# Extracting Images and Displacements from Raw Spectral Domain Optical Coherence Tomography Data

Jaeho Cho

*ECE444 - Bioinstrumentation and Sensing*
*The Cooper Union for the Advancement of Science and Art*
New York City, NY
jaeho.cho@cooper.edu

*Abstract*—**Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere.**

*Index Terms*—**SD-OCT, SDPM, Medical Imaging**

## I. INTRODUCTION

Optical coherence tomography (OCT) is an optical imaging modality that can perform high resolution, crosssectional imaging by measuring echoes of backscattered light. The unique features of OCT make it a powerful imaging modality for many fundamental research and clinical applications. Coupled with catheter, endoscopic, laparoscopic, or needle delivery devices, OCT could enable new minimally invasive surgical procedures.

OCT performs internal imaging by measuring the magnitude and echo time delay of backscattered light. Crosssectional images are generated by performing multiple axial measurements of echo time delay (A-scans) and scanning the incident optical beam transversely. This produces a two-dimensional data set, which represents the optical backscattering in a cross-sectional plane through the sample (B-scan).

Repeated A-scans at the same location as a function of time can be used to generate a time-resolved image (M-scan). This technique called Spectral domain phase microscopy (SDPM) is useful for measuring the displacement of structures within the sample.

Spectral domain optical coherence tomography (SD-OCT) is a type of OCT that uses a spectrometer to capture the interference pattern of light reflected from a sample and a reference arm. In SD-OCT, the source is broadband and continuous-wave, the reference arm length is fixed at a position approximately corresponding to the position of the sample, and the spectral interference pattern between the light returning from the reference arm and all depths in the sample is dispersed by a spectrometer and collected simultaneously on an array detector such as a photodiode array or charge-coupled device (CCD). The spectral information is then used to reconstruct depth-resolved images of the sample.

The detector current ($I_D$) can be expressed as a function of wavenumber ($k$), commonly known as the "spectral interferogram":

$$I_D(k) = \frac{\rho}{4}[S(k)(R_R + R_{S1} + R_{S2} + ...)]$$

$$+ \frac{\rho}{2}\left[S(k)\sum_{n=1}^{N}\sqrt{R_R R_{S_n}}\cos\left(2k\left(z_R - z_{S_n}\right)\right)\right]$$

$$+ \frac{\rho}{4}\left[S(k)\sum_{n\neq m=1}^{N}\sqrt{R_{S_n}R_{S_m}}\cos\left(2k\left(z_{S_n} - z_{S_m}\right)\right)\right] \tag{1}$$

where $\rho$ is the responsivity of the detector, $S(k)$ is the spectral dependence of the light source, $R_R$ is the power reflectivity of the reference reflector, $R_{\{S_n\}}$ is the power reflectivity of the sample reflector, and $z_R$ and $z_{\{S_n\}}$ are the pathlength variables in the reference and sample arms measured from the beamsplitters, respectively. The first component is the "DC Terms", the second is the "Cross-correlation Terms" and the third is the "Auto-correlation Terms" [1].

The internal reflectivity profile of the sample can be estimated from the inverse Fourier transform of $I_D(k)$, generating an A-scan expressed as:

$$i_D(z) = \frac{\rho}{8}[\gamma(z)[R_R + R_{S1} + R_{S2} + ...]]$$

$$+ \frac{\rho}{4}\sum_{n=1}^{N}\sqrt{R_R R_{S_n}}\left[\gamma\left[2\left(z_R - z_{S_n}\right)\right] + \gamma\left[-2\left(z_R - z_{S_n}\right)\right]\right]$$

$$+ \frac{\rho}{8}\sum_{n\neq m=1}^{N}\sqrt{R_{S_n}R_{S_m}}\left[\gamma\left[2\left(z_{S_n} - z_{S_m}\right)\right] + \gamma\left[-2\left(z_{S_n} - z_{S_m}\right)\right]\right] \tag{2}$$

where $\gamma(z)$ is inverse Fourier transform of the source spectrum [1].

