)\right] \\
+ \frac{\delta(\delta^2 + C\hat{C}u^2)}{\sqrt{2\pi}(\delta^2 + C^2u^2)^{3/2}} \exp \left[-\frac{(C - \hat{C})^2}{2(\delta^2 + C^2u^2)/\delta^2}\right] \\
\times \text{erf} \left[\frac{\delta(\delta^2 + C\hat{C}u^2)}{u(2\pi(\delta^2 + C^2u^2))^{1/2}}\right], \] (11)

which has an infinite variance owing to the well-known Cauchy tail ~1/π2 at large C. Instead of variance, a rigorous measure is the confidence interval, e.g., the interval [Ĉ−a/2, Ĉ+a/2] given by equation

\[
\int_{\hat{C}-a/2}^{\hat{C}+a/2} \text{PDF}(C) dC = 0.99, \]

which has a 99% probability of containing the true value of glucose concentration.

If Σ uncertainties uij are small: max(uij)/||Σ|| ≪ 1, a cut-off can be imposed to keep only the Gaussian component and an approximated variance can be used:

\[ \text{variance}(C) = (\Sigma^T \Sigma)^{-1} \delta^2 \left[1 + \frac{\Delta \mu_u^2}{\delta^2} u^2 (\Sigma^T \Sigma)^{-1}\right], \] (12)

where u=\text{max}(uij). The uncertainty of C is degraded from Eq. (9) by an extra term (Δµu0/δ)u2(ΣTΣ)−1, where Δµu0 is Δµu mean square. The simple and intuitively obvious implication of Eq. (12) is that for any system, there is no need to measure Δµu with much less error than that of Σ and vice versa. More importantly, Σ is essentially a spectrum, and arbitrary applications of techniques such as spectral smoothing and other digital filtering, while producing appealing-looking calibration spectra, violates its integrity and may induce spectral distortion that affects the confidence of Ĉ. A merit and objective of laser-based photometry is to obtain Σ and Ĉ with high confidence that is causally and rigorously linked to the system performance.

2.2 Experimental Setup

2.2.1 Optical configuration and experimental condition

The experiments were set up to measure sample absorption in the transmission mode, using a tunable MIR semiconductor lasers around 9.6 μm (~1040 cm−1) and a reference laser at 5.4 μm (~1850 cm−1). A schematic diagram is shown in Fig. 1. The sample is a thin liquid film held between two CaF2 windows by surface tension. One window is fixed, and the other is movable with to a stepping-motor-driven micrometer. The sample thickness was varied 50 to 250 μm. The laser beam was focused to a spot ~100 μm diameter. The sample volume was ~0.4~2×10−6 mL, which is in the range for microassay or microfluidic applications.

The sample temperature was not measured, but the room temperature was controlled at ~25±0.5°C. Owing to very low laser power (compare Sec. 2.2.3), the instantaneous temperature rise due to laser adiabatic heating is calculated to be ~0.01 K, and the steady-state average temperature rise is <0.005 K (based on water-specific heat of 4.186 J/g K and conductivity of 0.58 W/m K). Thus, the laser-induced temperature rise is insignificant compared to the room temperature fluctuation. There was no effort to study temperature-dependence absorption.

2.2.2 Sample preparation

The anhydrous glucose (Avocado Research Chemicals, Heysham, England) was used to prepare the glucose deionized water (DW) water solution samples. Glucose concentrations...
ranged from 0 to 500 mg/dL. The Intralipid® solution of 10% (Baxter, Deerfield, Illinois) was diluted by DI water and mixed with stock glucose solution. The uncertainty of glucose concentration was calculated to be $\sim$1.5 mg/dL based on the errors of glucose weight and water volume measurement. Intralipid® concentration estimated uncertainty was ±0.1%, based on errors of measured liquid volume for mixing.

### 2.2.3 Laser and detector

The MIR tunable laser was operated in single mode, covering from 9.5 to 9.8 μm (1020 to 1050 cm$^{-1}$) as illustrated in Fig. 2, with a superposition of glucose and water absorption spectrum. The other laser has a fixed wavelength at 5.4 μm, and serves as a reference. The lasers were operated in pulse mode with repetition rate from 25 to 50 kHz; the pulse width ranged from 200 to 400 ns. The pulse peak power was from 5 to 10 mW. The duty cycle was 0.5 to 1%, and the average power was only 25 to 100 μW. A MCT photodiode (Kolmar Technologies) detector with a transimpedance amplifier (TIA) was used to measure the transmitted power.

### 2.2.4 Data acquisition, signal processing, and system noise

An automated data acquisition system was used. The computer generated a sweep of sample thickness and the transmission was measured. The signal was digitized at 40-MHz sampling rate. The overall photometric system noise was determined from three sources: laser noise, detector noise, and electronic noise. The laser RIN power spectral density was measured to be $-95$ dB/Hz, which appear to be white-noise like without $1/f$ component down to 1 Hz. The combined detector and electronic noise equivalent power spectral density was measured to be $\sim 10^{-23}$ W$^2$/Hz.

### 3 Experimental Results

The experimental results include: 1. demonstration of large dynamic range and sample thickness for accurate measurement of absorption coefficients $\mu_a$, 2. calibration for specific absorption coefficient $\sigma_{a,\text{gla}}$ and its accuracy, 3. the effect of IL as an interferent substance, and 4. application to the case of fixed and unknown sample optical thickness.

#### 3.1 System Dynamic Range and Measurements of Absorption Coefficient

Figure 3 shows a typical transmittance that varies over 4 orders of magnitude as a function of sample thickness, which is essential for the accuracy and precision of $\mu_a$ obtained as the slope of the curve. The lower curve was for a glucose solution, and the upper curves were for pure water as a reference. The full signal range was actually over 6 orders of magnitude, but only results for thickness $\geq 120$ μm were shown. This signal dynamic range and sample thickness is much larger than those in FTIR-based works. Speed is also a figure of merit. The time for a typical scan was a few seconds, limited by the computer control and acquisition time. The actual laser on-time was only $\sim 120$ ms and its total energy was $<1$ mJ for the entire scan.

