ulti-layered Surface Estimation for Low-cost Optical Coherence Tomography

Rapp, Joshua; Mansour, Hassan; Boufounos, Petros T.; Koike-Akino, Toshiaki; Parsons, Kieran

Abstract

Optical coherence tomography (OCT) has broad applicability for 3D sensing, such as reconstructing the surface profiles of multi-layered samples in industrial settings. However, accurately determining the number of layers and their precise locations is a challenging task, especially for low-cost OCT systems having low signal-to-noise ratio (SNR). This paper introduces a principled and noise-robust method of detection and estimation of surfaces measured with OCT. We first derive the maximum likelihood estimator (MLE) for the position and reflectivity of a single opaque surface. We next derive a threshold that uses the acquisition noise variance and the number of measurements available to set a target probability for false acceptance of spurious surface estimates. The threshold and MLE are then incorporated into an algorithm that sequentially detects and estimates surface locations. We demonstrate reconstruction of fine details in samples with optical path lengths around 1 mm and depth error down to 1.5 mm despite SNRs as low as –10 dB.

IEEE Transactions on Computational Imaging 2024

Multi-layered Surface Estimation for Low-cost Optical Coherence Tomography

Joshua Rapp, *Member, IEEE,* Hassan Mansour, *Senior Member, IEEE,* Petros Boufounos, *Fellow, IEEE,*
Toshiaki Koike-Akino, *Fellow, Optica,* and Kieran Parsons, *Senior Member, IEEE*

Abstract—**Optical coherence tomography (OCT) has broad applicability for 3D sensing, such as reconstructing the surface profiles of multi-layered samples in industrial settings. However, accurately determining the number of layers and their precise locations is a challenging task, especially for low-cost OCT systems having low signal-to-noise ratio (SNR). This paper introduces a principled and noise-robust method of detection and estimation of surfaces measured with OCT. We first derive the maximum likelihood estimator (MLE) for the position and reflectivity of a single opaque surface. We next derive a threshold that uses the acquisition noise variance and the number of measurements available to set a target probability for false acceptance of spurious surface estimates. The threshold and MLE are then incorporated into an algorithm that sequentially detects and estimates surface locations. We demonstrate reconstruction of fine details in samples with optical path lengths around 1 mm and depth error down to 1.5 μm despite SNRs as low as –10 dB.**

Index Terms—**Surface estimation, detection, optical coherence tomography, 3D imaging, probability of false acceptance.**

I. INTRODUCTION

OPTICAL coherence tomography (OCT) is an imaging technique that uses a sequence of low-coherence interferometry (LCI) measurements to recover three-dimensional (3D) structures through partially scattering media [2]. In the common Fourier-domain (FD) OCT configuration, interference patterns are measured as a function of wavelength, and depth information is recovered via a Fourier transform. Although best known for its use in opthalmology since its initial demonstration of ocular imaging [3], OCT has also been applied to numerous other applications that require nondestructive microscopic 3D measurements [4]. For instance, OCT (or LCI) has been extensively used in industrial settings for inspecting and characterizing ceramics [5], polymer strain [6], wear damage [7], optical components [8], wafer topography [9], paper properties [10, 11], varnish or paint thickness [12, 13, 14, 15], printed electronics [16, 17], wind turbine delamination [18], and identification document security features [19]. In addition to 3D measurement, OCT can also be used to separate direct and global illumination, and measure material properties such as scattering, dispersion, or birefringence [20, 21].

A major factor limiting the widespread adoption of OCT for industrial inspection tasks is the financial cost of the low-noise, swept-source lasers used in medical or metrological settings.

Recently, Yurdakul et al. introduced an affordable spectral-domain (SD) OCT configuration that uses commodity components such as complementary metal-oxide-semiconductor (CMOS) camera sensors and visible-wavelength light-emitting diodes (LEDs) [22]. The OCT system further uses a line-field configuration, which distributes the illumination from a point into a line field, allowing for parallelized scanning. Similar OCT system designs could enable the proliferation of fast, low-cost 3D component inspection.

However, the use of low-cost components has inevitable downsides. The weaker illumination from the LED yields measurements with lower signal-to-noise ratio (SNR), and the line-field configuration further decreases the illumination power [23]. Measurement noise adversely affects 3D reconstruction by causing spurious surface detections that degrade visualization and downstream analysis tasks such as part inspection. A particular challenge for industrial inspection with OCT is imaging multi-layered specimens, especially with an unknown number of layers. While OCT profilometry for opaque samples is straightforward because the number of layers is known to be 1, a multi-layered surface measurement with OCT is more complicated. When the number of layers is unknown a priori, both the number and location of the layers must be jointly determined. Existing methods for determining the number of layers are usually ad hoc [11, 16, 17, 18, 19] or require manual specification, e.g., pre-specified search regions for each layer [12, 13, 14]. While ad hoc methods may suffice at high SNR, a more principled approach is needed for low SNR, when falsely detecting a surface is much more likely.

In this paper, we propose a method for recovering precise, reliable surface reconstructions despite low-cost OCT hardware. We demonstrate automatic determination of the number and location of layers in multi-layered samples with no manual intervention or prior knowledge of the sample structure. We focus in particular on the low-SNR setting in which most heuristic methods fail. First, we derive the maximum likelihood estimator (MLE) of the depth and reflectivity of a surface under the OCT imaging model, assuming a single-layer sample and a heteroscedastic Gaussian approximation to Poisson noise. Next, we derive a detection threshold that sets a target probability of falsely accepting a surface point estimate due to noise. The threshold computation automatically accommodates multiple measurement frames of the same sample points, important for increasing the SNR and improving surface detection. We combine the threshold and MLE into an algorithm that sequentially detects each layer and estimates a continuous-valued depth estimate, assuming
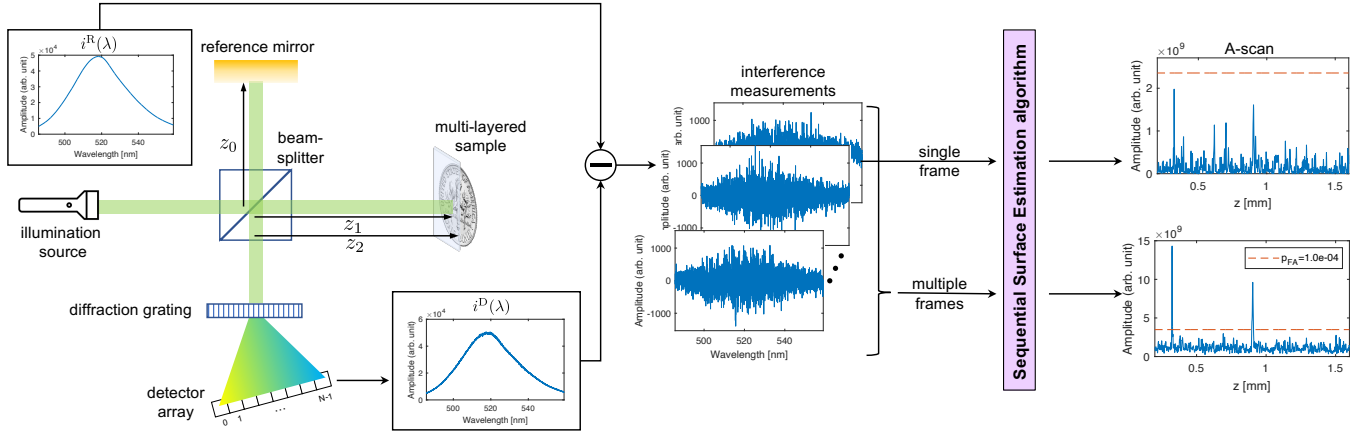
Fig. 1: OCT acquisition and processing pipeline. Light from the illumination source reflects off both the reference mirror and sample, and is recombined by the beamsplitter. The interfered light $i^{\mathrm{D}}(\lambda)$ is spectrally resolved by the diffraction grating and sampled by the detector array. Subtracting a reference-only calibration measurement $i^{\mathrm{R}}(\lambda)$ proportional to the power spectral density isolates the component of the interference measurements containing depth information. Our algorithm sequentially detects and estimates peaks in the A-scan above a threshold that maintains a fixed probability of false acceptance p_{FA} based on the noise statistics. For a single measurement frame, significant noise yields a high threshold that prevents false acceptance of A-scan peaks due to noise. Incoherent addition of multiple frames increases SNR and improves the ability to detect peaks corresponding to true surfaces.

layers have separation greater than the coherence length of the illumination. We validate our algorithm on several multi-layered samples imaged with the prototype low-cost SD-OCT system of Yurdakul et al. [22]. Although demonstrated on a line-field system, our approach is broadly applicable to scanning point or full-field FD-OCT, including swept-source (SS) OCT. An overview of our processing pipeline is shown in Fig. 1.

The rest of this paper is organized as follows. In Section II, we present the OCT measurement model for multi-layered samples. In Section III, we derive the MLE for depth estimation with OCT under a heteroscedastic Gaussian noise assumption and show how it compares to alternative depth estimation approaches. In Section IV, we derive the threshold used for surface detection. Section V outlines our algorithm, which takes advantage of the single-surface MLE and detection threshold. We then present surface reconstructions from experimental measurements recorded with our apparatus in Section VI. Finally, we draw conclusions about the approach in Section VII.

II. MEASUREMENT MODEL

The OCT measurement model is well established—see for example [24]—and we highlight here the key details. A sketch of the SD-OCT system is shown as part of Fig. 1. We first develop the model assuming a point illumination and later discuss how the model changes with non-uniformities in the line-scan illumination.

The light source emits a polychromatic plane wave and the electric field incident on the beamsplitter is

$$E_{\mathrm{I}}(\lambda, \omega) = s(\lambda, \omega) \exp\left[\mathrm{i}\left(\frac{2\pi z}{\lambda} - \omega t \right) \right], \qquad (1)$$

for wavelength λ, angular frequency ω, depth z, and time t. The beamsplitter splits the light into the two arms of the

interferometer, and the resulting field after reflection in the reference arm is

$$E_{\mathrm{R}}(\lambda, \omega) = \frac{s(\lambda, \omega)}{\sqrt{2}} r_0 \exp\left[\mathrm{i}\left(\frac{4\pi z_0}{\lambda} - \omega t \right) \right], \qquad (2)$$

where r_0 is the reflectivity of the reference mirror and z_0 is its depth. Similarly, the resulting field after reflection in the sample arm is

$$E_{\mathrm{S}}(\lambda, \omega) = \frac{s(\lambda, \omega)}{\sqrt{2}} \sum_{\ell=1}^{L} r_\ell \exp\left[\mathrm{i}\left(\frac{4\pi z_\ell}{\lambda} - \omega t + \phi_\ell \right) \right]. \quad (3)$$

In the sample arm, the electric field is a convolution of the incident light with the depth-dependent sample reflectivity profile. We consider a discrete reflector model for L reflectors with reflectivities $\{r_\ell\}_{\ell=1}^{L}$, depths[1] $\{z_\ell\}_{\ell=1}^{L}$, and homogeneous media between the reflectors. A phase offset ϕ_ℓ accounts for shifts in the measurement phase, e.g., due to transmission through or reflection from the sample [25], or due to vibration.

After reflecting from the reference mirror and sample layers, the electric fields are recombined by the beamsplitter. In an SD-OCT system, a diffraction grating separates the recombined light by wavelength. The light intensity incident on the detector is

$$
\begin{aligned}
I(\lambda) &= \frac{1}{2} \left\langle |E_{\mathrm{R}} + E_{\mathrm{S}}|^2 \right\rangle_t \\
&= \frac{S(\lambda)}{2} \Bigg\{ \frac{1}{2} \sum_{\ell=0}^{L} r_\ell^2 \\
&\quad + \sum_{\ell=1}^{L} r_0 r_\ell \cos\left[\frac{4\pi}{\lambda}(z_\ell - z_0) + (\phi_\ell - \phi_0) \right] \\
&\quad + \sum_{\ell=2}^{L} r_\ell \sum_{m=1}^{\ell-1} r_m \cos\left[\frac{4\pi}{\lambda}(z_\ell - z_m) + (\phi_\ell - \phi_m) \right] \Bigg\},
\end{aligned}
\qquad (4)
$$

[1] Since the refractive index changes between surfaces ℓ and $\ell+1$, we more precisely measure the *optical path lengths*. For simplicity, we refer to z as the depth.

where $\langle \cdot \rangle_t$ denotes temporal averaging (which acts as a low-pass filter to eliminate high-frequency components), and $S(\lambda) = \langle s(\lambda, \omega) \rangle_t$ is the power spectral density (PSD) of the illumination source.

