THE FRACTIONAL BISPECTRUM NOISE REDUCTION TECHNIQUE AND APPLICATIONS IN INTERFEROMETRY

by

Jonathan Babaie

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Optical Science and Engineering

Charlotte

2019

Approved by:

_____
Dr. Faramarz Farahi

_____
Dr. Angela Davies Allen

_____
Dr. Stuart Smith

_____
Dr. Alireza Tabarraei

ABSTRACT

JONATHAN BABAK BABAIE. The Fractional BiSpectrum
Noise Reduction Technique and Applications in Interferometry.
(Under the direction of DR. FARAMARZ FARAHI)

In the field of optical metrology, there are many existing techniques for measurement of the surface profile or film thickness of a sample of interest. However, many of these common techniques break down when there exists a high level of noise in the output optical signal due to the nature of certain types of samples. Sources of noise that can suppress the signal information can come from highly rough surfaces and turbulence within the medium if there is necessity for light to transmit through. This dissertation presents the results of a research effort to develop a technique to suppress the noise content in order to distinguish the signal. In particular, a noise-reduction technique, Fractional BiSpectrum, was developed for the purpose of properly identifying signal frequencies within the measurement data when these frequencies have a known relationship with each other. It is also shown how this technique can be implemented within common interferometry setups such as those measuring the thickness of films with turbid material as well as the surface profile and/or roughness of a sample and still make the proper measurement, i.e. identify the proper temporal or spatial frequency within a noisy signal. It is demonstrated that using this noise-reduction technique enables to produce a confident measurement – i.e. enhance SNR greater than 3dB. In some of the demonstrated cases, enhancement will lead to the detection of the correct measurement parameter – i.e. thickness or height value - where the existing techniques fail.

ACKNOWLEDGMENTS

It should firstly be stated that there were many times through to the completion of my research that due to sources of distractions my advisor pushed me to continue and ultimately complete my work. There is a high level of respect and appreciation for all the advice and support that was given to me from my advisor, Dr. Faramarz Farahi.

There were also many obstacles that through collaboration with my long-time colleague, Dr. Mehrdad Abolbashari, I was able to overcome. For this I am also very grateful.

I must also mention my colleagues within the lab that I would see on a normal basis to discuss the daily matters. Some of these people include Dr. Navid Farahi, Dr. Gelareh Babaie, and Dr. Nasim Habibi. I am thankful to be surrounded by these insightful people with unique perspectives and considerations for how to solve problems.

It goes without question that there were many faculty members within the optical science department in UNCC and those within the Center for Precision Metrology group that provided support, advice, and equipment that without, the completion of this research would not be possible.

And lastly, on a more personal level, I could not have finished my research without the constant support from my wife, parents, and the rest of my family that continuously gave me the strength to finish my work even with the many other distractions that existed in my life at the same time. To them I am very thankful.

DEDICATION

This dissertation is dedicated to my family members who have exposed me to all of nature's elements in life:

To my mom, who has been the solid ground under my feet, giving me the proper stability and calmness to approach each day.

To my dad, who has been the gust of wind behind my back, pushing me to achieve all that I can.

To my wife, who is my steady lit fire, who without, I could not traverse the steepest mountains in the darkest of nights.

And to my little Javanni, who is my perfect little rainbow, helping me to experience new wonders and beauty in life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

FT	Fourier transform
FBS	Fractional BiSpectrum
FBSS	Fractional BiSpectrum Separated
SNR	Signal-to-Noise Ratio
PDF	Probability Density Function
FOV	Field of View
WSI	Wavelength Scanning Interferometry
PSI	Phase-Shifting/Scanning Interferometry
OCT	Optical Coherence Tomography
CSI	Coherence Scanning Interferometry
WLI	White-Light Interferometry
CMM	Coordinate-Measuring-Machine
STM	Scanning Tunneling Microscopy
AFM	Atomic Force Microscopy

CHAPTER 1: INTRODUCTION

Measurement of surface profile and film thickness of a sample which contains a high degree of roughness, i.e. greater than 300 nm, or high turbidity within the medium, can be very challenging using optical methods. Due to the roughness of the surface, the light used for measurement scatters and therefore a lot of information is lost to the detector. In the case of a measurement using laser light, a speckle pattern is generated. This speckle pattern produces localized high and low intensity points that tend to dominate an interferometry signal as shown in Figure 1-1 [1,2].

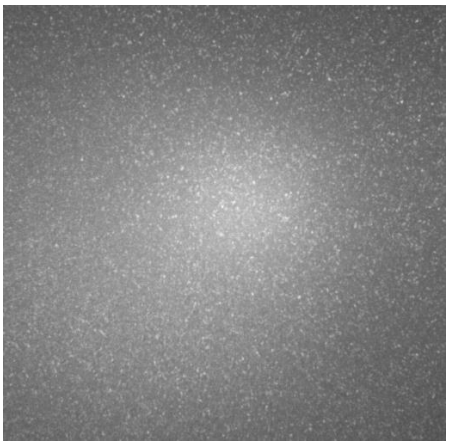

Figure 1-1: This demonstrates a typical speckle pattern generated from laser light scattered off of a rough surface and / or transmitted through a turbid medium.

When this speckle noise becomes comparable to the signal, i.e. the noise power it produces is within 3dB of the signal power, it becomes difficult to extract the phase or amplitude information from the sample under measurement. While techniques such as the Fourier transform (FT) statistically spread random noise based on the frequency content of the data, if the noise threshold is too great with respect to the signal frequency information, this technique will reach to its limit of applicability. Thus, it can be necessary to further

enhance the signal with respect to the noise power to provide a measurement with higher confidence, or in some cases, provide a correct measurement where use of the FT would commonly fail. Thus, a signal-enhancement technique was developed for this purpose. Hence, the goal of this research is not to outperform the typical metrology tools used in industry where the tool is already capable of measuring; the goal is instead to provide the ability of extending the measurement region or confidence where previously noise embedded within the signal made it too difficult to obtain a reliable measurement. This is achieved by enhancement of the SNR.

1.1 Literature Review

Since the scope of this research includes the measurement of thickness of film and surface profile by implementation of the newly introduced noise-enhancement technique, a discussion on the current metrology techniques will be explored for both genres. A brief discussion on some of the limitations of each technique will be demonstrated at the end of each respective section. A table will summarize some of the limitations specific to surface roughness and or medium turbidity in Section 1.1.3.

1.1.1 Film Thickness Measurement

There are many types of techniques to measure the thickness of a film or coating. Thus, only the most popular metrology techniques that are applicable for semi-transparent materials (with respect to visible and near-infrared wavelengths) will be analyzed. Based on the measurement thickness range of the demonstrated interferometry technique in Chapter 2 (10 μm to 5 mm), some of metrology techniques explored will be more viable

than others. Thus, the metrology techniques included in this section are Confocal Microscopy, Ellipsometry, Reflectometry, and Coherence Scanning Interferometry. [3,4]

Confocal Microscopy

Confocal Microscopy is performed by moving the sample under test in the Z direction until it comes into focus. However, the sample is illuminated such that only a small area of the sample is detected. Thus, the measurement is only performed over that area. The typical method for detecting the focus is by adding a pinhole in front of the detector as shown in Figure 1-2.

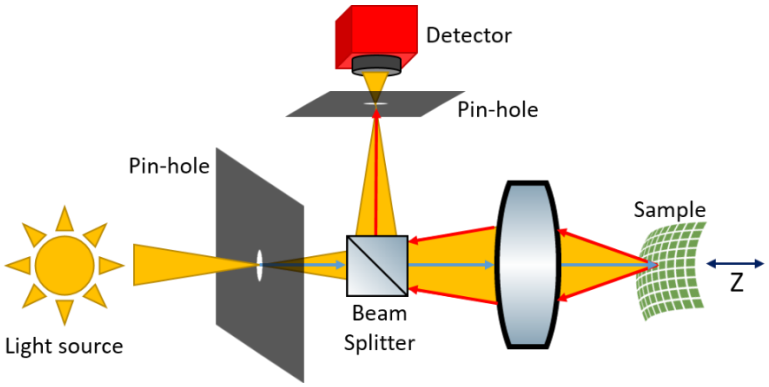

Figure 1-2: This demonstrates a schematic for a typical confocal microscopy measurement.

As is seen in Figure 1-2, if the point on the sample is in focus, it will correspond to the maximum amount of light being reflected back to the detector. In this way for a typical profile measurement, it is necessary to provide a lateral scan over the surface of the sample to find the heights at each location. For thickness measurement, the light must penetrate through the sample and reflect from the bottom surface. Each point on the surface must scan in Z to ensure all focus points have been discovered.

Some of the advantages of this technique are the accuracy in X, Y, and Z for samples which may be normally difficult to measure optically due to under or over saturated spots. Identification of the focal plane requires very little computational effort as only a one-dimensional signal comparing height with intensity is analyzed. Lateral resolution can go as low as 0.2 micrometers, while height resolution as low as 0.4 micrometers which relates to a pinhole diameter of 1 mm.

This measurement, however, is quite time consuming if one is interested in high spatial resolution across the sample as compared to its field of view. The height resolution can also be superfluous as scanning a thick film – i.e. greater than a mm - will require a high number of steps in Z direction. It should be mentioned that other confocal techniques exist that do not scan in Z such as 'Chromatic Confocal,' but this will not be covered in this dissertation. This type of measurement is also susceptible to errors due to drift of stage. For thickness measurement, it should be mentioned that discontinuities within the sample can lead to extra detected heights within the layer, or can blur the result on the bottom surface and lead to errors in height or no detection. An important consideration - since a confocal system already limits the power of light received by the detector due to the use of a pinhole - for turbid mediums is the necessity for a very high-power source, as much of light is lost to scatter. The theoretical limit of a typical confocal thickness measurement is on the order of a micrometer due to how close two "airy disks" or point-spread functions can be in the Z direction and still be detected in addition to spherical aberration that is present due to the lens system. [5–7]

Ellipsometry

Ellipsometry is a technique which uses the principle of light polarization to detect the thickness of the sample. Thus, the ellipsometry setup is required to be at a non-perpendicular angle with sample under measurement, since polarization information is somewhat identical at normal incidence. A typical setup for ellipsometry is shown in Figure 1-3.

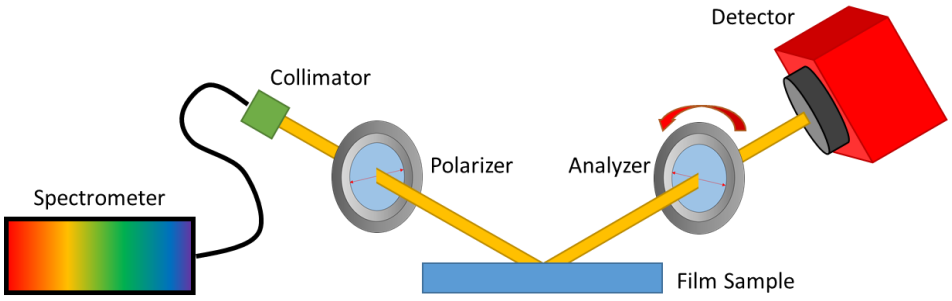

Figure 1-3: This demonstrates a schematic for an Ellipsometry measurement with a tunable light source.

As can be seen in Figure 1-3, there are two polarizers – one altering the incoming beam and one the outgoing beam. The first polarizer is used for creating a linear polarization from the source. For calculation simplification, this polarizer is usually set at a 45° angle with respect to TE or S polarization and TM or P polarization. This gives equal amplitudes in each respective polarization. If the index of reflection is known, then the angle of incidence is usually set to the Brewster's angle for further simplification; however, in the general sense of ellipsometry, this is not necessary. Based on the simplification of setting the polarizer at a 45° angle, the following ratio can be determined for the outgoing light.

$$\frac{r_P}{r_S} = tan(\psi)e^{j\Delta} \tag{1-1}$$

where r_P is the ratio of incoming and outgoing P polarized electric fields, r_S is the ratio of incoming and outgoing S polarized electric fields, tan(ψ) is the ratio of outgoing polarization electric field amplitudes, and Δ is the phase difference between the respective polarizations of the outgoing electric field.

To experimentally determine the polarization exiting the sample, the analyzer is used. This is typically elliptically polarized as there are different components of each polarization as well as a phase difference between polarizations. To find the elliptical polarization of the outgoing light, the analyzer is rotated. While only three angles are needed for defining the elliptical polarization - 0°, 45°, and 90° respectively – additional measurements are usually taken for increasing precision.

Once the ellipticity of the polarization is determined, the amplitude parameter, ψ, and phase lag parameter, Δ, can be determined as shown in Eqn. 1-2 and 1-3.

$$cos(2\psi) = \frac{I_0(90°) - I_0(0°)}{I_0(90°) + I_0(0°)} \tag{1-2}$$

$$sin(2\psi)\, cos(\Delta) = \frac{2I_0(45°)}{I_0(90°) + I_0(0°)} \tag{1-3}$$

where I_0 is the intensity of light detected by the detector at particular angles of the analyzer. It should be noted that in the case Δ is determined to be most nearly 0° or 180° which corresponds to cos(Δ) approximately equal to a magnitude of one, there can be a large inaccuracy in the resultant calculation. To avoid cases such as these, one can introduce a retarder in the setup to change the phase lag and thus reduce the error.

Once ψ and Δ have been determined, different physical parameters of the surface can be identified, such as index of refraction and film thickness. If one is interested in solving for index of refraction, using standard Fresnel equations for a boundary will give P and S polarization ratios as a function of angle and complex index. Since the relation

between P and S polarizations is already known based on ψ and Δ, these equations can be combined to solve for the complex index of the sample. This is shown in Eqn. 1-4.

$$\tilde{n}_1 = \frac{\tilde{n}_0 \sin(\theta_0)\sqrt{1 - 4\sin^2(\theta_0)\tan(\psi)e^{j\Delta} + 2\tan(\psi)e^{j\Delta} + \tan^2(\psi)e^{j\Delta}}}{\cos(\theta_0)[1 + \tan(\psi)e^{j\Delta}]} \tag{1-4}$$

where $\tilde{n}_1$ is the complex index of refraction of the sample of interest, $\tilde{n}_0$ is the complex index of refraction of the external medium, and θ_0 is the incident angle with respect to the normal of the sample. Thus, the physical parameter, index of refraction can be obtained from the ellipsometry setup.

It is also possible to determine film thickness based on the ellipsometry setup assuming index of refraction is known for all three mediums – ambient, film, and substrate. In this case, since there are many internal reflections, the polarization of the outgoing beam must take into account all of them. Thus, the overall ratio of output electric fields polarizations is changed to the following based on the limit of infinite reflections.

$$\frac{R_P}{R_S} = \frac{r_{P1}r_{S0}r_{S1}e^{-j4\beta} + (r_{P0}r_{S0}r_{S1} + r_{P1})e^{-j2\beta} + r_{P0}}{r_{P0}r_{P1}r_{S1}e^{-j4\beta} + (r_{P0}r_{P1}r_{S0} + r_{S1})e^{-j2\beta} + r_{S0}}, \beta = 2\pi\frac{d}{\lambda}\tilde{n}_1\cos(\theta_1) \tag{1-5}$$

where r is the ratio of incoming to outgoing electric fields for respective polarization and boundary, d is the thickness of the film, λ is wavelength of source, and R_P/R_S is the overall ratio of outgoing electric fields with respect to polarization. This equation can then be set equal to Eqn.1-1 above and be used to solve for d. It should be noted that since the equation is quadratic in nature, there will be two complex solutions for thickness. One should choose the value with the lower imaginary component since thickness should be a completely real value.

The last item to note with respect to thickness measurement is that for thicker films, the phase difference between the respective polarizations can become greater than 2π

radians. If this is the case, there is an ambiguous thickness associated with film unless a thickness range is known prior to measurement. To determine an accurate thickness for this case, one should use a spectrometer as the tunable source and vary the input wavelength. For each wavelength across the scan, one should calculate the thickness independently and compare the results to find the unique result with a consistent thickness. Using a spectrometer is the standard technique in ellipsometry because it allows for averaging thickness results and verification of a unique thickness.

To characterize ellipsometry relating to thickness measurement, there are many benefits to the system. Firstly, the measurement can be very fast, since the only data acquired is from different angles of the analyzer and the detector only needs one intensity value per angle. The resolution is very good (< 0.01 nm), especially for very thin films on the order of tens of nanometers with the ability to measure thicknesses below 1 nm. Since a scanning source can be fast and used for robustness of the measurement – as mentioned above, the measured thickness will be reliable as there are redundant thicknesses calculated for each respective wavelength which can also be used to determine if a measurement is not conclusive.

One of the major limitations of this technique includes whether the sample itself alters the polarization of the of light which would not allow for the correct measurement. Additionally, there is only a single thickness data point over the surface as opposed to a thickness profile. Another limitation relates to the level of turbidity within the medium – as a turbid layer would induce a dominant scattering component as compared to the signal, which would not allow for proper polarization determination. Lastly, with respect to the developed system used in Chapters 2 and 3, this technique has an upper limit in thickness

measurement on the order of 1 to 5 micron due to the high number of cycles in the delay between the different polarizations. Thickness samples greater than this create a high error in Δ even with use of a scanning wavelength source. [8–10]

Reflectometry

Reflectometry is another technique that uses the optical parameters of the sample to detect thickness. It contrasts from ellipsometry in that the incident light is usually directed normal to the surface as shown in Figure 1-4. This is to simplify the complexity of the Fresnel equations and thus reduce the computation for calculating the correct thickness.