It appears in Fig. 3 that $\mu_a$ can be obtained via Beer’s law: $\ln T = -\mu_a d$, where $T$ is the transmittance and $d$ the sample optical thickness. However, Beer’s law is not necessarily correct, especially for laser measurements. The reason is that a focused laser beam contains many transverse components, and the effective path length is longer than the actual. Proper analysis is needed and the exact theory is given in the Appendix in Sec. 5. Deviation from Beer’s law was indeed observed for a tightly focused spot and small sample thickness in this work. Calculation was performed to ensure that the results here were obtained for sufficiently large spot size and sample thickness to correctly infer the absorption coefficient.
An example of \( \mu_a \) measurement is the upper curve of Fig. 3, which actually consists of five single-scan results but are indistinguishable on this scale. The slopes yield \( \mu_{a,\text{water}} \) (9.54 \( \mu \text{m} \)) ranging from 530.17 to 532.53 \( \text{cm}^{-1} \), with regression standard errors (SE) varying from 0.12 to 0.29 \( \text{cm}^{-1} \), which is \( \sim 2 \times 10^{-4} \) in relative terms, and the regression \( R^2 \) ranges from 0.999996 to 0.999999.

More important is the residual results shown in Fig. 4. Ideally, they should be consistent with photometric error as discussed in Eq. (9). Large nonrandom or systematic residuals might indicate the presence of uncontrolled variables. The residuals in Fig. 4 appear random with magnitude and behavior consistent with expected photometric system noise, which is plotted as the dashed curve in Fig. 4. It is based on the transmittance signal expressed as:

\[
S = \bar{S} + \text{noise} = \bar{S}(1 + \text{noise}/\bar{S})
\]

\[
\ln(S) - \ln(\bar{S}) = \ln(1 + \text{noise}/\bar{S}) = \text{noise}/\bar{S} = [n_{\text{det}} + (\bar{S}n_{\text{RIN}})^2]^{1/2}/\bar{S},
\]

where \( n_{\text{det}} \) and \( n_{\text{RIN}} \) are the detector noise and laser RIN, respectively. At small sample thickness, the transmitted signal is strong and the noise is dominated by laser RIN, since \( \bar{S}n_{\text{RIN}} \gg n_{\text{det}} \) and the noise is

\[
-\bar{S}n_{\text{RIN}}[N_{\text{RIN}}(f) \times \Delta f]^{1/2} \sim 2 - 3 \times 10^{-4},
\]

where \( N_{\text{RIN}}(f) \) is the laser RIN power spectral density mentioned in Sec. 2.2.4, and \( \Delta f = 250 \text{ Hz} \) is the bandwidth. At large thickness, the reverse is true: the signal is weak and dominated by detector noise. The dashed curve and data are in excellent agreement, verifying the causal relation between \( \mu_a \) measurement accuracy and the basic photometry performance.

However, more extensive measurements with repeated scanning indicated another source of error unrelated to photometry. It was the mechanical variation of the sample thickness control assembly. This mechanical error was intermittent but sufficiently frequent to degrade \( \mu_a \) precision to \( \sim 1 \) to 2 \( \times 10^{-3} \). It could be traced to the \( \sim 90-\text{nm} \) scan-to-scan variation of the micrometer, independently observed via interferometry. This variation was actually within the manufacturer specification of \( \sim 100-\text{nm} \) repeatability. Another potential source of errors is the thermal expansion of the assembly, which was built on a 5-cm-long steel frame with a thermal expansion of 60 \( \text{nm/K} \). If repeated scanning heated the frame, or room temperature varied by \( \sim 1^\circ \text{C} \), this could also contribute to errors, although evidence suggested it was less likely than the former.

Observation of this mechanical error actually underscores the importance of establishing a baseline photometric accuracy. This type of micrometer is usually adequate for many applications, but the reason for its error to be detected was because of the high sensitivity and accuracy of photometric measurements. An implication is that the full advantage of laser accuracy can be exploited for cases without moving parts or with high mechanical precision, such as microfluidic channels or optical waveguide probes.

### 3.2 Specific Absorption Coefficient of Glucose (without Intralipid\(^\text{\textregistered}\))

Calibration for specific glucose absorption coefficient \( \sigma_{a,\text{glu}} \) was obtained for solutions without IL to establish the baseline glucose accuracy without interferent. It was derived from the differential absorption \( \Delta \mu_{a,\text{glu}} = \mu_a(C_{\text{cal}}) - \mu_a(0) \) as described by Eqs. (3)–(7). A challenge is that \( \Delta \mu_{a,\text{glu}} \) is a just small difference (\( \leq 6\% \)) between two large absorption slopes, as seen in Fig. 3. Figure 5(a) shows \( \Delta \mu_{a,\text{glu}}(\lambda) \) as a function of...
$C_{\text{glu}}$ from 0 to 500 mg/dL at $\lambda = 5.4$, 9.63, and 9.54 $\mu$m. The resulting regression statistics based on Eq. (3) are listed in Table 1. There are two main results: the validity of the linear model, and the accuracy of $\sigma_{a,\text{glu}}$ calibration.

The linear model discussed in Sec. 2.1 is clearly valid, as part A of Table 1 shows that the nonlinear terms ($C_{\text{glu}}^2$, dependency) for both the 9.63- and 9.54-$\mu$m results are insignificant, as indicated in column “$P(H_0:\text{coef}=0)$” (which indicates the probability for the null $H_0$ hypothesis that the nonlinear coefficient=0).