Strong Reference Reflectivity Assumption In (4), the three terms are often referred to as the "DC (direct current)," "cross-correlation," and "auto-correlation" components [24]. Note that only cross-correlation encodes the depth relative to the reference z_0. We make the common assumption that the reference reflectivity is much stronger than the sample layer reflectivities [24, Ch. 2.4], i.e., $r_0 \gg r_\ell$, $\ell = 1, \ldots, L$, so we can eliminate the small auto-correlation component and simplify the DC component to get a linear forward model approximation. In the Supplement, we show that this assumption contributes negligible error when $r_0 \gg r_\ell$ holds. For convenience, we also define $z_0 = 0$, $\phi_0 = 0$, and $a_\ell = r_\ell / r_0$ relative to the reference, so the incident intensity is approximated as

$$I(\lambda) \approx \frac{S(\lambda) r_0^2}{2} \left[\frac{1}{2} + \sum_{\ell=1}^{L} a_\ell \cos\left(\frac{4\pi}{\lambda} z_\ell + \phi_\ell \right) \right]. \quad (5)$$

Measurement Statistics Let η be the efficiency of the detector converting photons into photo-electrons. At detector element $n = 1, \ldots, N$, the number of detected photo-electrons is a Poisson random variable whose rate is proportional to the incident intensity at wavelength λ_n, i.e., $i_n \sim \text{Poisson}[\eta I(\lambda_n)]$. However, most cameras do not detect individual photo-electrons directly. Instead, the actual detector measurement includes effects of the camera gain, dark current, read noise, and quantization [26]. Assuming the noise is dominated by the signal-dependent Poisson component (i.e., the measurement is shot-noise limited), we need only consider the effect of the gain. Define β as the scaling factor from photo-electrons to bits in the camera readout after analog-to-digital conversion, so $i_n^{\text{D}} = \beta i_n$ is the raw detector measurement.

Rather than deal with the statistics of a scaled Poisson distribution, we assume in this paper that the noise can be effectively modeled as an independent, heteroscedastic additive Gaussian, where the variance is proportional to the mean. For simplicity, we combine the many unwieldy scaling factors into a single scalar γ and a normalized PSD $\Psi(\lambda_n)$, where

$$\sum_{n=1}^{N} \Psi(\lambda_n) = 1, \quad (6)$$

so that

$$\gamma \Psi(\lambda_n) = \beta \eta r_0^2 S(\lambda_n) / 2. \quad (7)$$

Continuing with the assumption that $r_0 \gg r_\ell$ such that a_ℓ is small, the mean and variance of the detector measurement are dominated by the DC component. Thus we can approximate the detector measurement of the scaled, noise-corrupted intensity to be

$$i_n^{\text{D}} \approx \gamma \Psi(\lambda_n) \left[\frac{1}{2} + \sum_{\ell=1}^{L} a_\ell \cos\left(\frac{4\pi}{\lambda_n} z_\ell + \phi_\ell \right) \right] + w_n, \quad (8)$$

where $w_n \sim \mathcal{N}\left(0, \sigma_n^2\right)$ is Gaussian noise that is mutually independent for all n, and $\sigma_n^2 = \beta \gamma \Psi(\lambda_n)/2$.

DC Removal We observe from the derivation of (8) that the DC component contributes noise despite not containing any depth information. Fortunately, the DC component is straightforward to estimate [24, Ch. 2.4]. In a separate acquisition with the sample arm blocked, we can capture the light intensity from the reference arm only, which is given as $I^{\text{R}}(\lambda) = \frac{1}{2}\langle |E_{\text{R}}|^2 \rangle_t$. Measurement of the reference intensity can be performed offline, so we assume

$$i_n^{\text{R}} = \beta \eta I^{\text{R}}(\lambda) = \gamma \Psi(\lambda_n)/2, \quad (9)$$

where the shot noise is suppressed through temporal averaging. Finally, the measurement data is considered to be the vector $\mathbf{y} = [y_1, \ldots, y_N]$, where

$$\begin{aligned} y_n &= i_n^{\text{D}} - i_n^{\text{R}} \\ &= \gamma \Psi(\lambda_n) \sum_{\ell=1}^{L} a_\ell \cos\left(\frac{4\pi}{\lambda_n} z_\ell + \phi_\ell \right) + w_n. \end{aligned} \quad (10)$$

Note that $i^{\text{R}}(\boldsymbol{\lambda})$ also serves as an estimate of the wavelength-dependent noise variance, up to the scaling factor β. We discuss calibration of β in Section VI-A2.

Multiple Measurements To increase the SNR, F separate measurement frames $\mathbf{y}_1, \ldots, \mathbf{y}_F$ may be captured in sequence at the same illumination position. We assume the phase $\phi_{\ell,f}$ is independent and identically distributed (i.i.d.) uniformly on $[0, 2\pi)$, and the noise values $w_{n,f}$ are independent across frames $f = 1, \ldots, F$.

In a line-field acquisition, each row of pixels in the camera simultaneously captures the spectrally-resolved interference measurement from a different point across the line-field illumination. We make the assumption that each measurement is independent, i.e., crosstalk between camera rows is negligible [27]. Although we assume the normalized spectral density $\Psi(\lambda_n)$ is identical for each row, the scaling factor γ may vary across rows due to non-uniform intensity of the point source after spreading into a line. Note that this differs from other OCT configurations: scanning-point OCT typically has γ constant for all points, whereas full-field OCT may have γ that varies across both lateral dimensions.

III. SINGLE LAYER ESTIMATION

Simultaneously estimating the number and location of the surfaces is a challenging task. We thus focus first on estimating the depth when it is known that there is exactly 1 layer. Later, we harness the single-layer estimator for multi-layer, multi-frame estimation.

A. Conventional Estimation Methods

For a single surface ($L = 1$) and single frame ($F = 1$), the linear forward model (5) is exact, as there are no sample inter-reflections. Then the measurements are given by

$$y_n = a \gamma \Psi(\lambda_n) \cos\left(\frac{4\pi}{\lambda_n} z + \phi \right) + w_n. \quad (11)$$

Since there is a single sinusoid, conventional estimation methods use Fourier analysis to identify a peak in the depth domain. Indeed, the MLE of the frequency of a single complex sinusoid in additive white Gaussian noise is well known to be the peak of the periodogram, i.e., the squared magnitude of the discrete-time Fourier transform [28]. OCT depth estimation thus usually applies the discrete Fourier transform (DFT) via the fast Fourier transform (FFT) algorithm and then refines the discrete peak position. However, the periodogram-based approach neglects several factors that can cause small errors in the final estimates. Naïve application of the DFT ignores the fact that the measurements are usually real-valued, rather than complex. For estimating the parameters of real sinusoids, the positive and negative frequencies interfere with each other [29], resulting in a slight bias in the peak position. In other words, the periodogram does not correspond to the correct likelihood function to maximize for a real sinusoid.

Furthermore, OCT systems typically sample the interfered light at uniform wavelengths[2] λ_n, which means the samples are non-uniformly spaced in wavenumber $k_n = 2\pi/\lambda_n$. However, the FFT requires uniform wavenumber samples. The most widespread approach for achieving uniform k samples is to interpolate the data and resample [30, 31]. However, interpolation amounts to shift-variant low-pass filtering, reducing the sensitivity to greater depths (i.e., high-frequency interference components) and propagates noise to the unsampled wavenumbers, reducing robustness. Alternatively, the non-uniform (NU) DFT or FFT computes the Fourier transform for arbitrarily spaced points in wavenumber and/or depth by interpolation with guarantees on the precision [31, 32, 33, 34, 35, 36, 37]. To avoid the problems caused by interpolation, alternative hardware systems have been proposed that yield samples uniformly spaced in wavenumber rather than wavelength [30, 38, 39, 40, 41]. Although the FFT can be directly applied to the uniform-k samples, these systems require more complicated and often expensive hardware.

B. Heteroscedastic Gaussian MLE

Although periodogram maximization can often achieve reasonable surface estimates, it does not exactly correspond to the MLE based on the single-surface OCT model. Here, we start from first principles, expanding our earlier work [1] to derive the exact MLE for single-surface OCT in heteroscedastic Gaussian noise by including the PSD, non-uniform wavenumber sampling, and real-valued measurements. We then propose an implementation that identifies a good initialization and directly maximizes the exact likelihood.

1) Derivation: Define the vectors $\mathbf{c} = [c_1, \ldots, c_N]^\mathsf{T}$ and $\mathbf{s} = [s_1, \ldots, s_N]^\mathsf{T}$, where $c_n = \gamma\Psi(\lambda_n)\cos(4\pi z/\lambda_n)$ and $s_n = \gamma\Psi(\lambda_n)\sin(4\pi z/\lambda_n)$. Given the data vector $\mathbf{y} = [y_1, \ldots, y_N]^\mathsf{T}$ and assuming heteroscedastic Gaussian noise,

---

[2]In SD-OCT, uniform wavelength samples arise because a diffraction grating has an approximately linear relationship between wavelength and diffraction angle, and the detector pixels are uniformly spaced. In SS-OCT, the laser may have a linear sweep in wavelength versus time or some other sweep patterns that do not yield uniform wavenumber samples.

we show in Appendix A how the MLE is the depth z that maximizes

$$\Lambda(\mathbf{y}, z) = \mathbf{y}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{H}(\mathbf{H}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{H})^{-1}\mathbf{H}\mathbf{\Sigma}^{-1}\mathbf{y}, \quad (12)$$

where we define the matrices $\mathbf{H} = [\mathbf{c}, \mathbf{s}]$ and $\mathbf{\Sigma} = \text{diag}([\sigma_1^2, \ldots, \sigma_N^2]^\mathsf{T})$. Although $\mathbf{H}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{H}$ is usually approximated as diagonal in order to compute its inverse [1, 42], it is in fact straightforward to compute the exact inverse in closed form. We can thus rewrite the full objective to maximize as

$$\Lambda(\mathbf{y}, z) = \frac{1}{G}\Big[(\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y})^2 + (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{c})(\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y})^2$$
$$- 2(\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y})(\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y})\Big], \tag{13}$$

where $G = (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{c})(\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s}) - (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})^2$, and the depth MLE is

$$\hat{z}_{\text{ML}} = \arg\max_z \Lambda(\mathbf{y}, z). \tag{14}$$

We further define the functions:

$$\Omega(\mathbf{y}, z) = \frac{1}{G}\Big\{[(\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y}]^2$$
$$+ [(\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{c})\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y}]^2\Big\}^{\frac{1}{2}} \tag{15}$$

$$\Phi(\mathbf{y}, z) =$$
$$\arctan\left(-\frac{(\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{c})\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y}}{(\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s})\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{y}}\right). \tag{16}$$

Then given $\hat{z}_{\text{ML}}$, we can compute the MLEs of the reflectivity $\hat{a}_{\text{ML}} = \Omega(\mathbf{y}, \hat{z}_{\text{ML}})$ and phase offset $\hat{\phi}_{\text{ML}} = \Phi(\mathbf{y}, \hat{z}_{\text{ML}})$.