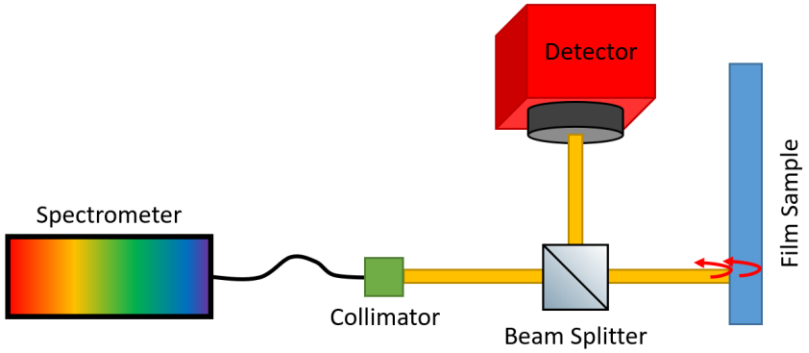

Figure 1-4: This demonstrates a schematic for a typical reflectometry measurement with a spectrometer as a source.

In general, reflectometry looks for the overall reflectance coming back from the surface at a set number of wavelengths. The first component considered in reflectometry is the variation of index of refraction with wavelength. For this, typically the Cauchy approximation is used. This approximation is shown in Eqn. 1-6.

$$n(\lambda) = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} \tag{1-6}$$

where *A*, *B*, and *C* are constants that must be determined through the measurement. This approximation, while not a correct assessment of index over a wide range of wavelengths, usually holds valid in the visible spectrum for dielectrics. For semiconductors, amorphous materials, and other more complex mediums, other approximations must be used with more terms. After defining the approximation for the material, it should be determined whether the index has an imaginary component or is absorptive. In the case where the material is absorptive, a similar approximation can be made for $k(\lambda)$. However, for simple materials, the absorption component can be assumed to be zero with respect to the real component of index.

Next, for finding thickness, reflectivity must be related to index of refraction for a thin film. Using Fresnel equations for an incident angle perpendicular to the sample, the following equation to determine thin film reflection can be shown [11]:

$$r_{tf} = \frac{r_1 + r_2 e^{-2j\beta}}{1 + r_1 r_2 e^{-2j\beta}}, \beta = 2\pi \frac{d}{\lambda} n_1 \tag{1-7}$$

where r_{tf} is the thin film reflection, r_1 is the reflectance of the first interface, and r_2 is the reflectance of the second interface. It should be noted the similarity of Eqns. 1-5 and 1-7 with the difference due to the angle of incidence as well as the complex index reduced to the real component. Below shows the relationship between reflectance at a boundary and index assuming no absorption and perpendicular incidence:

$$r = \frac{n_0 - n_1}{n_0 + n_1} \tag{1-8}$$

Thus, r_1 and r_2 can be calculated for each boundary as a function of the three indexes. With respect to the detected reflectance of the sensor, the final form can be shown:

$$R_{tf} = \left|r_{tf}\right|^2 = \frac{{r_1}^2 + 2r_1 r_2 \cos(2\beta) + {r_2}^2}{1 + 2r_1 r_2 \cos(2\beta) + {r_1}^2 {r_2}^2} \tag{1-9}$$

where R_{tf} is the observable reflectance. Since there are six unknowns in this particular case – A, B, C, n_0, n_1, and n_2 – data will need to be taken for at least six wavelengths. For improvement of accuracy – similar to Ellipsometry – data is usually swept over a broad range of wavelengths in the visible spectrum.

This technique is very accurate (< 1 nm) with the assumption that the material is well defined – i.e. index of refraction and absorption coefficient are well known and material is homogeneous. Based on current tools in industry, this technique can measure on the order of tens of nanometers up to approximately 150 microns. For very thin films (under 15 nm), since the incidence angle is perpendicular to the surface, it is difficult to measure with reflectometry since there is not much variation in the reflection and thus the fitting function is not accurate. This technique does very well if there are well-defined equations – i.e. Eqn. 1-6 - with not too many coefficients. However, with the more complicated materials, defining them properly based on the many coefficients becomes complicated and multiple solutions can arise from a measurement. If this becomes the case, incorrect thicknesses are determined. Thus, for non-crystalline materials or for materials with internal inconsistencies, this technique is normally not used.

There are also techniques that combine ellipsometry with reflectometry for furthering range or redundancy, but this will not be explored in this document. [10,12,13]

Coherence Scanning

Coherence scanning is a widely used technique in profile measurement, but can also apply to thickness measurement. Coherence scanning is an interferometry technique which looks at the contrast of the fringes in each location as the stage is scanning in the Z direction. Figure 1-5 depicts a typical coherence scanning setup.

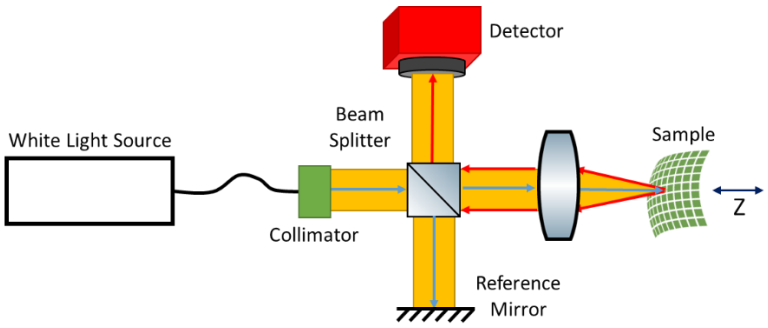

Figure 1-5: This demonstrates a schematic for a typical coherence interferometry measurement setup.

As can be seen in the Figure 1-5, for this technique it is important to have a wide-band source. Typical systems use a standard white light source which gives rise to the commonly used name, 'White Light Interferometry.' Since the source has a broad spectrum, interference fringes will only be detected on the interferogram if the optical pathlengths of the sample path and reference path are most nearly matched. This means that if the sample has a profile variation, certain locations on the sample will have a minimal optical path length difference while other locations will not, resulting in a low contrast fringe pattern. An example interferogram of a spherical surface is shown in the Figure 1-6.

As can be seen in Figure 1-7, when varying the sample path with respect to the reference path, high contrast fringes will appear at location of minimal optical path difference. Thus, if one analyzes the pixel performance over a scan, for each pixel there will be a maximum contrast associated with a perfect balance between paths. To determine the maximum contrast, the pixel data over the scan is reduced to its envelope.

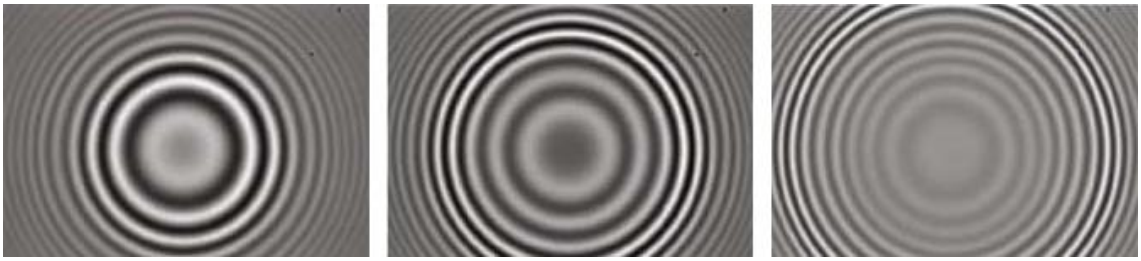

Figure 1-6: This demonstrates the variation in coherence scanning interferograms for multiple distances from a spherical sample.

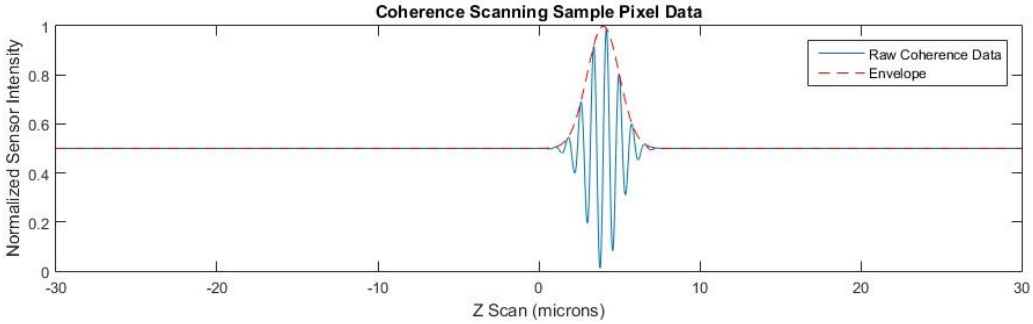

Figure 1-7: This demonstrates the typical envelope for a single pixel's intensity over a coherence scan.

As can be seen from both Figures 1-6 and 1-7, for the particular coherence of the light source, there is a small range where fringe contrast is seen and a peak location can be determined by the envelope for maximum contrast. Many algorithms have been developed for correctly identifying the envelope and thus obtaining a very accurate maximum which relates directly to the profile information. Since the scan is performed over the whole area of the image, identifying each pixel peak will lead to a profile measurement over the surface.

With regards to thickness measurement, the scan must cover both the top and bottom surface to ensure two peaks in pixels measurement and thus a thickness profile. For example, a point measurement for a 20-micron thickness is shown below in Figure 1-8.

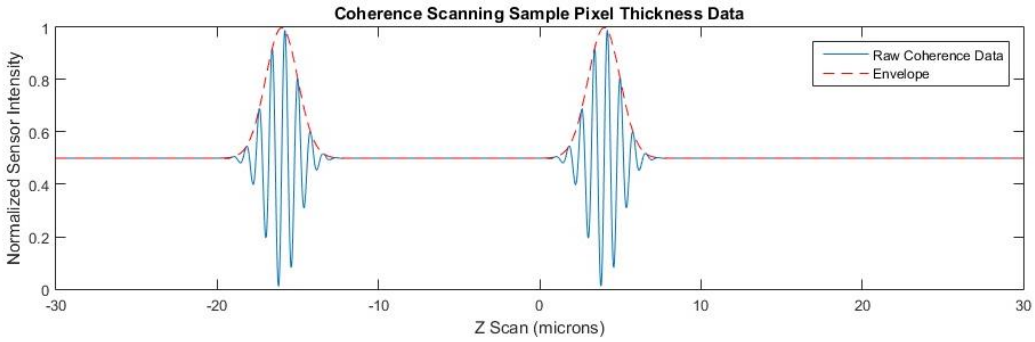

Figure 1-8: This demonstrates the coherence scan envelope seen for a single pixel's intensity on a thick film.

It should be noted that coherence scanning techniques for measuring film thickness requires knowledge of index of refraction for properly identifying the actual thickness. This technique is very accurate (<0.1 nm) for providing a thickness profile for a surface regardless of the material and geometry and can produce repeatable results down to tens of nanometers. There are a few constraints however.

Firstly, the scan itself must take enough data points in Z to correctly identify the envelope. This becomes problematic when the thickness gets larger and a very big scan is required. Thus, more data points require more computation and thus more time. This is why typical coherence scan tools usually restrict the user to less than 100-micron scans unless operator knows approximate thickness of sample.

A second constraint is that of the field of view. The surface area corresponding to a pixel on the camera is assumed to be flat with respect to the reference itself. However, if the area increases and there is variation in the slope of the sample, the contrast in pixel intensity will be adversely affected due to the fact that over one pixel, the intensity is averaged. Thus, depending on the flatness of the sample, the overall field of view is restricted and typically a high-power optic is required to restrict the field of view. Thus, if

one were to desire thickness measurement over a large field of view, one would need to stitch multiple measurements which leads to a timely and less accurate process.

A third constraint is that if the material is rough or has internal discontinuities, the light that must transmit through the sample to the bottom surface gets scattered and thus eliminates the desired fringes.

The last constraint is that the thickness of the film must not be too thin where the two envelopes overlap. In this case, most envelope determining algorithms will breakdown and no longer accurately identify the two peaks. Thus, this leads to the thickness range of this technique from greater than one micrometer to a few millimeters approximately. [14–16]

1.1.2 Surface Measurement

With regards to surface measurement, the techniques that will be explored will only include those applicable for rough surfaces – i.e. hundreds of nanometers and greater. There will be two categories of surface metrology to discuss. The first is the ability to measure roughness of the surface without measuring the surface profile. Methods included in this category are Specular Reflection Method, Diffuse Reflection Method, and Speckle Pattern Technique. The second category will include those that have the ability to measure the surface profile of rough surfaces. The important difference between the two categories is that while the first only has the ability to provide roughness information, the second can compute roughness information from the raw profile that is generated. Therefore, the second type of metrology is more extensive and thus desired. Techniques in this category

that will be discussed include Mechanical Stylus Method, Scanning Tunneling Microscopy (STM), and Coherence Scanning. [17,18]

Specular Reflection Method

This technique is used for finding the roughness of the surface of a sample without finding the profile. In particular, this technique detects is the amount of light reflected from a sample as compared to that of an ideally flat sample with the same properties. A depiction of the setup is shown in Figure 1-9.

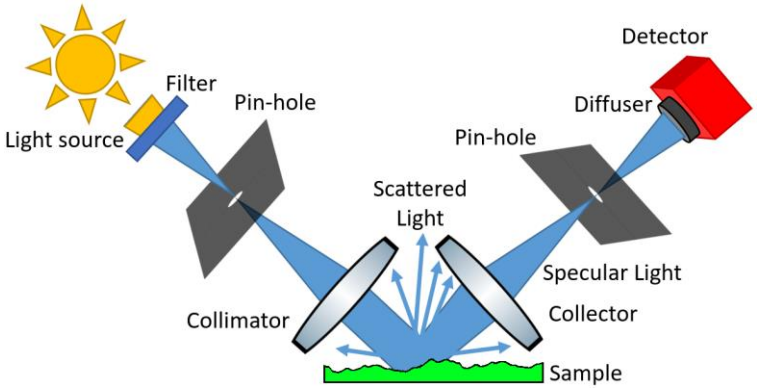

Figure 1-9: This depicts the specular reflection technique for measuring roughness.

As one can see from Figure 1-9, some of the scattered light due to surface roughness do not reach the detector. The degree of roughness can be characterized by Eqn. 1-10.

$$\frac{R_{SP}}{R_0} = e^{-\left[\frac{4\pi\sigma\cos\theta_i}{\lambda}\right]^2} \approx 1 - \left[\frac{4\pi\sigma\cos\theta_i}{\lambda}\right]^2 \qquad (1\text{-}10)$$

where σ is the surface roughness, R_{SP} is the detected specular reflection, R_0 is best-case specular reflection, θ_i is the incident angle of light, and λ is the wavelength of the light. To calculate what R_0 should be, one can use the Fresnel equations for a boundary as was discussed in the ellipsometry section where for a non-perpendicular angle, absorptive

medium, and multiple polarizations of light, the calculation can be quite complex and approximations may need to be made. However, if incident light is perpendicular to the surface and the sample is assumed to have negligible absorption, specular reflection can be estimated to be $R_0 = \left(\frac{n_0 - n_1}{n_0 + n_1}\right)^2$, where n_0 could be reduced to one if in a normal environment. Once this reflection has been defined for the particular setup, the roughness calculation can be deduced directly based on the detected reflection.

There are a few concerns with this particular measurement technique however. Firstly, the technique is sensitive to errors in the assumed index of refraction. One would need to have a relatively accurate estimate of this parameter. Secondly, there is no verification of the detected result as there is only one data point for reflectance. This setup could become more robust if the incident angle could be varied and multiple reflectance values were measured, however this would demand for a complex geometric setup as both source and detector must be simultaneously adjusted. Lastly, the derivation of Eqn. 1-10 requires that the sample roughness be less than one-tenth of the wavelength used. Thus, if a rougher surface were to be analyzed, this technique would not work due to a break-down in the initial assumption or a more complex equation would need to be used to properly characterize the surface. [19–22]

Diffuse Reflection Method

The Diffuse Reflection Method takes the previous method one step further by measuring the lost scattered light in addition to the specular light, as shown in Figure 1-10. This can drive a more robust result than the previous method because roughness can be determined from both reflections respectively. The difficulty with this method is capturing

all the stray light and detecting it. For this, a diffuse, highly reflective sphere is used as shown in Figure 1-10.

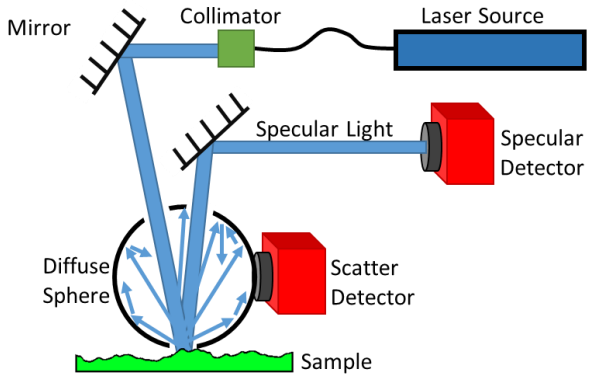

Figure 1-10: This depicts the diffuse reflection technique for measuring roughness.

The specular light can pass through the sphere without getting altered, and thus the theory from the previous section holds valid for the detected specular reflectance. The scattered light trapped in the diffuse sphere is detected in the "scatter detector." Using intuition based on Eqn. 1-10, it can be deduced that the total internal scatter detected with relation to roughness, σ, is as follows:

$$\frac{R_{SC}}{R_0} = \left[\frac{4\pi\sigma \cos\theta_i}{\lambda}\right]^2 \tag{1-11}$$

where R_{SC} is the total reflection due to all scattered light. While this technique is a more robust technique for calculating roughness, the limitations still remain the same from the previous technique with respect to the degree of roughness that can be measured. This technique can also be used for defect detection, as scattered light reflectance with respect to the specular reflectance as the sample is scanned in one direction can give a lot of information for the surface of the sample. [20,22–24]

Speckle Pattern Technique

A speckle pattern is a pattern generated from a scattering surface which causes local bright and dark spots for a coherent source. While speckle mostly destroys an optical measurement, the speckle can be characterized to determine the degree of roughness of the sample of interest. Figure 1-11 shows a typical speckle pattern due to a relatively rough surface.