The linear regression results are shown in part B of Table 1, with $\hat{\sigma}_{a,\text{glu}}(\lambda) = 5.697 \times 10^{-2}$ and $3.588 \times 10^{-2}$ cm$^{-1}$/mg/dL for 9.63 and 9.54 $\mu$m, respectively. Their uncertainties are given in the SE column, and in terms of relative unit, 0.76 and 0.62%. Ideally, $\hat{\sigma}_{a,\text{glu}}$ SE should be determined only by $\Delta \mu_a$ photometric error. Unfortunately, the mechanical error dominated these results. As evidenced in Fig. 5(b), the root-mean-squared (RMS) residual error of $\Delta \mu_{a,\text{glu}}(\lambda)$ is $\sim 0.29$ cm$^{-1}$ for $\lambda = 9.5-9.6 \mu$m, which is consistent with mechanical errors for ten scan averages. The photometric error contribution should be $\sim 0.05$ cm$^{-1}$. Without the mechanical error, the expected relative SE of $\hat{\sigma}_{a,\text{glu}}(\lambda)$ should have been $\sim 1.4 \times 10^{-3}$ and $1.1 \times 10^{-3}$.

If this calibration is used to infer glucose concentration in unknown samples based on the simplified Eq. (8): $\hat{C} = [\hat{\sigma}_{a,\text{glu}}]^{-1} \Delta \mu_a$, the confidence level can be rigorously established based on Eq. (11), and a calculation result is shown in Fig. 6. The upper solid curves represent the confidence level for an actual worst-case mechanical error, and the lower dashed curves are for the hypothetical photometric-only error. For both, the calculation was for a single-scan error of $\Delta \mu_a$. The results show that the given mechanical error, glucose could be determined within $\pm 21$ mg/dL with 90% confidence and $\pm 62$ mg/dL with 99% confidence. For hypothetical photometry-only errors, the results would be $\pm 6$ and $\pm 9$ mg/dL for 90 and 99% confidence, respectively.

It is also clear in part B in Table 1 that the linear coefficient for 5.4 $\mu$m is within the SE, suggesting that glucose has no significant absorption at this wavelength. Thus, it can be used as a reference for samples of fixed but unknown thickness. This is further discussed in Sec. 3.4. Also evident in Fig. 5(b) is that the RMS error for $\lambda = 5.4 \mu$m is smaller than that of 9.5 to 9.6-$\mu$m by a factor proportional to $\mu_a$ at these wavelengths, which are 233 cm$^{-1}$ at 5.4 $\mu$m and $\sim 541$ cm$^{-1}$ at 9.6 $\mu$m. This behavior indicates that $\Delta \mu_a$ error is not a constant-amplitude additive term, but is consistent with the mechanical error, which is proportional to $\mu_a$.

### 3.3 Effect of Intralipid® Interference

A matrix of samples with glucose concentration ranging from 0 to 500 mg/dL and IL concentrations from 0 to 8% were measured at $\lambda = 9.63$ and 5.4 $\mu$m. The differential absorption coefficients $\Delta \mu_{a,\text{glu}}$ are listed in Table 2 and plotted in Fig. 7.

---

**Table 1** Regression results of differential absorption coefficients of glucose in water.

<table>
<thead>
<tr>
<th>Wavelength ($\mu$m)</th>
<th>Coefficient</th>
<th>Standard error (SE)</th>
<th>Student Test</th>
<th>$P(H_0:\text{coef}=0)$</th>
<th>$R^2$ Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.63</td>
<td>Linear</td>
<td>0.05834</td>
<td>0.00147</td>
<td>39.7</td>
<td>$1.8 \times 10^{-10}$</td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>$-3.6 \times 10^{-6}$</td>
<td>$3.7 \times 10^{-6}$</td>
<td>$-0.98$</td>
<td>0.356</td>
</tr>
<tr>
<td>9.54</td>
<td>Linear</td>
<td>0.05288</td>
<td>0.00144</td>
<td>36.8</td>
<td>$3.2 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>$2.4 \times 10^{-6}$</td>
<td>$3.4 \times 10^{-6}$</td>
<td>0.72</td>
<td>0.512</td>
</tr>
<tr>
<td>9.63</td>
<td>Linear</td>
<td>0.05697</td>
<td>0.000433</td>
<td>131.5</td>
<td>$4.4 \times 10^{-16}$</td>
</tr>
<tr>
<td>9.54</td>
<td>Linear</td>
<td>0.05388</td>
<td>0.000333</td>
<td>161.5</td>
<td>$1.7 \times 10^{-10}$</td>
</tr>
<tr>
<td>5.4</td>
<td>Linear</td>
<td>0.00017</td>
<td>0.00018</td>
<td>0.945</td>
<td>0.369</td>
</tr>
</tbody>
</table>

---

**Fig. 6** Calculated confidence interval for inferred (predicted) glucose concentration. The upper solid curves are for actual errors; the lower curves are hypothetical confidence intervals if there were only photometric errors.
Each measurement resulted from a ten-scan average. The symbols are experimental data, and the planes are linear regression fits. It is clear that two substances have nearly or-thogonal absorption behaviors at these wavelengths. Glucose shows strong absorption at 9.63 μm, while IL has little ab-sorption; the reverse is true at 5.4 μm. Interestingly, IL has a negative absorption coefficient at both wavelengths. This can be interpreted that the IL volume that displaces water has a lower absorption coefficient than water. Similar negative effects were reported for blood at 4.6 and 5.4 μm in a previous work. The analysis according to Eq. (7) was applied to obtain matrix $\Sigma$ of Eq. (6). The main results include: 1. the validity of the linear model in Eq. (4); 2. calibration matrix $\Sigma$; and 3. the effect of IL on the precision and accuracy of glucose determination.