2) Proposed Implementation: The function $\Lambda(\mathbf{y}, z)$ in (13) is multi-modal, so a good initial value is needed to find the global maximum. We show in Appendix A that using the approximations $\mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{c} = \mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s} \approx 0$ and $\mathbf{c}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{c} \approx \mathbf{s}^\mathsf{T}\mathbf{\Sigma}^{-1}\mathbf{s} \approx \gamma/\beta$ reduces (13) to the periodogram

$$\Lambda(\mathbf{y}, z) \approx \frac{4}{\beta\gamma}\left|\sum_{n=1}^N y_n \exp\left(-\mathrm{i}\frac{4\pi}{\lambda_n}z\right)\right|^2, \tag{17}$$

which has a simple discrete implementation via the NUDFT. We thus use a two-step process similar to Rife and Boorstyn [28], in which we compute the approximate MLE for initialization on a discrete grid and then refine our estimate via an iterative algorithm to get the off-grid MLE.

a) Discrete Initialization: We define the partial NUDFT matrix $\mathbf{A} \in \mathbb{C}^{N \times M}$, where the (n, m) element is

$$[\mathbf{A}]_{n,m} = \exp\left(\mathrm{i}\frac{4\pi}{\lambda_n}z_m\right), \tag{18}$$

at depths z_m for $m = 1, \ldots, M$, which are typically uniformly-spaced with separation δ_z. To avoid missing the true peak, a good rule of thumb is to set δ_z less than the coherence length, which is the resolution of an OCT measurement based on the illumination source PSD [43]. We then compute (17) for the vector of discrete depths $\mathbf{z} = [z_1, \ldots, z_M]^\mathsf{T}$ via the element-wise squared magnitude of the back-projection

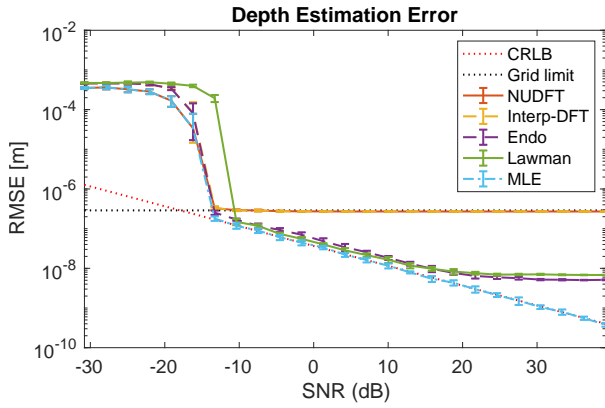$$B(\mathbf{y}) = |\mathbf{A}^\mathsf{H}\mathbf{y}|^2, \tag{19}$$

Fig. 2: Depth estimation performance comparison. The MLE is more robust to additive noise and more accurate than approximate methods, achieving the CRLB above $-13\,\mathrm{dB}$ SNR.

where $[\cdot]^{\mathsf{H}}$ denotes the Hermitian or conjugate transpose. The position of the maximum $\widetilde{m} = \arg\max_m B(\mathbf{y})$ yields the discrete approximate depth MLE $\hat{z}_{\mathrm{BP}} = z_{\widetilde{m}}$. Note that a matrix-free NUFFT implementation may also be used, although we find the explicit matrix multiplication to be more efficient when parallelizing back-projection for line-field measurements.

b) Continuous Optimization: Given a sufficiently small discrete spacing δ_z, the objective $\Lambda(\mathbf{y}, z)$ is locally unimodal around $\hat{z}_{\mathrm{BP}}$. We can then refine the discrete estimate via iterative maximization to get a continuous estimate $\hat{z}_{\mathrm{ML}}$. Fast convergence can be achieved via the RTSAFE algorithm, which combines Newton's method with bisection for robust, bounded optimization [44, Ch. 9.4,]. The Newton iterations are initialized with $z^{(0)} = \hat{z}_{\mathrm{BP}}$ and continue as

$$z^{(t+1)} = z^{(t)} - \frac{\Lambda'(z^{(t)})}{\Lambda''(z^{(t)})}, \qquad (20)$$

unless the step would escape the bounding interval, in which case interval bisection occurs. The algorithm continues until $|z^{(t+1)} - z^{(t)}| < \epsilon$ for some tolerance ϵ. The first and second derivatives of $\Lambda(\mathbf{y}, z)$ are found in Appendix B. For the bounding interval, we use $\hat{z}_{\mathrm{BP}} \pm \delta_z$.

C. Performance Comparison

1) Simulation Setup: For heteroscedastic noise, the signal-to-noise ratio (SNR) varies for each element of the measurement vector. For the signal component in (11), the signal power of the nth element is $\mu_n^2 = \gamma^2 \Psi^2(\lambda_n)\, a^2/2$. Given the noise power is $\sigma_n^2 = \beta\gamma\Psi(\lambda_n)/2$, then the SNR for element n is

$$\mathrm{SNR}_n = \frac{\mu_n^2}{\sigma_n^2} = \frac{a^2\gamma}{\beta}\Psi(\lambda_n). \qquad (21)$$

The average SNR across the measurement is then

$$\mathrm{SNR} = \frac{1}{N}\sum_{n=1}^{N}\mathrm{SNR}_n = \frac{a^2\gamma}{\beta N}. \qquad (22)$$

To evaluate the performance of the MLE implementation as a function of average SNR, we simulated an OCT experiment with fixed a, β, and N while varying γ. The OCT system

TABLE I: Depth estimation runtime comparison

Method	NUDFT	Interp-DFT	Endo	Lawman	MLE
Runtime (s)	0.0173	0.0199	0.0518	0.041	0.0266

emulates the experimental setup described in Section VI, with a center wavelength of $530\,\mathrm{nm}$ and an approximately Gaussian PSD with wavelength bandwidth of $35\,\mathrm{nm}$. The noise is simulated as Poisson, and we set $\beta = 1$.

2) Comparison Methods: The comparison methods are implemented as follows:

- **NUDFT** implements our discrete initialization as outlined in Section III-B2a, i.e., it estimates the depth by finding the maximum of the back-projection with the NUDFT matrix.
- **Interp-DFT** applies the uniform DFT to measurements that have been resampled to have uniform wavenumber spacing after spline interpolation. The depth estimate is then the location of the DFT peak.
- **Endo** refines the Interp-DFT estimate by computing the centroid around the DFT peak [8].
- **Lawman** resamples the measurement, and then applies a Hann window prior to the uniform DFT [14]. The discrete estimate is refined by fitting a Gaussian function to the DFT peak.
- **MLE** uses the NUDFT estimate as initialization and refines the estimate via Newton's method, as described in Section III-B2b.

3) Results: Fig. 2 shows the error as a function of the average SNR for depth estimation of a single opaque surface. Because the NUDFT and Interp-DFT estimates are confined to a discrete grid with $\delta_z = 1\,\mu\mathrm{m}$ depth resolution, the root mean squared error (RMSE) assuming uniformly distributed depths is limited by the quantization error to $\sqrt{\delta_z^2/12}$ (i.e., "Grid limit") [45]. The Endo, Lawman–Liang, and MLE methods all use continuous refinement to surpass the grid limit. Our MLE implementation outperforms all comparison methods and achieves the Cramér–Rao Lower Bound (CRLB, derived in Appendix C) for SNR values above $-13\,\mathrm{dB}$. At high SNR in particular, improvement comes from maximizing the likelihood directly, rather than using curve fitting approximations.

Table I compares the runtimes of the various methods. Computing the NUDFT is slightly faster than Interp-DFT because it skips the interpolation step. Of the methods that use continuous refinement, the MLE is the fastest and not much slower than the NUDFT, suggesting that Newton's method is an efficient implementation for maximizing the likelihood and surpassing the grid limit.

IV. NEYMAN–PEARSON DETECTION THRESHOLD

In Section III, we explored surface estimation, assuming the presence of a single layer in the measurement. However, detecting whether a layer is indeed present is non-trivial, especially for samples with an unknown or varying number of layers across the specimen. Here we develop a principled rule for layer detection, based on how the noise statistics are affected by the estimation procedure.
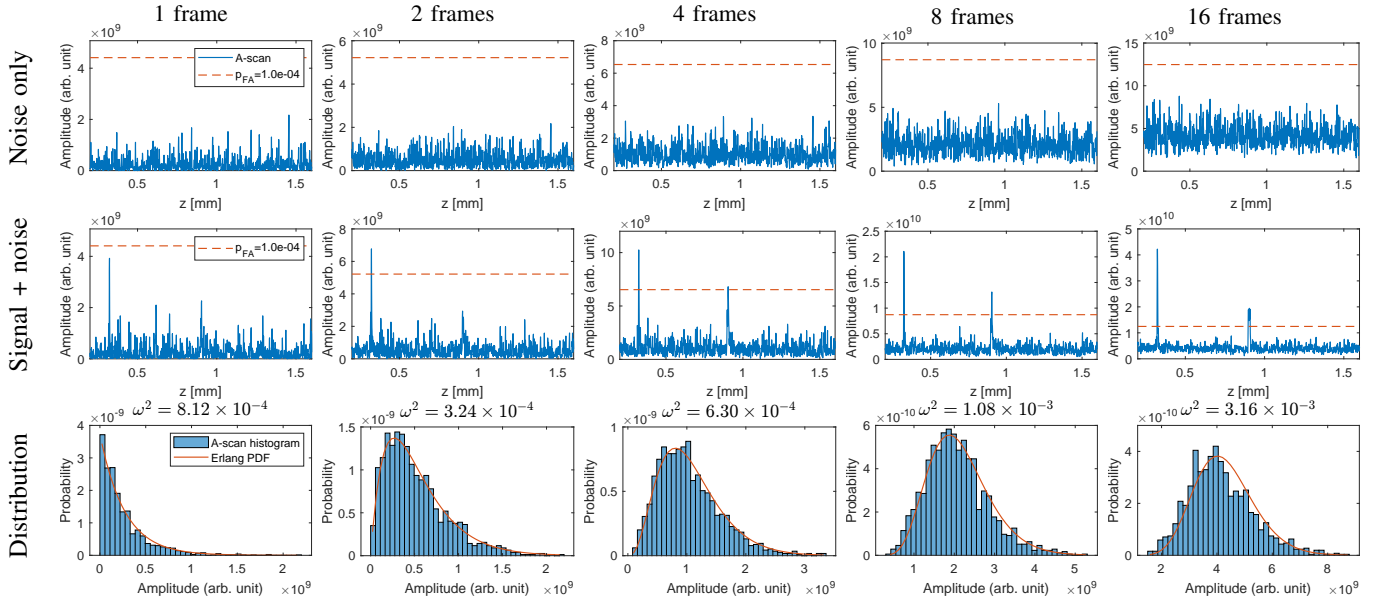
Fig. 3: Demonstration of the accuracy of the Erlang density noise model for experimental data. The top row shows A-scans dominated by noise, whereas the second row contains noise plus two signal peaks. A threshold for false alarm probability $p_{FA} = 1 \times 10^{-4}$ is shown as the dashed lines. The bottom row shows a histogram of the A-scan values and our Erlang density approximation.

As detailed in Section III-B2, a coarse estimate can be computed via back-projection with the NUDFT matrix. Given F independent measurement frames, define the A-scan as the vector

$$\mathbf{b} = \sum_{f=1}^{F} |\mathbf{A}^{\mathsf{H}}\mathbf{y}|^2, \tag{23}$$

which sums the squared magnitudes of the back-projections for each frame. Define the peak A-scan amplitude as $\rho = \max(\mathbf{b})$.

The detection rule must effectively determine whether the peak is actually due to noise. The distribution of surface reflectivities is unknown, hence we derive a Neyman–Pearson detection criterion based on a probabilistic model of the noise alone. Namely, given a fixed probability of accepting a spurious detection p_{FA}, we determine a threshold τ_{FA} on the A-scan amplitude. Only if the peak amplitude $\rho \geq \tau_{FA}$ do we consider the peak to be a layer detection and proceed with estimating its parameters.

To determine τ_{FA}, we suppose the interference measurement consisted of only noise. Let $\mathbf{w}_f \in \mathbb{R}^N$ be a zero-mean, jointly Gaussian random vector, with element w_n having variance σ_n^2. Following Djukanović and Popović-Bugarin [46], the back-projection of the noise vector $\boldsymbol{\nu}_f = \mathbf{A}^{\mathsf{H}}\mathbf{w}_f$ is approximated as an i.i.d. circularly-symmetric complex Gaussian random vector with both $\mathrm{Re}\{\nu_m\}$ and $\mathrm{Im}\{\nu_m\} \sim \mathcal{N}(0, \sigma_\nu^2/2)$, where

$$\sigma_\nu^2 = \sum_{n=1}^{N} \sigma_n^2. \tag{24}$$

Then the magnitude of the back-projection $\zeta_m = |\nu_m|$ is a Rayleigh($\sqrt{\sigma_\nu^2/2}$) random variable, and the squared magnitude ζ_m^2 is an Exponential(σ_ν^2) random variable. The mth A-

scan element from F independent noise frames $\mathbf{w}_1, \ldots, \mathbf{w}_F$ is

$$\psi_m = \sum_{f=1}^{F} \zeta_m^2 \sim \mathrm{Erlang}(F, \sigma_\nu^2). \tag{25}$$

For a given detection threshold τ_{FA}, the probability of falsely accepting a detection due to noise is

$$p_{FA} \approx 1 - \prod_{m=1}^{M} \Pr[\psi_m < \tau_{FA}] \tag{26}$$

$$= 1 - \left[1 - \exp\left(-\frac{\tau_{FA}}{\sigma_\nu^2} \right) \sum_{f=0}^{F-1} \frac{1}{f!} (\tau_{FA}\sigma_\nu^2)^f \right]^M, \tag{27}$$

which uses the closed-form definition of the Erlang cumulative distribution function. For multi-frame measurement, we can use root-finding methods to numerically solve (27) for the threshold τ_{FA} given a target p_{FA}. Moreover, a closed-form solution for the threshold exists in the case of $F = 1$:

$$\tau_{FA} = -\sigma_\nu^2 \log\{1 - [1 - p_{FA}]^{1/M}\}. \tag{28}$$

Unlike Djukanović and Popović-Bugarin [46], who set a threshold for sinusoid detection based on the statistics of a single DFT bin from a single measurement, our threshold derivation not only considers the joint probability of false acceptance across all elements of $\boldsymbol{\nu}_f$ but also easily allows for an arbitrary number of measurement frames.