## II. METHODOLOGY

### A. Data Acquisition

The raw data were acquired in air, which has of an index of refraction ($n$) of 1, using a ThorLabs Telesto 3, which uses an SLD with a center wavelength ($\lambda_0$) of 1310 nm, bandwidth ($\Delta\lambda$) of 100 nm; and an LSM-03 objective lens with numerical aperture (NA) of 0.055. Raw OCT data were collected in the form of A-scans, with each A-scan containing 2048 pixels each representing a depth of 3.6 $\mu$m.

Assuming that the lateral resolution is limited by the confocal geometrical optics, and the axial resolution is limited by

the low-coherence interferometer; the lateral resolution ($\delta x$) can be calculated as,

$$\delta x = 0.37 \frac{\lambda_0}{\text{NA}} \approx 8.81 \ \mu\text{m} \qquad (3)$$

and the axial resolution ($\delta z$),

$$\delta z = l_c = \frac{2 \ln(2)}{\pi} \frac{\lambda_0^2}{\Delta \lambda} \approx 7.57 \ \mu\text{m} \qquad (4)$$

where $l_c$ is the coherence length [1].

To improve the signal-to-noise ratio (SNR) and reduce artifacts, background scans were collected along with the sample scans to perform background subtraction.

As the sample was kept entirely to one side of the zero path length, the mirror image artifacts can be simply removed by displaying only the positive distances [1]. This correction results in the pixel aspect ratio of the Bscans to be double that of images without accounting for the mirror image artifacts; approximately 0.271 instead of 0.136.

The B-scan data consisted of $10,000$ A-scans of the sample taken along a line segment in optical coordinates, and 175 backgrounds. Two M-scans of a speaker were acquired, M-scan 1 and M-scan 40, where the speaker was playing 1 tone and 40 tones, respectively. The M-scans were acquired by collecting A-scans at a rate of 97656.25 Hz, for approximately 1 second. The raw data for M-scan 1 consisted of 104448 A-scans and M-scan 40 consisted of 106496 A-scans, each with 320 background scans.

### B. Image Processing

All processing and analysis were performed using MATLAB.