### 3.3.1 Probability of nonlinear terms

The regression results for nonlinear terms are shown in Table 3, part A. For 9.6 μm, a t-test clearly shows that various nonlinear terms can be neglected. However, the results for 5.4 μm indicates a probable quadratic term for $C_{\text{IL}}^2$ with $P(H_0: \text{coef}=0)=1.2 \times 10^{-8}$. This quadratic term persisted even when glucose is eliminated from the regression model, as shown in the last two rows of Table 3, part A. Further analysis, however, suggests that this nonlinear term may not be as probable as suggested by the statistics. The regression residuals for data of different glucose concentrations are not random, but systematically form clusters with respect to each value of IL concentration. A possibility for this peculiar result is that each set of samples with the same IL but different glucose concentrations were measured under similar room temperature and mechanical conditions. But sets with different IL concentrations were measured at different times and conditions, and thus had different errors. This systematic error would generate false quadratic behavior only for IL but none for glucose. Also, for both wavelengths, the cross-correlation terms $C_{\text{glu}}C_{\text{IL}}$ are insignificant, indicating that their absorp-tions are simply additive to each other.

### Table 2 Water-subtracted absorption coefficients (cm$^{-1}$) for various glucose and Intralipid® concentrations (absorption of pure water in parentheses).

<table>
<thead>
<tr>
<th>λ</th>
<th>Glucose concentration (mg/dL)</th>
<th>0%</th>
<th>2%</th>
<th>5%</th>
<th>8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.63 μm</td>
<td>(water $\mu_a$=541.1 cm$^{-1}$) 0</td>
<td>0.00</td>
<td>-0.96</td>
<td>-2.92</td>
<td>-4.01</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>6.08</td>
<td>4.8</td>
<td>2.67</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>11.67</td>
<td>10.33</td>
<td>8.58</td>
<td>7.04</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>17.48</td>
<td>15.73</td>
<td>14.00</td>
<td>12.12</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>22.74</td>
<td>21.33</td>
<td>19.45</td>
<td>17.44</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>28.26</td>
<td>26.57</td>
<td>25.43</td>
<td>24.34</td>
</tr>
<tr>
<td>5.4 μm</td>
<td>(water $\mu_a$=233.91 cm$^{-1}$) 0</td>
<td>0.00</td>
<td>-7.98</td>
<td>-17.58</td>
<td>-25.88</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.15</td>
<td>-8.30</td>
<td>-17.45</td>
<td>-26.15</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.26</td>
<td>-8.14</td>
<td>-17.55</td>
<td>-26.24</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.11</td>
<td>-8.18</td>
<td>-17.48</td>
<td>-26.46</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0.01</td>
<td>-8.21</td>
<td>-17.69</td>
<td>-26.35</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>-0.03</td>
<td>-8.53</td>
<td>-17.29</td>
<td>-26.45</td>
</tr>
</tbody>
</table>

![Fig. 7 Water-subtracted absorption coefficient as a function of glucose and Intralipid® concentration for two wavelengths. The symbols represent measurements, and the planes are linear regression fits.](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)
### Table 3 Regression results for glucose and Intralipid® absorptions.

#### Part A - With nonlinear coefficients

<table>
<thead>
<tr>
<th>Wavelength (μm)</th>
<th>Coefficient</th>
<th>Standard error (SE)</th>
<th>Student t</th>
<th>P(H0: coef=0)</th>
<th>r²</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.63</td>
<td>$\sigma_{glu}$</td>
<td>0.05772</td>
<td>0.00113</td>
<td>51.3</td>
<td>$7.7 \times 10^{-22}$</td>
<td>0.99963</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{il}$</td>
<td>-0.639</td>
<td>0.0737</td>
<td>-8.68</td>
<td>$4.9 \times 10^{-8}$</td>
<td>0.99953</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{\mu}^{(2)}$</td>
<td>$-2.1 \times 10^{-6}$</td>
<td>$2.5 \times 10^{-6}$</td>
<td>-0.821</td>
<td>0.422</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma_{\mu}^{(2)}$</td>
<td>0.0147</td>
<td>0.00952</td>
<td>1.54</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>$-1.5 \times 10^{-4}$</td>
<td>$1.1 \times 10^{-4}$</td>
<td>-1.39</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>$\sigma_{glu}$</td>
<td>-0.00047</td>
<td>0.00122</td>
<td>-0.38</td>
<td>0.706</td>
<td>0.99960</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{il}$</td>
<td>-4.035</td>
<td>0.0802</td>
<td>-50.3</td>
<td>$1.1 \times 10^{-21}$</td>
<td>0.99950</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{\mu}^{(2)}$</td>
<td>$2.8 \times 10^{-7}$</td>
<td>$2.7 \times 10^{-6}$</td>
<td>0.1</td>
<td>0.919</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma_{\mu}^{(2)}$</td>
<td>0.0986</td>
<td>0.0104</td>
<td>9.52</td>
<td>$1.2 \times 10^{-8}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>$-4 \times 10^{-5}$</td>
<td>$1.2 \times 10^{-4}$</td>
<td>-0.335</td>
<td>0.741</td>
<td></td>
</tr>
</tbody>
</table>

#### Part B - With linear coefficient only

<table>
<thead>
<tr>
<th>Wavelength (μm)</th>
<th>Coefficient</th>
<th>Standard error (SE)</th>
<th>Student t</th>
<th>P(H0: coef=0)</th>
<th>r²</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.63</td>
<td>$\sigma_{glu}$</td>
<td>0.05632</td>
<td>0.000305</td>
<td>184.5</td>
<td>$1.4 \times 10^{-36}$</td>
<td>0.99954</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{il}$</td>
<td>-0.5500</td>
<td>0.0192</td>
<td>-28.7</td>
<td>$6.3 \times 10^{-19}$</td>
<td>0.99950</td>
</tr>
</tbody>
</table>