An example of threshold selection is demonstrated for experimental data in Fig. 3. The top two rows show a series of A-scans for different numbers of frames: the first row is computed for a measurement containing only noise, and the second row is for a measurement that also contains two surface reflections. The bottom row shows the empirical noise density, i.e., a histogram of the noise-only A-scan values, as well as

the Erlang probability density function (PDF) based on our noise model. The fit of the PDF is quantified by the Cramér–von Mises ω^2 criterion [47]. The Erlang density shows a very good fit to the histogram, yielding an effective threshold that prevents false acceptance of noise spikes, although the increase in the ω^2 criterion for 16 frames suggests a slight model mismatch that could reduce the effectiveness of the threshold.

While our threshold derivation assumes heteroscedastic Gaussian noise, we further show in the Supplement that the Erlang model likewise holds for measurements with Poisson noise. By the central limit theorem, the back-projection of Poisson noise with sufficiently large N leads to $\boldsymbol{\nu}$ that is still well-approximated as a circularly-symmetric complex Gaussian random vector.

V. MULTI-LAYER DETECTION AND ESTIMATION

While surface detection and estimation are straightforward for a single layer, extension to multiple layers is complicated by the unknown number of layers. Existing methods of multi-layer estimation typically apply single-surface estimation methods sequentially, using ad hoc selection methods or prior knowledge to determine how many layers need to be estimated. For instance, Lawman and Liang require pre-specified regions in which to search for each interface location and then use post-processing to eliminate obvious errors [12, 13, 14]. Atalar et al. determine the number and location of targets by setting a minimum reflectivity threshold as a fixed fraction of the largest peak height in each A-scan [48]; however, this approach assumes that sidelobes are the main source of spurious layer detection, ignoring the effect of noise. Alternatively, a standard practice in OCT layer detection is to use a global threshold on the B-scan magnitude [10, 11, 16, 17, 18, 19]. However, the threshold value generally requires hand-tuning and does not account for the possibility of spatially-varying amplitude statistics across the B-scan. More sophisticated methods taking advantage of the lateral component of cross-sectional B-scans likewise assume a known number of surfaces that span the B-scan [49, 15], but in practice—as in the experimental specimens we consider later—layers may not be contiguous and may change in number across the lateral dimension. Other related approaches encourage sparsity in volumetric OCT reconstruction via ℓ_1-norm regularization [50, 51], but these convex optimization methods have high computational complexity that scales poorly for large data volumes.

We propose a multi-layered estimation approach that combines our principled single-surface detection and estimation methods. Our algorithm, which we call Sequential Surface Estimation (SSE), first computes a detection threshold τ_{FA} and the A-scan $\mathbf{b}$. The algorithm then sequentially compares each peak in the A-scan to the threshold in order of peak height. By assuming the layers are separated by at least the coherence length of the illumination source, the signals from different layers do not significantly affect each other. Thus, we can use the same threshold for all layers and simply remove a small region around each detected depth from the search space for the next layer detection. If a peak is sufficiently large, we compute a continuous depth estimate for that detected layer

---

Algorithm 1 Sequential Surface Estimation (SSE)

Input: measurement frames $\mathbf{y}_1, \ldots, \mathbf{y}_F$; wavelengths $\boldsymbol{\lambda}$; PSD $\gamma\Psi(\boldsymbol{\lambda})$; discrete depths $\mathbf{z} = [z_1, \ldots, z_M]^{\mathsf{T}}$; noise variances $\sigma_1^2, \ldots, \sigma_N^2$; false acceptance probability p_{FA}; minimum layer separation $d_{\min}$; maximum layer count $L_{\max}$

Output: depths $\{\hat{z}_{\text{SSE}}^{(t)}\}_{t=1}^{T}$ and reflectivities $\{\hat{a}_{\text{SSE}}^{(t)}\}_{t=1}^{T}$

1: Set up NUDFT matrix $\mathbf{A}$ as in (18)
2: Form A-scan $\mathbf{b}$ as in (23)
3: Compute σ_ν^2 as in (24)
4: Solve for threshold τ as in (27)
5: **Initialize:**
　　$t = 1$, $\mathcal{M}(1) = \{1, \ldots, M\}$
6: **while** converged = False and $t \leq L_{\max}$ **do**
7:　　// Detection
8:　　$\rho(t) = \max_{m \in \mathcal{M}(t)} \mathbf{b}$
9:　　**if** $\rho(t) > \tau$ **then**
10:　　　// Estimation
11:　　　$\widetilde{m} = \arg\max_{m \in \mathcal{M}(t)} \mathbf{b}$
12:　　　$\hat{z}_{\text{BP}}^{(t)} = z_{\widetilde{m}}$
13:　　　Calculate $\hat{z}_{\text{SSE}}^{(t)}$ as in (30)
14:　　　$\hat{a}_{\text{SSE}}^{(t)} = \frac{1}{F} \sum_{f=1}^{F} \Omega(\mathbf{y}_f, \hat{z}_{\text{SSE}}^{(t)})$
15:　　　$t = t + 1$
16:　　　Update $\mathcal{M}(t)$ as in (29)
17:　　**else**
18:　　　converged = True
19:　　　$T = t - 1$

---

via the MLE. Else if a peak does not meet the threshold, then we assert that all layers have been detected. In particular, we ensure continuous-valued parameter estimation from multiple captured frames by harnessing the single-layer MLE derived above.

The basic SSE algorithm is summarized in Algorithm 1 and detailed as follows.

Inputs: The algorithm requires a set of measurement frames $\mathbf{y}_1, \ldots, \mathbf{y}_F$, the sampled wavelengths $\boldsymbol{\lambda}$, the PSD at each wavelength $\gamma\Psi(\boldsymbol{\lambda})$, a vector of discrete candidate depths $\mathbf{z} = [z_1, \ldots, z_M]^{\mathsf{T}}$, the noise variances $\sigma_1^2, \ldots, \sigma_N^2$, false acceptance probability p_{FA}, and a minimum layer separation $d_{\min}$. Setting a maximum number of layers $L_{\max}$ is optional.

Minimum Layer Separation: The minimum layer separation can be introduced by modifying the index set over which the A-scan maximum is found. For iteration t, the set of indices to be included in the search is

$$\mathcal{M}(t; d_{\min}) = \bigcap_{i=1}^{t-1} \left\{ m \in 1, \ldots, M : |z_m - \hat{z}_{\text{SSE}}^{(i)}| > d_{\min} \right\}. \tag{29}$$

The indices $m \in \mathcal{M}(t)$ correspond to depths that are farther than $d_{\min}$ from previously detected layers $1, \ldots, t-1$.

Layer Detection: The A-scan is computed as in (23), and the peak A-scan amplitude $\rho(t) = \max_{\mathcal{M}(t)} \mathbf{b}$ is used to compare to the threshold τ. If $\rho(t) > \tau$, then a layer has been detected.

Layer Estimation: For each detected layer, the index of the A-scan peak $\widetilde{m} = \arg\max_{\mathcal{M}(t)} \mathbf{b}$ is used to assign an initial

depth estimate $\hat{z}_{\mathrm{BP}}^{(t)} = z_{\widetilde{m}}$ based on the discrete set of candidate depths $\mathbf{z}$. The depth estimate is then refined via the continuous optimization:

$$\hat{z}_{\mathrm{SSE}}^{(t)} = \arg\max_{z} \sum_{f=1}^{F} \Lambda(\mathbf{y}_f, z) \tag{30}$$
$$\text{s.t.} \quad z_{\widetilde{m}-1} \leq z \leq z_{\widetilde{m}+1},$$

which uses the single-surface log-likelihood function from (13) and constrains the search to be within one bin of the discrete initialization. We likewise compute the reflectivity as the average of the ML estimates $\hat{a}_f = \Omega(\mathbf{y}_f, \hat{z}_{\mathrm{SSE}}^{(t)})$ for each frame.

The algorithm continues until no more layers are detected, i.e., $\rho(t) \leq \tau$, or the optional maximum layer count has been reached ($t = L_{\max}$).

VI. EXPERIMENTAL VALIDATION

A. Experimental Setup

1) Hardware: All experiments were performed on the low-cost prototype line-field OCT system of Yurdakul et al. [22], pictured in Fig. 4(a). The illumination is a partially-coherent narrowband green LED (Thorlabs M530L4) with central wavelength $\lambda_0 \approx 530\,\text{nm}$ and wavelength bandwidth $\Delta\lambda \approx 35\,\text{nm}$, which yields an approximately $3.5\,\mu\text{m}$ full-width at half-maximum (FWHM) axial resolution [43]. The LED is diffused, focused into a line, and further filtered by a mechanical slit to create a line-field illumination. After coherent mixing of the reference and sample arm illumination via a beam-splitter, a transmission grating (Wasatch Photonics, $1800\,\text{lines/mm}$ at $532\,\text{nm}$) diffracts the light quasi-linearly. The intensity of the diffracted light is then captured by a CMOS camera (FLIR, BFS-U3-70S7M-C) which has 2200×3208 pixels.

Each row (3208 pixels) captures the interference spectrum for one point on the scan line, so the spectra for 2200 points along the line can be recorded simultaneously. In practice, the illumination intensity near the edges is much weaker than the center, so we crop the 200 rows on either edge. After processing the interference pattern, an entire B-scan can be recovered from a single line-scan measurement (x- and z-directions). To achieve full 3D scanning, the sample is mechanically scanned in the lateral (y) direction orthogonal to the scan-line via a motorized XY stage (Zaber, ASR100B120B-T3A).

2) Noise Calibration: Our model assumes the reference measurement is $i_n^{\mathrm{R}} = \gamma\Psi(\lambda_n)/2$ and the noise variance is $\sigma_n^2 = \beta i_n^{\mathrm{R}}$. To robustly estimate β, we use a simplified version of the method of Foi et al. [52]. First, we identify level sets of the smooth reference measurement i^{R}, and then compute the variance of detected interference measurement i^{D} on each level-set. Then we perform linear regression to estimate $\hat{\beta}$ from the full camera measurement. In practice, we find $\hat{\beta} \approx 3.2$ for our camera. Although in principle, a single offline calibration can be used to determine β, we observe slightly better results by calibrating for each scan line.

B. Comparison Methods

We compare our SSE algorithm to three baseline methods for multi-layer surface estimation described in the literature.

For each method, we compute the A-scan via back-projection with the partial NUDFT matrix so as to highlight the differences based on detection strategy alone.

- **Atalar**: The A-scan is computed as in (23), and the peak amplitude is ρ. As described by Atalar et al. [48], the threshold is set as a fixed fraction of ρ (e.g., $\tau = 0.5\rho$), and all peaks above the threshold are accepted. In our evaluation, we retain only the 5 largest peaks with reflectivities above the threshold.
- **Lawman**: The A-scan is computed as in (23), and surfaces are assumed to exist within pre-specified depth ranges, as described by Lawman [14]. For each depth range, the A-scan peak is assumed to correspond to a surface, and no amplitude-based thresholding is used. This approach requires prior knowledge to hard-code the search range and is only applicable to an entire B-scan when the number of layers is consistent and when the layers can be completely separated in depth.
- **Percentile**: Numerous approaches implement a global threshold for layer detection in a B-scan, and while the methods differ (or details are omitted), the threshold generally requires hand-tuning [10, 11, 16, 17, 18, 19]. Here we implement global thresholding based on a percentile of B-scan magnitudes, hand-selected to optimize the trade-off between the false acceptance and true rejection rates. Note that unlike the proposed SSE method or that of Atalar et al. [48], the same threshold is applied across all A-scans.

C. Results

For all experiments, we use $p_{\mathrm{FA}} = 1 \times 10^{-4}$, $d_{\min} = 20\,\mu\text{m}$, and $L_{\max} = 5$. We show the estimated reflectivities in grayscale (in dB) and the depth in false color. Fig. 4 demonstrates a characteristic result for our method. The prototype SD-OCT system is pictured in Fig. 4(a). Fig. 4(b) shows a sample we call 'dime+tape', which is a US ten-cent coin covered on one side by Scotch tape. A full 3D point cloud from 4 measurement frames per scan line is shown in Fig. 4(c), and the reconstructions for individual scan lines are highlighted in Fig. 4(d). The tape adheres to the coin at the highest features but not in the deepest crevices, thus certain parts of the reconstruction show three distinct layers, including both the top and bottom surfaces of the tape.

