Noise and sensitivity in optical coherence tomography based vibrometry

SANGMIN KIM,1 JOHN S. OGHALAI,2 AND BRIAN E. APPLEGATE2,3,* (ID)

1Department of Biomedical Engineering, Texas A&M University, College Station, TX 77843, USA
2Caruso Department of Otolaryngology – Head & Neck Surgery, University of Southern California, Los Angeles, CA 90033, USA
3Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089, USA
*brianapp@usc.edu

Abstract: There is growing interest in using the exquisite phase sensitivity of optical coherence tomography (OCT) to measure the vibratory response in organ systems such as the middle and inner ear. Using frequency domain analysis, it is possible to achieve picometer sensitivity to vibration over a wide frequency band. Here we explore the limits of the frequency domain vibratory sensitivity due to additive noise and consider the implication of phase noise statistics on the estimation of vibratory amplitude and phase. Noise statistics are derived in both the Rayleigh \((s/n = 0)\) and Normal distribution \((s/n > 3)\) limits. These theoretical findings are explored using simulation and verified with experiments using a swept-laser system and a piezo electric element. A metric for sensitivity is proposed based on the 98% confidence interval for the Rayleigh distribution.

© 2019 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

There is growing interest in using optical coherence tomography (OCT) for highly sensitive vibrational measurements in the auditory system. This is motivated by the ability to spatially resolve the vibratory response with subnanometer sensitivity. The detailed functional images generated have been used as a tool to probe fundamental auditory biomechanics and investigated as potential clinical diagnostics. Most commonly, a pure-tone sine wave stimulus is applied and the vibration of the tissue of interest at the stimulus frequency is measured. However, multi-tones and/or broad-band stimuli can also be used to probe non-linear biological phenomenon. These applications and more promise to drive the continued development of this nascent field.

Similar approaches have been published under a variety of names such as phase-sensitive optical coherence tomography for vibrometry [1], optical coherence tomographic vibrography [2], and volumetric optical coherence tomography vibrometry (VOCTV) [3]. They all utilize the interferometric phase of a low-coherence interferometer to extract periodic small-scale motion in the sample of interest. It is the low-coherence counterpart to the well-known technique, Laser Doppler Vibrometry (LDV). LDV has been used for many years in a variety of applications including materials testing [4,5] and biomechanical measurements [6–8]. The primary advantage of OCT based vibrometry is the spatial resolution. An LDV only reports the vibratory response of the dominant reflector; as a consequence in many applications a strongly-reflective object is placed on the object of interest. OCT based vibrometry can spatially resolve the position of the vibrating structure deep within the sample with a resolution of 10s of microns and sensitivity to vibrations with amplitudes under 10 pm [9].

A clear understanding of the fundamental noise statistics is key to interpreting results and pushing the sensitivity limit of this technology for scientific and clinical applications. While various aspects have been investigated and reported in the literature there has not been a detailed and complete derivation with consistent notation and assumptions. The sensitivity for vibrometry
is fundamentally linked to the noise statistics of the OCT signal, hence we start from the spectral interferogram and derive the noise statistics through the entire processing chain.

Before beginning we offer a brief review of the literatures surrounding this topic. Most of the work is not directly linked to OCT based vibrometry, but rather other related techniques that make use of the interferometric phase in OCT. The noise statistics for optical coherence tomography have been derived previously [10–12] to demonstrate the advantages of spectral interferometry over time-domain interferometry for OCT. These were derived for shot-noise, defining the fundamental limit to OCT sensitivity. Shortly thereafter, researchers began to develop techniques which exploited the phase-sensitivity inherent to spectral interferometry. In the context of spectral domain phase microscopy, Choma et al. [13] noted that the displacement sensitivity is proportional to $(2SNR)^{-1/2}$, where SNR is the signal to noise ratio and we have corrected the factor of 2 multiplier. Joo et al. [14] showed a similar result along with experimental verification of the distribution. Both results are for the time-domain phase. The noise statistics in the frequency domain were in part derived by Szkulmowski et al. [15] who noted the Rician distribution that arises for the magnitude of the phase for flow velocity estimation. Many of the relevant statistical arguments we will make have also been addressed in the context of Magnetic Resonance Imaging [16].

In this paper the effect of additive noise on vibrometry was analyzed by considering the noise statistics through the entire processing chain with results for both the time-domain and frequency-domain phase. These are shown to be consistent with both numerical simulation and experiments with a swept-source OCT system.

2. Derivation of noise statistics

We will start by assuming all of the noise is additive since the main noise sources such as shot-noise, thermal noise, and relative intensity noise (RIN) are additive [17–19]. This is slightly more general, but consistent with prior derivations which assumed additive shot-noise. We further assume that the additive noise term, $n_k(k, t)$, is independent, identically distributed (i.i.d) white Gaussian noise with zero mean and a standard deviation of $\sigma_k$. In order to determine the effect of additive noise on the extraction of the vibratory response, it is necessary to understand how the additive noise in the $k$-domain propagates through each mathematical operation, starting with the spectral interferogram $H(k, t)$ and ending with the magnitude, $|\phi(z, f)|$, and phase, $\phi(z, f)$, of the vibratory response. A flow chart showing the processing steps along with the statistical results, e.g. the mean and standard deviation, is shown in Fig. 1. We define signal to noise ratio (SNR) for the OCT system in the conventional way, signal power divided by the variance of the noise, i.e. $SNR = (s/n)^2$ or on a dB scale $SNR = 20\log(s/n)$, where $s$ is the signal amplitude and $n$ is the standard deviation of the noise.

A time series of spectral interferograms, $H(k, t)$ is collected synchronously with sound output. We assume the DC term is subtracted from $H(k, t)$ since $H(k, t)$ is commonly preprocessed to remove it. Considering a single vibrating reflector and neglecting any autocorrelation,

$$H(k, t) = \frac{\rho}{b} S(k) \sqrt{R_R R_S} \cos(2kn \Delta z + 2k_0 n \delta z(t)) + n_k(k, t)$$

(1)

where $k$ is the wavenumber, $n$ is a refractive index, $k_0$ is the center wavenumber, $\rho$ is the detector responsivity, $S(k)$ is the power spectrum of the light source, $R_R$ and $R_S$ are the reflectivities of sample and reference arms of the interferometer, respectively, $\Delta z$ is the pathlength difference between sample and reference arms, and $\delta z(t)$ is the subresolution displacement of the reflector. The additive noise term, $n_k(k, t)$, is independent, i.i.d white Gaussian noise with zero mean and a standard deviation of $\sigma_k$. It is also assumed that its statistical properties do not change over time, i.e. it is a stationary process.