### ANOVA

<table>
<thead>
<tr>
<th>Degree of freedom</th>
<th>Sum of square</th>
<th>Mean of square</th>
<th>F ratio</th>
<th>P(H0: coef=0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>5753</td>
<td>2876.47</td>
<td>23871</td>
</tr>
<tr>
<td>Error</td>
<td>22</td>
<td>2.651</td>
<td>0.1205</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>5756</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Covariance matrix

$10^{-4} \begin{bmatrix} 9.3 \times 10^{-4} & -0.038 \\ -0.038 & 3.68 \end{bmatrix}$

#### 5.4

<table>
<thead>
<tr>
<th>Wavelength (μm)</th>
<th>Coefficient</th>
<th>Standard error (SE)</th>
<th>Student t</th>
<th>P(H0: coef=0)</th>
<th>r²</th>
<th>Adjusted r²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_{glu}$</td>
<td>-0.00169</td>
<td>0.000742</td>
<td>-2.28</td>
<td>0.033</td>
<td>0.99755</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{il}$</td>
<td>-3.3084</td>
<td>0.0466</td>
<td>-71</td>
<td>$1.7 \times 10^{-27}$</td>
<td>0.99732</td>
</tr>
</tbody>
</table>

### ANOVA

<table>
<thead>
<tr>
<th>Degree of freedom</th>
<th>Sum of square</th>
<th>Mean of square</th>
<th>F ratio</th>
<th>P(H0: coef=0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>6366</td>
<td>3183</td>
<td>4472</td>
</tr>
<tr>
<td>Error</td>
<td>22</td>
<td>15.7</td>
<td>0.712</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>6381</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Covariance matrix

$10^{-3} \begin{bmatrix} 5.5 \times 10^{-4} & -0.022 \\ -0.022 & 2.17 \end{bmatrix}$

Guo et al.: Mid-infrared laser measurements of aqueous glucose

Journal of Biomedical Optics March/April 2007 /L50098Vol. 12/H208492/H20850024005-8
3.3.2 Linear coefficients
With the nonlinear terms excluded, the linear regression results are shown in Table 3, part B. For both, the $F$-ratio statistics indicated that the linear model alone was quite sufficient to account for the data. The coefficients $\hat{a}_{\text{glu}}$, $\hat{a}_{\text{IL}}$ along with the contour of 95% joint-confidence ellipse (based on $F$-ratio distribution) and individual 84% confidential intervals (based on $t$-distribution) are plotted in Fig. 8. For 9.63 $\mu$m, the $\hat{a}_{\text{glu}}$ value of 0.0563 cm$^{-1}$/mg/dL is comparable to the value 0.05697 for no-IL obtained in Sec. 3.2. For 5.4 $\mu$m, $\hat{a}_{\text{glu}}$ is clearly not significant.

3.3.3 Inferred glucose and intralipid concentration
With calibrated coefficients $\Sigma$, the reverse inference of glucose and IL concentration can be obtained via Eq. (8). There was no need to prepare unknown samples given the available extensive number of samples. The results are shown in Fig. 9. The filled circles represent known values, the open squares with center dots represent inferred values. The residuals of all 24 concentrations are shown in the scatter plot of Fig. 10.

It appears from Figs. 9 and 10 that the residuals of IL concentration do not seem to be random around the expected values. All inferred IL concentrations in Fig. 9 are either higher or lower than expected values, forming two distinctive clusters in Fig. 10. This reflects the systematic error discussed before for the 5.4-$\mu$m IL data.

More important are the glucose results. The RMS of all residual glucose errors shown in Fig. 10 is 5.7 mg/dL, indicated by the two thick dashed lines. Figure 11 plots the glucose RMS error for each IL. Here, the result in Sec. 3.2 is just a subset with $C_{\text{IL}}=0\%$. These RMS errors are consistent with the mechanical errors, and there is no obvious correlation between glucose RMS and $C_{\text{IL}}$, indicating that IL has no effect on glucose error. The result is hardly surprising, considering that the two substances have orthogonal absorption behaviors at the two wavelengths.

Ellipsoid confidence contours can be calculated for inferred glucose and IL concentrations; however, they are not...
relevant with regard to IL, owing to the IL systematic errors, and IL is also not of interest. Therefore, the bivariate measurements of aqueous glucose

![Graph showing confidence interval for inferred glucose concentration](image)

Fig. 12 Calculated confidence interval for inferred glucose concentration with an arbitrary IL concentration between 0 to 8%. The upper solid curves are for a single scan; the lower curves are for a ten-scan average.

3.4 Approach for Fixed and Unknown Thickness

In applications such as monitoring a microfluidic channel, microassay array, or diffuse scattering from tissues, the sample effective optical thickness $d$ is usually fixed or unknown. The system must infer the concentrations without varying the thickness. For these cases, the algorithm is slightly modified from that in Sec. 2.1 by treating $d$ as an additional unknown variable. The details are described in Appendix B in Sec. 6. The essential result is similar to Eq. (8):

$$\hat{C}_X = - (\Sigma_X^T \Sigma_X)^{-1} \Sigma_X^T \ln S,$$

where $\hat{C}_X$ is modified from $\hat{C}$ by including $d$, $\Sigma_X$ is modified by including water absorption, and a key difference is to use $\ln S$ instead of $\Delta \mu_w$, where $S$ is the optical signal that can be transmitted or backscattered.