In Fig. 5, we show performance comparisons with alternative methods for the 'dime+tape' sample as we increase the number of measurement frames. Fig. 6 shows similar comparisons for the 'dime+glue' sample (photograph in Fig. 9) consisting of a dime covered in translucent glue. The optical path length through the glue is up to $\approx 1\,\text{mm}$, so the reflections from both the glue and coin are weak, leading to extremely low-SNR measurements. For both samples, the B-scans are fairly noisy, and surface extraction is non-trivial. The method of Atalar et al. [48], which sets a threshold relative to the strongest peak in each A-scan, performs poorly in regions without strong surface reflections. Since Atalar et al. [48] assume at least one surface per A-scan, if the strongest peak is quite low, the threshold is set too low, and many spurious
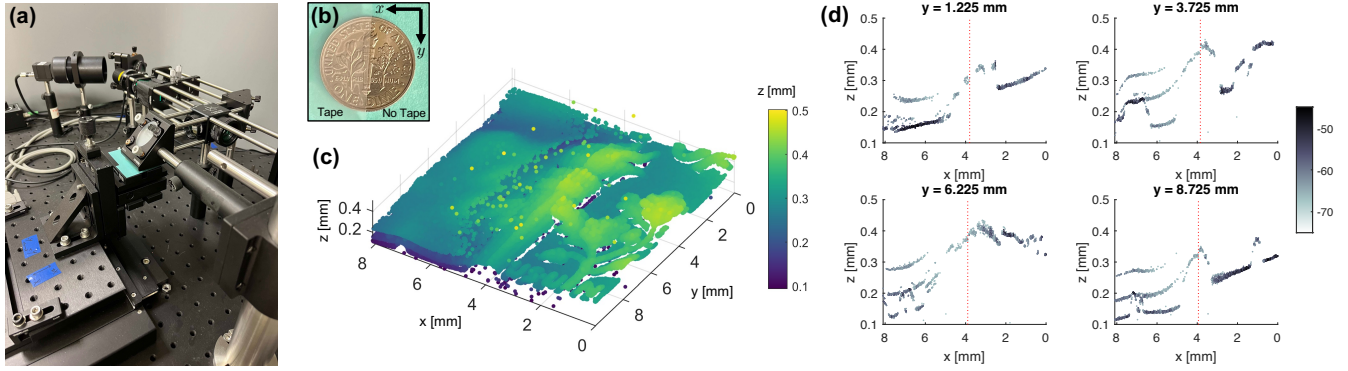
Fig. 4: Surface estimation using the prototype OCT apparatus (a) to scan the 'dime+tape' sample (b). The full 3D point cloud (c) and results for individual y scan lines (d) are shown for 4 frames per scan line. The tape covers the coin roughly between 4 and 8 mm, indicated by the red dotted line in (d). The bare side of the coin ($x < 4\,\text{mm}$) shows a single layer, while the tape-covered side ($x > 4\,\text{mm}$) shows up to three distinct layers: i.e., the top and bottom tape surfaces, and the coin surface below the tape.
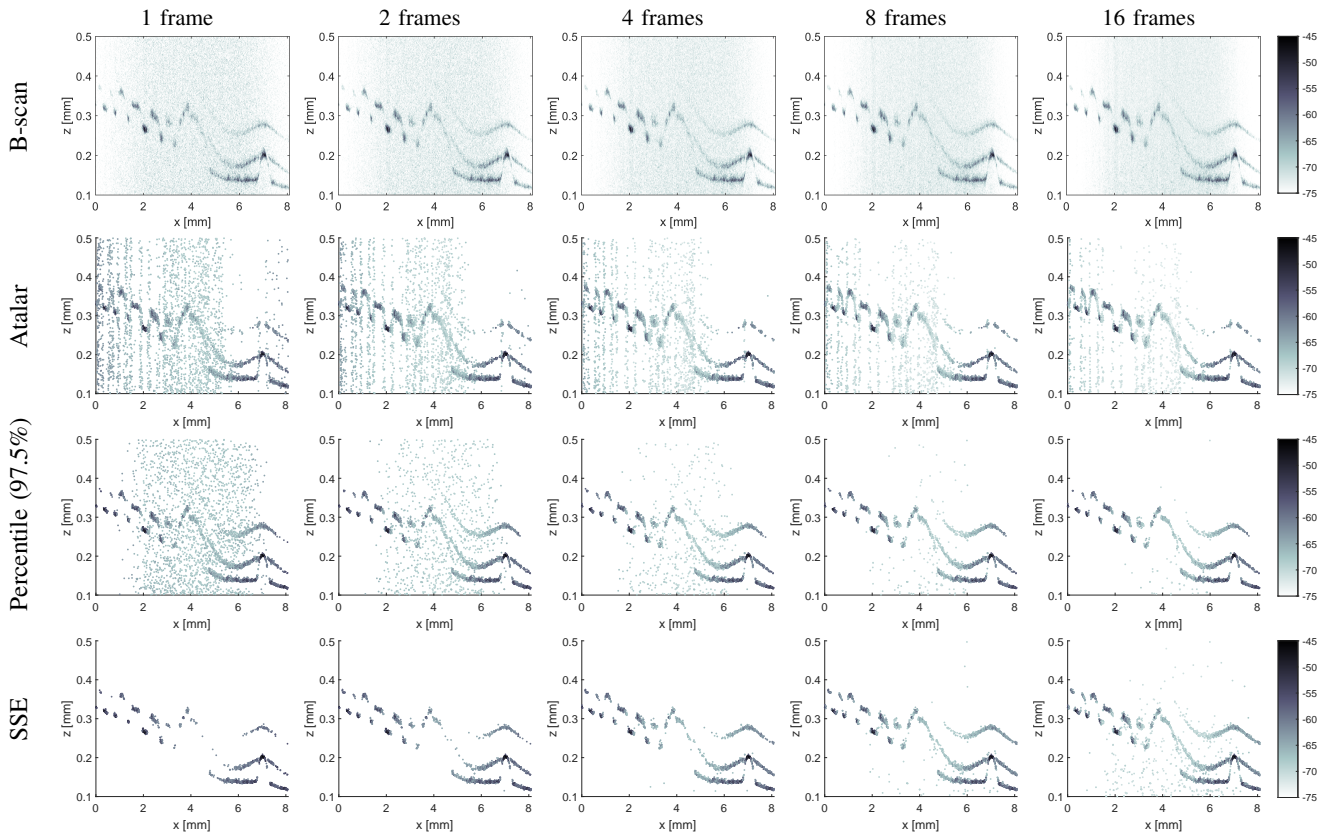

Fig. 5: Reconstruction comparisons of the 'dime+tape' sample (scan-line $y = 9.975\,\text{mm}$) for an increasing number of frames. The method of Atalar et al. [48] sets a threshold relative to the peak at each x position, leading to many spurious detections when the peak reflectivity is low. The percentile method sets a global threshold on the B-scan magnitude, which performs better as the number of frames increases. The method of Lawman and Liang [13] is not shown because the layers cannot be separated by simple depth segmentation.

surfaces are falsely accepted. The percentile method, which sets a global threshold for the entire B-scan has uneven performance across the scan-line, as the signal and noise statistics are spatially varying. The performance of the percentile method improves as the number of frames increases and achieves acceptable results for 16 frames in Fig. 5. However, results for the 'dime+glue' method are poor, and the percentile level must be tuned by hand for each dataset. The method of Lawman and Liang [13] requires pre-specifying search regions for each

layer and is thus impractical for the 'dime+tape' sample, which has varying numbers of layers for each A-scan, and the layers are not easily separated in depth. For the 'dime+glue' sample, even with the prior assumption that there are exactly two layers, separated in depth above and below $-0.7\,\text{mm}$, the method of Lawman and Liang [13] accepts too many false surface detections. Post-processing would be required to remove spurious surfaces.

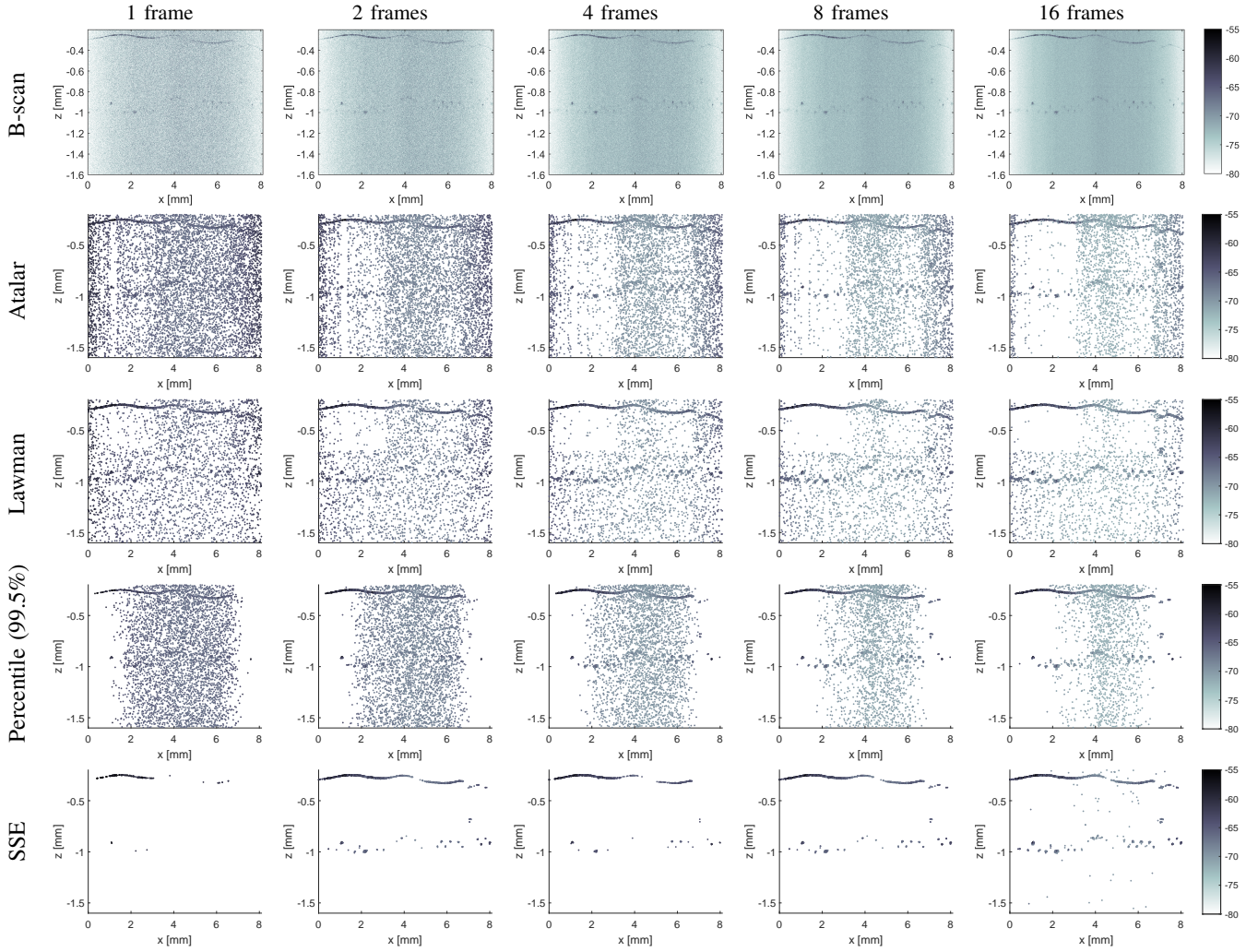In comparison, the SSE shows much better robustness to

Fig. 6: Reconstruction comparisons of the 'dime+glue' sample (scan-line $y = 8.9\,\mathrm{mm}$) for an increasing number of frames. The Atalar et al. [48], Lawman–Liang [13], and percentile methods accept a large number of spurious detections for low-SNR measurements. The proposed SSE method only accepts surface points unlikely to be artifacts due to noise, resulting in far fewer outliers.
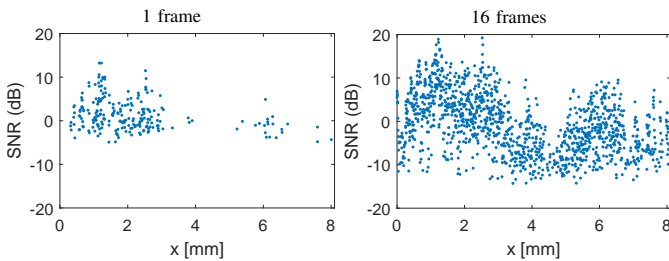

Fig. 7: SNR estimates for the SSE results in Fig. 6. The SNR is estimated shown only for those A-scans in which at least one surface was detected. Results show a sensitivity to measurements with an SNR as low as $-10\,\mathrm{dB}$.