The constant, $b$ takes on the value 1 if both sides of the interferometer are collected and 2 if only one side is detected. In practice most swept-source based systems collect both sides using a
balanced detector which has the added benefit of canceling the majority of the DC component and autocorrelation artifact. The remaining small DC component is typically subtracted in the preprocessing step noted above. Most spectrometer based systems only collect one side. In this case, the autocorrelation artifact remains and a large DC component must be subtracted in preprocessing.
Next we take the inverse Fourier transform of the interferometric signal, Eq. (1), to generate the complex signal, \( h(z, t) \), where \( z \) is the optical pathlength difference, related to the tissue depth \( (\Delta z) \) by \( z = 2n \Delta z \). The magnitude, \( |h(z, t)| \), is the time series of A-scans or line images which are typically plotted against \( \Delta z \). We would commonly take the mean along time and display the resulting line image \( \bar{h}(z) \) as the structural image.

The remaining processing steps are completed identically at each depth, hence for simplicity let us consider the interferometric signal at a single depth. Then, the \( L \) point discrete Fourier transform (DFT) yields,

\[
h(z = 2n \Delta z, t) = \frac{e^{i\phi(z, t)}}{2\pi} \Phi(z, t) + n_z(z = 2n \Delta z, t)
\]

(2)

where \( s(0) = \sum_{k=1}^{N} S(k) \), \( N \) is total number of spectral channels, \( n_i(z, t) \) is the complex noise in \( z \)-domain transformed from \( n_i(k, t) \), and \( i(z, t) \) is ideal \( h(z, t) \) without the complex noise. The noise statistics of the magnitude of Eq. (2) will follow a Rician distribution, as shown previously [15,16]. The shape of the Rician distribution changes rapidly with the relative magnitude of signal over noise. When the \( s/n \) is greater than 3, the distribution is approximately Normal, leading to the mean and standard deviation of \( |h(z, t)| \) shown in Fig. 1. This is the limit in which the OCT SNR is commonly derived, see for instance Eq. (7) in [10]. However when the signal is exactly zero, i.e. \( R_S \) is zero, the distribution \( n_i(z) \) has a Rayleigh distributed magnitude and a random phase with probability distribution that is uniform over \( 2\pi \) radians. The Rayleigh distribution arises because the real and imaginary components of \( n_i(z) \) are independent and normally distributed about 0 with a variance of \( N \sigma^2_k \) [20]. For completeness the mean and standard deviation of \( |h(z, t)| \) in the Rayleigh limit are also shown in Fig. 1. These results can be arrived at by simple variable substitution into Eq. (3) of [16], hence a detailed derivation is omitted.

Small scale vibrations (sub-resolution) of the reflector \( (\delta z) \) induce small changes in the phase of \( h(z) \) over time. The effect of the noise, \( n_i(z) \), on the phase of \( h(z) \), can be illustrated by complex domain analysis as shown in Fig. 2. It can be seen from Fig. 2 that \( n_z(z) \) distorts the ideal phase with the phase noise, \( \phi_{\text{noise}}(z) \), given by

\[
\phi_{\text{noise}}(z) = \tan^{-1}\left( \frac{|n_i(z)| \sin(\phi_{\text{ideal}}(z))}{|i(z)| + |n_i(z)| \cos(\phi_{\text{ideal}}(z))} \right)
\]

(3)

In the limit that \( |h(z, t)| \gg |n_i(z, t)| \) Eq. (3) simplifies to

\[
\phi_{\text{noise}}(z) \approx \frac{|n_i(z)| \sin(\phi_{\text{ideal}}(z))}{|i(z)|}
\]

(4)

Assuming that the intensity of the reflector of interest does not change over time, the mean of the phase noise in the time-domain is exactly 0, i.e.

\[
E[\phi_{\text{noise}}(z, t)] \approx E \left[ \frac{|n_i(z, t)| \sin(\phi_{\text{ideal}}(z, t))}{|i(z, t)|} \right] = E\left[ |n_i(z)| \right] E[\sin(\phi_{\text{ideal}}(z, t))] = 0
\]

(5)

In Eq. (5), expectation operator \( E[\cdot] \) is dropped in \( \cos(\phi_{\text{ideal}}(z, t)) \) and \( \sin(\phi_{\text{ideal}}(z, t)) \) in the last equation since they are deterministic and independent between \( |n_i(z, t)| \), \( \cos(\phi_{\text{ideal}}(z, t)) \) and \( \sin(\phi_{\text{ideal}}(z, t)) \) makes their expectations separable. Since \( n_z(z, t) \) and \( \phi_{\text{ideal}}(z, t) \) are stationary in the statistical
Fig. 2. Graphical phase noise analysis in the complex domain. $|i(z)|$ and $\phi_i(z)$ are the magnitude and the phase of the ideal A-scan without noise. $|n_z(z)|$ and $\phi_n(z)$ are the magnitude and the phase of the noise, $n_z(z)$. $\phi_{n-i}(z)$ is the phase difference between the noise and ideal A-scan.

sense, the time dependence is dropped in calculating expectations of $|n_z(z, t)|$, $\cos(\phi_n(z, t))$, and $\sin(\phi_n(z, t))$. The noise phase is distributed uniformly over $2\pi$ [16], hence expected values of $\cos(\phi_n(z, t))$ and $\sin(\phi_n(z, t))$ are zero, leading to zero mean for the phase noise.

The standard deviation of the phase noise in the time-domain is expressed as

$$\sqrt{E[\phi_{\text{noise}}^2(z, t)]} \approx \sqrt{\frac{E[|n_z^2(z, t)|\sin^2(\phi_{n-i}(z, t))]}{|i(z)|^2}}$$

$$= \sqrt{\frac{E[|n_z^2(z, t)|]-E[|n_z^2(z, t)|]E[\cos(2\phi_n(z, t))-2\phi_{n}(z, t)]}{2|\epsilon(z)|^2}}$$

$$= \frac{\sqrt{E[|n_z^2(z)|]}}{2|\epsilon(z)|^2} = \sqrt{\frac{N\sigma_{\text{r}}^2}{2|\epsilon(z)|^2}} = \frac{1}{\sqrt{2\text{SNR}_z}}$$

where $|i(z)|^2$ and $N\sigma_{\text{r}}^2$ are powers of the signal and noise in z-domain, respectively, and $\text{SNR}_z$ is defined by the ratio of signal power to noise power. The noise power is calculated using values extracted at the region where there is no signal. Since the $n_z(z)$ is Rayleigh distributed in this region, its power is given by $\mu_z^2+\sigma_z^2 = N\sigma_{\text{r}}^2$. The uniformly distributed phase of the noise $\phi_n(z, t)$ causes the cosine term to become zero, leaving the averaged noise power in the numerator. The averaged noise power is easily derived from the second moment of the Rayleigh distribution [16].

As seen in Eq. (5), the phase noise is expressed as the weighted sum of real and imaginary components of $n_z(z)$ which are $n_z(z, t)\cos(\phi_n(z, t))$ and $n_z(z, t)\sin(\phi_n(z, t))$, respectively. Because those real and imaginary components are stationary zero mean uncorrelated white Gaussian random processes, the probability density function (PDF) of the phase noise is also stationary zero mean uncorrelated white Gaussian random process since it is the weighted sum of those processes [21], having a variance of $1/(2\text{SNR}_z)$. These results are in agreement with those of Joo et al. [14] in terms of the distribution derived in the context of spectral domain optical coherence phase microscopy.