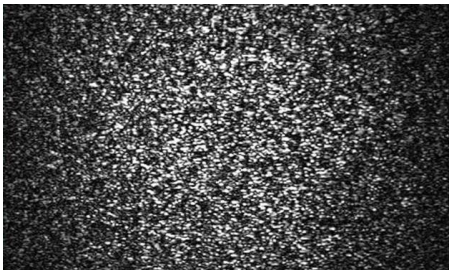

Figure 1-11: This shows the typical pattern generated from laser light on a rough surface.

As can be seen in Figure 1-11, the image consists of a random intensity pattern with very bright and dark spots. The spots can be analyzed for spatial frequency, contrast, and spot size. The overall spot size of the pattern is also a function of wavelength as well as incident angle, which directly effects speckle spots sizes and other parameters. Using these relationships, a model can be developed for predicting the roughness. If the setup is properly characterized, this technique can measure surfaces with a roughness well over tens of microns. This allows for characterization of rougher samples than the methods discussed above, which were restricted to roughness measurement of hundreds of nanometers with the use of light in the visible spectrum. [25–27]

Mechanical Stylus Method

This mechanical technique is a very common technique in metrology. A stylus or sharp tip is laterally scanned over the sample of interest. Depending on the necessity of the precision of measurement, stylus tips may include a capacitive or optical sensor, or if high precision is not necessary, Linear Variable Differential Transformers (LVDT) are used. While scanned, a constant pressure to the surface is generated and one such method is by using a magnet and a force coil. It is vital to the performance of this technique that a constant force be applied to the stylus during measurement. When using an LVDT for example, the system follows the absolute movement past a null point by evaluating the induced voltage from a primary coil to two secondaries. Using a linear stage, the stylus is transferred across the sample to get a line of data points. This process can be iterated in the Y direction (if X direction is scanned) to obtain data for a complete surface profile.

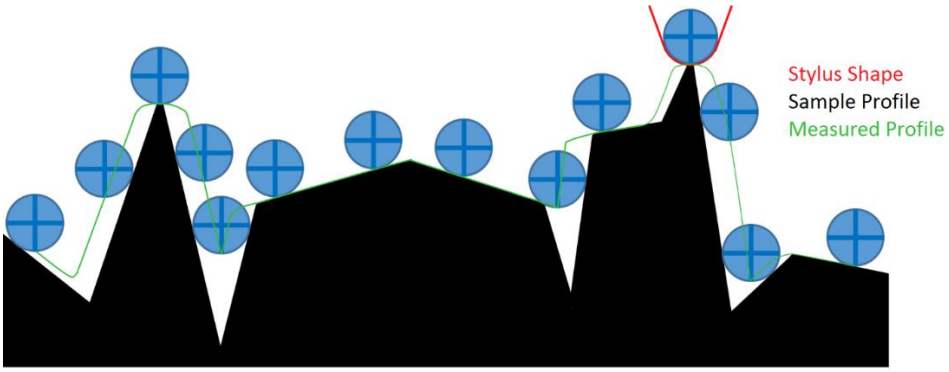

Figure 1-12: This shows the exaggerated error due to the mechanical stylus measurement technique.

While this technique is very accurate in Z (< 1 nm for smooth surfaces), it has a few restrictions. Firstly, the best resolution in X and Y is limited to 1 to 2 microns due to the optical encoders used for measuring displacement. Secondly, based on the cone angle and radius of tip of the stylus, limitations occur related to the inherent spatial frequencies

present on the sample surface. For example, if the surface is very smooth, a resolution of 0.1 nm can be achieved while a resolution above 1nm can be achieved for rougher surfaces. The error due to this phenomenon is shown in the Figure 1-12. A third restriction is that of time. Initial setup of location of scan for a sample can be time consuming and even more so is the scan time of measurement as compared acquisition time of an optical measurement which captures the whole field of view in one shot – i.e. a 100px by 100px measurement dataset at a sample rate of 100Hz would take on the order of 100s. For measurements of higher density, scan time can take on the order of hours to days. A final restriction is due to the fact that this measurement requires direct contact with the surface of the sample under measurement. Thus, this technique is not used for highly sensitive objects as the measurement is considered destructive. Thus, if a measurement is required to be non-contact, this measurement technique cannot be used. [28–30]

Scanning Tunneling Microscopy (STM)

Scanning Tunneling Microscopy uses the same technique by scanning the sample in one direction to detect the height changes, but it uses a phenomenon known as tunneling current. When two metals come within nanometers of each other, a tunneling current is generated due to the proximity of the electrons. The current detected by this technique is exponential with respect to the distance between the conductors. This variation in current provides a signal that is very sensitive to the location of the object without direct contact of the sample's surface. Also, since the variation in current is exponential, this gives very good detail for identifying the actual location. For atomically flat surfaces it is observed that the variation on the surface due to individual atoms can be identified.

There are two variations for measurement with this method. The first technique is to keep a constant height over the surface and identify the variation in current. While this technique seems to be good due to less motion, it is only practical if the surface analyzed is very smooth since the tip must be on the order of 10nm from the surface to detect current. The second technique is more applicable for rougher surfaces. The tool will move the stylus to keep the current constant (distance from surface) as it scans the surface laterally.

This technique claims better than 0.1 nm vertical resolution and lateral resolution of 1 nm when operating in a vacuum. The STM probes can have a tip radius of less than 50nm. While this technique is extremely accurate, it is slow similar to the previous technique. A typical field of view for an STM system is on the order of 100 by 100 microns. The tool also requires enough conductivity in the sample to generate a few nA of current. With this restriction, it should be mentioned that the sample must be either a metal or have a degree of conductivity. [31–33]

It should be noted that Atomic Force Microscopy (AFM) uses a similar measurement method but instead measures the force generated from the overlapping of electron clouds of the stylus tip and the sample. Force generated must be at least greater than 0.2nN. With this technique, there is no requirement for the sample material. [34]

Coherence Scanning

Although Coherence Scanning Interferometry was already covered in the thickness measurement section, it needs to be stated that coherence scanning is also a widely used technique for the measurement of roughness. For profile measurement, the sample only needs to contain one null fringe envelope over the entire Z scan. Since each pixel in the field of view is analyzed independently, this technique avoids the restrictions other laser

interferometric techniques would have due phase-unwrapping requirements. The only limitation in this technique is when variation in height within one pixel is on the order of the coherence length of the source, the fringe contrast diminishes and thus adversely affect the signal to noise. For example, if the coherence length of the source used is approximately 1 μm, depending on the physical area of one pixel, as roughnesses approach this threshold, the visibility of the coherence envelope will decrease. To resolve this problem one of three implementations would need to take place - using a better X, Y resolution in the measurement while sacrificing the overall field of view and cost of optics, increasing the overall resolution of the sensor while sacrificing time of measurement and cost of sensor, or stitching multiple measurements with risk of lateral inaccuracies as well as time of measurement. [35–37]

Optical Techniques not Suitable for Roughness Measurement

It should be mentioned that standard highly coherent interferometers configured in Michelson, Twyman-Green, Mach-Zehnder, and Fizeau configurations are typically not used for roughness measurements as the strength of these techniques is the precision of measurement. A moderate roughness ($> \lambda/4$) would generate phase jumps in the raw information which would lead to an ambiguity in the actual profile of the object. This is also the case for phase-shifting interferometry, which while providing extremely high precision (hundreds of picometers) for very flat surfaces, cannot provide any reliable height data greater than hundreds of nanometers. Any technique such as this requiring the use of a phase unwrapping technique to generate a height map is typically not used for roughness measurements due to the restriction of allowable slopes present within the sample. [38–42]

1.1.3 Summary of Limitations for Rough Surface and Turbid Film Measurement

This section summarizes the limitations of each of the previous respective sections. Limitations presented are specific to measurement of a high level of surface roughness and / or a highly turbid medium.

Table 1-1: Limitations in film thickness measurement due to rough top surface and / or turbid layer.

Film Thickness Measurement			
Technique	**Resolution**	**Measurement Range**	**Rough Surface / Turbid Medium Limitations or Challenges**
Confocal Microscopy	0.4 µm	5 µm - 2 mm	Laser power must be high due to use of pinhole. With turbid layers, due to scattering, little optical power comes back to sensor. Could potentially detect incorrect peaks in middle of turbid layer due fluctuations intensity observed.
Ellipsometry	< 0.01 nm	< 1 nm - 5 µm	Polarization modelling for fitting breaks down due to speckle. Upper limit for thickness measurement limited to a few microns. Material cannot affect polarization.
Reflectometry	< 1 nm	15 nm - 150 µm	Modelling for fitting index of refraction breaks down due to speckle. Limited materials can be used as index variation with wavelength must be predictable.
Coherence Scanning	< 0.1 nm	5 µm - 5 mm	Inner surface contrast fringes blurred due to top surface roughness and or medium turbidity. Technique cannot handle much scatter. Slopes present within sample limit FOV for rough surfaces.

Table 1-2: Upper limit limitations in roughness measurement due to rough surfaces.

Roughness Measurement			
Technique	**Resolution**	**Roughness Upper Limit**	**Rough Surface Limitation**
Mechanical Stylus Method	0.1 nm to 1 nm (dependent on surface)	Microns	Contact method is destructive. Limit dependent on stylus probe radius and slopes present within sample. Time consuming for larger FOV.
STM and AFM	< 0.1 nm	250 nm	Due to precision, limit in axial range of hundreds of nm. Very time consuming for larger FOV.
Confocal Microscopy	0.4 µm	Tens of microns	Laser power must be high due to the little optical power that comes back to sensor. This is due to the dual issue of use of pinhole in combination with a high-level of scattering. Lower magnification objectives desired to properly characterize rougher surfaces.
Specular / Diffuse Reflection Technique	Nanometers	Hundreds of nanometers	Simple statistical approximation for specular and diffuse reflection breaks down for roughness greater than 300 nm. Cannot give information on surface profile.
Speckle Pattern Technique	10 to 100 nm (dependent on technique - i.e. laser or white light)	Tens of microns	With proper calibration and characterization of surface, is capable of measuring very rough surfaces, however, it cannot give information on surface profile.
Coherence Scanning	< 0.1 nm	1-10 microns	Limited FOV due to presence of steep slopes in surface. Would require stitching for larger FOV. Coherence length of source restricts upper roughness limit.
Standard Laser / Phase-Shifting Interferometry	< 1 nm	λ / 4	Speckle dominates phase map. Change in phase is not distinguishable over scan in reference.
Laser Triangulation and Structured Light	> 2 µm	Hundreds of microns	Resolution not sufficient for characterizing less than 10-micron roughness. Both experience HDR problems.

1.2 BiSpectrum Technique

As can be seen in the previous section for thickness measurement of films with a high degree of inhomogeneity and with many scattering centers, and profile measurement of rough surfaces, optical techniques, generally, suffer from common problems of scattering and speckle noise. If the noise component cannot be easily separated from the wanted information, failure in the measurement will occur. Thus, a signal-to-noise ratio, or SNR, enhancement technique was devised for the purpose of extending the confidence level of the measurement. The developed signal enhancement technique was given the term 'Fractional BiSpectrum' technique, or FBS [43,44]. This algorithm was conceptualized due to some of the limitations that exist in the BiSpectrum technique.

The BiSpectrum technique was developed to enhance the signal to noise ratio for data which had an inherent double frequency within the signal. As can be seen in astronomy and cosmology, when analyzing optical data from a distant source, there is both a natural frequency coupled with an additional inherent frequency which in turn creates a frequency at the sum of the two frequencies [45]. Since the information was greatly lost to a high threshold of noise, the BiSpectrum was developed using this relation embedded within the signal. The noise-reduction technique has since been applied to other fields such as radar, telecommunication, geophysics, biomedicine, and digital image processing [46–50]. Furthermore, extensive research has been done to properly reconstruct the phase and amplitude information directly from the data within the BiSpecrtum domain [50–59].

The BiSpectrum can be defined as the Fourier Transform, or FT, of the third order cumulant [46]. The third order cumulant is a term used in statistics when dealing with a probability density function, or PDF. If it were to be assumed that there is a zero-mean

process, meaning the expectation value for the sample set goes to zero, the third order cumulant is the same as the third order moment or skewness and thus equivalent to a triple autocorrelation function or TAF. This function is defined as follows for a one-dimensional, discrete signal:

$$c_x^{(3)} = m_x^{(3)} = \langle x(i)x(i+k)x(i+l)\rangle \tag{1-12}$$

where x is the value at the index of i, and k and l are shift indices. The BiSpectrum is then defined as the FT of this function or:

$$B(p,q) = \sum_{l=0}^{N-1}\sum_{k=0}^{N-1} c_x^{(3)}(l,k)\, exp[-j2\pi(pl+qk)] = F\left\{c_x^{(3)}(l,k)\right\} \tag{1-13}$$

where $F\{.\}$ is the FT and N is the length of the signal. It should be noted before further derivation that the major benefit of the BiSpectrum over the power spectrum or other algorithms is the ability to perform a spectral component correlation as well as maintaining phase relationships between multiple Fourier spectrum components. For the computational simplicity of a real discrete signal it can be shown that:

$$B(p,q) = X(p)\cdot X(q)\cdot X(-p-q) \tag{1-14}$$

where $B(p,q)$ is the complex BiSpetrum which can be separated into its magnitude component and phase component, p and q are the indices in BiSpectrum space that range from $-N/2+1$ to $N/2-1$, and $X(.)$ is the FT of $x(.)$. Some of the properties of the BiSpectrum that come about due to the TAF are as follows:

1) For a stationary Gaussian process with zero mean, $B(p,q) = 0$.
2) For a single simple harmonic function, i.e. $x(i) = A_0\, cos(\,2\pi fi + \phi)$,$B(p,q) = 0$.
3) The BiSpectrum is a periodic function where $B(p,q) = B(p+2\pi, q+2\pi)$.
4) The BiSpectrum has hexagonal symmetry where $B(p,q) = B(q,p) = B(-p-q,q)$ $=B(p,-p-q)$ $=B(-p-q,p)$ $=B(q,-p-q)$. Thus, the only limits needed for full

definition of BiSpectrum space is for $q \geq 0, p \geq q, p + q \leq N/2 - 1$. This can dramatically lessen computation time.

5) The BiSpectrum information has the ability to recover the phase and amplitude information of the Fourier spectrum.
6) There exists an invariance property where a spatial or temporal shift in the initial signal has no impact on the BiSpectrum space.
7) Maybe the most important property for the purposes of this research, if there exists three inherent frequencies within a signal where the sum of two of the frequencies results in the third, a non-zero result will occur in the BiSpectrum.

Based on the properties of the BiSpectrum, for an additive Gaussian noise introduced to the system, the BiSpectrum trends to zero, but does not trend to zero if two frequencies have an additive relation containing a third. Thus, using the BiSpectrum, one can identify signal frequencies where a much higher noise power exists as compared to Fourier space. Since in astronomy applications, there exists a coupled third frequency relation based on the 7th property above within the detected signal, it was possible to extract the signal using the BiSpectrum.

To demonstrate the advantage of the BiSpectrum, a simulation was developed. The simplest relationship where three frequencies, f_1, f_2, f_3, hold the 7th property the BiSpectrum above occurs when $f_1 = f_2$, and $f_3 = 2f_1$. Thus, let us assume there are two signal frequencies present within the spatial-domain as shown in Figure 1-13. For the use of the BiSpectrum, we will set one signal frequency to be twice of the other.

The inherent frequencies of 15 cycles per sample length and 30 cycles per sample length are shown in Figure 1-13. Both sinusoids in the simulation have an amplitude of one. These embedded frequencies are summed up together and combined with white

Gaussian noise containing a noise power of nine times that of each respective signal frequency power. The combined raw signal is shown in Figure 1-14.

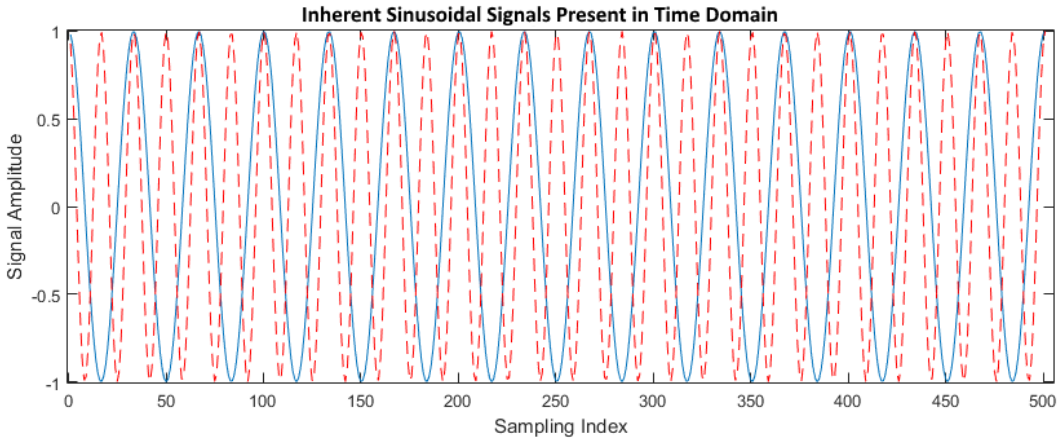

Figure 1-13: This shows an example of frequencies embedded within the signal for the purpose of demonstrating the advantage of the BiSpectrum.