a) *A-scan:* As the raw data is initially captured as interferometric measurements given by the detector current, to reconstruct the internal sample reflectivity profile, the detector current first needs to be transformed from their original wavelength-based form into a uniformly spaced wavenumber (k-space) domain, which can be accomplished by a matrix multiplication. As can be seen in Fig. 1, the transformation to the k-domain does not significantly change the actual signal. Next, the signal undergoes background subtraction, where the average of the background scans is subtracted from the sample scan. This step is effective for removing the DC component in equation (1). As Fig. 1 shows, the background subtracted signal has noticeably more high frequency fluctuations. Fig. 2 also shows the difference between the raw detector currents of the sample and background scans. As the signal needs to be transformed into the spatial domain, a windowing function is applied to minimize spectral leakage and reduce artifacts. The windowing function is applied in the k-space domain, and the resulting signal is then deconvolved to compensate for system-induced distortions, which can further enhance resolution and image quality. Finally, a Fourier transform is performed to convert the signal from the k-space domain into spatial coordinates, yielding the depth-resolved reflectivity profile.
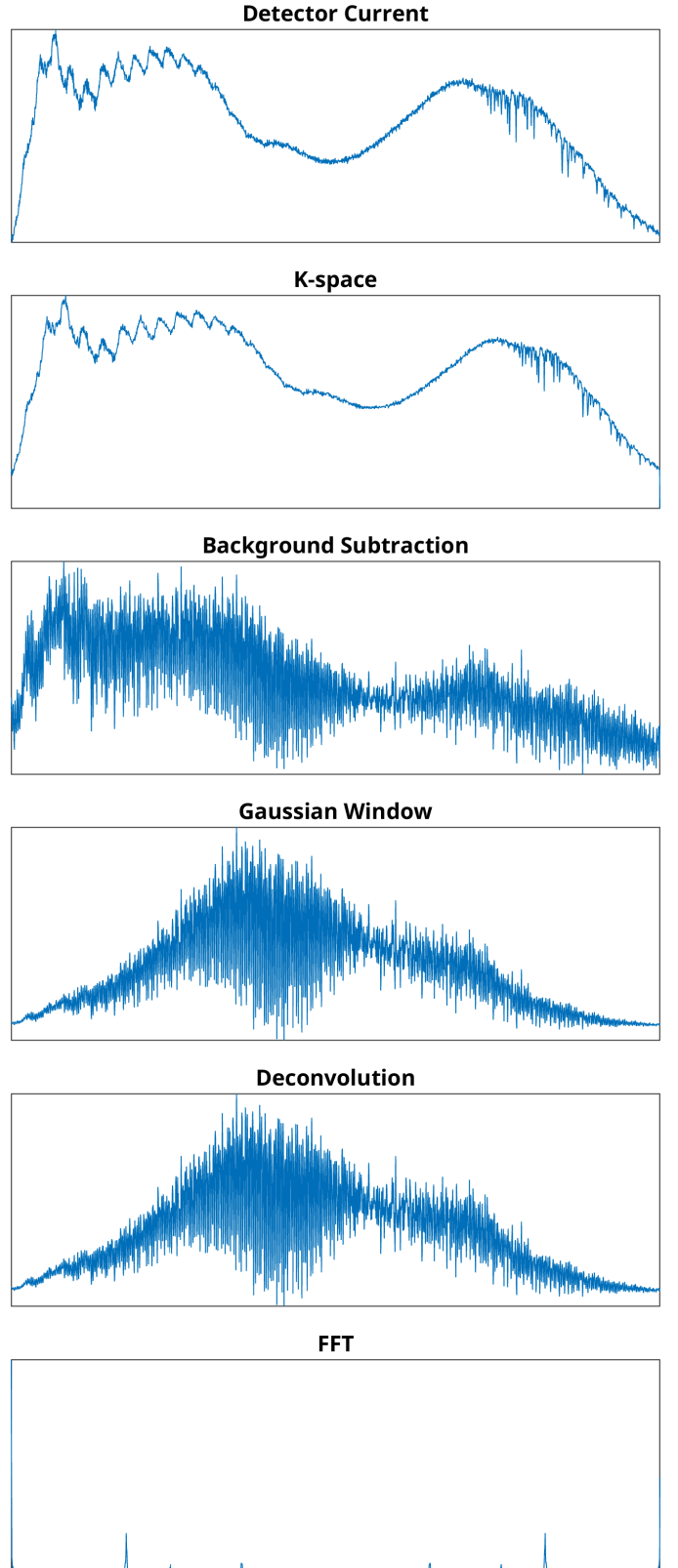


Fig. 1: Generation of an A-scan from Detector Current to reflectivity. The raw data is transformed from wavelength to wavenumber, followed by background subtraction, windowing, and deconvolution. The final step involves a Fourier transform to obtain the depth-resolved reflectivity profile.
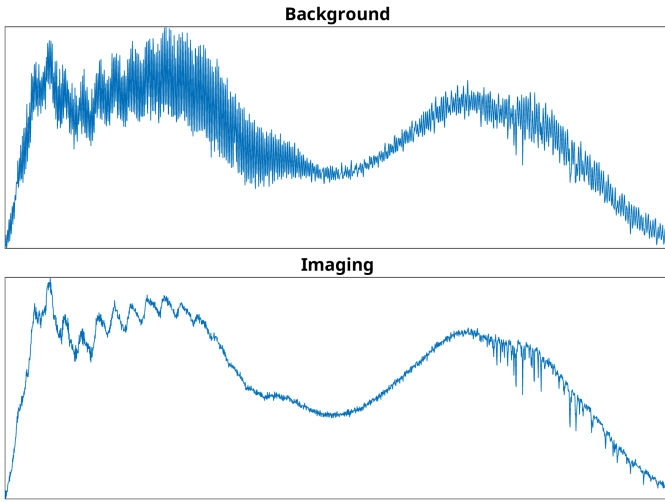
Fig. 2: Comparison of raw detector currents from the sample and background scans.