Fig. 8: Precision estimation for the SSE applied to the prototype OCT system in 'height' and 'depth'-map configurations and either 1 or 8 frames. The RANSAC algorithm [53] is used to compute a robust linear fit for one scan line of a flat metal specimen. The number of inliers found by RANSAC is shown in (a), and the root-mean-square error (b) is computed for the inliers relative to the linear fit. The depth error is around $1.5\,\mu\mathrm{m}$ at a distance of $0.1\,\mathrm{mm}$ and increases to $3-4\,\mu\mathrm{m}$ at $1\,\mathrm{mm}$.

noise. In particular, the SSE correctly accounts for noise statistics that vary within and across datasets. The SSE shows good rejection of false layer acceptance for 1, 2, or 4 frames. Further increasing F appears to cause a slight mismatch between the assumed noise model and the actual measurements, causing the computed threshold to be too low, and increasing the number of spurious detections. Still, the SSE results far surpass those from comparison methods. In principle, the SSE is similar
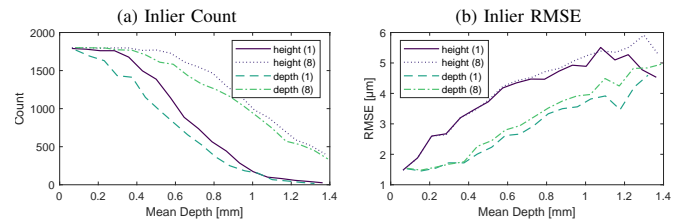
to other threshold detection methods (Atalar et al. [48], and percentile) in that a detected layer is defined as a peak in the A-scan above a particular threshold magnitude. However, the results in Fig. 5 and Fig. 6 demonstrate that our principled threshold selection achieves a better tradeoff between a low false acceptance rate and a high true detection rate.
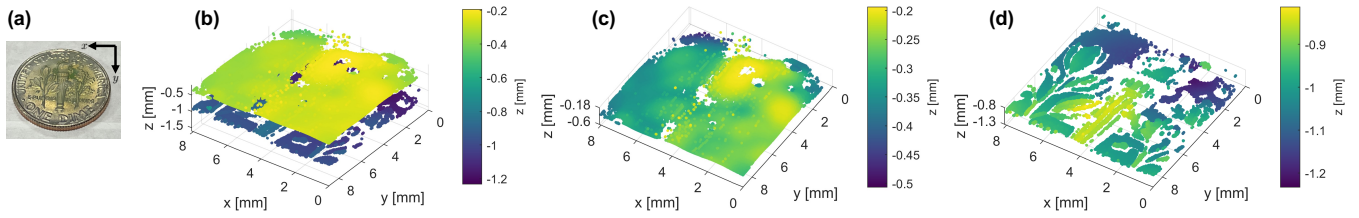
Fig. 9: Effect of spatial regularization for surface estimation of the 'dime+glue' specimen (a). The resulting 3D point cloud (b) can easily be separated into the top (c) and bottom (d) layers for better inspection of the surface detail.

SNR Estimation: Careful threshold selection is unnecessary for high-SNR measurements, as a broad range of thresholds can easily distinguish between peaks due to signal and noise. Here we estimate the SNR of our measurements to underscore why principled threshold selection is so important for low-cost low-SNR OCT systems. Figure 7 shows estimates of the SNR for the sample measurement in Fig. 6. The total power P_{tot} is computed from the total A-scan intensity. Then after identifying and removing likely signal components, the noise power P_{noise} is computed from the total remaining A-scan intensity. The SNR is estimated as the ratio $\mathrm{SNR} = (\mathrm{P}_{tot} - \mathrm{P}_{noise})/\mathrm{P}_{noise}$. Figure 7 shows only the estimates for which the SNR is nonzero (i.e., a surface reflection was identified). The number of A-scans with a surface detection clearly increases as the number of frames increases. For small numbers of frames, the reflectivity threshold is conservative, accepting surface detections only if the SNR is roughly $0\,\mathrm{dB}$ or greater. As the number of frames increases, the amplitude peak corresponding to a surface increases with the number of frames faster than the Neyman–Pearson threshold, which allows for surface detection from measurements with SNR estimated to be below $-10\,\mathrm{dB}$.

Precision Estimation: To evaluate the depth estimation precision of the OCT system and our SSE algorithm, we use a series of line scans of a planar aluminum test target. Line scans of the flat target were made at varying distances from the reference plane $z = 0$ in two configurations: (1) as a 'height' map ($z > 0$); and (2) as a 'depth' map ($z < 0$). For each distance, the SSE algorithm was applied to data from either 1 or 8 frames. To get a proxy for the ground truth position at each x-position in the line scan, the RANSAC algorithm [53] was used to fit a line to the estimated depths, rejecting outliers. Then the average error with respect to the line fit was used to calculate the precision from the inliers. The mean depth was the average of the estimated inlier depths. As shown in Fig. 8(a), the number of inliers increases with the number of frames, in accordance with the increase in SNR. Fig. 8(b) shows the depth error is around $1.5\,\mu\mathrm{m}$ at a distance of $0.1\,\mathrm{mm}$ from the reference plane—less than the axial resolution for the OCT system—and the error increases to $3-4\,\mu\mathrm{m}$ at $1\,\mathrm{mm}$ from the reference plane. However, the precision estimates are roughly the same for 1 or 8 frames, suggesting that, though surface detection improves with increasing numbers of frames (and increasing SNR), the depth error is limited by other factors such as the illumination bandwidth.

Optional Point Cloud Denoising: To further improve performance beyond the pixelwise detection and estimation strategy of the SSE, point cloud post-processing can be used to reject outliers or produce smoother surface estimates based on information from neighboring pixels. Classical point cloud processing approaches often harness signal processing on graphs [54, 55, 56, 57] or use local polynomial approximations [58]. A more recent trend is to use neural networks to learn and enforce priors on manifold structures for point cloud denoising [59, 60, 61, 62].

For a simple demonstration of the effect of post-processing, we show results in Fig. 9 with outlier rejection and total variation (TV) denoising via the method of Schoenenberger et al. [54] applied to the SSE output. The graph-based method proceeds in two steps: 1) remove outliers by forming an ϵ-nearest neighbor graph (i.e., connecting all points that are within a ball of radius ϵ from each other) and removing vertices that have few small edge weights; 2) apply total variation (TV) regularization to the graph coordinates (i.e., the remaining surface locations after outlier rejection), which promotes spatial smoothness. The graph-based regularization is implemented via the GSP graph signal processing toolbox [63]. Outlier rejection uses neighborhood size $\epsilon = 3 \times 10^{-5}\,\mathrm{m}$, kernel variance $\theta = 3 \times 10^{-5}\,\mathrm{m}$, and degree threshold $\tau = 2$; TV regularization uses neighborhood size $k = 25$ and regularization parameter $\gamma = 1 \times 10^{-6}$ (see [54] for details). We apply the post-processing separately to each 2D line-scan to avoid long computations for processing large 3D graphs.

Fig. 9 shows the results of combining the individual line-scans into a full 3D point cloud representation of the 'dime+glue' sample. In the Supplement, we further show how each of the denoising steps affects the SSE results for individual scan-lines. Using a larger number (i.e., 16) frames with additional post-processing enables good recovery of the true surface points with few outliers. Because the top glue surface and bottom coin layer are well-separated in depth, we can use a depth threshold of $-0.7\,\mathrm{mm}$ to individually inspect the reconstruction of the top and bottom layers. The top glue surface is almost completely recovered, while most of the coin surface is recovered with a few gaps.

VII. DISCUSSION AND CONCLUSION

Simultaneously determining both the number and position of layers is a challenging task at low SNR, since noise can cause spikes easily mistaken for surface profiles. We therefore introduce an algorithm for multi-layer surface reconstruction

with OCT that automatically detects layers using a threshold to limit the probability of falsely accepting noise spikes as true signal. Compared to existing methods, the number of spurious surface detections is greatly reduced, improving the ability to visualize and further analyze recovered surfaces.

While our approach effectively limits the number of detected outliers due to noise, some true signal peaks are inevitably missed. Several promising approaches could reduce the number of missed surface points, which are left for future work. One such approach would be to incorporate spatial information into the initial surface estimate directly rather than simply as post-processing, so that weak reflections are easier to detect. Another approach would be to better suppress or model the sources of noise. For instance, cross-talk is known to occur in line-scan or full-field OCT systems [27, 64, 65], and could potentially be removed computationally. Alternatively, we observe that the noise in our SD-OCT system is dominated by the DC measurement component. A useful direction of future research would explore low-cost configurations and noise statistics for balanced detectors, which remove common-mode noise. Finally, our method could be adapted for multi-layer estimation in other modalities such as lidar.

APPENDIX A
MLE DERIVATION

We consider the case of a sample known to have a single opaque layer and seek to estimate the depth z, reflectivity a, and phase shift ϕ relative to the reference mirror. Given the single-frame interference measurement vector $\mathbf{y} = [y_1, \ldots, y_N]^\mathsf{T}$ and assuming jointly Gaussian but heteroscedastic noise with variance σ_n^2 at element n, the likelihood is

$$p(\mathbf{y}; a, z, \phi) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma_n^2}}$$
$$\times \exp\left\{-\frac{1}{2\sigma_n^2}\left[y_n - \gamma\Psi(\lambda_n)a\cos\left(\frac{4\pi}{\lambda_n}z + \phi\right)\right]^2\right\}. \quad (31)$$

The negative log-likelihood is

$$\mathcal{L}(\mathbf{y}; a, z, \phi) = \frac{1}{2}J_\mathbf{y}(a, z, \phi) - \log\left\{\prod_{n=1}^{N}\frac{1}{\sqrt{2\pi\sigma_n^2}}\right\}, \quad (32)$$

where

$$J_\mathbf{y}(a, z, \phi) = \sum_{n=1}^{N}\frac{1}{\sigma_n^2}\left[y_n - \gamma\Psi(\lambda_n)a\cos\left(\frac{4\pi}{\lambda_n}z + \phi\right)\right]^2. \quad (33)$$

Maximizing the likelihood with respect to a, z, ϕ requires minimizing $J_\mathbf{y}(a, z, \phi)$. Expanding the cosine in (33) and defining $\alpha_1 = a\cos(\phi)$ and $\alpha_2 = -a\sin(\phi)$ yields the quadratic equation

$$J'_\mathbf{y}(\alpha_1, \alpha_2, z) = \sum_{n=1}^{N}\frac{1}{\sigma_n^2}\left[y_n - \alpha_1\gamma\Psi(\lambda_n)\cos\left(\frac{4\pi z}{\lambda_n}\right) \right.$$
$$\left. - \alpha_2\gamma\Psi(\lambda_n)\sin\left(\frac{4\pi z}{\lambda_n}\right)\right]^2. \quad (34)$$

We define the column vectors $\mathbf{c} = [c_1, \ldots c_N]^\mathsf{T}$ and $\mathbf{s} = [s_1, \ldots, s_N]^\mathsf{T}$ as functions of some depth z, where $c_n = \gamma\Psi(\lambda_n)\cos(4\pi z/\lambda_n)$ and $s_n = \gamma\Psi(\lambda_n)\sin(4\pi z/\lambda_n)$. We also define the parameter vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]^\mathsf{T}$ and the matrix $\mathbf{H} = [\mathbf{c}, \mathbf{s}]$. The matrix $\boldsymbol{\Sigma}^{-1} = \text{diag}([1/\sigma_1^2, \ldots, 1/\sigma_N^2]^\mathsf{T})$ contains the reciprocals of the noise variances on its diagonal. Then we can rewrite (34) as the quadratic function

$$J'_\mathbf{y}(\alpha_1, \alpha_2, z) = (\mathbf{y} - \mathbf{H}\boldsymbol{\alpha})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\alpha}), \quad (35)$$

which is well-known to be minimized by

$$\hat{\boldsymbol{\alpha}} = (\mathbf{H}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{H})^{-1}\mathbf{H}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}. \quad (36)$$