It is sufficient to consider an example to illustrate the principle, and this is summarized in Table 4. The first two rows show the numerical values of $\Sigma_X$, $(\Sigma_X^T \Sigma_X)^{-1}$ and $\ln S(\lambda)$, which are obtained from the transmittance data of an aqueous glucose sample set at a fixed assembly thickness. The last two rows show inferred value $\hat{d} = 150.6 \mu m$ and $\hat{C}_{glu} = 570$ mg/dL, to be compared with known values 150$\mu m$ and 500 mg/dL, respectively. There is good agreement for $\hat{d}$, but $\hat{C}_{glu}$ is a bit off, which is likely due to some unaccounted optical loss (Appendix B in Sec. 6).

A crucial aspect of this result is the choice of wavelengths. A requirement of Eq. (14) is that $(\Sigma_X^T \Sigma_X)^{-1}$ is not badly conditioned, which in principle is no different from Eq. (8). In practice, however, it is a more critical requirement, since $\Sigma_X$ contains $\mu_{w,water}(\lambda)$, which provides spectral contrast to infer $d$. Since $\mu_{w,water}(\lambda)$ is large, the $\lambda$’s must be chosen such that $\mu_{w,water}(\lambda)$ has variation much larger than their uncertainty. In the second row of Table 4, the small-font numbers for $(\Sigma_X^T \Sigma_X)^{-1}$ indicate their uncertainty. Their small uncertainty values ensure that $(\Sigma_X^T \Sigma_X)^{-1}$ is meaningful, owing to the large water absorption difference between 5.4 and 9.6 $\mu m$. In comparison, any wavelength set in the laser tuning range 9.5 to 9.8 $\mu m$ was not as useful, since water absorption varies too slowly over this range. Both this and the IL results suggest that large wavelength diversity can be advantageous for accuracy and confidence, as discussed in Appendix B Sec. 6.

**Table 4** Example of inferring unknown optical thickness.

<table>
<thead>
<tr>
<th>$\lambda_1 = 5.4 \mu m$, $\lambda_2 = 9.63 \mu m$</th>
<th>$\Sigma_X = \begin{bmatrix} \mu_{w,water}(\lambda_1) &amp; \sigma_{a,glu}(\lambda_1) \ \mu_{w,water}(\lambda_2) &amp; \sigma_{a,glu}(\lambda_2) \end{bmatrix}$</th>
<th>$(\Sigma_X^T \Sigma_X)^{-1}$</th>
<th>$\ln S(\lambda)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$233.91$</td>
<td>$0.17$</td>
<td>$541.1$</td>
<td>$56.97$</td>
</tr>
</tbody>
</table>

\[ \begin{bmatrix} 0.0185 & -0.176 \\ 10^{-3} & 3.9 \times 10^{-4} \end{bmatrix} \pm \begin{bmatrix} 5 \times 10^{-2} \\ -0.176 \end{bmatrix} \pm \begin{bmatrix} 1.98 \\ 0.07 \end{bmatrix} \]

<table>
<thead>
<tr>
<th>$\frac{d}{\hat{C}_{glu}} = - (\Sigma_X^T \Sigma_X)^{-1} \Sigma_X^T \ln S$</th>
<th>$\frac{d}{\hat{C}_{glu}} = 150.6 \mu m$</th>
<th>$\hat{d} = 150.6 \mu m$</th>
</tr>
</thead>
</table>

With proper unit conversion

\[ \frac{d}{\hat{C}_{glu}} = 150.6 \mu m, \quad \frac{\hat{d}}{C_{glu}} = 570 \text{ mg/dL} \]

\[ \frac{\hat{d}}{C_{glu}} = 500 \text{ mg/dL} \]
4 Summary and Conclusion

This work describes the use of mid-IR semiconductor lasers to measure aqueous glucose concentration and examines the issue of accuracy. Compared with many previous works\textsuperscript{7-14} that focused mainly on spectroscopy, this work aims at photometric engineering to demonstrate the merit of the laser-based technique as opposed to others such as FTIR. Specifically, the work focuses on the laser properties of high signal-to-noise ratio, large dynamic range, and speed. The laser capability to focus onto a small sample volume (~4 \times 10^{-7} \text{ mL}) is also suitable for microfluidic and microarray bioassay applications.

The approach measures the absorption coefficients of aqueous glucose and Intralipid\textsuperscript{®} solutions of various concentrations for different wavelengths (5.4, 9.54, and 9.63 \textmu m). Linear regression is applied to obtain the specific absorption coefficients for each solute, which are then applied to infer their concentrations and compare with known values.

The important issue is accuracy, and the errors are shown to be causally and deterministically related to the basic instrument noises and system errors. There are no unaccounted sources of error. This causal relation allows inferring measurement confidence level based on rigorous system understanding instead of phenomenological assessment. The results show that glucose concentration could be determined ±21 and ±32 mg/dL with 90 and 99% confidence, respectively, in a single measurement. These values improve to ±9.5 and ±15 mg/dL with ten measurements. These uncertainties were not limited by photometry, but by a larger mechanical error of the sample thickness control assembly. This mechanical error is not fundamental, and in applications without this problem, the expected uncertainty due to photometry-only error would be ±0 and ±9 mg/dL for 90 and 99% confidence level, respectively, for a single measurement. Intralipid\textsuperscript{®} up to 8% concentration had no effect on the glucose measurement.

The choice of wavelengths is also evidently important. The large difference between 5.4 and 9.63 \textmu m offers large spectral contrast between glucose, IL, and water that contributes to the measurement confidence. In general, for the laser method, fewer \lambda's entails a simpler system, whereas FTIR offers the entire spectrum. However, an insight from these results is that with regard to accuracy, measurements with many wavelengths but noisy is not necessarily better than those with fewer wavelengths but low noise. Other recent MIR works\textsuperscript{10-13,24} even with FTIR also suggest that a few wavelengths but low noise. Other recent MIR works\textsuperscript{10-13,24} even with FTIR also suggest that a few wavelengths might be sufficient. The laser approach can thus be applicable and practical.