b) *B-scan:* The B-scan image was generated by stacking the A-scans together horizontally, and cropping to account for the mirror image artifact. The B-scan was then processed to improve the image quality by simply clipping the range of pixel values to around the max and minimum magnitudes of the region of interest illustrated in Fig. 3.
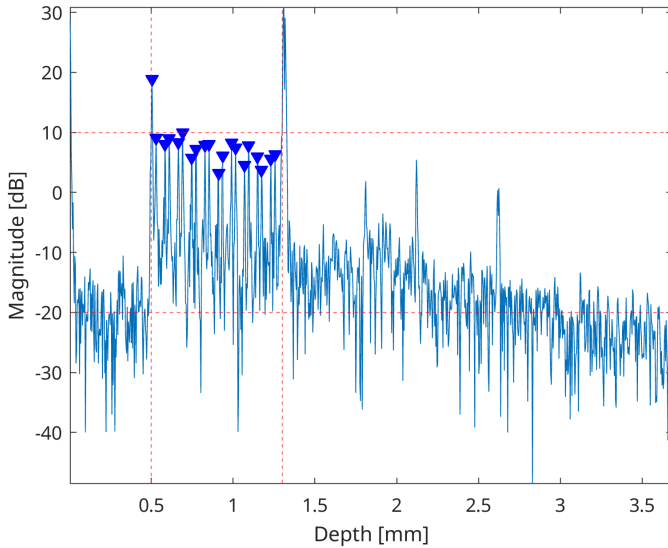


Fig. 3: A-scan at the center of the B-scan data, with the thresholds for the region of interest marked. The peaks within the region of interest are marked and were used to mark the edges of the tape layers.

Using the A-scan given in Fig. 3, the thickness of the tape layers were measured as the difference between the pairs of peaks and averaged. It was assumed that the larger difference between pairs of peaks corresponded to the tape layers, and the smaller difference corresponded to the distance between the tape layers.

c) *M-scan:* The M-scans were generated by collecting the A-scans together into a time-series, and individual pixel displacements were analyzed by tracking the phase of the pixel over time:

$$d(t) = \frac{\lambda_0 \angle i_D(t)}{4\pi n} \tag{5}$$

where $n$ is the index of refraction of the medium [2], and $i_D(t)$ is the complex value of the A-scan at a given pixel varying over time $t$

## III. RESULTS

### A. A-scan

The processed and generated A-scans for the B-scan and M-scans are shown in Fig. 4. These each represent the depth-resolved reflectivity profiles reconstructed from the raw data. They are expectedly symmetric with the negative depths mathematically representing the cross-correlation terms and the positive depths representing the mirror image artifacts.
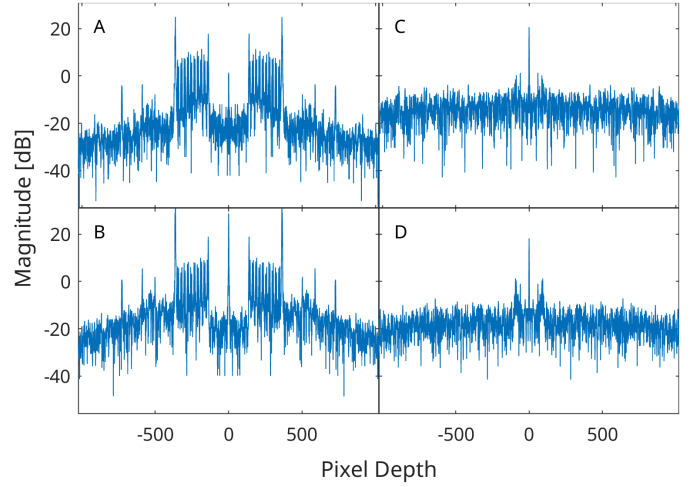


Fig. 4: A-scans of the raw data. A) the first A-scan of the Bscan data B) A-scan at the center of the Bscan data C) A-scan at the center of the Mscan1 data D) A-scan at the center of the Mscan40 data