Replacing $\boldsymbol{\alpha}$ with its estimate, we can expand

$$J'_\mathbf{y}(\hat{\alpha}_1, \hat{\alpha}_2, z) = \mathbf{y}^\mathsf{T}\boldsymbol{\Sigma}^{-1}[\mathbf{I} - \mathbf{H}(\mathbf{H}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{H})^{-1}\mathbf{H}\boldsymbol{\Sigma}^{-1}]\mathbf{y}. \quad (37)$$

The MLE $\hat{z}_{\text{ML}}$ is then the value of z that maximizes

$$\Lambda(\mathbf{y}, z) = \mathbf{y}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{H}(\mathbf{H}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{H})^{-1}\mathbf{H}\boldsymbol{\Sigma}^{-1}\mathbf{y}. \quad (38)$$

Using the exact inverse

$$(\mathbf{H}\boldsymbol{\Sigma}^{-1}\mathbf{H})^{-1} = \frac{1}{G}\begin{bmatrix} \mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s} & -\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s} \\ -\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s} & \mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c} \end{bmatrix}, \quad (39)$$

where $G = (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s}) - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})^2$, we can rewrite the full objective to maximize as

$$\Lambda(\mathbf{y}, z) = \frac{1}{G}\Big[(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 + (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2$$
$$- 2(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})\Big]. \quad (40)$$

and $\hat{z}_{\text{ML}} = \arg\max_z \Lambda(\mathbf{y}, z)$. We can likewise expand (36) based on the inverse in (39):

$$\begin{bmatrix}\hat{\alpha}_1 \\ \hat{\alpha}_2\end{bmatrix} = \frac{1}{G}\begin{bmatrix}(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} \\ (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c})\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}\end{bmatrix} \quad (41)$$

We then define the functions

$$\Omega(\mathbf{y}, z) = \frac{1}{G}\Big\{[(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}]^2$$
$$+ [(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c})\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}]^2\Big\}^{\frac{1}{2}} \quad (42)$$

$$\Phi(\mathbf{y}, z) =$$
$$\arctan\left(-\frac{(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c})\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}}{(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y} - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}}\right). \quad (43)$$

Given $\hat{z}_{\text{ML}}$, the ML reflectivity is then

$$\hat{a}_{\text{ML}} = \sqrt{\hat{\alpha}_1^2(\hat{z}_{\text{ML}}) + \hat{\alpha}_2^2(\hat{z}_{\text{ML}})} = \Omega(\mathbf{y}, \hat{z}_{\text{ML}}), \quad (44)$$

and the phase offset is

$$\hat{\phi}_{\text{ML}} = \arctan(-\hat{\alpha}_2(\hat{z}_{\text{ML}})/\hat{\alpha}_1(\hat{z}_{\text{ML}})) = \Phi(\mathbf{y}, \hat{z}_{\text{ML}}). \quad (45)$$

A. MLE Approximations

The following approximations can be used to simplify the exact MLE formulations:

$$
\begin{aligned}
\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c} &= \mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s} \\
&= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2} \cos\left(\frac{4\pi z}{\lambda_n}\right)\sin\left(\frac{4\pi z}{\lambda_n}\right) \\
&= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{2\sigma_n^2}\left[\sin(0)+\sin\left(\frac{8\pi z}{\lambda_n}\right)\right] \\
&\approx 0
\end{aligned}
\tag{46}
$$

$$
\begin{aligned}
\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s} &\approx \mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c} \\
&= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2} \cos^2\left(\frac{4\pi z}{\lambda_n}\right) \\
&= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{2\sigma_n^2}\left[1+\cos\left(\frac{8\pi z}{\lambda_n}\right)\right] \\
&\approx \frac{\gamma}{\beta},
\end{aligned}
\tag{47}
$$

where (47) uses $\sigma_n^2 = \beta\gamma\Psi(\lambda_n)/2$. Then

$$
\begin{aligned}
\Lambda(\mathbf{y},z) &\approx \frac{(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 + (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2}{\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c}} \\
&\approx \frac{\beta}{\gamma}\left|\sum_{n=1}^{N}\frac{y_n}{\sigma_n^2}\gamma\Psi(\lambda_n)\exp\left(-\mathrm{i}\frac{4\pi z}{\lambda_n}\right)\right|^2 \\
&= \frac{4}{\beta\gamma}\left|\sum_{n=1}^{N} y_n\exp\left(-\mathrm{i}\frac{4\pi z}{\lambda_n}\right)\right|^2,
\end{aligned}
\tag{48}
$$

and

$$
\begin{aligned}
\Omega(\mathbf{y},z) &\approx \frac{\sqrt{(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 + (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2}}{\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c}} \\
&\approx \frac{\beta}{\gamma}\left|\sum_{n=1}^{N}\frac{y_n}{\sigma_n^2}\gamma\Psi(\lambda_n)\exp\left(-\mathrm{i}\frac{4\pi z}{\lambda_n}\right)\right| \\
&= \frac{2}{\gamma}\left|\sum_{n=1}^{N} y_n\exp\left(-\mathrm{i}\frac{4\pi z}{\lambda_n}\right)\right|.
\end{aligned}
\tag{49}
$$

APPENDIX B
MLE DERIVATIVES

The exact MLE is the value of z that maximizes (13). The first and second derivatives of (13) are useful for fast maximization via Newton's method. We thus define the vectors $\underline{\mathbf{c}} = \mathbf{c}\oslash\boldsymbol{\lambda}$, $\underline{\mathbf{s}} = \mathbf{s}\oslash\boldsymbol{\lambda}$, and $\underline{\mathbf{y}} = \mathbf{y}\oslash\boldsymbol{\lambda}$, where the symbol $\oslash$ implies elementwise (Hadamard) division, i.e., $\underline{c}_n = c_n/\lambda_n$.

Let the cost function be $\Lambda(\mathbf{y},z) = g(\mathbf{y},z)/h(\mathbf{y},z)$. The derivatives with respect to z are

$$
\begin{aligned}
g'(\mathbf{y},z) = 8\pi\Big\{ &(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 \\
&- (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&+ (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&- (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 \\
&- (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&+ (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&- (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&+ (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})\Big\}
\end{aligned}
\tag{50}
$$

$$
\begin{aligned}
g''(\mathbf{y},z) = 32\pi^2\Big\{ &2(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&- 2(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&- 2(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&+ 2(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&- 4(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&- 4(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})] \\
&+ 4(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) \\
&+ (\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}} - \underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}})\big[(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 - (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2\big] \\
&+ (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c})\big[(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 - (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{y}})\big] \\
&+ (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\big[(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})^2 - (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{y}})\big] \\
&+ (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})\big[(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{y}}) \\
&+ 2(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y}) + (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{y})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{y}})\big]\Big\}
\end{aligned}
\tag{51}
$$

$$
\begin{aligned}
h'(\mathbf{y},z) = 8\pi\{&\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c}(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}}) - \mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s}(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}}) \\
&- \mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s}(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}} - \mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}})\}
\end{aligned}
\tag{52}
$$

$$
\begin{aligned}
h''(\mathbf{y},z) = 32\pi^2\Big\{&(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{c} - \mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}} - \underline{\mathbf{s}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}}) \\
&+ 2(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}}) + 4(\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{s})(\underline{\mathbf{c}}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}}) \\
&- (\mathbf{c}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})^2 - (\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{s}})^2 - 4(\mathbf{s}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\underline{\mathbf{c}})^2\Big\}.
\end{aligned}
\tag{53}
$$

The cost function derivatives are

$$
\Lambda'(\mathbf{y},z) = \frac{g'(\mathbf{y},z)h(\mathbf{y},z) - g(\mathbf{y},z)h'(\mathbf{y},z)}{[h(\mathbf{y},z)]^2}
\tag{54}
$$

$$
\Lambda''(\mathbf{y},z) = \frac{g''(\mathbf{y},z) - h''(\mathbf{y},z)\Lambda(\mathbf{y},z) - 2h'(\mathbf{y},z)\Lambda'(\mathbf{y},z)}{h(\mathbf{y},z)}.
\tag{55}
$$

The Newton iterations are initialized with $z^{(0)} = \hat{z}_{\mathrm{BP}}$ and continue as

$$
z^{(t+1)} = z^{(t)} - \frac{\Lambda'(\mathbf{y},z^{(t)})}{\Lambda''(\mathbf{y},z^{(t)})}.
\tag{56}
$$

APPENDIX C
SINGLE-SURFACE CRAMÉR–RAO LOWER BOUND

The Cramér–Rao Lower Bound (CRLB) places a lower bound on the variance of an unbiased estimator [42, Ch. 3]. The CRLB is computed by first forming the Fisher information matrix (FIM) and computing the inverse. The Fisher information for a vector parameter $\boldsymbol{\theta}$ is

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\} \quad (57)$$

The CRLB for parameter i is given by the (i,i) element of the inverse Fisher info matrix:

$$\mathrm{Var}(\hat{\theta}_i) \geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{ii}. \quad (58)$$

We consider the CRLB for the vector parameter $\boldsymbol{\theta} = [a, z, \phi]^{\mathsf{T}}$, assuming a single reflecting surface and using the heteroscedastic Gaussian noise model. Given the likelihood in (31), the elements of the FIM are found to be

$$[\mathbf{I}(\boldsymbol{\theta})]_{11} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{y};\boldsymbol{\theta})}{\partial a^2}\right\}$$
$$= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2}\cos^2\left(\frac{4\pi z}{\lambda_n}+\phi\right) \quad (59)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{12} = [\mathbf{I}(\boldsymbol{\theta})]_{21} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{y};\boldsymbol{\theta})}{\partial a\partial z}\right\}$$
$$= -\sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2}\frac{4\pi}{\lambda_n}a\cos\left(\frac{4\pi z}{\lambda_n}+\phi\right)\sin\left(\frac{4\pi z}{\lambda_n}+\phi\right) \quad (60)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{13} = [\mathbf{I}(\boldsymbol{\theta})]_{31} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{y};\boldsymbol{\theta})}{\partial a\partial\phi}\right\}$$
$$= -\sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2}a\cos\left(\frac{4\pi z}{\lambda_n}+\phi\right)\sin\left(\frac{4\pi z}{\lambda_n}+\phi\right) \quad (61)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{22} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{y};\boldsymbol{\theta})}{\partial z^2}\right\}$$
$$= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2}\left(\frac{4\pi}{\lambda_n}\right)^2 a^2\sin^2\left(\frac{4\pi z}{\lambda_n}+\phi\right) \quad (62)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{23} = [\mathbf{I}(\boldsymbol{\theta})]_{32} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{y};\boldsymbol{\theta})}{\partial\phi\partial z}\right\}$$
$$= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2}\frac{4\pi}{\lambda_n}a^2\sin^2\left(\frac{4\pi z}{\lambda_n}+\phi\right) \quad (63)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{33} = -\mathbb{E}\left\{\frac{\partial^2 \log p(\mathbf{y};\boldsymbol{\theta})}{\partial\phi^2}\right\}$$
$$= \sum_{n=1}^{N} \frac{\gamma^2\Psi^2(\lambda_n)}{\sigma_n^2}a^2\sin^2\left(\frac{4\pi z}{\lambda_n}+\phi\right). \quad (64)$$

REFERENCES

[1] J. Rapp, H. Mansour, P. Boufounos, P. Orlik, T. Koike-Akino, and K. Parsons, "Maximum likelihood surface profilometry via optical coherence tomography," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2022, pp. 1561–1565.

[2] A. F. Fercher, "Optical coherence tomography – development, principles, applications," *Z. Med. Phys.*, vol. 20, no. 4, pp. 251–276, Nov. 2010.

[3] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.

[4] D. Stifter, "Beyond biomedicine: a review of alternative applications and developments for optical coherence tomography," *Appl. Phys. B*, vol. 88, no. 3, pp. 337–357, Aug. 2007.

[5] M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," *Opt. Lett.*, vol. 25, no. 8, pp. 545–547, Apr. 2000.

[6] D. Stifter, P. Burgholzer, O. Höglinger, E. Götzinger, and C. Hitzenberger, "Polarisation-sensitive optical coherence tomography for material characterisation and strain-field mapping," *Appl. Phys. A*, vol. 76, no. 6, pp. 947–951, Apr. 2003.

[7] M. Dufour, G. Lamouche, V. Detalle, B. Gauthier, and P. Sammut, "Low-coherence interferometry, an advanced technique for optical metrology in industry," *World Conf. Non-Destr. Test.*, vol. 47, no. 4, pp. 216–219, Apr. 2005.

[8] T. Endo, Y. Yasuno, S. Makita, M. Itoh, and T. Yatagai, "Profilometry with line-field Fourier-domain interferometry," *Opt. Express*, vol. 13, no. 3, pp. 695–701, Feb. 2005.