The next step is to compute the Fourier transform of the phase $\angle h(z, t) = \phi(z, t)$ along time to yield the magnitude and phase of the vibratory response as a function of depth and frequency. Again, since the same processing is done for each depth, we will explicitly consider the phase at only a single depth, $(z_r)$ i.e. $\phi(z_r, t)$, without losing any generality. The time dependent phase at $z_r$ is then

$$\phi(t) = 2k_0n\delta z_0(t) + n_z(t) = 2k_0nA_{\text{vib}}\cos(2\pi f_{\text{vib}}t + \theta_{\text{vib}}) + n_z(t)$$

(7)
where \( t \) is the time index, \( A_{vib}, f_{vib}, \) and \( \theta_{vib} \) are vibrational amplitude, frequency, and phase, respectively, \( n_i(t) \) is the time domain phase noise, and we have dropped the \( z_t \) to simplify notation.

The time domain phase in Eq. (7) is described with a sinusoidal function to express an ideal vibration [22]. Generally, any sound can be represented by a sum of sinusoidal waves, hence this does not significantly limit the applicability of the derivation. As shown above, \( n_i(t) \) is a stationary white Gaussian noise that has a zero mean and standard deviation of \((2\text{SNR}_z)^{-1/2}\).

The next step is to take the Fourier transform of \( \phi(t) \) to yield the complex vibratory response, \( \phi(f) \) at \( z_t \) as shown in Fig. 1, i.e. \( \tilde{\phi}(\phi(z,t)) \). Assuming \( M \) samples in the time series, the frequency domain vibrational signal is expressed as

\[
\phi(f) = 2k_0n\delta(z(f)) + n_f(f) = Mk_0nA_{vib}\delta(f \pm f_{vib})e^{j\theta_{vib}} + n_f(f)
\]  

(8)

where \( n_f(f) \) is the frequency domain phase noise transformed from \( n_i(t) \), and \( \delta(f \pm f_{vib}) \) is the Dirac delta function in the frequency domain. In this equation, \( M \) is added after the Fourier transform. \( M \) represents total measurement time given a sampling period, so more measurement time is achieved by larger \( M \), resulting in longer vibrational signal. Since Fourier transform compresses a sinusoidal signal into an impulse, a stronger frequency component is obtained with longer vibrational signal. The real and imaginary components of \( n_f(f) \) are also stationary zero mean uncorrelated white Gaussian since they are simply a weighted sum of \( n_i(t) \). Their variances are \( M\sigma_f^2/2 \) [20] where \( \sigma_f^2 \) is the variance of \( n_i(t) \). At the frequency of the stimulus \( f_{vib} \) the vibrational amplitude \( (A_{vib}) \) and \( (\theta_{vib}) \) are given by

\[
\frac{\phi(f_{vib})}{Mk_0n} = A_{vib}e^{j\theta_{vib}} + \frac{n_f(f_{vib})}{Mk_0n} = \left[A_{vib}\cos(\theta_{vib}) + \frac{\text{Re}[n_i(f_{vib})]}{Mk_0n}\right] + j\left[A_{vib}\sin(\theta_{vib}) + \frac{\text{Im}[n_i(f_{vib})]}{Mk_0n}\right]
\]

(9)

where we have scaled the phase by the factor \( Mk_0n \) to convert to nanometers from radians, and \( \text{Re}[n_i(f)] \) and \( \text{Im}[n_i(f)] \) are the real and imaginary component of \( n_i(f) \).

The measured vibrational amplitude, \( A_{\text{meas}} \), is the magnitude of Eq. (9). As can be seen in Eq. (9), the measured amplitude differs from true amplitude, \( A_{vib} \), due to noise, i.e.

\[
A_{\text{noise}} = A_{\text{meas}} - A_{vib}
\]

(10)

where \( A_{\text{noise}} \) is the amplitude detection noise.

The statistical properties of \( A_{\text{noise}} \) depend on the magnitude of \( A_{vib} \). When \( A_{vib} \) is equal to 0, \( A_{\text{meas}} \) is Rayleigh distributed [16], hence \( A_{\text{noise}} \) follows the Rayleigh distribution. The sensitivity of an OCT vibrometry system is typically measured in this limit by observing the noise of a frequency band where there is no vibration. In this case, the root mean square (RMS) value of \( A_{\text{noise}} \) is

\[
\sqrt{E[A_{\text{noise}}^2]} = \sqrt{\mu_f^2 + \sigma_f^2} = \frac{\sigma_f}{k_0n\sqrt{2M}} \sqrt{\pi \frac{4 - \pi}{2}} = \frac{\sigma_f}{k_0n\sqrt{M}}
\]

(11)

where \( \mu_f = \sqrt{\frac{\pi}{2} \frac{\sigma_f}{k_0n\sqrt{2M}}} \) and \( \sigma_f = \sqrt{\frac{4 - \pi}{2} \frac{\sigma_f}{k_0n\sqrt{2M}}} \) are the mean and the standard deviation of \( A_{\text{noise}} \) as shown in the right bottom of Fig. 1. Equation (11) can be used to predict the performance of vibratory measurements given the \( \text{SNR}_z \) in z-domain since \( \sigma_f \) is equal to \((2\text{SNR}_z)^{-1/2}\).

When \( A_{vib} \neq 0 \), \( A_{\text{meas}} \) is Rician distributed, however it can be approximated as a Gaussian distribution in the limit that the signal is significantly larger than the standard deviation of the noise. It has been shown previously that a \( s/n > 3 \) (i.e. \( A_{vib} \geq 3\sigma_f / k_0n\sqrt{2M} \)) was sufficient to allow this approximation [16]. In OCT vibrometry, a measured displacement signal is typically in the range of few to one hundred nanometers [1,3,23,24] while noise standard deviations are commonly less than one nanometer [3,24]. Practically, the majority of measurements will have \( s/n \) that satisfy this criteria. Thus, the distribution of the detected magnitude \( (A_{\text{meas}}) \) in OCT
vibrometry can be considered to be Gaussian whose mean and standard deviation are the square root of \( A_{\text{vib}}^2 + \sigma_f^2/2Mk_0^2n^2 \) and \( \sigma_f^2/2Mk_0^2n^2 \), respectively [16], leading to the result in Eq. (12).

\[
\begin{align*}
\mu_{f_n} &= E[A_{\text{noise}}] = E[A_{\text{mea}} - A_{\text{vib}}] = E[A_{\text{mea}}] - A_{\text{vib}} = \sqrt{A_{\text{vib}}^2 + \frac{\sigma_f^2}{2Mk_0^2n^2}} - A_{\text{vib}} \\
\sigma_{f_n} &= \sqrt{E[(A_{\text{noise}} - E[A_{\text{noise}}])^2]} = \sqrt{E[(A_{\text{mea}} - \mu_{f_n})^2]} = \frac{\sigma_f}{\sqrt{k_0N\sqrt{2M}}} 
\end{align*}
\]

The noise in extracting the phase of the vibration, \( \theta_{\text{vib}} \), in the frequency domain is derived using Eqs. (5) and (6) since the time domain phase noise is additive zero mean white Gaussian, yielding

\[
\begin{align*}
\mu_{\theta_{\text{vib}}} &= E[\theta_{\text{noise}}] = 0 \\
\sigma_{\theta_{\text{vib}}} &= \sqrt{E[\theta_{\text{noise}}^2]} = \sqrt{\frac{1}{2SNR_p}} = \sqrt{\frac{M\sigma_t^2}{2(Mk_0nA_{\text{vib}})^2}} = \sqrt{\frac{\sigma_f^2}{2Mk_0^2n^2A_{\text{vib}}^2}} = \frac{\sigma_f}{A_{\text{vib}}k_0n\sqrt{2M}} 
\end{align*}
\]

where \( \theta_{\text{noise}} \) is the phase detection noise. The phase detection noise is also Gaussian distributed, and its standard deviation is the same as that of \( A_{\text{noise}} \), Eq. (12), except for the \( A_{\text{vib}} \) term in the denominator.