### B. B-scan

The comparison between differently processed B-scan images is shown in Fig. 5. The processed B-scan image appears significantly darker than the unprocessed image, but expectedly so, as a large portion of the signal is below the threshold magnitude illustrated in Fig. 3. Comparing the images with and without deconvolution, it becomes apparent that the deconvolution process does not significantly improve the image quality. The image also appears to get darker at higher depths, which may be a demonstration of how the magnitude of the source light falls off as it propagates through the sample.
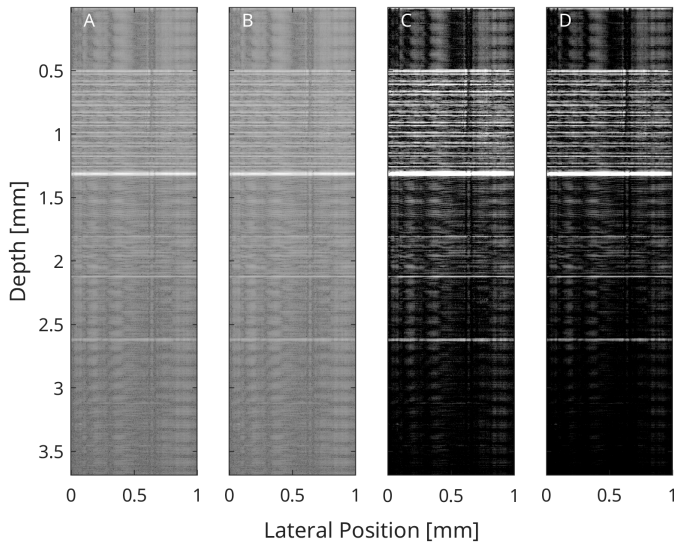
Fig. 5: B-scan images scaled to match real aspect ratio and cropped to display only half of the full image, correcting for mirror image artifacts. A) Unprocessed B-scan B) unprocessed B-scan without deconvolution C) processed B-scan D) processed B-scan without deconvolution

A close up of the layers of tape is shown in Fig. 6 where 10 layers of tape can be distinguished. Using the distances between peaks in Fig. 3, the thickness of the tape layers was on average 0.0552 mm; and the distance between the layers was 0.0256 mm.
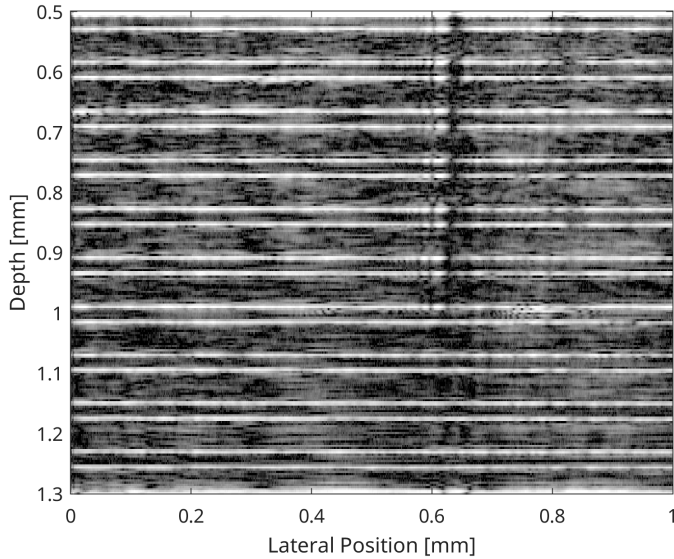


Fig. 6: A magnified view of the tape layers from the processed B-scan image. The large gaps are interpreted as the tape layers, and the smaller gaps are interpreted as the distance between the tape layers.
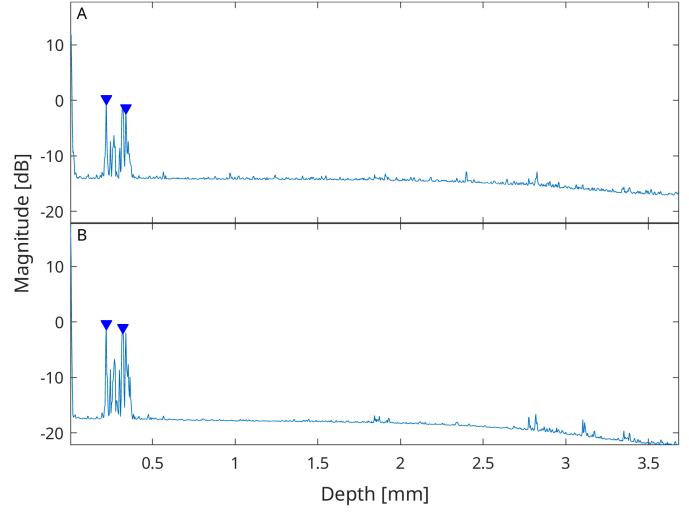
### C. M-scan