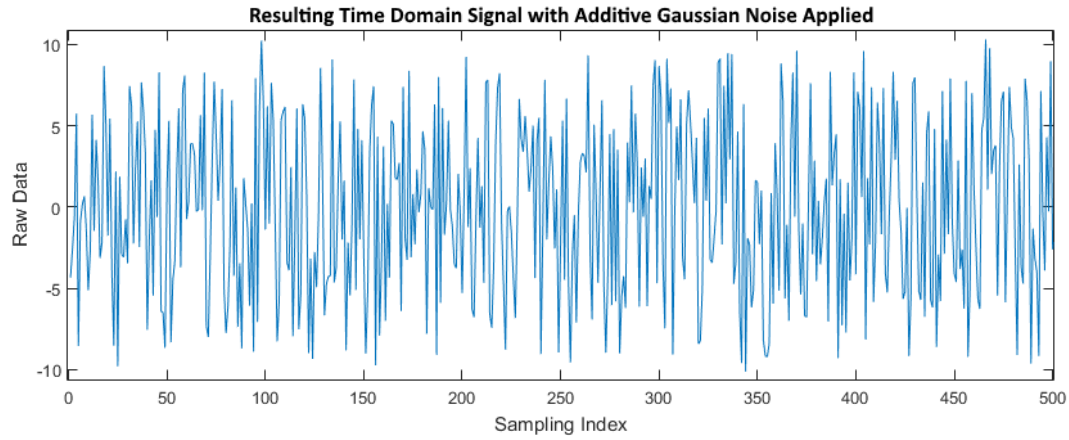

Figure 1-14: This shows the raw time-domain signal data of the combination of two inherent frequencies with white Gaussian noise for demonstration of the BiSpectrum.

As is seen in Figure 1-14, the frequency content cannot be distinguished without the use of a further noise reduction technique such as using the FT for determining whether any inherent frequencies are indeed present within the data. Thus, the FT is performed on the spatial signal to discover whether the frequencies of 15 or 30 can be distinguished.

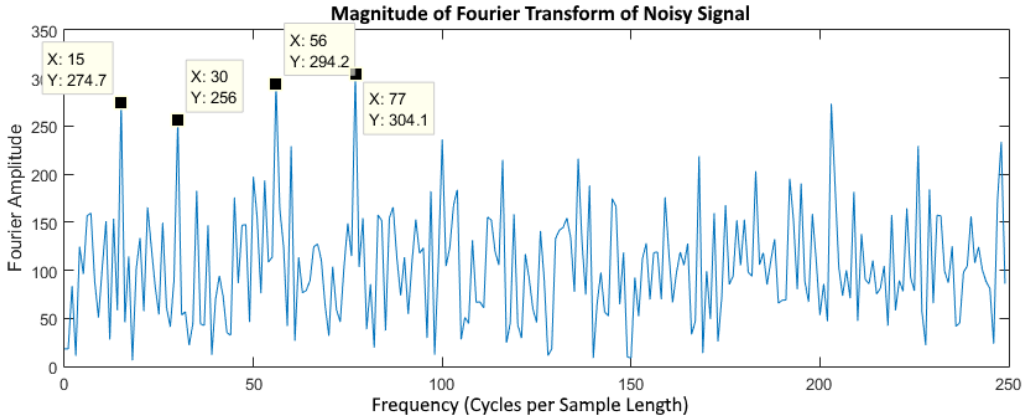

Figure 1-15: This shows the Fourier transform of the noisy data presented in Figure 1-18. The inherent signal frequencies are not the most dominant frequencies in Fourier space.

Figure 1-15 shows the Fourier transform of the noisy data and while frequencies of 15 and 30 have strong amplitudes, they are not the highest amplitude frequency components detected. Thus, the FT would fail to correctly identify the correct frequency component present within the data. However, if there were knowledge that the double frequency was present (or the sum of two inherent frequencies were present), the BiSpectrum could be performed for further enhancement of the inherent signal frequencies.

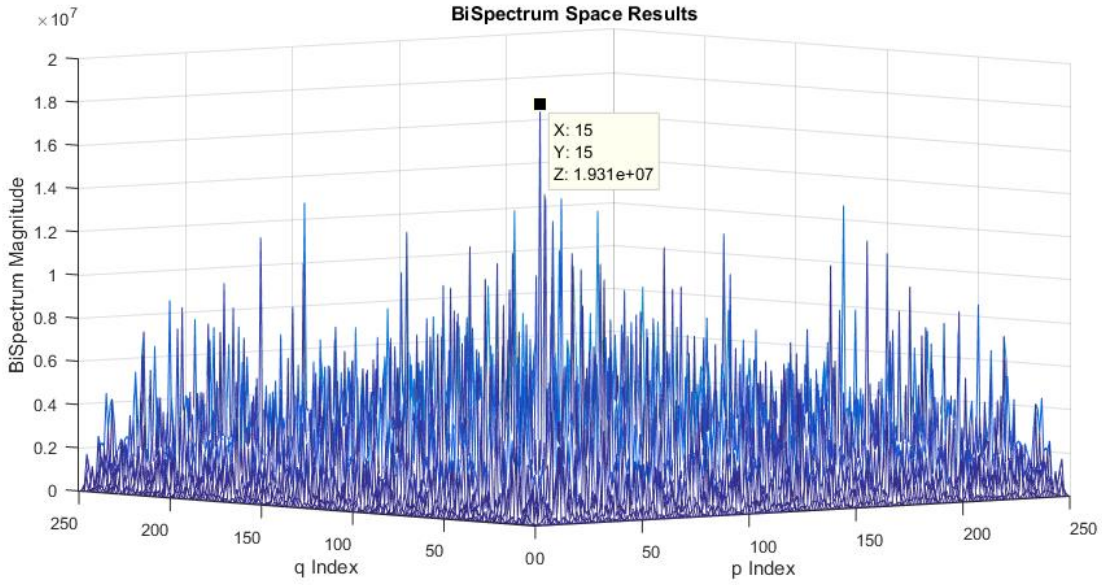

Figure 1-16: This shows two-dimensional BiSpectrum space of the noisy signal data. The BiSpectrum is two-dimensional because the signal is one-dimensional.

The magnitude of the BiSpectrum for one of the identical six BiSpectrum regions is shown in the Figure 1-16. As is seen, the peak magnitude in the BiSpectrum space occurs when indices *q* and *p* are equal to 15. This occurs due to Property 7 for the BiSpectrum which states that when the sum of each index frequency is equal to that of another signal frequency present within the signal, the BiSpectrum will not trend to zero. Since both frequencies 15 and 30 are present, this will give rise to a spike occurring at this particular pixel in BiSpectrum space. If there is additional knowledge that a double frequency is present in the signal, a further filtering can be performed in BiSpectrum space where an analysis on the $p = q$ line could be performed for the reasoning that *p* and *q* must be identical for the double frequency presence to be detected. Data on this line for the BiSpectrum space is shown in Figure 1-17.

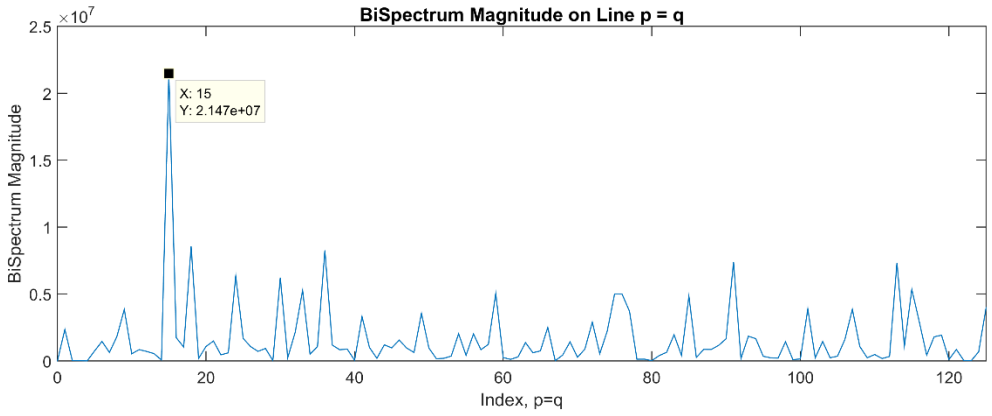

Figure 1-17: This shows the BiSpectrum data on the line $p = q$. The inherent signal frequency of 15 is easily distinguishable.

As is seen in Figure 1-17, while the initial signal to noise ratio, or SNR, of the noisy spatial content was closer to 1:1 in the full BiSpectrum space, a great increase in SNR can

be seen in the selected line of the BiSpectrum. This gives great confidence that the inherent frequency has been properly identified as compared to the FT of the same data. This process demonstrates the strength of the BiSpectrum with respect to enhancing the SNR for a noisy measurement.

1.2.1 BiSpectrum Restriction

While the BiSpectrum analysis has many applications in science and is a good technique for recovering phase and frequency information for a noisy time/spatial-domain signal, it can only enhance the SNR for the specific scenarios mentioned above. For example, if two inherent frequencies are not coupled together or a double frequency is not also present in the time/spatial domain, there is no filtering that can be performed in BiSpectrum space that can help to extract the present signal frequencies. Therefore, if another known relation exists between multiple frequencies in the time/spatial domain, the BiSpectrum would not provide any advantages over other techniques with respect to enhancing signal frequencies as compared to those related to the noise.

1.3 Outline of the study

In Chapter 2, we introduce the Fractional BiSpectrum technique, which is the innovative algorithm used for enhancing the SNR of a noisy measurement. This technique provides flexibility in inherent frequency relationships as compared to that of the BiSpectrum. We implement the FBS technique on a common interferometry setup to determine the film-thickness of thick films. We first characterize the system and measure

films that generate a very low amount of optical noise, followed by the measurement of rough/turbid films.

In Chapter 3, we introduce a subsequent noise reduction technique similar to that of the FBS, but where the multiple frequency content is separated in respective images. We apply this technique to the same film-thickness measurement system and compare results from the normal FBS technique with the newly introduced technique.

In Chapter 4, we implement our FBS signal-enhancing algorithm using a highly explored interferometry technique, Wavelength Scanning Interferometry, and apply it to profilometry measurement. After introducing the background theory for the interferometry technique, we provide some profilometry results on roughness standards.

In Chapter 5, we discuss some addition noise-reduction techniques and strategies based off of the FBS.

Chapter 6 will conclude the contents of this dissertation.

CHAPTER 2: OPTICAL FILM THICKNESS MEASUREMENT OF TURBID MATERIALS USING THE FRACTIONAL BISPECTRUM NOISE-REDUCTION TECHNIQUE[1]

Abstract

A novel approach for deriving the thickness information from a rough and/or turbid thick film or coating is demonstrated. For the application where the roughness or inhomogeneity of the thickness layer produces a dominant speckle pattern, laser interferometry is often not a suitable technique for reliable thickness measurements. The implementation of a signal-to-noise ratio (SNR) enhancement algorithm, which can relate the spatial frequency information produced by the sample when illuminated by multiple lasers is demonstrated. Development of a metrology system using a specific laser interferometry setup to incorporate the noise-reduction technique is also presented. With respect to the detected information of the measurement system, this SNR enhancement technique allowed for the reliable identification of the proper frequency and thus the thickness of the sample.

2.1 Introduction

Thickness measurement of films made from turbid mediums and/or having rough surfaces have many applications in improving the manufacturing process in the automotive, aerospace, defense, metal packaging, flexible packaging, industrial electronics, paper, and medical industries, to name a few. Types of layers or coatings of interest include primers, base coats, top coats such as paints, anodized layers, and adhesives. In each respective industry, however, there is a lack of fast, non-contact

[1] This paper is published in the Optics Communications journal.

measurement systems able to provide a reliable thickness measurement of the rough and turbid layers common to the industry. In such industries, having quick feedback on thickness measurement would allow for reduction of material used – saving in waste and money, ensure reliability and quality of manufactured layer – reducing issues that would arise after products are in households, and in some cases can alert the manufacturer of potential issues in manufacturing where they were previously not aware. The system developed in this manuscript would allow for the thickness measurement to be performed from one side of the layer – does not need access to both sides – which does not require any alteration of the sample and could be used for internal layer measurement. The speed of the measurement is very fast and can potentially be used for live samples in production. The difficulty of the measurement lies in the rough surface or turbidity of the material and when measuring these types of films using a laser light source, a speckle pattern is generated [1].

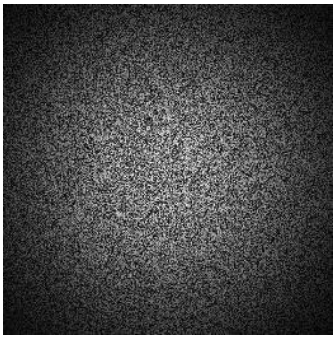

Figure 2-1: This shows the Fractional BiSpectrum data on the line $p = q$. The inherent frequency of 16 cycles per sample length is easily distinguishable.

When a surface or turbid medium generates a high enough noise due to scattering or speckle, it may produce a Signal-to-Noise ratio less than 3dB in the power spectrum. That is, if the Fourier Transform (FT) is taken for the purpose of filtering out the noise, the

average amplitude of the noise must not be comparable to that of the signal amplitude or it becomes unclear how to separate the signal information from the noise. Thus, for these types of interference patterns, using the FT for the purpose of deriving thickness information becomes unreliable. There are other non-Fourier related techniques that are currently used for thickness measurement [2-9], but these techniques also break down due to the high level of optical scatter.

Film thickness measurement has been subject to extensive research for many decades [2,3]. A very popular technique is ellipsometry that uses the ratio and lag of respective incoming and outgoing S and P polarizations of light to determine characteristics of the sample medium such as index of refraction and thickness [4,5]. While this technique can be very precise and robust when used with a spectrometer, it also breaks down under optically noisy conditions. This type of measurement requires the medium's physical parameters to be well characterized. If there is a high level of turbidity or top surface roughness, the medium is not well characterized and a high level of fluctuation in the outgoing polarization will occur. Thus, the analyzer will pick up incorrect amplitudes at respective angles and the proper ratio will be incorrectly determined. It should also be noted that due to the precision related to the lag term in the measurement, this technique is typically not used for thicknesses greater than 1μm to 5μm. Since the focus of this research is to measure thicker samples – tens of microns to millimeters - this technique would not be applicable.

Reflectometry is another technique for measuring film thickness, but instead of using polarization states of light it only uses the intensity from the interference of beams [4,6,7]. The beam is projected at a normal incidence to the sample and scanned over a set

of wavelengths. If the sample index is well defined by a model such as the Cauchy equation or another more complex equation, the combination of this equation and the standard Fresnel reflection equations can be used to solve for thickness. Since there are typically 6 to 15 unknown coefficients depending on the approximation equation used, the source must scan at least that number of wavelengths, but usually will scan a much higher number for robustness purposes. This technique will breakdown if the model used is not a proper approximation of the medium which is normally the case for these highly rough or turbid mediums. Also, interference intensity data used for best determining the coefficients will be skewed due to the high level of noise occurring from scattering.

Confocal Microscopy is another technique where the focused light is projected on to the sample and is scanned in the perpendicular direction [8,9]. When the sensor detects a maximum brightness, the position of the scanning stage is known, and a height is determined. For thickness measurement, the sample must be scanned such that the top and bottom surfaces pass the optimum focus point. The thickness of the sample can be calculated by its refractive index and the height difference between two recorded stage positions. When the top surface of the sample has a relatively high degree of roughness, the data from the bottom surface will be undeterminable since the path that light takes is no longer well defined. If the sample medium has a high level of turbidity, there will be other source of reflections within the sample creating unwanted peaks. These unwanted peaks will generate false heights which makes this technique unreliable for such samples.

As another technique, Coherence Scanning Interferometry is a precise measurement technique that can be used for the purposes of thickness measurement [10,11]. This technique scans the sample in the perpendicular direction and looks at the

interference data between a reference arm and the sample. Since a non-coherent source is used, which only gives distinct fringes at a small offset between sample surface and reference arm, identifying when the fringe has the highest contrast determines the absolute height of the sample. Similar to confocal microscopy, for the purpose of thickness measurement, this technique must scan vertically to pass through both the top and bottom surfaces of the sample. While this technique is precise (sub nm), for the purposes of thickness measurement, the contrast fringes get blurred if there is a high degree of roughness on the top surface. This is in addition to reading false heights within highly turbid mediums instead of identifying the actual bottom surface – similar to confocal microscopy. Lastly, if the slopes due to roughness on the bottom or top surfaces are too steep compared to the pixel area or coherence length of the source, no coherence fringe will be present, and thus no measurement can be made.

Optical Coherence Tomography, or OCT, is an extension of Coherence Scanning Interferometry that uses low coherence light and scans laterally and perpendicular to observe the sub-surface structural data and is typically used in the medical field [12-14]. There are two main techniques used within the field of OCT, time-domain and spectral-domain [13]. While this technique is very robust and can handle speckle data, there are three items of interest. Scanning depths of greater than 100μm can be time consuming while achieving a decent accuracy and this technique also breaks down at a few millimeters below the surface due to the low-level signal returning. Since the signal is captured incrementally in depth and intensities are assigned, it needs post-processing to identify each layer thickness. Lastly, relative to the setup proposed in this paper, this technique is much more complex due to either the laser scanning or motion required.

Non-optical techniques can be used, but will suffer from either alteration or destruction of the sample – in the case of step height measurement – or the need for a very well-defined chemistry composition of the medium – in the case of Ion Beam thickness measurement techniques [15]. Thus, using a novel noise-reduction technique based on the BiSpectrum, we have investigated the ability to extract the information from a signal contaminated with a higher noise level than discernable in the Fourier domain such that confident measurement results could still be obtained for the scenario in which the speckle and scattering components are significant.