3. Simulation and experimental results

3.1. Simulation results

Next we decided to numerically simulate interferograms of a vibrating reflector, that is, raw data acquired by M-scan that records the motion of a sample. This allowed us, in a very controlled setting where we know for certain that all of the noise is additive, to verify that the equations we have derived hold. The simulation was executed in MATLAB (Mathworks, Inc).

The set of interferograms were modeled using Eq. (1) such that \( \rho S(k)\sqrt{R_RR_K}/b = 1, AZ = 1.713 \) mm, \( k_0 = 2\pi/1310 \) rad. \( \text{nm}^{-1} \), and \( n = 1 \). The wavelength bandwidth and the total number of spectral channels (N) were set to be 100 nm and 2000, respectively. The vibrational signal \( \delta z(t) \) was modeled using Eq. (7) with \( A_{\text{vib}} = 10 \) nm, \( \theta_{\text{vib}} = 30 \) degree. And the vibrational frequency, \( f_{\text{vib}} \), was defined to be the digital frequency, 0.1, i.e. \( f_{\text{vib}} = 0.1 \times A\text{-scan rate} \). The number of points for the vibrational signal (M) was set to be 1000. The vibrational signal was embedded in the modeled interferograms to make one set of M-scan data as shown in Eq. (1). The k-domain noise \( n_k(k, t) \) in Eq. (1) was shaped to be additive Gaussian which has zero mean and standard deviation \( \sigma_k \) ranging from 0 to 0.1, and to be stationary so that each interferogram has the same statistical properties. The change in the standard deviation \( \sigma_k \) was intended to see the effects of additive noise on vibratory measurements.

Figure 3(a) shows one of the modeled interferograms, \( H(k, t) \), in the time-series with a k-domain noise standard deviation of 0.1. Figure 3(b) displays the magnitude response of the fast Fourier transform of that signal, \( |h(z, t)| \), where a reflector is located at index 200 and has a SNR of 40.97 dB calculated from \( 20\log_{10}(s/h) \). From the complex signal the time-domain phase, \( \phi(200, t) \), can be extracted at the location of the reflector to yield the trace in Fig. 3(c). At this stage the sinusoidal vibration of the reflector is apparent. A fast Fourier transform yields the magnitude response of the vibration \( |\phi(200, f)| \) shown in Fig. 3(d). The vibrational amplitude and phase are extracted at the normalized digital frequency of 0.1, \( \phi(200, f = 0.1) \), showing 9.99 nm and 29.86, respectively. Note, the displacement in Fig. 3(d) was placed on a log-scale in order to make the noise more visible. The k-domain noise was generated with the MATLAB function of \text{randn} multiplied by the desired standard deviation to generate Gaussian distributed random numbers with zero mean. For one M-scan data set, the noise was generated 1000 times with the same mean and standard deviation to be added to each modeled interferogram as seen in Eq. (1). These M-scan data were iteratively acquired to calculate statistical properties of experimental
amplitude and phase detection noises. Then the theoretical mean and standard deviation of the amplitude, $|\phi(z,f)|$, were calculated for comparison to simulation results.

First, the amplitude detection noise ($A_{\text{noise}}$) was extracted in the frequency region ranging from 0.25 to 0.5 where $A_{\text{vib}}$ was equal to 0 and its second moment was compared to Eq. (11) as shown in Fig. 4. In this figure, the top row shows the results from 300 M-scans while the bottom one displays from 3000 M-scans where each M-scan has 251 samples for $A_{\text{noise}}$. Hence 75,300 ($251 \times 300$) and 753,000 ($251 \times 3000$) samples were used for the top row and the bottom row, respectively. As seen in Figs. 4(a) and 4(c), the second moment of $A_{\text{noise}}$ is on the rise as the k-domain noise standard deviation ($\sigma_k$) increases, matching the theoretical line clearly. This increase was caused by the rise of the standard deviation of the time domain phase noise ($\sigma_t$) and reduced SNR. Not only the second moment but also the distribution of $A_{\text{noise}}$ was checked by comparing the histogram (blue box) to the fitted Rayleigh distribution (red line) obtained by the MATLAB function of `fitdist` as shown in Figs. 4(b) and 4(d). As would be expected, the histogram more closely matches the fit as the number of samples increases. The mean and standard deviation of the amplitude detection noise ($A_{\text{noise}}$) at $\phi(200,f=0.1)$ was calculated from the modeled M-scans and plotted in Fig. 5 as a function of k-domain noise standard deviation, $\sigma_k$. The bottom x-axis on each plot is $\sigma_k$ with the corresponding SNR plotted on the top x-axis except Figs. 5(c) and 5(f). Note that the SNR is always well above 9.5 dB ($s/n = 3$), so we are in the normal distribution limit. The y-axis is the noise mean in Figs. 5(a) and 5(d) and noise STD in Figs. 5(b) and 5(e). In Figs. 5(c) and 5(f), the x-axis and y-axis are $A_{\text{noise}}$ and its frequency, respectively. The top and bottom rows in the figure differ by the number of M-scans used to generate the statistics, 300 top and 3000 bottom. A qualitative comparison of the top and bottom rows shows the expected trend: as we increase the number of samples used to generate the statistic, the results of the simulation more closely match the theory. The noise mean oscillates above and below the theoretical values. The standard deviation likewise oscillates above and below the theoretical line and clearly follows the same trend, showing the increase in a linear scale along $\sigma_k$. Also, it is shown that mean and standard deviation of amplitude detection noise
increases as expected with an increase in the k-domain noise standard deviation. This is due to an increase in the standard deviation ($\sigma_t$) of the time domain phase noise $n_t(t)$ in Eq. (7) caused by a decrease in SNR in z-domain. The histogram of $A_{\text{noise}}$ was compared to the normal distribution fit in Figs. 5(c) and 5(f) calculated from the data of $\sigma_k = 0.02$, showing that the histogram matches the normal distribution as derived in Eq. (12).