Fig. 7: Average A-scan magnitude of M-scans, two prominent peaks are marked on each plot, which are further analyzed. A) M-scan 1 with peaks at 0.22 mm (pixel 61) and 0.34 mm (pixel 94) B) M-scan 40 with peaks at 0.22 mm (pixel 61) and 0.32 mm (pixel 89).
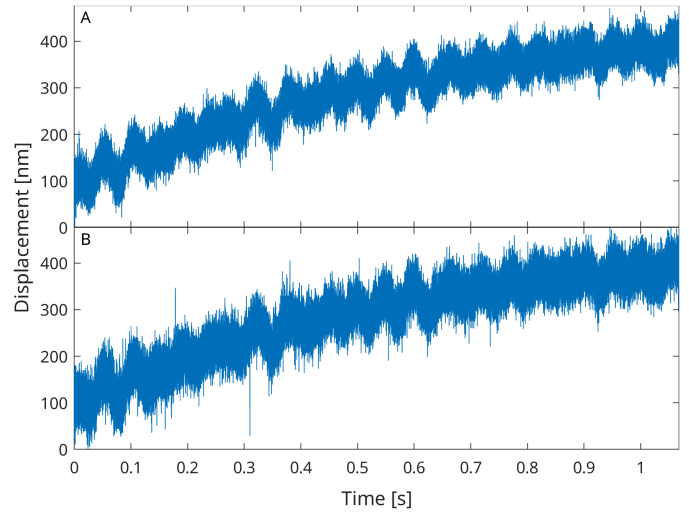


Fig. 8: Displacement over time for two prominent peaks in M-scan 1. The displacement axis was shifted to make the change in displacement clearer. A) first peak at 0.22 mm B) second peak at 0.34 mm.
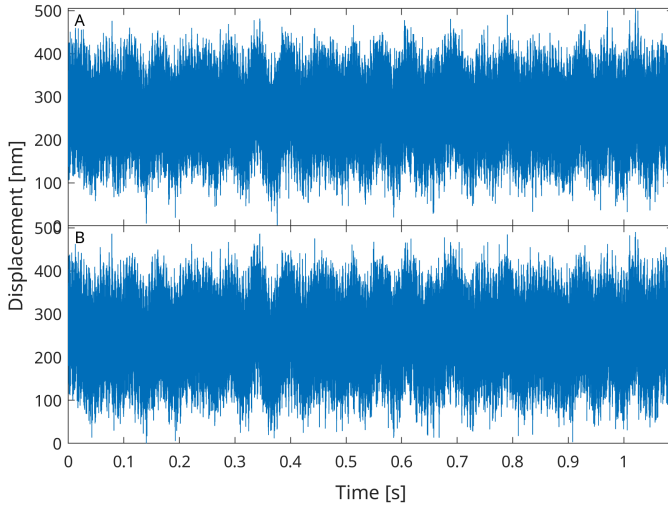
Fig. 9: Displacement over time for two prominent peaks in M-scan 40. The displacement axis was shifted to make the change in displacement clearer. A) first peak at 0.22 mm B) second peak at 0.32 mm.
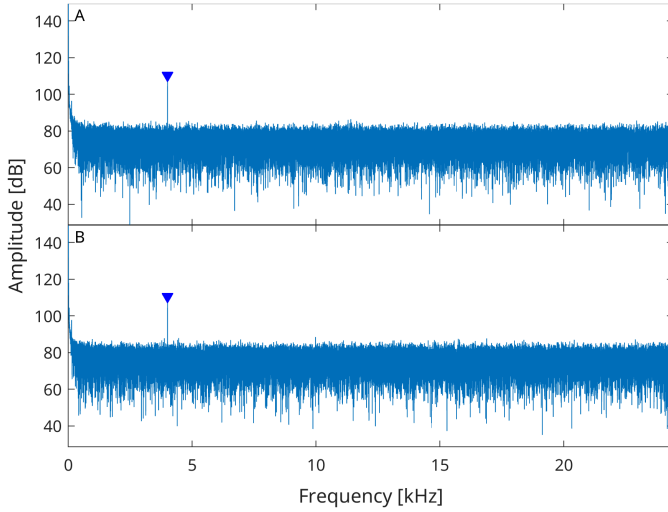