In this paper, we will briefly discuss this noise-reduction algorithm in section 2.2. Section 2.3 will demonstrate the experimental setup including the physical configuration as well as the automated software used in measuring film thickness of rough samples. In section 2.4, measurement tool accuracy will be demonstrated followed by section 2.5 which will discuss the measurement of rough and turbid mediums. Section 2.6 will summarize the results of this work.

2.2 Fractional BiSpectrum Technique

The BiSpectrum has been used in the field of astrophysics for noise reduction [16-19]. The BiSpectrum can be defined as the Fourier Transform (FT) of the third order cumulant [20]. The third order cumulant is a term used in statistics pertaining to the probability density function, or PDF. Based on the properties of the BiSpectrum, for a Gaussian process, the BiSpectrum is zero; that is, the BiSpectrum statistically eliminates the data related to an additive Gaussian noise. Furthermore, the BiSpectrum does not trend to zero if two frequencies have an additive relation containing a third [20]. Thus, in using the BiSpectrum, we can identify specific frequencies present within the signal where a

higher variance in additive Gaussian noise is present as compared to the Fourier Transform. However, these specific frequencies are identifiable solely based on a specific relationship with other frequencies within the signal. For example, for the application of identifying a binary star, this property can be exploited as the two masses are moving in harmony. However, there is high level of noise within the signal due to the turbulent atmospheric conditions in addition to trying to resolve information up to the diffraction limit of the telescope [16,17]. Thus, the BiSpectrum is an ideal technique to derive the signal from the noise in the acquired data. The definition of the discrete BiSpectrum for a finite, real-valued, discrete dataset is shown in Eqn. 2-1 [20]. A 'discrete dataset' is defined as a time-domain or spatial-domain set of datapoints, i.e., $x[i]$, for i=0,1,2,…,N-1.

$$B[p,q] \equiv X[p] \cdot X[q] \cdot X[-p-q] \tag{2-1}$$

$B[p,q]$ is the complex BiSpectrum which can be separated into its magnitude component and phase component, p and q are the frequency indices in BiSpectrum space that range from $-N/2$ to $N/2-1$ where N is the finite data length of the discrete dataset, and $X[.]$ is the discrete FT of $x[.]$ [20]. After properties of the BiSpectrum became well established, there was great interest in many other disciplines such as radar, telecommunication, geophysics, biomedicine, and digital image processing [20-24].

The Fractional BiSpectrum was then developed to provide more flexibility than the BiSpectrum when signal frequencies inherent within a discrete dataset have a specific relationship with other signal frequencies. The discrete Fractional BiSpectrum or FBS for a finite, real-valued, discrete dataset is defined as [25]

$$FBS[p,q;k] \equiv X[p] \cdot X[q] \cdot X[-p-k \cdot q] \tag{2-2}$$

where the variable *k* is a fractional number, and is used to relate inherent signal frequencies that may exist within the discrete dataset. If signal frequencies are related properly, *k* in effect enhances these frequencies with respect to the additive Gaussian noise present. There are a few key properties that should be identified [25].

1) For a stationary Gaussian process with zero mean, $FBS[p, q; k] = 0$.
2) For a single simple harmonic function, i.e. $x[i] = A_0 \cos(2\pi f i + \varphi)$, $FBS[p, q; k] = 0$.
3) The FBS of a real-valued, discrete dataset has dual symmetry where $FBS[p, q; k] = FBS^*[-p, -q; k]$.
4) If there exist two signal frequencies within the discrete dataset and the relationship between the two is known, a single *k* value can be chosen such that a non-zero result will occur in the FBS.

The important properties for the FBS with respect to enhancing the SNR are properties 1 and 4, where a Gaussian process trends to zero and a non-zero result will occur for a discrete dataset with two inherent signal frequencies with a relationship related to *k*. Suppose there is a known relationship between the first frequency, f_1, and the second frequency, f_2, such that $f_1 = \mu f_2$ and f_1 and f_2 are both integer frequencies. It is known that $X[f_1]$ and $X[f_2]$ should result in peak amplitudes in Fourier space. If it was the case that for a particular set of indices in FBS space that $p = f_1$, $q = f_1$, and $-f_1 - kf_1 = -f_2$, then all three Fourier terms in the FBS would result in non-zero amplitudes assuming the discrete dataset is composed of real values. Thus, a *k* value chosen where $k = f_2/f_1 - 1$ will lead to a non-zero result in the FBS space at $FBS(f_1, f_1; k)$. It should be noted that in the case where $f_2 = 2f_1$, the result is the same as that of the BiSpectrum where $k = 1$.

For the purposes of the thickness measurement system specifically, it is designed such that there will be only two inherent signal frequencies in the same spatial direction

within the image. From that standpoint, all data within discrete dataset that is unrelated to these frequencies will be 'noise.' Each respective inherent signal frequency will be referred to as 'inherent frequency' or 'signal,' which is a single simple harmonic function. The amplitude of these inherent frequencies within each respective frequency domain – Fourier and FBS – will be compared with the maximum amplitude of the noise which will be defined as SNR_I, or the signal-to-noise ratio of interest. If SNR_I, or the ratio between the signal amplitude and the maximum noise amplitude, is greater than $\sqrt{2}$ in the respective domains, there is considered to be a high enough confidence in the measurement. While this does not relate power of noise present to the signal, this comparison gives better clarity to how well the technique has determined the thickness. For simplicity, 'dataset' will be used when discussing the discrete spatial-domain or time-domain data. For thickness measurement, these two inherent frequencies are used to enhance exactly one frequency magnitude which we will relate to the thickness detected. Since the criterion for *k* creates a peak value in the FBS domain only on the line $p = q$, and only the magnitude component of the Fourier domain is needed to determine the single peak frequency, the FBS can be reduced to a more basic equation, Eqn. 2-3, for thickness measurement.

$$|FBS[p;K]| = |X[p]|^2 \cdot |X[K \cdot p]| \tag{2-3}$$

K is similar to *k* in that it is relating the two inherent frequencies. Since symmetry can be assumed for the magnitude about $f = 0$ for a real-valued dataset, analysis of only the positive frequencies are adequate for both respective domains, FT and FBS. If we have the condition that f_1 is less than f_2, in this case, $K = f_2/f_1$, which is a value between 1 and ∞ since *f* must be positive. Thus, the SNR will be enhanced with respect to the lower inherent frequency magnitude by using the FBS. It should be noted that if it was desired

to alter the FBS domain using a different K value, the higher frequency magnitude could in-fact be enhanced. This corresponding factor would be exactly $1/K$. A simulation was performed to show the comparison between the standard FT and FBS techniques where the initial dataset contained two inherent frequencies. For the simulation, the following signal was generated.

$$x[i] = a\sin(2\pi f_1 i + \varphi_1) + a\sin(2\pi f_2 i + \varphi_2) + \alpha[i] \tag{2-4}$$

where a = 1, φ_1 and φ_2 are random initial phases, and $\alpha[i]$ is one realization of random values with a normal distribution, $N(0, \sigma^2)$, where σ^2 was set to 5. The process for properly comparing the FT and FBS is shown in Figure 2-2.

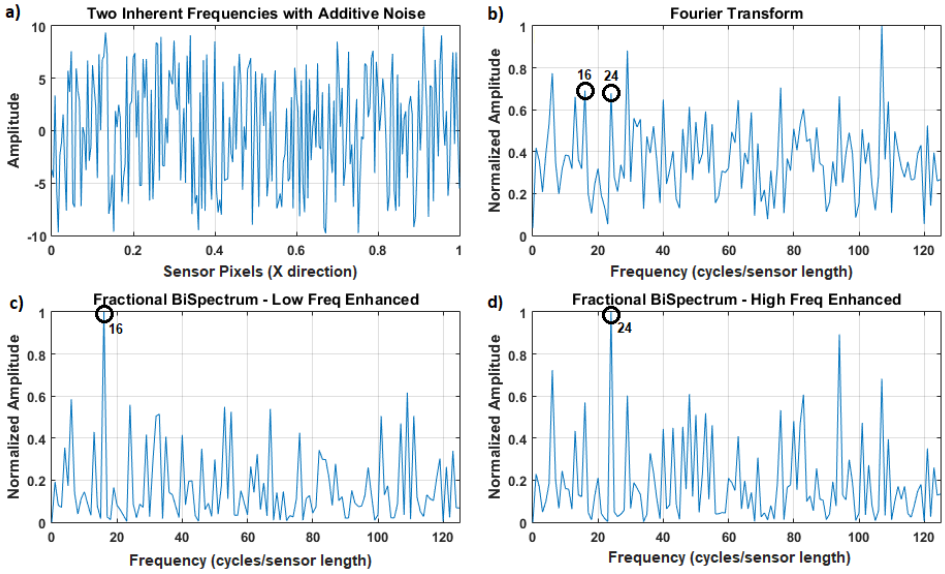

Figure 2-2: (a) Initial sensor dataset has two inherent frequencies of 16 and 24 combined with additive Gaussian noise with a noise power of ten time that of the sinusoidal signals. The remaining plots show the comparison of FT result (b) with FBS result for enhancing lower frequency (c) and higher frequency (d).

Figure 2-2a represents the dataset with two inherent frequencies – in this case 16 and 24 cycles per sensor length respectively – with an additive Gaussian noise applied to the dataset with a noise power of ten times that of the two sinusoidal functions in the

dataset. As is seen, there is no discernable sinusoidal variation related to the inherent frequencies, as the additive Gaussian noise component is too high. In order to clarify the signal frequencies within the dataset, the FT is taken as shown in Figure 2-2b. What is seen, however, is that the proper frequency content – 16 and 24 cycles per sensor length - are still at a level less than that of the noise. In this situation, the proper frequency could not be identified and thus either no result or an incorrect result would be produced from the data.

In Figures 2-2c and 2-2d, it is seen with reasonable certainty that frequencies of 16 and 24 cycles per sensor length are identified using the FBS technique. As expected, using the proper *K* values, the amplitude of one of the two frequencies was enhanced as compared to the noise threshold. Thus, using this technique, the proper frequency would be identified which relates to the proper thickness identification as described in Appendix A. If the result for the FBS is still unclear ($SNR_I < \sqrt{2}$)– as is seen in Figure 2-2d - it should be mentioned that both FBS datasets can be combined by performing a correlation computation. Since this cross-correlation result uses both the FBS datasets, it naturally will give a result with a higher confidence. Thus, for all the measurements in our film-thickness system, we provide only this final cross-correlation result of the two FBS outputs.

To have a comprehensible comparison of the FT and FBS techniques, an additive Gaussian noise was applied to the two sinusoidal functions as referenced in Eqn. 2-4 where the power of noise was varied from 1 to 18 as compared to the signal. A Monte-Carlo simulation was used to create 10000 random datasets for each respective Gaussian noise power and compare how well the FBS and FT can identify the proper signal frequency. This is shown in Figure 2-3.

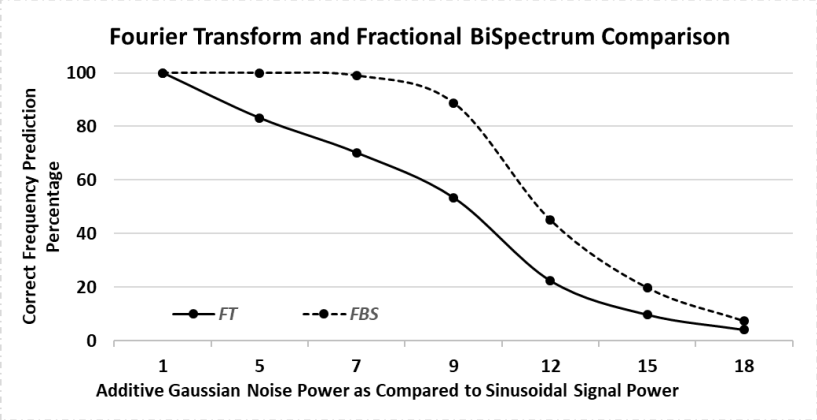

Figure 2-3: Comparison between FBS and FT based on percentage of correct frequencies predicted. 10,000 randomized datasets were generated for each respective noise power.

As is seen in Figure 2-3, the performance of the FBS technique outperforms that of the FT especially at a higher noise power. When the additive Gaussian noise power was in the range of 5 to 9 times that of the sinusoidal signal power in the dataset, there was seen to be a drastic improvement in the correct frequency identification of the FBS technique as compared to that of the FT. It is at this relative level of noise power that we will see the most significant enhancement for this technique.

In the next section we describe a thickness measurement system which implements the FBS as a technique to measure the thickness of films.

2.3 Film Thickness Measurement

For measuring the thickness of a film, there are many techniques to choose from [2-11]. Here we describe a system that uses the noise-reduction technique provided in the previous section in a simple and standard multiple reflection interferometry technique [26]. This system is depicted in Figure 2-4.

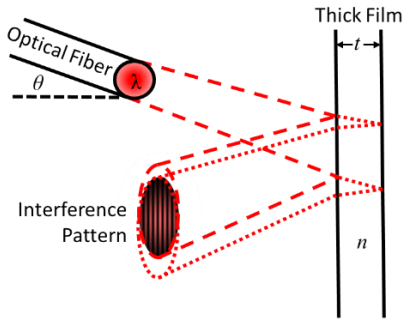

Figure 2-4: Depiction of film thickness setup used. Laser light of a specific wavelength is incident on the thickness sample. Based on the two major reflections from front and back surface of film, an interference pattern is created at the intersection of the two beams.

As can be seen, an optical fiber is used to guide an incident laser beam at a particular wavelength, λ, and angle, θ (with respect to the normal to film surface), onto a thick film. If the film allows for partial transmission of the particular wavelength, there will be two major reflections - one from the top surface and one from the bottom surface of the film. Based on the thickness, t, and index, n, of the sample, there will be a spatial separation between the centers of the beams as well as the diameters and curvatures of the beams. Since the two beams are mutually coherent, where they overlap will create an interference based on this shift in beam centers and curvatures. Due to the lateral shift of the beam centers, this interference pattern will be dominated by a spatial frequency in one direction. Using Gaussian beam analysis, Eqn. 2-5 can be formulated relating this spatial frequency with the physical parameters shown in the Figure 2-4. A derivation of this equation is shown in Appendix A. An approximation was made assuming that the propagation of the beam to the sample is much larger than the Rayleigh length and the assumption that the linear phase component is dominant if the displacement of beam centers is much less than the width of the respective beams.

$$P \approx \frac{\lambda \cdot l\sqrt{n^2 - \sin^2\theta_i}}{2t \cdot \sin\theta_i \cos\theta_i} \tag{2-5}$$

In Eqn. 2-5, P is the spatial period or the inverse of the spatial frequency and l is the total distance traveled by the beam. As can be seen in the equation, there is a direct relation between the spatial period of the observable interference pattern and the wavelength of the laser used. This will be used for the determination of K in the FBS. A simulation of the interference pattern for 532nm and 633nm wavelengths incident on a film of 100μm at a 50º angle was generated and is shown in Figure 2-5. The size of the interference pattern at the image plane is 11mm, the sample is placed at 2cm from the fiber output tip, and the interference pattern is captured 4cm away from the sample. Both respective beams have an NA of 0.12.

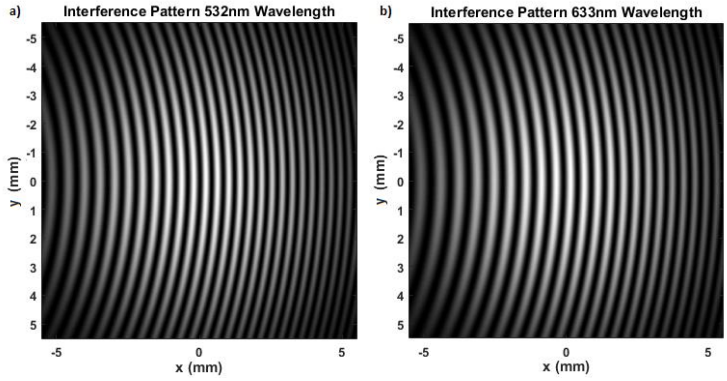

Figure 2-5: Simulated interference pattern generated from a 100μm thick film for 532nm (a) and 633nm (b) source wavelengths respectively. (a) shows much denser fringes than that of (b) due to the relationship related to wavelength as defined in Eqn. 2-5.

As can be seen, there is a dominant spatial frequency in the x-direction of both patterns. It is seen that this spatial frequency is directly proportional to wavelength of the lasers used as predicted by Eqn. 2-5. Since multiple lasers produce fringe patterns each with a dominant frequency, if multiple lasers are used, all frequencies will be present within

the dataset and the *K* value can be readily calculated from the respective laser wavelengths. Thus, the FBS noise-reduction technique can be used to obtain the thickness. In this case, a film with relatively rough surfaces or that of turbidity within the layer would create a noisy spatial dataset. A schematic for the experimental setup is shown in Figure 2-6.

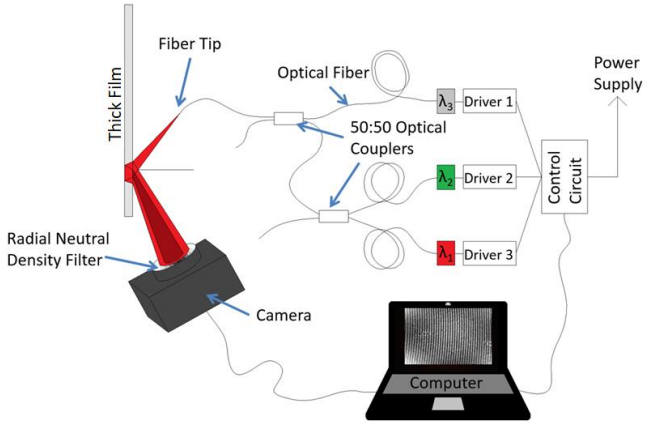

Figure 2-6: Schematic of setup used for thickness measurement. Outputs of three lasers with wavelengths λ_1, λ_2, and λ_3 are combined using single mode fiber directional couplers, which all can be delivered, at the same time, onto the target film with a single fiber.