![Image](image-url)

**Fig. 4.** Simulation results for $A_{\text{noise}}$ when $A_{\text{vib}} = 0$. Top row is for 300 simulated M-scans and bottom row for 3000 M-scans. (a, c) are the measured and theoretical second moment as a function of the STD of the noise in the spectral interferogram, $\sigma_k$. (b, d) are histograms (blue) with Rayleigh distribution fits (red line) of frequency domain noise from the data of $\sigma_k = 0.02$.

The phase detection noise ($\theta_{\text{noise}}$) was calculated by subtracting the true vibrational phase ($\theta_{\text{vib}}$) from the measured one, $\theta_{\text{meas}} = \angle \phi(200, f = 0.1)$. Its mean and standard deviation were plotted in Fig. 6 as a function of $\sigma_k$. As seen in the results of $A_{\text{noise}}$ in Fig. 5, the mean and standard deviation of $\theta_{\text{noise}}$ match the theory more closely as the number of used samples increases. Both mean and standard deviation show the same trend while fluctuating slightly about the theory. However, the mean oscillates about 0 while the standard deviation fluctuates about the theoretical line. This shows the same trend as the amplitude detection noise because time domain phase noise directly affects the standard deviation while it has no effect on the mean as seen in Eq. (13). As seen in the amplitude detection noise plots, the increase in the standard deviation of $\theta_{\text{noise}}$ results from the decrease in SNR since this brings about rise up of $\sigma_t$. The distribution of $\theta_{\text{noise}}$ was verified by comparing the histogram acquired from the data of $\sigma_k = 0.02$ to the normal distribution fit as seen in Figs. 6(c) and 6(f). It is shown from these plots that the histogram of $\theta_{\text{noise}}$ follows the shape of the fitted distribution as derived in Eq. (13). Therefore, these MATLAB simulation results give us confidence that the derived equations are correct within the assumptions of additive noise and the statistical stationary state.
Fig. 5. Simulation results for noise in the measure of $A_{\text{vib}}$. Top row is for 300 simulated M-scans and bottom row for 3000 M-scans. (a, d) are the measured and theoretical mean ($\mu_{fR}$) as a function of the STD of the noise in the spectral interferogram, $\sigma_k$. (b, e) are the measured and theoretical STD ($\sigma_{fR}$) as a function of the STD of the noise in the spectral interferogram, $\sigma_k$. (c, f) are histograms (blue) with normal distribution fits (red line) of $A_{\text{noise}}$ from the data of $\sigma_k = 0.02$.

3.2. Experimental results

In order to verify the derived equations under normal experimental conditions, the movement of a piezo electric element was measured with a swept laser system as described in [25] where a stereo-microscope was used instead of the endoscope. The piezo was driven sinusoidally at 4 kHz and measured with A-scan (laser sweep) rate of 128.04 kHz over 50 ms, acquiring 6403 A-scans per M-scan. Two neutral density (ND) filters were used to change the incident light power on the piezo, allowing us to acquire A-scans with 3 different SNRs such as 56.33 dB, 49.25 dB, and 18.09 dB. True vibrational amplitude and phase were estimated by averaging 100 sets of M-scan data with the highest A-scan SNR (56.33 dB) which are 2.788 nm and $-144.556^\circ$, respectively.

Figure 7 shows averaged A-scans and frequency domain vibrational responses from 100 M-scans measured with the piezo. And 1st, 2nd, and 3rd columns of Fig. 7 display results from 3 different SNRs such as 56.33 dB, 49.25 dB, and 18.09 dB respectively. Figures 7(d)–7(f) show frequency domain vibrational responses calculated from A-scans seen in Figs. 7(a)–7(c). To obtain frequency domain vibrational responses, time domain vibrational signals were extracted at peak locations red-arrowed in Figs. 7(a)–7(c). The unwrapped phase was fit to a sixth-order polynomial and subtracted in order to reduce low frequency drift. Those subtracted signals were fast Fourier transformed after applying the Hann window for the frequency domain representation. As can be seen in Fig. 7, the more attenuation, the more the peak intensity of the A-scan reduces. This is accompanied by an increase in optical pathlength due to the glass thickness of the ND filters. Multiple reflections in the ND filter also led to a few extra peaks, particularly apparent in the lowest SNR signal. The second moment of the frequency domain noise was acquired for each case with magnitudes ranging from 6 kHz to 10 kHz extracted along 100 M-scan data set by...
Fig. 6. Simulation results for noise in the measure of $\theta_{\text{vib}}$. Top row is for 300 simulated M-scans and bottom row for 3000 M-scans. (a, d) are the measured and theoretical mean ($\mu_{\theta_{\text{vib}}}$) as a function of the STD of the noise in the spectral interferogram, $\sigma_k$. (b, e) are the measured and theoretical STD ($\sigma_{\theta_{\text{vib}}}$) as a function of the STD of the noise in the spectral interferogram, $\sigma_k$. (c, f) are histograms (blue) with normal distribution fits (red line) of $\theta_{\text{noise}}$ from the data of $\sigma_k = 0.02$

calculating $\sqrt{\mu^2 + \sigma^2}$ with their mean and standard deviation. As shown in Figs. 7(d)–7(f), the predicted RMS noise increases, 6.40 pm, 9.89 pm, and 288.44 pm, as the SNR decrease.

As a first step, frequency domain noise (or $A_{\text{noise}}$ when $A_{\text{vib}} = 0$) was analyzed as shown in Fig. 8 to verify Eq. (11). In this figure, the second moments of $A_{\text{noise}}$ were calculated for each SNR using data from different frequency regions such as $6\sim10$ kHz, $22\sim32$ kHz, and $50\sim60$ kHz. Those frequency regions were selected to see how $A_{\text{noise}}$ changes with frequency since low frequency noise in Figs. 7(d) and 7(e), is clearly not Gaussian white. Above 10 kHz (not shown) the noise appears to converge toward Gaussian white.

The experimental second moments were compared in Fig. 8(a) to the theoretical ones calculated from Eq. (11), following the trend of the theory. To investigate the difference between experimental and theoretical results in more detail, percent errors were computed using $|\text{Experimental second moment} - \text{Theoretical second moment}|/\text{Theoretical second moment} \times 100$ and displayed in Fig. 8(b). In this plot, the second moments from the data with an SNR of 18.09 dB show similar errors between the three frequency ranges, as might be anticipated from the plot in Fig. 7(f). However, the two data sets with higher SNR shows larger error as the frequency decreases. At relatively low SNR or higher SNR and frequencies above $\sim10$ kHz the assumption of Gaussian white noise appears to hold. Below the 10 kHz threshold a $\sim1/f$ noise component dominates leading to poorer agreement with the theoretical values. Nevertheless, as seen in Figs. 8(c)–8(e) the histograms for the highest SNR of 56.33 dB are well fit by a Rayleigh distribution. The theoretical distribution matches the experimental distribution at all frequencies and converges toward the theoretical second moment as characterized by the width of the distributions as the frequency increases.
Fig. 7. Averaged A-scans and frequency domain vibrational responses measured from 100 M-scan with a piezo electric element driven by 4 kHz. (a, d) are A-scan and 100 frequency domain vibrational responses, respectively when no ND filter is used while (b, e) and (c, f) are from the use of 1 ND and 2 ND filters, respectively. In (d, e, f), thick black line shows the RMS value of the frequency domain noise.