Fig. 10: The frequency spectrum of the displacement over time for M-scan 1 at A) 0.22 mm with a peak at 4.01 kHz B) 0.34 mm with a peak at 4.01 kHz.
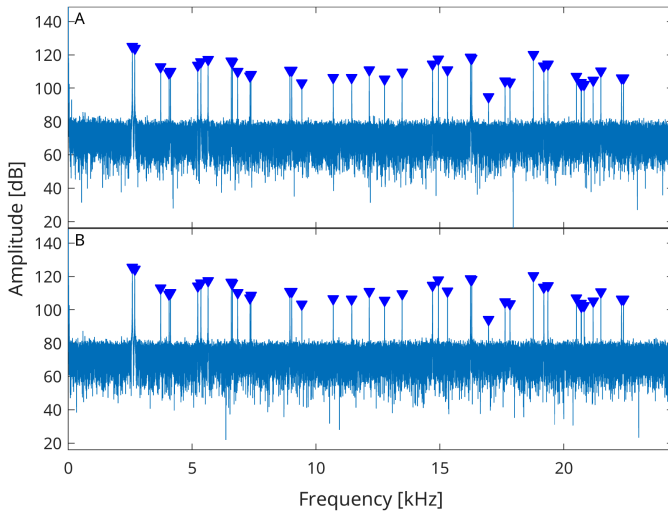


Fig. 11:

## IV. DISCUSSION

The effect of deconvolution and background subtraction on the A-scan generation is shown in Fig. 12 where it can be seen that background subtraction reduces the DC component while minimally effecting the signal, as the region of the signal where the tape layers are located are nearly perfectly overlapped. It also unexpectedly appears to reduce the auto-correlation terms as the signal near the zero pathlength is noticeably lower. The deconvolution process appears to shift the magnitude of the signal down, which is expected as the deconvolution process essentially normalizes the signal to the smoothed background.
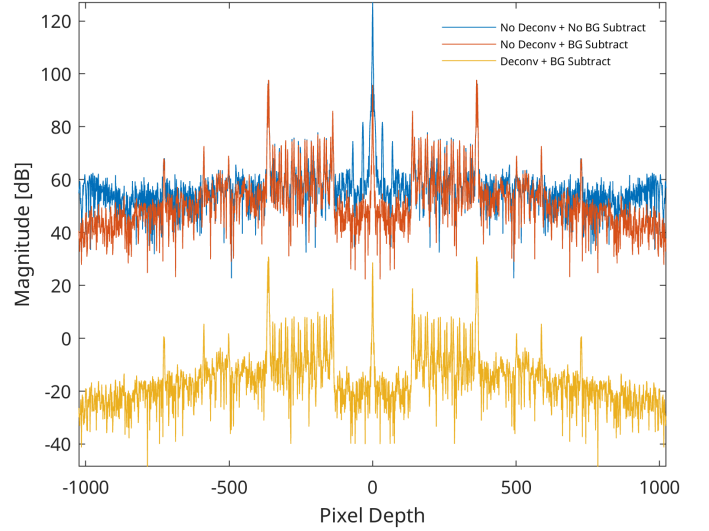


Fig. 12: Comparison of A-scan generation with and without deconvolution or background subtraction

The processing times for the different processing steps are shown in Table I, which shows that the B-scan processing times are significantly shorter compared to the M-scans, which is understable as the M-scans contain significantly more data. The time it takes to transform the data from wavelength to wavenumber surprisingly dominates the overall processing time, given that it is a necessary but simple step. However, it is understandable as it is a large matrix multiplication, compared to the computationally less intensive steps of column subtraction, convolution, element-wise division, and the Fast Fourier Transform.