5 Appendix A: Deviation from Beer’s Law for Focused Laser Beam

In transmission or diffuse backscattering, the laser light coherence property may require the full optical solution for accurate interpretation of the result. For an absorptive thin film with parallel interfaces, i.e., a lossy etalon, the transmittance is:

\[ t_{\text{TM}} = \{1 + [1 - \exp(2i d k_2)] (k_1 - k_2)^2 / 4k_2 k_3 \}^{-1} \exp(id k_2) \]

where TE and TM denote the transverse electric and magnetic polarization, respectively, and \( d \) is the sample thickness. For all other quantities, the subscripts 1 and 2 correspond to the window and sample, respectively, \( e = (n + i \kappa)^2 \) is the complex dielectric constant; \( k = \sqrt{\kappa_0^2 - \beta^2} \) with \( \kappa_0 = 2 \pi / \lambda \), and \( \beta \) is the transverse wave number of the beam parallel to the interface.

For small \( \beta \), the paraxial approximation can be used:

\[ k = \sqrt{\kappa_0^2 - \beta^2} = \sqrt{\kappa_0^2 - \kappa^2} = (n + i \kappa k_0) - \beta^2 / 2n^2 k_0 
\]

\[ \approx (n + i \kappa) k_0 - \beta^2 / 2n^2 k_0, \]

\[ k = n k_0 (1 - \beta^2 / 2n^2 k_0) + i \kappa k_0 \left(1 + \beta^2 / 2n^2 k_0 \right) \]

where \( \kappa \) is the beam propagation angle in the medium. Substituting the imaginary component in \( \exp(i d k_2) \) of Eq. (15) results in the absorption term:

\[ \exp(-d k_0 / \cos \theta) = \exp \left(-\frac{1}{2} \mu_a d / \cos \theta \right); \quad (\mu_a = 2n k_0), \]

which indicates that the beam experiences a larger loss factor \( 1 / \cos \theta \) associated with a longer path length. A focused laser beam is a superposition of many plane waves with nonvanishing transverse \( \beta \), therefore it experiences a higher loss than a normal-incidence plane wave.

In this work, the index of refraction of the CaF\textsubscript{2} window is very close to that of water, and there is little difference between TE and TM transmissions. Thus the net beam transmittance is:

\[ T = \int d \beta, d \beta_x A(\beta_x) A(\beta_y) |r_{\text{TM}}|^2. \]

where \( A(\beta_x), A(\beta_y) \) are the power spectral density functions of transverse spatial frequencies \( \beta_x, \beta_y \). For an \( x-y \) Gaussian beam:

\[ A(\beta) = \frac{w}{\sqrt{2\pi}} \exp(-w^2 \beta^2 / 2). \]

With paraxial approximation, applying Eq. (19) with Eq. (20) yields:

\[ T = \exp(-\mu d) \frac{w_x}{(w_x^2 + \mu_d (k_2 n_2^2)^{1/2})^{1/2}} \frac{w_y}{(w_y^2 + \mu_d (k_2 n_2^2)^{1/2})^{1/2}}, \]

and

\[ T \approx \exp(-\mu d) \frac{w_x}{(w_x^2 + \mu_d (k_2 n_2^2)^{1/2})^{1/2}} \frac{w_y}{(w_y^2 + \mu_d (k_2 n_2^2)^{1/2})^{1/2}}. \]

\[ t_{\text{TE}} = \{1 + [1 - \exp(2i d k_2)] (k_1 - k_2)^2 / 4k_2 k_3 \}^{-1} \exp(id k_2) \]
\[ T \approx \exp(-\mu_d d) \frac{k_0^2 w_1^2 w_2^2}{\mu_d} \text{ for small } w_1, w_2. \]  

From Eq. (21), Beer’s law would be obtained for large beam waists: \( w_1^2, w_2^2 \gg \mu_d d / k_0^2 \delta^2 (= \mu_d d \lambda^2 / 4 \pi n_2^2) \). But for small beam waists, Eq. (22) shows a deviation by the term \( k_0^2 w_1^2 w_2^2 / \mu_d d \), which would lead to substantial error of \( \mu_d \) if not included.

In this work, the semiconductor laser beam is not Gaussian and has \( x-y \) asymmetry. A numerical calculation for a modeled laser beam produces results shown in Figs. 13–15. Figure 13 shows that the transmittance is superficially similar to Beer’s law on a large scale. But as evidenced in Fig. 14, the residuals from the exponential fit clearly deviate from Beer’s law, even if the interference fringes are ignored. The deviation is worse for smaller spot size as expected. The “effective” absorption coefficient deviates from the actual coefficient as shown in Fig. 15. Considering that the largest differential glucose absorption is only \( \sim 28.5 \text{ cm}^{-1} \) in this work, a deviation of \( 4 \text{ cm}^{-1} \) would result in an error of 14%.

This error does not just occur on the slope of the log transmittance, but also on its amplitude. Thus, for cases with fixed-thickness samples, the error on transmittance amplitude is similarly as severe and correction is necessary.

6 Appendix B: Approach for Fixed or Unknown Sample Effective Optical Thickness

In general, the transmission or diffuse backscattered return signal can be expressed as:

\[ S(\lambda) = Q(\lambda) \exp[-\mu_d(\lambda)d], \]  

where \( Q(\lambda) \) is a loss factor that may or may not be known, and \( \mu_d(\lambda) \) is a sum of absorptions:

\[ \ln S(\lambda) = -\mu_{d,m}(\lambda)d - \sum_{i=1}^{k} \sigma_{a,i}(\lambda) C_i d + \eta(\lambda), \]  

where \( \eta(\lambda) = \ln Q(\lambda), \mu_{d,m}(\lambda) \) is the baseline absorption coefficient of the medium (which can be a solvent other than water), \( d \) can represent single or round-trip, and \( C_i \) for \( i=1 \) to \( k \) denotes concentration of substance of interest.