[9] W. J. Walecki, A. Pravdivtsev, K. Lai, M. Santos, II, G. Mikhaylov, M. Mihaylov, and A. Koo, "Fiber-optics low-coherence integrated metrology for in-situ non-contact characterization of novel materials and structures," in *AIP Conf. Proc.*, ser. Characterization and Metrology for ULSI Technology, vol. 788, Richardson, Texas, USA, 2005, pp. 338–342.

[10] E. Alarousu, L. Krehut, T. Prykäri, and R. Myllylä, "Study on the use of optical coherence tomography in measurements of paper properties," *Meas. Sci. Technol.*, vol. 16, no. 5, pp. 1131–1137, Mar. 2005.

[11] T. Prykäri, M. Tuominen, E. Alarousu, J. Czajkowski, and R. Myllylä, "Investigating coating layer of paper by optical coherence tomography," in *Int. Top. Meeting Opt. Sens. Artif. Vision (OSAV)*, Saint Petersburg, Russia, 2008, pp. 410–417.

[12] S. Lawman and H. Liang, "Fourier domain optical coherence tomography for high-precision profilometry," in *O3A: Optics for Arts, Architecture, and Archaeology II*, vol. 7391. Proc. SPIE, Jul. 2009, art. no. 73910H.

[13] ——, "High precision dynamic multi-interface profilometry with optical coherence tomography," *Appl. Optics*, vol. 50, no. 32, pp. 6039–6048, Nov. 2011.

[14] S. J. Lawman, "Optical and material properties of varnishes for paintings," PhD Dissertation, Nottingham Trent University, Nottingham, England, Aug. 2011.

[15] Y. Dong, S. Lawman, Y. Zheng, D. Williams, J. Zhang, and Y.-C. Shen, "Nondestructive analysis of automotive paints with spectral domain optical coherence tomography," *Appl. Optics*, vol. 55, no. 13, pp. 3695–3700, May 2016.

[16] J. Czajkowski, T. Prykäri, E. Alarousu, J. Palosaari, and R. Myllylä, "Optical coherence tomography as a method of quality inspection for printed electronics products," *Opt. Rev.*, vol. 17, no. 3, pp. 257–262, May 2010.

[17] J. Czajkowski, T. Fabritius, J. Ulański, T. Marszałek, M. Gazicki-Lipman, A. Nosal, R. Śliż, E. Alarousu, T. Prykäri, R. Myllylä, and G. Jabbour, "Ultra-high resolution optical coherence tomography for encapsulation quality inspection," *Appl. Phys. B*, vol. 105, no. 3, pp. 649–657, Nov. 2011.

[18] P. Liu, R. M. Groves, and R. Benedictus, "3D monitoring of delamination growth in a wind turbine blade composite using optical coherence tomography," *NDT & E Int.*, vol. 64, pp. 52–58, Jun. 2014.

[19] M. J. Marques, R. Green, R. King, S. Clement, P. Hallett, and A. Podoleanu, "Sub-surface characterisation of latest-generation identification documents using optical coherence tomography," *Sci. Justice*, vol. 61, no. 2, pp. 119–129, Mar. 2021.

[20] I. Gkioulekas, A. Levin, F. Durand, and T. Zickler, "Micron-scale light transport decomposition using interferometry," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 1–14, Jul. 2015, art. no. 37.

[21] A. Kotwal, A. Levin, and I. Gkioulekas, "Interferometric transmission probing with coded mutual intensity," *ACM Trans. Graph.*, vol. 39, no. 4, pp. 1–16, Aug. 2020, art. no. 74.

[22] C. Yurdakul, D. S. Millar, H. Mansour, P. Wang, K. Kojima, T. Koike-Akino, K. Parsons, and P. V. Orlik, "Line-field coherent sensing with LED illumination," arXiv:2202.12161v1, Feb. 2022.

[23] S. Lawman, Z. Zhang, Y.-C. Shen, and Y. Zheng, "Line field optical coherence tomography," *Photonics*, vol. 9, no. 12, p. 946, Dec. 2022.

[24] J. A. Izatt and M. A. Choma, "Theory of optical coherence tomography," in *Optical Coherence Tomography: Technology and Applications*, ser. Biological and Medical Physics, Biomedical Engineering, W. Drexler and J. G. Fujimoto, Eds. Berlin, Heidelberg: Springer, 2008, pp. 47–72.

[25] C. Li, S. Chen, and Y. Zhu, "Maximum likelihood estimation of optical path length in spectral interferometry," *J. Lightwave Technol.*, vol. 35, no. 22, pp. 4880–4887, Nov. 2017.

[26] C. Li and Y. Zhu, "Cramer–Rao bound for frequency estimation of spectral interference and its shot noise-limited behavior," *IEEE J. Sel. Topics Quantum Electron.*, vol. 23, no. 2, pp. 410–416, Mar. 2017.

[27] L. Han and K. Bizheva, "Correcting spatial-spectral crosstalk and chromatic aberrations in broadband line-scan spectral-domain OCT images," *Biomed. Opt. Express*, vol. 14, no. 7, pp. 3344–3361, Jul. 2023.

[28] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.

[29] S. Djukanović, "An accurate method for frequency estimation of a real sinusoid," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 915–918, Jul. 2016.

[30] M. Wojtkowski, V. J. Srinivasan, T. H. Ko, J. G. Fujimoto, A. Kowalczyk, and J. S. Duker, "Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation," *Opt. Express*, vol. 12, no. 11, pp. 2404–2422, May 2004.

[31] K. Zhang and J. U. Kang, "Graphics processing unit accelerated non-uniform fast Fourier transform for ultrahigh-speed, real-time Fourier-domain OCT," *Opt. Express*, vol. 18, no. 22, pp. 23 472–23 487, Oct. 2010.

[32] A. Dutt and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SIAM J. Sci. Comput.*, vol. 14, no. 6, pp. 1368–1393, Nov. 1993.

[33] S. S. Sherif, C. Flueraru, Y. Mao, and S. Chang, "Swept source optical coherence tomography with nonuniform frequency domain sampling," in *Proc. Biomed. Opt.* Optical Society of America, Mar. 2008, p. BMD86.

[34] K. Wang, Z. Ding, T. Wu, C. Wang, J. Meng, M. Chen, and L. Xu, "Development of a non-uniform discrete Fourier transform based high speed spectral domain optical coherence tomography system," *Opt. Express*, vol. 17, no. 14, pp. 12 121–12 131, Jul. 2009.

[35] D. Hillmann, G. Hüttmann, and P. Koch, "Using nonequispaced fast Fourier transformation to process optical coherence tomography signals," in *Proc. SPIE-OSA Biomed. Opt.*, ser. Opt. Coher. Tomog. Coher. Tech. IV, vol. 7372. SPIE, Jul. 2009, pp. 100–105.

[36] K. K. H. Chan and S. Tang, "High-speed spectral domain optical coherence tomography using non-uniform fast Fourier transform," *Biomed. Opt. Express*, vol. 1, no. 5, pp. 1309–1319, Dec. 2010.

[37] S. Vergnole, D. Lévesque, and G. Lamouche, "Experimental validation of an optimized signal processing method to handle non-linearity in swept-source optical coherence tomography," *Opt. Express*, vol. 18, no. 10, pp. 10 446–10 461, May 2010.

[38] Y. Yasuno, V. D. Madjarova, S. Makita, M. Akiba, A. Morosawa, C. Chong, T. Sakai, K.-P. Chan, M. Itoh, and T. Yatagai, "Three-dimensional and high-speed swept-source optical coherence tomography for in vivo investigation of human anterior eye segments," *Opt. Express*, vol. 13, no. 26, pp. 10 652–10 664, Dec. 2005.

[39] Z. Hu and A. M. Rollins, "Fourier domain optical coherence tomography with a linear-in-wavenumber spectrometer," *Opt. Lett.*, vol. 32, no. 24, pp. 3525–3527, Dec. 2007.

[40] A. Morosawa and C. Chong, "Optical coherence tomography system and optical coherence tomography method," U.S. patent 7835010B2, Nov. 2010.

[41] X. Zhang, J. Pouls, and M. C. Wu, "Laser frequency sweep linearization by iterative learning pre-distortion for FMCW LiDAR," *Opt. Express*, vol. 27, no. 7, pp. 9965–9974, Apr. 2019.

[42] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[43] C. Akcay, P. Parrein, and J. P. Rolland, "Estimation of longitudinal resolution in optical coherence imaging," *Appl. Optics*, vol. 41, no. 25, pp. 5256–5262, Sep. 2002.

[44] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, 1992.

[45] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, 1998.

[46] S. Djukanović and V. Popović-Bugarin, "Efficient and accurate detection and frequency estimation of multiple sinusoids," *IEEE Access*, vol. 7, pp. 1118–1125, 2019.

[47] T. W. Anderson, "On the distribution of the two-sample Cramér-von Mises criterion," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1148–1159, Sep. 1962.

[48] O. Atalar, D. S. Millar, P. Wang, T. Koike-Akino, K. Kojima, P. V. Orlik, and K. Parsons, "Spectrally sparse optical coherence tomography," *Opt. Express*, vol. 28, no. 25, pp. 37 798–37 810, Dec. 2020.

[49] D. Williams, Y. Zheng, P. G. Davey, F. Bao, M. Shen, and A. Elsheikh, "Reconstruction of 3D surface maps from anterior segment optical coherence tomography images using graph theory and genetic algorithms," *Biomed. Signal Process. Control*, vol. 25, pp. 91–98, Mar. 2016.

[50] H. L. Seck, Y. Zhang, and Y. C. Soh, "Optical coherence tomography by using frequency measurements in wavelength domain," *Opt. Express*, vol. 19, no. 2, pp. 1324–1334, Jan. 2011.

[51] R. Kobayashi, G. Fujii, Y. Yoshida, T. Ota, F. Nin, H. Hibino, S. Choi, S. Ono, and S. Muramatsu, "Sparsity-aware OCT volumetric data restoration using optical synthesis model," *IEEE Trans. Comput. Imaging*, vol. 8, pp. 505–520, 2022.

[52] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.

[53] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[54] Y. Schoenenberger, J. Paratte, and P. Vandergheynst, "Graph-based denoising for time-varying point clouds," in *Proc. 3DTV-Conf., True Vis.-Capture, Transmiss. Display 3D Video*, Jul. 2015, pp. 1–4.

[55] S. Deutsch, A. Ortega, and G. Medioni, "Robust denoising of piece-wise smooth manifolds," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.*, Apr. 2018, pp. 2786–2790.

[56] W. Hu, X. Gao, G. Cheung, and Z. Guo, "Feature graph learning for 3D point cloud denoising," *IEEE Trans. Signal Process.*, vol. 68, pp. 2841–2856, 2020.

[57] C. Dinesh, G. Cheung, and I. V. Bajic, "Point cloud denoising via feature graph Laplacian regularization," vol. 29, pp. 4143–4158, 2020.

[58] Z. Xu and A. Foi, "Anisotropic denoising of 3D point clouds by aggregation of multiple surface-adaptive estimates," *IEEE Trans. Visual Comput. Graphics*, vol. 27, no. 6, pp. 2851–2868, Jun. 2021.

[59] C. Duan, S. Chen, and J. Kovacevic, "3D point cloud denoising via deep neural network based local surface estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.*, Brighton, United Kingdom, May 2019, pp. 8553–8557.

[60] M. Gadelha, R. Wang, and S. Maji, "Deep manifold prior," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops*, Oct. 2021, pp. 1107–1116.

[61] S. Luo and W. Hu, "Score-based point cloud denoising," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.* IEEE Computer Society, Oct. 2021, pp. 4563–4572.

[62] F. Pistilli, G. Fracastoro, D. Valsesia, and E. Magli, "Learning robust graph-convolutional representations for point cloud denoising," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 2, pp. 402–414, Feb. 2021.

[63] N. Perraudin, J. Paratte, D. Shuman, L. Martin, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "GSPBOX: A toolbox for signal processing on graphs," Mar. 2016, arXiv:1408.5781 [cs.IT].

[64] T. Bonin, G. Franke, M. Hagen-Eggert, P. Koch, and G. Hüttmann, "In vivo Fourier-domain full-field OCT of the human retina with 1.5 million A-lines/s," *Opt. Lett.*, vol. 35, no. 20, pp. 3432–3434, Oct. 2010.

[65] P. Stremplewski, E. Auksorius, P. Wnuk, L. Kozoń, P. Garstecki, and M. Wojtkowski, "In vivo volumetric imaging by crosstalk-free full-field OCT," *Optica*, vol. 6, no. 5, pp. 608–617, May 2019.