As can be seen, a computer is used both for setting up and triggering the laser diodes as well as receiving the information from the camera for the measurement. There are three laser diodes each with a respective current driver such that noise-reduction technique can be applied. The three lasers used have respective center wavelengths of 520nm, 642nm, and 847nm. We used three lasers because different thickness materials have different spectral reflectivity and with three lasers, we can choose up to three different two-laser combinations which will extend the capability of the thickness measurement system. Two 50:50 optical couplers are used to combine all three lasers into one output which is projected from the optical fiber tip to the film or coating under measurement. The optical fiber used was rated to be single mode for 633nm wavelength with a 125μm cladding

diameter and a 4µm core diameter. The NA given based on the respective indices of the cladding and the core was 0.12. To determine the proper placement of the fiber-tip, sample, and camera, firstly the optimum angle of incidence needed to be determined.

For our setup, it was determined that maximizing the number of fringes within the FOV was of the highest importance since it limits the minimum thickness detectable. Thus, the angle optimization is to maximize the number of fringes within the FOV. We isolated Eqn. 2-5 with respect to θ and swept it for determining which angle corresponds to a minimum period. Since n could not be removed in the calculation, we first set n = 1.5, which is a typical index for the mediums we expect to measure. This is shown in Figure 2-7a. For completeness, this process was done over a range of indices – from 1.1 to 3.0 – to see how index affected the optimum angle of the setup - as shown in Figure 2-7b.

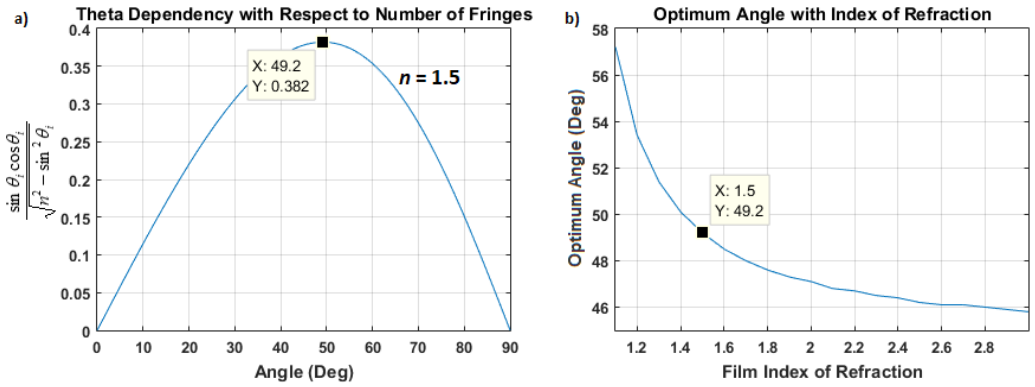

Figure 2-7: (a) depicts how the angle of incidence affects the number of fringes in measurement for n = 1.5 and (b) shows the optimum angle for a range of indices.

As is seen in Figure 2-7, a 49.2-degree angle of incidence will produce the highest density of fringes in the measurement. Although the index was varied drastically, a tight range of 46 to 57 degrees was observed. Since sample indices typically used for measurement are closer to 1.5, an angle of 50 degrees was chosen for the setup.

The next point to consider for the setup was how closely to position each component – sample and camera sensor – with respect to the fiber tip. Since it is assumed the distance to the sample and camera is much greater than the Rayleigh distance of the beam, a constant beam divergence angle of 6.9 degrees related to the NA of 0.12 can be assumed. Since the alignment for a lens in the direct path of a Gaussian beam must have a very tight tolerance with respect to perpendicular alignment and lens center position in order to have minimal aberrations, it was decided not to use a lens. Thus, the width of the camera sensor (11mm) limited the total light path length to 6cm. At this distance, the sensor would be able to acquire 75 percent of the beam spot - recovering most of the fringes. The location of the sample was chosen to be at 2cm based on geometry limitations giving a laser spot size of 5mm diameter on the sample and an observable FOV of 3.7mm by 3.7mm.

Lastly, an anti-Gaussian optical filter was used for the purpose of producing uniform intensity distribution across the beam such that the dynamic range of the camera would be sufficient for measurement. Based on the geometry of the setup, wavelengths of lasers, and limitation regarding the approximation of Eqn. 2-5, a thickness range of 17μm to 1300μm was determined for this setup. The details for this determination are described in the Appendix B. Figure 2-8a shows the physical experimental setup.

As is seen in Figure 2-8, the fiber tip, thickness sample, and camera were placed in the proper locations and angles as described above with a custom-made fixture. The software used for the development of this metrology setup was Labview that has the ability to directly communicate through RS232 to a custom developed board for controlling the lasers. It also allowed for communication with the monochromatic (Prosilica GT) camera

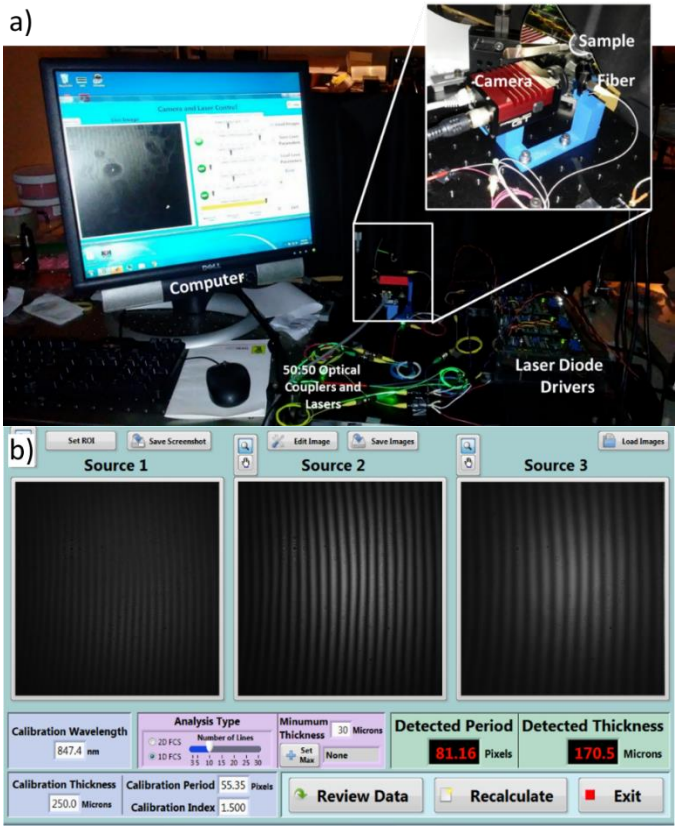

Figure 2-8: (a) shows the lab setup of film thickness measurement system and (b) shows the software interface of captured interference patterns and calculated thickness. The zoomed in window in (a) depicts the fixture developed for setting proper angles and lengths while patterns in (b) shows source 1 (520nm), source 2 (642 nm), and source 3 (847nm) have spatial periods in the x-direction related to their respective wavelengths.

used through GigE protocol. Lastly, the Labview software allowed for processing of the images in the measurement for calculation of thickness of sample using techniques described in section 2.2. Figure 2-8b depicts the final interface for viewing the result of the measurement, including the interference patterns and the measured thickness.

2.4 Measurement Tool Accuracy

Firstly, it was important to determine whether the developed film thickness setup could make an accurate measurement. To determine this, quartz wafer standards of 75µm, 125µm, 250µm, 325µm, and 500µm were used with a given index of 1.55. Each of the

samples were given an uncertainty of 25μm within their rated thickness from the manufacturer. Since the limits of the developed system range is up to over 1mm thickness, additional samples including a cell-phone cover and glass wafer with unknown thickness were also used. A refractive index of 1.5 was used for these samples.

It should be mentioned that since multiple wavelengths are used throughout the visible spectrum and into the near-infrared, the absolute index of refraction of the material will vary depending on the source which will introduce a slight shift in the absolute thickness measured. For example, for plastics the maximum variation in index, Δn, over the wavelength range used is approximately ± 0.025 [27]. If compared to the absolute index of refraction of approximately 1.5, this will introduce a shift in the measurement of $\pm 1.7\%$. Since the layers of interest are highly rough and/or turbid, this amount of shift is usually acceptable considering the measurement tolerance for such layers. If it were desired to reduce this shift and the material index dispersion with wavelength is well known, calculating a more exact relation between the spatial periods can be accomplished by using the exact index of refraction for each respective wavelength in Eqn. 2-5.

For each sample, independent measurements were taken using other metrology tools such as white-light interferometry for thin samples, confocal for intermediate thickness, and for thick samples, a micrometer was used. The results from these techniques were compared with the measured results from the developed system in Figure 2-9.

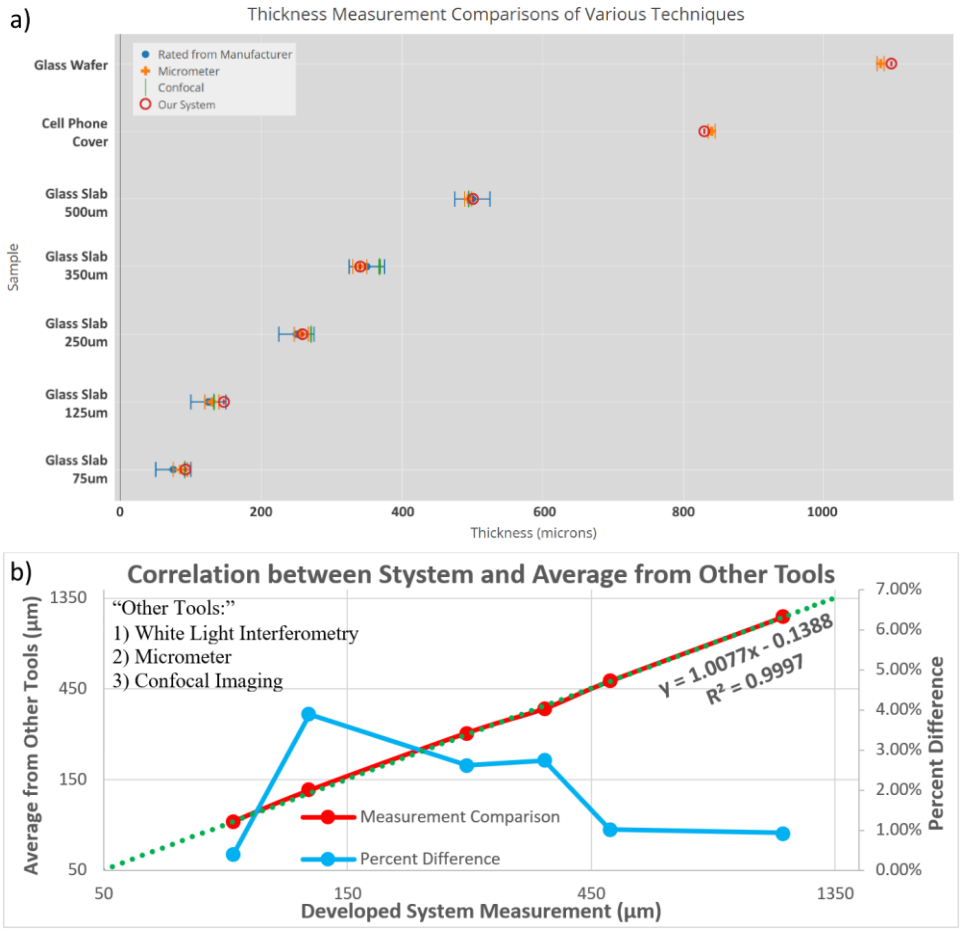

Figure 2-9: Comparison of developed system with metrology tools in industry. Shows confidence range for each measurement. Samples 1- 5 were quartz wafer standards, sample 6 was a cell phone cover, and sample 7 was a glass wafer with unknown thickness. (a) shows measurements for each respective metrology tool and (b) shows correlation coefficient closely equal to one for comparison of other systems as compared to the developed system.

As is seen in Figure 2-9, measurement made by the developed system on each quartz wafer fell within the specified uncertainty by the manufacturer. The developed system's measurements were very accurate with respect to the commonly used tools in the industry as shown in Figure 2-9b. The correlation coefficient between developed system and other tools was nearly equal to one using the entire measurement range of the system. Some of the discrepancies between the measurements can be related to the specific location of measurement on the sample - as thickness could vary, in addition to the refractive index

value used for samples 6 and 7, which may lead to a difference in result as compared to the absolute measurement of the caliber.

Lastly, to qualify the measurement system, each sample was measured 30 times. The maximum deviation in the respective measurements were +/- 2μm of the average thickness. This can be perceived as the repeatability of the measurement system.

2.5 Rough Surface Thickness Measurement

To realize the advantage of our system related to the novel SNR enhancement technique used, it was necessary to use relatively rough and inhomogeneous films to compare the results with the standard FT technique. For this, two samples were used. The first sample analyzed was the coating present on a Blu-ray disk. For protective purposes, all Blu-ray disks have an approximate coating or cover layer of 98μm above the substrate with and index of refraction of 1.55 [28,29] . In order to add noise to the measurement, simply adding dust, scratches, and smudges drastically skewed the interference pattern.

Prior to performing a measurement with high optical disturbance, it was first desired to measure at a clean location of the sample for comparison. Thus, the measurement data for a clean location is shown in Figure 10a. It should be noted that although there is an underlying Gaussian component in the frequency domain result due to the intensity variation across the beam, this was readily removed.

As is seen from the measurement of the clean Blu-ray disk sample, the raw image shows a predictable interference pattern in the x-direction with both green and red interference fringes present. There are two obvious peaks in the frequency domain which correspond to spatial frequencies produced by the film for two wavelengths of 520nm and 642nm. The horizontal axes in Figures 2-10 b and c are scaled to directly represent the film

thickness. Since the frequency range is referencing the red wavelength, the lower thickness, 97µm, corresponds to the red source and the actual thickness of the Blu-ray disk, and the higher thickness, 117µm, corresponds to the green source which is a scaled thickness related to the ratio between the wavelengths of the sources. For a clean sample, the signals in the FT are easily observable. Taking the FBS of the raw image for both the low and high frequencies, and then applying a cross-correlation of the datasets, it is seen that only one peak of 96µm exists which is related to both sources simultaneously. Thus, the thickness of the Blu-ray disk is close to that of the expected rated thickness.

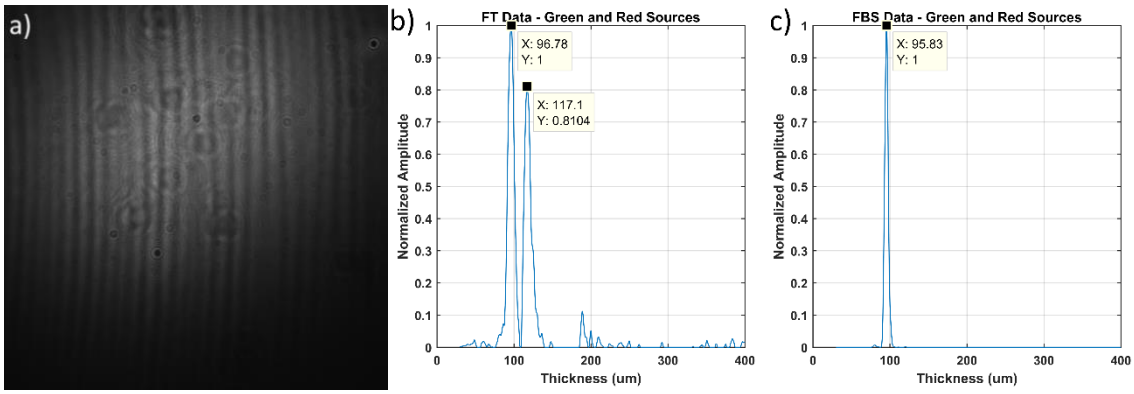

Figure 2-10: Thickness results of clean Blu-ray disk sample for FT (b) and FBS (c) based on the interference image (a) using green and red lasers with respective wavelengths of 520nm and 642nm. The frequency domain has been scaled in terms of film thickness.

A comparable measurement was then performed at an altered location on the sample of interest. The raw image for the two-laser combination is shown in Figure 2-11a.

As can be seen in the raw image, there seems to be scratches and smudges in the field of view. Since the fringes should appear in the x-direction on the image, there is very little discernable interference fringes such as is seen in Figure 10a due to the high noise level. The various techniques – FT and FBS - were then used to determine the thickness of the sample. Figures 2-11 b and c show the results for each respective technique.

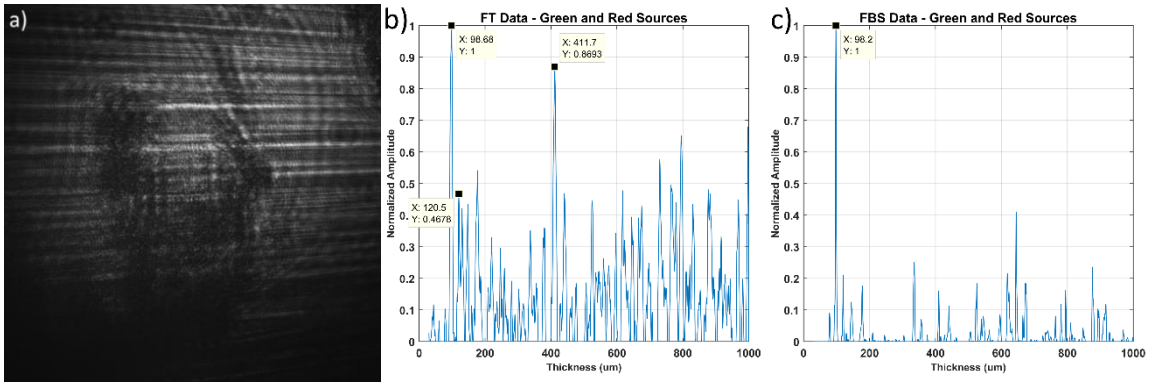

Figure 2-11: (a) shows raw interference image for combination of green (520nm) and red (642nm) lasers for Blu-ray disk measurement. (b) shows the Fourier Transform thickness result of the acquired image while (c) shows the Fractional BiSpectrum thickness result. The frequency domain has been scaled in terms of film thickness.