Fig. 8. Second moments and histograms of amplitude detection noise extracted in the frequency region where $A_{vib} = 0$. (a, b) show theoretical and experimental second moments of $A_{noise}$ and their percentage errors, respectively, where the second moments were measured in three frequency regions, 6~10 kHz, 22~32 kHz, and 50~60 kHz. (c, d, e) display histograms (blue) with Rayleigh distribution fits (red line) of the data with a SNR of 56.33 dB acquired from 6~10 kHz, 22~32 kHz, and 50~60 kHz, respectively.
Experimental mean and standard deviation of amplitude and phase detection noises are calculated from the 3 sets of 100 M-scan data and compared to theoretical ones as displayed in Fig. 9. In this figure, experimental results follow the trend of theoretical ones, showing slightly larger values than them. Also, histograms of measured $A_{\text{noise}}$ and $\theta_{\text{noise}}$ from the data with a SNR of 56.33 dB are compared to their normal distribution fits in Figs. 9(c) and 9(f), showing again that they follow the correct distribution.

![Fig. 9. Mean and standard deviation of amplitude and phase detection noises in accordance with A-scan SNR. (a, b) show in picometer scale mean and standard deviation of amplitude detection noise while (d, e) display in degree those of phase detection noise. Blue lines and red circles represent theoretical results calculated using Eqs. (12) and (13), and experimentally measured ones, respectively. (c, f) show histograms (blue box) with normal distribution fits (red line) of measured amplitude and phase detection noises from the data with a SNR of 56.33 dB.](image)

To investigate the difference between theoretical and experimental results in Fig. 9 in more detail, the percentage errors of $A_{\text{noise}}$ and $\theta_{\text{noise}}$ were calculated using $\frac{|\text{Theoretical RMS} - \text{Experimental RMS}|}{\text{Theoretical RMS}} \times 100$ and displayed in Fig. 10. In this figure, the second moments ($\sqrt{\mu^2 + \sigma^2}$) were used to consider the mean and the standard deviation simultaneously. It is shown in Fig. 10 that the percentage errors increase with increasing SNR, meaning that experimental results deviate more from theoretical ones. In this experiment, theoretical results were computed using the measured SNR in z-domain and Eqs. (6), (12), and (13). Since the SNR was determined with the magnitudes of the A-scans, any noise source that only impacts the phase would not be accounted for. Hence, other noise sources, which may not be additive white noise, were added to interferometric phases. For this reason, as seen in Figs. 7(d) and 7(e), clearly the total noise is not white, but rather more heavily weighted in the low frequency range. On the other hand, the measurement with the lowest SNR exhibits a noise profile that qualitatively looks white as can be seen in Fig. 7(f). Based on this observation, additive white noise in the k-domain is dominant when SNR is low whereas other non-white noise contributes significantly at higher SNRs. This results in greater deviation from the theoretical predictions for the higher SNR experiments and agrees with the finding from Fig. 8.
Fig. 10. RMS errors of amplitude and phase detection noises as function of A-scan
SNR. (a) and (b) show percentage errors of amplitude detection noise and phase detection
noise, respectively. RMS error was calculated by \(|\text{Theoretical RMS} - \text{Experimental RMS}| / \text{Theoretical RMS} \times 100\) where RMS was computed by the square root of mean
squared plus standard deviation squared.

4. Discussion

The sensitivity of OCT for vibrometry is clearly tied to the signal to noise ratio of the OCT system,
i.e. \(\sigma_r = (2\text{SNR})^{-1/2}\). The ultimate limit to the OCT system SNR is shot-noise, \(\sigma_k = (2\rho R_S B)^{1/2}\),
where B is the noise equivalent bandwidth and S is the source average power. Using the magnitude
of Eq. (2) and recognizing that \(s(0) = NS\) for an ideal flat-top laser power profile, the SNR can be
shown to be, \(\left\langle |i(z)|^2/E|n(z)|^2 \right\rangle^2 = N\rho SR_s / (4b^2B)\). This is equivalent to the result in [10] for \(b = 1\).
Assuming an anti-aliasing filter has been used, then B is half of the sampling frequency and the
SNR can be rewritten in terms of the sweep frequency \(f_{\text{sweep}}\) and duty cycle \(D\) or sweep time \(\Delta t\),
\[
\text{SNR} = \frac{\rho SR_s D}{2b^2 f_{\text{sweep}}} = \frac{\rho SR_s}{2b^2} \Delta t
\]
(14)

A similar equation can be derived for spectrometer based systems [10–12] where \(\Delta t\) is the camera
integration time. The SNR of the OCT system goes up with higher power on the sample \((S)\) and
slower sweep rate (higher integration time, \(\Delta t\)), hence the sensitivity of the vibrometer improves
as well. For every 10 dB improvement in the system SNR, the time-domain phase noise, \(n(t)\), is
reduced by \(1/\sqrt{10}\) for a \(\sim 3\) fold improvement. Equivalently, for a 3-fold improvement in \(s/n\), the
phase noise is reduced by \(1/3\).

While we have focused on revealing the statistical properties of the signals in OCT based
vibrometry we have not directly addressed sensitivity. In our work [3,26,27] we typically use
\(\mu_{f_k} + 3\sigma_{f_k}\) as a threshold. In other words, any signal below this threshold is neglected. This
metric has the advantage that it can readily be measured in each acquisition by considering the
mean and standard deviation in a frequency region near the stimulus frequency where there is no
stimulus and it corresponds to 98% confidence interval of the Rayleigh distribution. In terms of
the relationships of Fig. 1 the sensitivity is,
\[
sens = \frac{\mu_{f_k} + 3\sigma_{f_k}}{M \cdot \text{SNR}^{-1/2}} = \left(\frac{\sqrt{\pi} + 3\sqrt{2} - \pi}{2\sqrt{2}k_0n}\right) (M \cdot \text{SNR})^{-1/2}
\]
(15)

where \(M\), the number of time samples, and SNR are the only variables for a given system. For
every 100 fold increase in \(M\), the sensitivity is improved by a factor of 10. Likewise a 20 dB
improvement in SNR, improves the sensitivity by a factor of 10. If we consider the shot noise
limit, Eq. (14) can be substituted into Eq. (15) to get

\[
sens = \left(\sqrt{\pi} + \frac{3\sqrt{4-\pi}}{2\sqrt{2k_0n}}\right) \sqrt{\frac{2b^2}{\rho SR_\Delta tM}}
\]  

Selectively attenuating the reference arm power in the OCT interferometer reduces RIN [17] allowing the system to approach the shot-noise limit. Fundamentally, an increase in the sample reflectivity \((R_S)\) or the total acquisition time \((\Delta tM)\) improves the system sensitivity. Nominally, while it is not practical to increase the sample reflectivity for \(in vivo\) experiments, this reinforces the idea that one should select the brightest point in a region of interest in order to measure the vibratory response with the best sensitivity. Also, it can be helpful to acquire multiple M-scans at the same location to average the time domain phase signal in Eq. (7) before computing the Fourier transform. This will reduce uncorrelated noise at a rate of \(\sqrt{m}\), where \(m\) is the number of traces averaged. There is also a fundamental advantage to using a shorter wavelength of light (larger \(k_0\)) for vibrometry. Obviously, that advantage is offset by the larger scattering cross-section at shorter wavelengths which will result in lower SNR at depth in the tissue.