In controlled applications with fixed, known \( d \) and \( \eta \), Eq. (24) can be written:
\[ \frac{1}{d} \begin{bmatrix} \eta(\lambda_1) - \ln[S(\lambda_1)] \\ \eta(\lambda_2) - \ln[S(\lambda_2)] \\ \vdots \\ \eta(\lambda_n) - \ln[S(\lambda_n)] \end{bmatrix} = \begin{bmatrix} \mu_{a,w}(\lambda_1) \\ \mu_{a,w}(\lambda_2) \\ \vdots \\ \mu_{a,w}(\lambda_n) \end{bmatrix} + \begin{bmatrix} \sigma_{a,1}(\lambda_1) & \sigma_{a,2}(\lambda_1) & \cdots & \sigma_{a,k}(\lambda_1) \\ \sigma_{a,1}(\lambda_2) & \sigma_{a,2}(\lambda_2) & \cdots & \sigma_{a,k}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{a,1}(\lambda_n) & \sigma_{a,2}(\lambda_n) & \cdots & \sigma_{a,k}(\lambda_n) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{bmatrix}, \]  

(25)

and a solution similar to Eq. (8) is obtained:

\[ \hat{C} = (\Sigma^\top \Sigma)^{-1} \Sigma^\top \left[ \frac{1}{d} (\eta - \ln S) - \mu_{a,w} \right]. \]  

(26)

If \( d \) is unknown, Eq. (24) can be written:

\[ \begin{bmatrix} \eta(\lambda_1) - \ln[S(\lambda_1)] \\ \eta(\lambda_2) - \ln[S(\lambda_2)] \\ \vdots \\ \eta(\lambda_n) - \ln[S(\lambda_n)] \end{bmatrix} = \begin{bmatrix} \mu_{a,w}(\lambda_1) \\ \mu_{a,w}(\lambda_2) \\ \vdots \\ \mu_{a,w}(\lambda_n) \end{bmatrix} + \begin{bmatrix} \sigma_{a,1}(\lambda_1) & \sigma_{a,2}(\lambda_1) & \cdots & \sigma_{a,k}(\lambda_1) \\ \sigma_{a,1}(\lambda_2) & \sigma_{a,2}(\lambda_2) & \cdots & \sigma_{a,k}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{a,1}(\lambda_n) & \sigma_{a,2}(\lambda_n) & \cdots & \sigma_{a,k}(\lambda_n) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{bmatrix}, \]  

(27)

If \( \eta \) is negligible, Eq. (27) leads to Eq. (14). If \( \eta \) is unknown, the problem cannot be solved. However, \( \eta(\lambda) \) usually represents optics effects such as light collection loss, Fresnel loss, scattering loss, or chromatic dispersion, and therefore can be a slowly varying function of wavelength. For the sake of discussion, suppose that it can be parametrized to the first order:

\[ \eta(\lambda) = \eta_0 + \rho_1 \lambda, \]  

(28)

then Eq. (27) can be modified to yield:

\[ \begin{bmatrix} \ln[S(\lambda_1)] \\ \ln[S(\lambda_2)] \\ \vdots \\ \ln[S(\lambda_n)] \end{bmatrix} = \begin{bmatrix} \mu_{a,w}(\lambda_1) & \sigma_{a,1}(\lambda_1) & \cdots & \sigma_{a,k}(\lambda_1) \\ \mu_{a,w}(\lambda_2) & \sigma_{a,1}(\lambda_2) & \cdots & \sigma_{a,k}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{a,w}(\lambda_n) & \sigma_{a,1}(\lambda_n) & \cdots & \sigma_{a,k}(\lambda_n) \end{bmatrix} \begin{bmatrix} d \\ C_1d \\ \vdots \\ C_kd \end{bmatrix}, \]  

(29)

which allows inferring all three unknown parameters \( d, \eta_0 \), and \( \rho_1 \).

A crucial requirement relevant to the wavelengths used in this work is that \((\Sigma^\top \Sigma)^{-1}\) should have low uncertainty. Let each row or column of \( \Sigma \) be denoted as a vector \( s \). If the wavelength range is too narrow, various vector \( s \) can become close to each other, \((\Sigma^\top \Sigma)^{-1}\) can be nearly degenerate with large uncertainty, and the measurements would have poor confidence. A nominal criterion for the vectors to be “far” from each other is to use the quasi-Mahalanobis distance:

\[ \min \left[ \frac{\| s_i - s_j \|}{\sqrt{2}}, \forall \gamma \right] \]

\[ = \min \left( \sqrt{\sum_{p=1}^{5} \frac{[s_{ip} - s_{jp}]^2}{u^2(s_{ip}) + \gamma^2 u^2(s_{jp})}} \right)^{1/2}, \forall \gamma \gg 1, \]

(30)

where \( m_{xp} \) are the vector elements, \( u^2(s_{ip}) \) are their uncertainties, and \( \gamma \) is a proportional factor.

Applying this criterion in this case explains the reason why the 5.4- and 9.6-\( \mu \)m wavelength combination produced the best results in Sec. 3.4. The quasi-Mahalanobis distance for \( \Sigma \) columns with 5.4 and 9.63 \( \mu \)m is 93, compared to only \( \sim 3.5 \) to 6 for any set of wavelengths in the 9.5 to 9.8-\( \mu \)m range.

**Acknowledgment**

This work was supported in part by the DARPA/ARO under contract DAAD-190010361, the Air Force Research Laboratory under contract F29601-00-2-0058, and the State of Texas through the THECB and the Texas Center for Advanced Materials and Center for Superconductivity. We thank the late Pradeep Johnson, C. Peng, G. P. Luo, and A. Seetharaman for valuable assistance.

**References**


24. J. Burd (Private communication).