TABLE I: A-SCAN PROCESSING TIMES

| Process | L2K | BGS | Window | Deconv | FFT |
|---|---|---|---|---|---|
| B-scan | 0.1877 | 0.0127 | 0.0144 | 0.0143 | 0.0215 |
| B-scan ~~Deconv~~ | 0.1882 | 0.0129 | 0.0132 | NA | 0.0219 |
| B-scan ~~BGS~~ ~~Deconv~~ | 0.2095 | NA | 0.0132 | NA | 0.0221 |
| M-scan1 | 1.8668 | 0.1335 | 0.1335 | 0.1373 | 0.2261 |
| M-scan40 | 1.9084 | 0.1360 | 0.1362 | 0.1361 | 0.2303 |

# V. Appendix

TABLE II: Displacement Frequency Tones (Mscan40)

| Tone # | Frequency (kHz) | Peak 1 (dB) | Peak 2 (dB) |
|---|---|---|---|
| 1 | 2.58 | 124.93 | 125.32 |
| 2 | 2.68 | 123.78 | 124.18 |
| 3 | 3.73 | 112.65 | 112.89 |
| 4 | 4.07 | 109.10 | 109.31 |
| 5 | 4.12 | 109.88 | 110.17 |
| 6 | 5.22 | 113.77 | 114.18 |
| 7 | 5.35 | 115.70 | 116.02 |
| 8 | 5.64 | 117.06 | 117.40 |
| 9 | 6.58 | 116.14 | 116.47 |
| 10 | 6.62 | 115.45 | 115.78 |
| 11 | 6.83 | 109.88 | 110.18 |
| 12 | 7.33 | 107.12 | 107.02 |
| 13 | 7.37 | 108.19 | 108.60 |
| 14 | 8.94 | 110.75 | 110.86 |
| 15 | 9.02 | 110.68 | 111.01 |
| 16 | 9.43 | 103.01 | 103.24 |
| 17 | 10.69 | 106.41 | 106.50 |
| 18 | 11.44 | 106.34 | 106.26 |
| 19 | 12.14 | 110.81 | 110.90 |
| 20 | 12.76 | 105.31 | 105.62 |
| 21 | 13.47 | 109.49 | 109.57 |
| 22 | 14.70 | 114.27 | 114.62 |
| 23 | 14.94 | 117.47 | 117.80 |
| 24 | 15.30 | 110.87 | 111.24 |
| 25 | 16.23 | 118.48 | 118.57 |
| 26 | 16.28 | 117.91 | 118.09 |
| 27 | 16.96 | 94.81 | 94.19 |
| 28 | 17.63 | 104.12 | 104.73 |
| 29 | 17.83 | 103.47 | 103.43 |
| 30 | 18.77 | 120.19 | 120.40 |
| 31 | 19.19 | 113.30 | 113.49 |
| 32 | 19.36 | 114.34 | 114.47 |
| 33 | 20.51 | 106.89 | 106.92 |
| 34 | 20.70 | 103.40 | 103.86 |
| 35 | 20.71 | 101.98 | 102.03 |
| 36 | 20.82 | 102.04 | 102.47 |
| 37 | 21.18 | 104.77 | 105.25 |
| 38 | 21.49 | 110.26 | 110.61 |
| 39 | 22.33 | 105.97 | 106.40 |
| 40 | 22.40 | 106.10 | 106.39 |

## References

[1] J. A. Izatt and M. A. Choma, "Theory of Optical Coherence Tomography," *Optical Coherence Tomography: Technology and Applications*. Springer, Berlin, Heidelberg, pp. 47–72, 2008. doi: 10.1007/978-3-540-77550-8_2.

[2] B. L. Frost, "Optical Coherence Tomography Techniques for Contextualizing and Reconstructing Displacement Responses in the Mammalian Cochlea."