OCT systems are commonly specified by the signal to noise ratio expected for a perfect reflector \((R_s = 1)\) on a decibel scale. For instance, the maximum SNR of an OCT system might be reported as 107 dB, with a corresponding theoretical shot-noise limit of 110 dB. Developing a similar metric for vibrometry performance would be helpful in comparing systems with different architectures and newly developed algorithms for extracting the vibratory response. Assuming \(n = M = 1\) and using Eqs. (15) and (16), this system has a potential sensitivity of 8.4 pm based on measured SNR and 5.9 pm in the shot-noise limit. Tissue reflectivity is typically in the range of \(-40\) to \(-80\) dB, hence in order to realize the 8.4 pm sensitivity \(M\) would need to be set in the range \(10^4\)–\(10^8\). For our 128.04 kHz system, that corresponds to an acquisition time of 78 ms – 13 min.

Small changes in the optical pathlength due to a wide range of sources, e.g. air-currents and building vibrations, introduce phase-noise to the system. These tend to be concentrated in the low frequency range, producing an \(\sim 1/f\) dependence. Obviously, this deviates from our assumption of Gaussian white noise, however at some higher frequency, the mean and standard deviation will become approximately constant with increasing frequency. Beyond this limit our experimental measures match well with the derived equations because the noise is approximately Gaussian white. It is possible to suppress this low frequency phase noise using an interferometer where the reference and sample arm largely co-propagate so that the small scale vibrations are common to both, i.e. a common mode interferometer.

Larger scale changes in the optical pathlength can be introduced by patient/sample motion during data acquisition. Again, this will introduce phase-noise and could violate the condition of a stationary state utilized in the derivation of noise statistics. As we have noted elsewhere [9], under these circumstances, a broad band increase in phase-noise can be observed, however time-domain averaging of the interferometric phase substantially lowers the phase-noise since the motion artifact is incoherent.

A window function \((w)\) is typically multiplied with the time-domain phase before taking the Fourier transform, i.e. \(w\cdot\phi(z, t)\), in order to suppress sidelobes in the frequency domain. Strong sidelobes can obscure weaker signals and produce errors in measured vibrational amplitude and phase. The window has an impact on statistical the properties of amplitude and phase detection noises since it changes intensities of signal and noise in the time domain. The equations in Fig. 1 were derived without using a window, which is equivalent to assuming a rectangular window. More generally we would want to choose a window function based on our tolerance for sidelobe intensities and need for frequency resolution. The frequency domain equations in Fig. 1 can be generalized by replacing \(M\) with \(\frac{W_0^2(t)}{E_w}\), where \(E_w\) is the total energy of the window.
function calculated from $E_w = \sum_{i=1}^{M} (w(t_i))^2$ and $W(0)$ is the DC component of $W(f)$ calculated from $W(0) = \sum_{i=1}^{M} w(t_i)$.

For example, consider Eq. (11) with an arbitrary window and then assuming the Hann window

$$\sqrt{\mu^2_{fr} + \sigma^2_{fr}} = \sqrt{E[A^2_{\text{noise}}]} = \sqrt{\frac{E_{w}\sigma^2_f}{k_0^2n^2W^2(0)}} = \frac{\sigma_f\sqrt{E_w}}{k_0nW(0)} \tag{17}$$

$$= \frac{1.225\sigma_f}{k_0n\sqrt{M}} \text{ for Hann window} \tag{18}$$

Equation (17) shows the second moment of $A_{\text{noise}}$ for general case while Eq. (18) displays when the Hann window is used. Compared to Eq. (11), the second moment of using the Hann window increases by 1.225 times because the power of the interesting phase signal (2$k_0n\delta x(t)$) is reduced more than that of the time domain phase noise ($\eta_f(t)$) by the Hann window. Statistical properties of $A_{\text{noise}}$ and $\theta_{\text{noise}}$ are easily converted by replacing $M$ with $W^2(0)/E_w$.

It is shown from the derivation that the time domain phase noise ($\eta_f(t)$) in Eq. (7) is white Gaussian and the magnitude of the frequency domain noise ($\eta_f(f)$) is Rayleigh distributed when $A_{\text{vib}}$ is equal to 0. Since the noise power is preserved between these domains, their second moments have the relationship of $E[\eta_f^2(t)] = 1/M*E[\eta_f^2(f)]$, meaning that $E[\eta_f^2(f)]$ can be predicted by calculating $E[\eta_f^2(t)]$. This relationship assumes that other noise sources that induce small changes in the optical pathlength are not included. Thus, it is important to process data to reduce the effects of other noise sources not only to have the consistent relationship between $n_f(t)$ and $n_f(f)$ but also to increase measurement accuracy.

The preprocessing (second step in Fig. 1) typically done is the subtraction of a background signal to remove the constant term, which will appear at $z = 0$. The side-lobes from a strong signal at $z = 0$ can sometimes overlap with signals of interest. We typically measure the background as the mean over a large number of acquisitions and bandpass filter. It is essentially noise free and therefore does not appreciably contribute to the noise. We also typically preprocess the time-domain phase, $\phi(z, t)$, before calculating the Fourier transform to get $\phi(z, f)$. At a minimum we calculate and then subtract the mean of $\phi(z, t)$, to suppress the signal at 0 Hz. This approach is fast and does not contribute to the noise above 0 Hz. In post-processing where computational time is not critical, we sometimes will instead fit $\phi(z, t)$ to a polynomial and subtract the fit from $\phi(z, t)$. This suppresses the signal at 0 Hz and removes some of the low frequency drift again without introducing noise. Some researchers have taken the derivative by computing $\Delta \phi(z, t) = \phi(z, t_2) - \phi(z, t_1)$, similar to what is done for Doppler flow measurements. This also removes the 0 Hz portion of the signal, however it introduces an additional $\sqrt{2}$ noise so that $\sigma_f = (4SNR)^{-1/2}$.

5. Conclusions

In this article, additive noise is analyzed to investigate how it affects the measurement of vibratory amplitude and phase in the frequency domain. The statistical properties were derived in the limit of 0 vibratory signal and for vibratory signals with $s/n > 3$ and shown to follow a Rayleigh and Gaussian distribution, respectively. The derived equations were verified by MATLAB simulation and experimentally with a vibrating piezo electric element measured with a swept-source OCT system. The results show that derived equations are accurate within the assumption of additive Gaussian white noise, thus providing a theoretical tool to predict the performance of an OCT vibrometry system. These equations also provide a theoretical framework to investigate alternate signal processing algorithms to improve the measurement sensitivity. A measure of sensitivity is proposed that corresponds to the 98% confidence interval for the Rayleigh distribution. The adoption of this metric would make it easier to compare different systems and algorithms.
Funding
National Institute on Deafness and Other Communication Disorders (R01DC013774, R01DC014450); National Institute of Biomedical Imaging and Bioengineering (R01EB027113).