In the Fourier data, the thickness axis is again referencing the red laser. Thus, the two thicknesses related to signal information are 99µm and 121µm respectively. While there is a strong amplitude in the FT at 98µm, there is a relatively strong amplitude at approximately 411µm as well. If there were no prior information on the thickness of the sample, this would leave an uncertainty in the measurement result based on the SNR_I criteria. The FBS technique, however, is able to measure the correct thickness based on the two datasets with proper *K* values (in this case 642/520 and 520/642) and produce a result at 98µm. It is observed that the noise present within the resultant dataset is drastically reduced and the SNR has been enhanced when comparing to the FT data. Thus, one can state the layer having a thickness of 98µm with the FBS measurement with a higher confidence than the 98µm claim in the FT measurement. This thickness is very similar to the value measured at a clean location on the same Blu-ray disk.

This process was carried out as well for a transparency sheet sample. Due to the application of a typical transparency sheet, the sheet has an opaqueness that introduces inhomogeneity within the paper creating noise when observing the sample with laser light.

The transparency film used was rated at a thickness of 4 mil (or 102µm) without any specifications for index of refraction or tolerance range. Since transparency paper is composed of the material cellulose acetate which has an index of refraction of 1.46 – 1.5, 1.5 was used for the index [30]. It should be mentioned that absolute thickness accuracy of the measurement was not the main focus of this experiment, but that a confident measurement could be made. If accuracy is also of high importance, then explicitly calculating the refractive index of the material could be achieved by acquiring a slab of the material with a known thickness. Thus, this calculated refractive index would be used in all subsequent thickness measurements. The raw image for the transparency film is depicted in Figure 12a. For this measurement, the red (642nm) and infrared (847nm) lasers were used.

As is seen in the Figure 12a, this image shows a dominant speckle pattern. The frequency-based noise-reduction techniques were again applied to this image. Figures 12 b and c show the results of each respective technique.

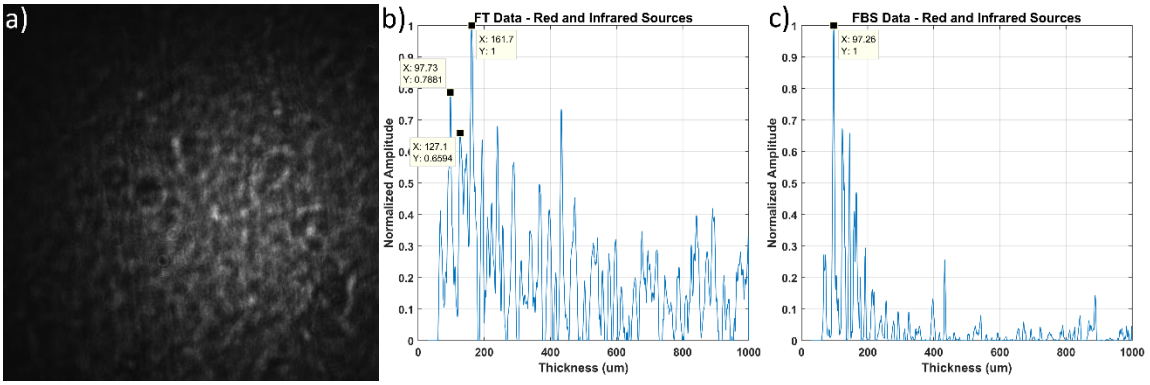

Figure 2-12: (a) shows raw interference image for combination of red (642nm) and infrared (847nm) lasers for transparency sheet measurement. (b) shows the Fourier Transform thickness result of the acquired image while (c) shows the Fractional BiSpectrum thickness result. The frequency domain has been scaled in terms of film thickness.

As can be seen in Figures 12 b and c, the thicknesses related to the signal frequencies (127μm for red and 98μm for infrared) did not correspond to the thickness with the greatest amplitude – in this case approximately 162μm. Thus, based on the FT alone, this measurement would in-fact predict the wrong thickness. Applying the FBS cross-correlation technique, however, leads to the proper measurement of 97μm. This thickness is most nearly the rated value of the transparency paper. Thus, in this case, not only did the FBS technique enhance the SNR of the data, but it was able to extract the proper result where the FT did not. It should be noted that in only applying the FBS for low or high frequency, there were other thicknesses above the SNR_I threshold amplitude other than that of the correct thickness, but applying the cross-correlation suppressed these other frequencies. Thus, utilizing both datasets helped to gain confidence in the measurement as predicted in the simulation section.

2.6 Conclusions

In this paper, a novel technique, Fractional BiSpectrum was introduced for enhancing the signal information in a noisy dataset if multiple inherent frequencies are present that have a known relationship with each other. This technique used the theory from the BiSpectrum as a starting point. The FBS technique can be used for the analysis of two inherent frequencies within a noisy dataset as long as the relationship between these two frequencies are known. Multiple simulations were performed to directly compare the FT to the FBS. The Monte-Carlo simulation showed that there was a strong SNR enhancement when the additive Gaussian noise power was five to nine times (-7dB to -9.5dB) that of the signal power.

It was then determined that a standard thickness measurement interference technique could be used to test the simulation results due to the dominant spatial frequency produced in one direction. A fully automated thickness measurement setup was constructed. The physical parameters of the system such as angle and spacing were identified and optimized. A fixture was developed for holding camera, sample, and fiber properly as well as a communication board for proper management of lasers from computer to drivers. The theoretical limits of the system were determined. Lasers with wavelengths of 520nm, 642nm, and 847nm were used for producing different combinations of two spatial frequencies to be analyzed.

The system's accuracy was verified using clean samples and comparing the thickness measurement to that of other metrology tools. For measurements over the whole theoretical range of the system, a five percent difference or less of inaccuracy was seen. The system was then used to make a measurement of layers with highly rough surfaces and turbid mediums.

Firstly, a clean location of the Blu-ray disk was measured and a thickness of 96μm was determined. The disk location was altered such that a noisy interference pattern would be generated. Both the FT and the FBS were able to calculate the proper coating thickness of the sample, however, while the uncertainty of the measurement using the FT was high, a confident measurement was taken with the FBS at 98μm.

Secondly, a transparency sheet was used as a sample under test. The rated thickness given by the manufacturer was 102μm. As was seen, a speckle-like pattern resulted from the transparency sheet. The FT of the measurement could not properly determine the thickness of the transparency sheet while the FBS could at 97μm.

It was shown that while this system may not outperform other thickness measurement systems in precision, the system would allow for higher certainty of measurement results as well as producing measurements where other systems would commonly break-down. The SNR enhancing technique illustrated in this article has shown value in extending the limitations in a standard measurement system and when applicable, can extend the measurement capability of other systems.

REFERENCES

[1] E. Hecht, The speckle effect, in: Optics, 5th ed., Pearson, Essex, 2017: pp. 634–635.

[2] O.S. Heavens, Optical properties of thin films, Reports Prog. Phys. 23 (1960) 1. http://stacks.iop.org/0034-4885/23/i=1/a=301.

[3] A. Bergauer, C. Eisenmenger-Sittner, Properties and characterization of thin films, in: Phys. Thin Film., Vienna, 2017: pp. 75–143.

[4] D.A. Miller, Optical properties of solid thin films by spectroscopic reflectometry and spectroscopic ellipsometry, The City University of New York, 2008.

[5] K. Pedersen, Ellipsometry, Aalborg, 2004.

[6] P. Hlubina, J. Lunacek, D. Ciprian, R. Chlebus, Spectral interferometry and reflectometry used to measure thin films, Appl. Phys. 92 (2008) 203–207.

[7] Filmetrics, Taking the mystery out of thin-film measurement, San Diego, 2012.

[8] R.H. Webb, Confocal optical microscopy, Reports Prog. Phys. 59 (1996) 427–471.

[9] C. Sheppard, Measurement of thin coatings in the confocal microscope, Micron. 32 (2001) 701–705.

[10] P.J.D. Groot, Coherence Scanning Interferometry, in: R. Leach (Ed.), Opt. Meas. Surf. Topogr., Springer, Middlesex, 2011: pp. 187–208.

[11] M.F. Fay, T. Dresel, Applications of model-based transparent surface films, Opt. Eng. 56 (2017).

[12] A.F. Fercher, W. Drexler, C. Hitzenberger, T. Lasser, Optical coherence tomography - Principles and applications, Reports Prog. Phys. 66 (2003) 239–303.

[13] A.G. Podoleanu, Optical coherence tomography, J. Microsc. 247 (2012) 1–11.

[14] M.I. Mokbul, Optical coherence tomography: Basic concepts and applications in neuroscience research, J. Med. Eng. 2017 (2017) 1–20.

[15] M. Mayer, Rutherford backscattering spectrometry (RBS), in: ICTP (Ed.), Work. Nucl. Data Sci. Technol. Mater. Anal., Garching, 2003: pp. 59–80.

[16] A. Glindemann, R. Lane, J.C. Dainty, Estimation of binary star parameters by model fitting the bispectrum phase, Opt. Soc. Am. 9 (1992) 543–548.

[17] C.L. Matson, Weighted-least-squares phase reconstruction from the bispectrum, Opt. Soc. Am. 8 (1991) 1905–1913.

[18] M.R. Raghuveer, C. Nikias, Bispectrum estimation: A parametric approach, IEEE Trans. Acoust. 33 (1985) 1213–1230.

[19] J.C. Marron, P.P. Sanchez, R.C. Sullivan, Unwrapping algorithm for least-squares phase recovery from the modulo 2pi bispectrum phase, J. Opt. Soc. Am. 7 (1989) 14–20.

[20] A. Totsky, V. Lukin, A. Zelensky, J. Astola, K. Egiazarian, G. Khlopov, V. Morozov, I. Kurbatov, P. Molchanov, A. Roenko, D. Fevralev, Bispectrum-based methods and algorithms for radar, telecommunication signal processing and digital image reconstruction, Tampere International Center for Signal Processing, 2008.

[21] G. Sundaramoorthy, M.R. Raghuveer, Bispectral reconstruction of signals in noise Amplitude reconstruction issues, IEEE Trans. Acoust. 38 (1990) 1297–1306.

[22] A.E. Cetin, An iterative algorithm for signal reconstruction from Bispectrum, IEEE Trans. Signal Process. 39 (1991) 2621–2628.

[23] T. Matsuoka, T.J. Ulrych, Phase estimation using the Bispectrum, Proc. IEEE. 72 (1984) 1403–1411.

[24] C. Nikias, A. Petropulu, Signal reconstruction from the phase of the Bispectrum, IEEE Trans. Signal Process. 40 (1992) 601–610.

[25] M. Abolbashari, S.M. Kim, G. Babaie, J. Babaie, F. Farahi, Fractional bispectrum transform definition and properties, IET Signal Process. 11 (2017) 901–908.

[26] L.A. Ageev, V.D. Yegorenkov, Lecture-room interference demo using a glass plate and a laser beam focused on it, Eur. J. Phys. 31 (2010) 801–809.

[27] N. Sultanovaa, S. Kasarovaa, I. Nikolov, Dispersion Properties of Optical Polymers, Acta Phys. Pol. A. 116 (2009) 585–597.

[28] Blu-ray Disc Association, White paper Blu-ray Disc format, 4th ed., Burbank Blu-ray Disc Association, Burbank, 2015.

[29] W.R. Mueller, Protective coatings for optical disc information recording media, and methods and apparatus for applying same, 5,935,673, 1999.

[30] J.F. Shackelford, Introduction to Materials Science for Engineers, 5th ed., McGraw-Hill, New York, 2000.

[31] E. Hecht, Gaussian laserbeams, in: Optics, 5th ed., Pearson, Essex, 2017: pp. 625–627.

2.7 Appendix A

For estimation of the period of the generated fringes, one must use Gaussian beam theory. Figure 2-13 shows the geometry of the system along with all the variables of interest.

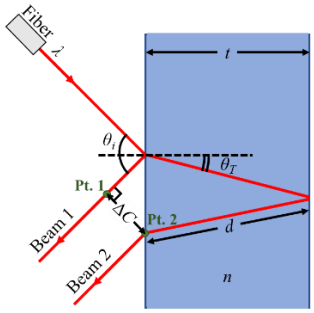

Figure 2-13: This shows the variables associated with geometry of developed system.

In Figure 2-13, θ_i is the incident angle, θ_T is the transmission angle, n is the refractive index of the sample, t is the thickness, $2d$ is the distance traveled by the beam internally, ΔC is the distance between beam centers, and Pts. 1 and 2 on respective beam paths interfere with each other. l_1 will represent the optical path length of light reflected from the top surface, starting from the tip of the fiber to Pt. 1, and l_2 will correspond to the optical path length of light reflected from the bottom surface, starting from the tip to Pt. 2. Δl will be defined as the optical path difference between the two beams. Calculating for Δl and ΔC can be done directly and is shown in Eqn. 2-6 and Eqn. 2-7.

$$\Delta l = \frac{2t}{\cos\theta_T}(n - \sin\theta_T \sin\theta_i) \tag{2-6}$$

$$\Delta C = 2t \tan\theta_T \cos\theta_i \tag{2-7}$$

From Gaussian beam theory [28], Eqn. 2-8 can be derived which determines the phase variation at the cross-section of a Gaussian beam at a particular distance from the waist of the beam.

$$\varphi_1(r,l) = kl + k\frac{r^2}{2R(l)} \tag{2-8}$$

where *r* is the radial distance from the center of the beam at a propagation length *l* from the waist of the beam. The waist of the beam, W_0 is defined as the smallest width of the beam. For the purposes of this experiment, the waist of the beam coincides with the fiber tip. The width of the beam at any particular propagation length *l* is defined as *W*, while curvature of the beam at a particular length is termed *R*. Finally, when the width of the beam becomes $\sqrt{2}W_0$, this specific propagation length is known as z_0. The basic Gaussian beam equations detailing the relationship of these parameters are shown in Eqn. 2-9, Eqn. 2-10, and Eqn. 2-11 [31].

$$R(l) = l\left[1 + \left(\frac{z_0}{l}\right)^2\right] \tag{2-9}$$

$$W(l) = W_0\sqrt{1 + \left(\frac{l}{z_0}\right)^2} \tag{2-10}$$

$$z_0 = \frac{\pi {W_0}^2}{\lambda} \tag{2-11}$$

where Eqn. 2-11 relates all the equations to the wavelength of the beam. As can be seen in Figure 2-13, Eqn. 2-8 can be related to each beam at l_1 and l_2. If one takes the difference between the two respective phases, this will deduce the expected interference pattern. Considering that z_0 is on the order of tens of microns for the wavelengths used in our application, for any reasonable propagation distance, *l*, of the beam, it can be assumed that $\frac{z_0}{l} \ll 1$. This will allow for Eqn. 2-9 and Eqn. 2-10 to reduce considerably giving rise to a simplification in Eqn. 2-8. Thus, the difference in phase giving rise to the interference pattern at l_1 and 1_2 is stated in Eqn. 2-12.

$$\Delta\varphi = k \cdot \Delta l + \frac{k}{2}\left[\frac{l_1 {r_2}^2 - l_2 {r_1}^2}{l_1 l_2}\right] \tag{2-12}$$

Since r_1 and r_2 are a function of radial spatial parameters, both will be characterized in terms of x and y with respect to r_1. To relate the two, ΔC must be used. If x is set to the axis from pt.1 to pt. 2 as depicted in Figure 2-13, Eqn. 2-12 can be restructured to Eqn. 2-13.

$$\Delta\varphi = k\left[\Delta l + \frac{\Delta C^2}{2l_2}\right] - \frac{k\Delta C}{l_2}x - \frac{k\Delta l}{l_1 l_2}(x^2 + y^2) \tag{2-13}$$

The first term of the Eqn. 2-13 relates to the initial phase of the spatial frequency and is not important for the purposes of our application. The second term relates to the dominant spatial frequency of interest that should be further analyzed. The third term gives rise to a wide bandwidth of frequencies that should be reduced as much as possible. To reduce this third term, it can be shown that the width of the beam at l_1, $W(l_1)$, should be much greater than the displacement in center of the two beams, ΔC.

Furthermore, to determine the period of the spatial frequency from the second term, one should find a phase difference equal to 2π. This is shown in Eqn. 2-14.

$$2\pi = \frac{k\Delta C}{l_2}P \tag{2-14}$$

where P is the spatial period in x-direction. Replacing ΔC with the terms in Eqn. 2-7 and replacing k with $2\pi/\lambda$, we can arrive at the final form for the spatial period.

$$P \approx \frac{\lambda \cdot l_2\sqrt{n^2 - \sin^2\theta_i}}{2t \cdot \sin\theta_i \cos\theta_i} \tag{2-15}$$

2.8 Appendix B

With respect to determination of the thickness measurement range limits, the following procedure was used.