Acknowledgements
We gratefully acknowledge critical reading of this manuscript by Amir Nankali, James Dewey, and Michael Choma.

Disclosures
The authors declare no conflicts of interest.

References
1. H. M. Subhash, N. H. Anh, R. K. K. Wang, S. L. Jacques, N. Choudhury, and A. L. Nuttall, “Feasibility of spectral-domain phase-sensitive optical coherence tomography for middle ear vibrometry,” J. Biomed. Opt. 17(6), 060505 (2012).
2. E. W. Chang, J. B. Kobler, and S. H. Yun, “Subnanometer optical coherence tomographic vibrography,” Opt. Lett. 37(17), 3678–3680 (2012).
3. H. Y. Lee, P. D. Raphael, J. Park, A. K. Ellerbee, B. E. Applegate, and J. S. Oghalai, “Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea,” P. Natl. Acad. Sci. USA 112, 3128–3133 (2015).
4. L. Mallet, B. C. Lee, W. J. Staszewski, and F. Scarpa, “Structural health monitoring using scanning laser vibrometry: II. Lamb waves for damage detection,” Smart Mater. Struct. 13(2), 261–269 (2004).
5. Y. C. Lin, H. Hocheng, W. L. Fang, and R. Chen, “Fabrication and fatigue testing of an electrostatically driven microcantilever beam,” Mater. Manuf. Processes 21(1), 75–80 (2006).
6. S. S. Narayan, A. N. Temchin, A. Recio, and M. A. Ruggiero, “Frequency tuning of basilar membrane and auditory nerve fibers in the same cochlea,” Science 282(5395), 1882–1884 (1998).
7. W. X. He, A. Fridberger, E. Porsov, K. Grosh, and T. Y. Ren, “Reverse wave propagation in the cochlea,” Proc. Natl. Acad. Sci. U. S. A. 105(7), 2729–2733 (2008).
8. T. Y. Ren, W. X. He, and P. G. Gillespie, “Measurement of cochlear power gain in the sensitive gerbil ear,” Nat. Commun. 2(1), 216 (2011).
9. W. Kim, S. Kim, S. Huang, J. S. Oghalai, and B. E. Applegate, “Picometer scale vibrometry in the human middle ear using a surgical microscope based Optical Coherence Tomography and Vibrometry system,” Biomed. Opt. Express 10(9), 4395–4410 (2019).
10. M. A. Choma, M. V. Sarunic, C. H. Yang, and J. A. Izatt, “Sensitivity advantage of swept source and Fourier domain optical coherence tomography,” Opt. Express 11(18), 2183–2189 (2003).
11. R. Leitgeb, C. K. Hitzenberger, and A. F. Fercher, “Performance of fourier domain vs. time domain optical coherence tomography,” Opt. Express 11(8), 889–894 (2003).
12. J. F. de Boer, B. Cense, B. H. Park, M. C. Pierce, G. J. Tearney, and B. E. Bouma, “Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography,” Opt. Lett. 28(21), 2067–2069 (2003).
13. M. A. Choma, A. K. Ellerbee, C. H. Yang, T. L. Creazzo, and J. A. Izatt, “Spectral-domain phase microscopy,” Opt. Lett. 30(10), 1162–1164 (2005).
14. C. Ioo, T. Akkin, B. Cense, B. H. Park, and J. E. de Boer, “Spectral-domain optical coherence phase microscopy for quantitative phase-contrast imaging,” Opt. Lett. 30(16), 2131–2133 (2005).
15. M. Szkulmowski, A. Szkulmowska, T. Bajraszewski, A. Kowalczyk, and M. Wojtkowski, “Flow velocity estimation using joint Spectral and Time Domain Optical Coherence Tomography,” Opt. Express 16(9), 6008–6025 (2008).
16. H. Gudbjartsson and S. Patz, “The Rician Distribution of Noisy MRI Data,” Magn. Reson. Med. 34(6), 910–914 (1995).
17. W. V. Sorin and D. M. Baney, “A Simple Intensity Noise-Reduction Technique for Optical Low-Coherence Reflectometry,” IEEE Photonics Technol. Lett. 4(12), 1404–1406 (1992).
18. B. M. Hoeling, A. D. Fernandez, R. C. Haskell, E. Huang, W. R. Myers, D. C. Petersen, S. E. Ungererma, R. Y. Wang, M. E. Williams, and S. E. Fraser, “An optical coherence microscope for 3-dimensional imaging in developmental biology,” Opt. Express 6(7), 136–146 (2000).
19. P. R. Morkel, R. I. Laming, and D. N. Payne, “Noise Characteristics of High-Power Doped-Fiber Superluminescent Sources,” Electron. Lett. 26(2), 96–98 (1990).
20. M. A. Richards, “The discrete-time Fourier transform and discrete Fourier transform of windowed stationary white noise,” Georgia Institute of Technology, Tech. Rep (2013).
21. A. Papoulis and S. U. Pillai, Probability, random variables, and stochastic processes (Tata McGraw-Hill Education, 2002).
22. D. J. Inman, Vibration with control (Wiley, 2006).
23. E. W. Chang, J. T. Cheng, C. Roosli, J. B. Kobler, J. J. Rosowski, and S. H. Yun, “Simultaneous 3D imaging of sound-induced motions of the tympanic membrane and middle ear ossicles,” *Hear. Res.* **304**, 49–56 (2013).

24. R. L. Warren, S. Ramamoorthy, N. Ciganovic, Y. Zhang, T. M. Wilson, T. Petrie, R. K. K. Wang, S. L. Jacques, T. Reichenbach, A. L. Nuttall, and A. Fridberger, “Minimal basilar membrane motion in low-frequency hearing,” *Proc. Natl. Acad. Sci. U. S. A.* **113**(30), E4304–E4310 (2016).

25. W. Kim, S. Kim, J. S. Oghalai, and B. E. Applegate, “Endoscopic optical coherence tomography enables morphological and subnanometer vibratory imaging of the porcine cochlea through the round window,” *Opt. Lett.* **43**(9), 1966–1969 (2018).

26. H. Y. Lee, P. D. Raphael, A. P. Xia, J. Kim, N. Grillet, B. E. Applegate, A. K. E. Bowden, and J. S. Oghalai, “Two-Dimensional Cochlear Micromechanics Measured In Vivo Demonstrate Radial Tuning within the Mouse Organ of Corti,” *J. Neurosci.* **36**(31), 8160–8173 (2016).

27. A. P. Xia, X. F. Liu, P. D. Raphael, B. E. Applegate, and J. S. Oghalai, “Hair cell force generation does not amplify or tune vibrations within the chicken basilar papilla,” *Nat. Commun.* **7**(1), 13133 (2016).