1) Maximum Thickness

As a constraint in our equations for minimizing the second order effects in our optical interference pattern, we have:

$$\frac{W(l_1)}{\Delta C} \gg 1 \tag{2-16}$$

Using Eqn. 2-9 and Eqn. 2-10 and assuming l_1 far from z_0, we can represent the numerator of Eqn. 2-16 as:

$$W(l_1) = \frac{\lambda \cdot l_1}{\pi W_0} \tag{2-17}$$

and based on the geometry of our setup, the denominator can be replaced as:

$$\Delta C = \frac{2t \cdot sin\theta_i \, cos\theta_i}{\sqrt{n^2 - sin^2\,\theta_i}} \tag{2-18}$$

where t is the thickness of the sample, θ_i is the incident angle of the beam with respect to the sample, n is the index of refraction of the sample, λ is the wavelength of the source, and W_0 is the radius of the beam at the tip of the fiber. Substituting these terms into Eqn. 2-18, we have:

$$\frac{\lambda \cdot l_1 \sqrt{n^2 - sin^2\,\theta_i}}{2\pi \cdot t \cdot W_0 \, sin\,\theta_i \, cos\,\theta_i} \gg 1 \tag{2-19}$$

Using simulation with no approximations, it was determined that the minimum allowable value for Eqn. 2-16 was 3 where the linear phase component still dominated the second-order term in Fourier space by a factor of 5. Rearranging this constraint to solve for t and setting the ratio equal to 3 for the absolute maximum allowable thickness, we have:

$$t_{max} \approx \frac{\lambda \cdot l_1 \sqrt{n^2 - sin^2\,\theta_i}}{6\pi \cdot W_0 \, sin\,\theta_i \, cos\,\theta_i} \tag{2-20}$$

Using our previously determined optimum parameters for the setup, θ_i = 50°, n = 1.5, W_0 = 1.37μm (from NA= 0.12), λ = 650nm, and l_1 = 2cm, we can solve for the maximum thickness to be approximately 1318μm. Thus, we set the upper limit of our thickness measurement to **1300μm**.

2) Minimum Thickness

It was determined in simulation that a minimum of 6 fringes within the FOV are needed to produce a reasonable resolution in thickness - which is directly related to the resolution limitations in the frequency domain. This limitation restricts the minimum thickness that can be measured and thus, we need to estimate the number of fringes. Since this period is occurring at the surface of the sample under measurement, we need to determine the number of fringes based on the spot size occurring at that location. This is shown below:

$$F_T = \frac{2W(l_1)}{P} \tag{2-21}$$

where F_T is the estimated total fringes in the interference pattern. Substituting the terms from Eqn. 2-17 and Eqn. 2-15, we get:

$$F_T \approx \frac{4t\ sin\,\theta_i\ cos\,\theta_i}{\pi W_0\sqrt{n^2 - sin^2\,\theta_i}} \tag{2-22}$$

Setting the fringe value to 6 for the minimum detectable thickness and rearranging in terms of t, we will get:

$$t_{min} \approx \frac{6\pi W_0\sqrt{n^2 - sin^2\,\theta_i}}{4\ sin\theta_i\ cos\,\theta_i} \tag{2-23}$$

Again, substituting the optimum parameters of $\theta_i = 50°$, $n = 1.5$, and $W_0 = 1.37\mu m$, we get a minimum thickness of approximately 16.9μm. Thus, we will set the lower limit of the thickness range to **17μm.**

3) Summary

It was determined that the upper limit of the thickness measurement was based on the beam spot size at the surface of the sample being much greater than the distance between the two interference beams. This is necessary to reduce the second order effects on the interference pattern. The lower limit in contrast was determined based on the lowest number of fringes that would give a reasonable resolution in thickness data. After applying the Gaussian optics theory and using the optimum parameters previously determined, an acceptable thickness range of **17μm to 1300μm** was determined.

CHAPTER 3: SNR ENHANCEMENT TECHNIQUE: A NEW APPROACH TO FRACTIONAL BISPECTRUM[2]

Abstract

When the power of noise within the time-domain or spatial domain signal is comparable to that of the power of the signal information itself, it is difficult to extract the desired information. Typically, the Fourier transform can be used that localizes the signal in the frequency domain and distributes the noise; however, the limit is always defined by the Signal-to-Noise Ratio (SNR). The Fractional BiSpectrum was developed to enhance the signal information to a higher degree than that of the Fourier transform if there exist multiple frequencies within the signal that has a well-known relationship. This technique has been demonstrated on a typical interferometry setup with two inherent frequencies. The previously demonstrated Fractional BiSpectrum technique used one image for capturing both sets of frequency content within the image. In this letter, we introduce a modified Fractional BiSpectrum technique using the same setup configuration but where these two spatial frequencies are separated into two respective images. We will show that this method further improves the SNR relative to earlier reported results.

3.1 Introduction

The Fourier Transform (FT) can be used for the purpose of identifying, and filtering out frequencies from a time-domain or spatial-domain signal. For the purpose of extracting a specific frequency or frequency content from a very noisy signal, the FT may be used, but will eventually break down if the power of the noise within the signal bandwidth (in

[2] This paper is submitted to the IET Electronics Letters journal.

Fourier domain) is comparable to that of the signal power contained within the respective bandwidth. For such scenarios, the BiSpectrum was developed, which would naturally suppress the Gaussian noise within the signal if a specific relationship existed between frequencies inherent within the signal [1,2]. This technique was used in many different areas of research, such as astronomy, radar, telecommunication, geophysics, biomedicine, and digital image processing to name a few [2]. This technique however, requires that inherent signal frequencies have a specific relationship with each other.

Thus, the Fractional BiSpectrum (FBS) technique was developed to enhance the Signal-to-Noise Ratio (SNR) if there are multiple signal frequencies present within the signal which have a known relationship but do not have the specific relationship requirement of the BiSpectrum [3,4]. For the purposes of this manuscript, we will consider only two inherent frequencies within the signal, f_1 and f_2. While the general form of the FBS contains a magnitude and phase component for a two-dimensional frequency space, for this manuscript, only the magnitude of the diagonal frequencies within the FBS are analysed as formulated in (1) [3].

$$|FBS[p;K]| = |X[p]|^2 \cdot |X[K \cdot p]| \tag{3-1}$$

where $X[.]$ is the FT of $x[n]$ - which is the discrete signal over $n = 0,1,\ldots,N\text{-}1$ and N is the signal length, p is the frequency index ranging from $-N/2$ to $N/2-1$, and $K = f_2/f_1$. As f_1 and f_2 are not known directly, K only requires the relationship between the two frequencies to be known. It was shown that when an additive Gaussian noise was introduced, the FBS was able to suppress the noise to a higher degree than that of the Fourier transform analysis [3].

For the purpose of demonstrating this SNR enhancement technique, we used a standard interferometry technique for measuring film thickness. Figure 3-1 depicts the geometry of the film thickness measurement technique.

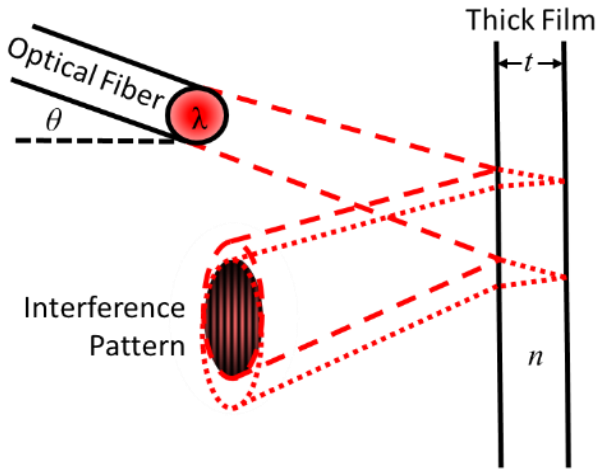

Figure 3-1: Depiction of interferometry setup used to measure film thickness

As is seen in Figure 3-1 for this interferometer, laser light reflected from the top and bottom surface of the thick-film will interfere where the two beams intersect. For a specific θ, incident laser angle with respect to the normal of the film, λ, wavelength of the laser light, and t, thickness of the thick-film, this interference will generate a fringe pattern with a dominant spatial period related to (2) [3],

$$P \approx \frac{\lambda \cdot l\sqrt{n^2 - \sin^2\theta_i}}{2t \cdot \sin\theta_i \cos\theta_i} \tag{3-2}$$

where l is the length of the light path from the fiber tip to the surface of the sensor plane. As can be seen, the spatial period is directly proportional to λ. Thus, if multiple lasers were used with different wavelengths, the ratio between spatial frequencies can be readily determined assuming all other variables remain fixed. Therefore, the FBS technique can be used in such a thickness measurement system for enhancing the SNR.

It should be noted that there are many other techniques for the purpose of measuring the thickness of films [5-9], but these techniques will commonly break down if the boundary surface is too rough or the film material is too turbid.

3.2 Fractional BiSpectrum Separated Technique

It was theorized that since two sources are used for film-thickness measurement simultaneously, both sources introduce a respective noise component related to scattering as seen in the image of the measurement. Thus, while the signal amplitude from each individual inherent frequency in the Fourier-domain is not being increased, the underlying noise power which is statistically spread throughout the spectrum is increased due to the presence of two sources. Therefore, there was a desire to separate the fringe patterns into separate images where the average noise power in each respective Fourier domain would remain lower while each inherent frequency amplitude would be the same. Separating inherent frequencies could be accomplished by cycling through each respective laser source. Thus, the magnitude of the discrete Fractional BiSpectrum Separated (FBSS) technique was formulated, shown in Eqn. 3-3.

$$|FBSS[p;K]| = |X_1[p]|^2 \cdot |X_2[K \cdot p]| \tag{3-3}$$

where X1 and X2 are the discrete Fourier transforms of each respective signal x1 and x2, where x1 and x2 each contain one respective inherent frequency. In this case, K is determined based on which signal is defined as X1 and X2. As these two signals can easily be interchanged, this can be used to increase the robustness in the measurement taken.

3.3 Experimental Results

In this section a thick-film measurement of an optically turbid Blu-ray disk coating was then performed. The thickness of this coating is used for protective purposes and has a nominal thickness of 98μm and refractive index of 1.55 [10].

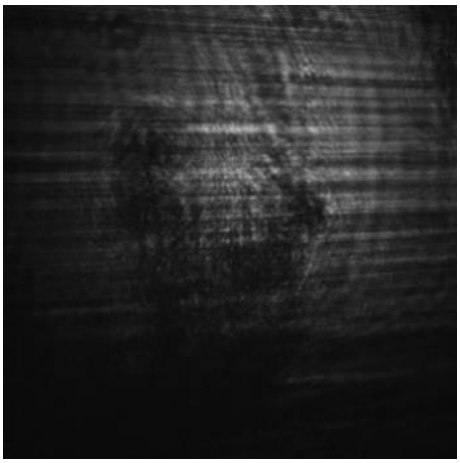

Figure 3-2: Blu-ray disk interference pattern with both green and infrared laser light illuminating sample.

While three sources were used for measurement – 520nm, 642nm, and 847nm – for robustness, a relatively high noise power was observed within the interferogram image when a combination of green and infrared lasers are used. Figure 3-2 depicts the raw image from the measurement. Figure 3-3 depicts the results from the FT and FBS for determination of thickness.

Figure 3-3a shows thicknesses related to signal frequencies, 96.3μm and 159μm are much lower in amplitude than other peaks resulted from the noise. In particular, the noise related to the frequency of 347μm is dominant. Using the FBS technique in Figure 3-3b, we see that since the initial SNR was very poor, the amplitude related to the actual thickness, 97.7μm is dominated by the noise.

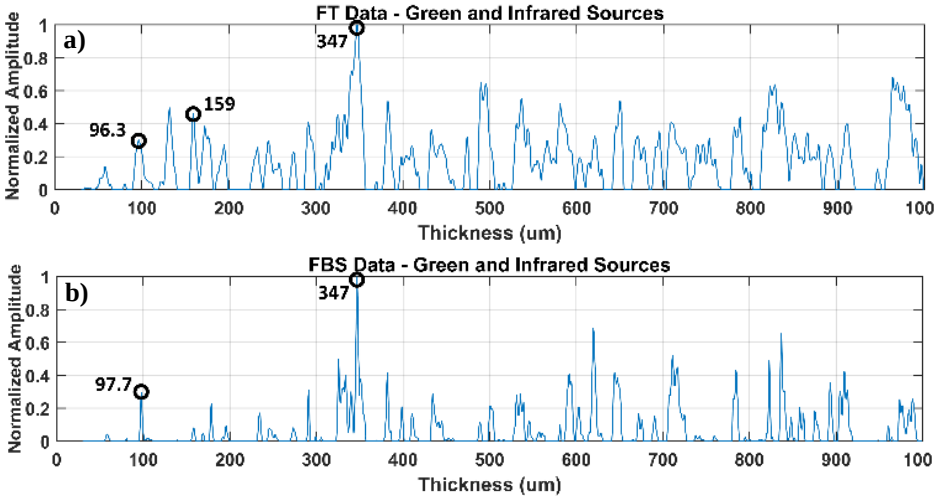

Figure 3-3: Thickness results of Blu-ray disk sample for FT and FBS based on the interference image using 520nm and 847nm lasers simultaneously. The frequency domain has been scaled in terms of film thickness. (a) FT of image in the x-direction, and (b) FBS of image in the x-direction.

We then alternated the green and infrared sources and took an image for each respective wavelength. These images are shown in Figure 3-4.

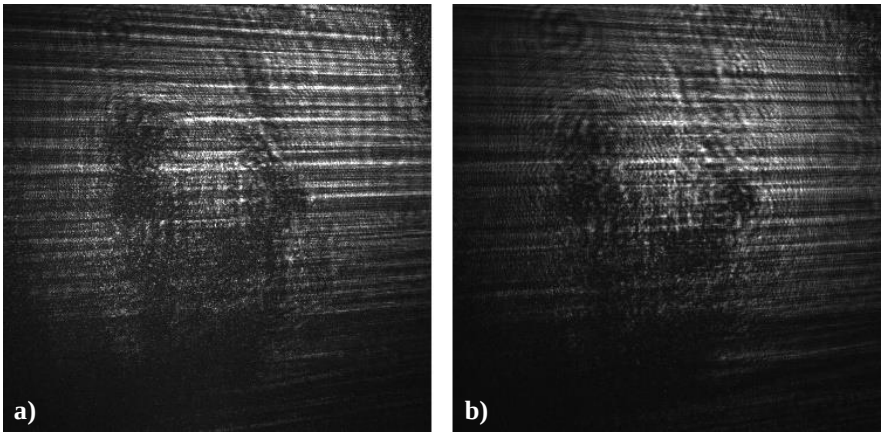

Figure 3-4: Separated interference patterns for 520nm and 847nm wavelengths. (a) Interference pattern for the green source, and (b) Interference pattern for the near infrared source

Figure 3-4a and 3-4b are the interference patterns – each containing an individual spatial-frequency component in the *x*-direction. Using the FBSS technique introduced in

section 3.2, we then attempted to measure the thickness of the turbid film. The results are depicted in Figure 3-5.

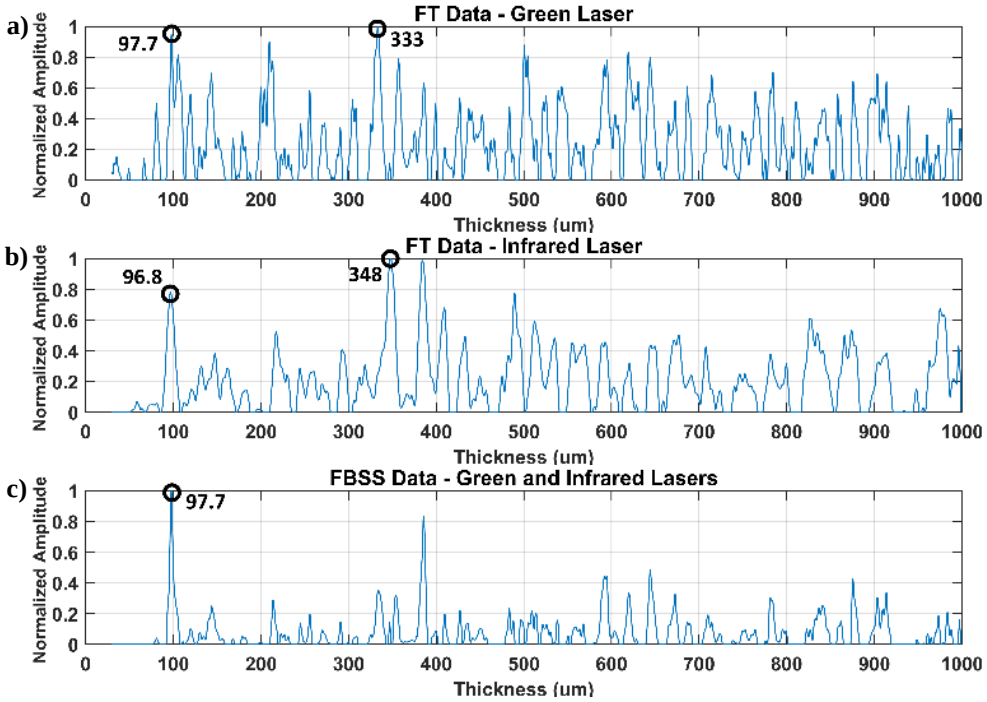

Figure 3-5: Thickness results of Blu-ray disk sample based on the separation of signals into multiple images. The frequency domain has been scaled in terms of film thickness. (a) FT result for 520nm of Blu-ray disk sample, (b) FT result for 847nm of Blu-ray disk sample, and (c) FBSS result combining images with each wavelength

As is seen in Figure 3-5a and 3-5b, the FT performed on each respective interference pattern again failed to identify the proper thickness as the relative power of noise (resulting from the surface scattering) was too high. However, when the FBSS technique was used, Figure 3-5c, the SNR was enhanced and thus the maximum amplitude observed in the respective domain occurred at a thickness of 97.7μm. This closely corresponds to the thickness of the Blu-ray disk coating of 98μm as specified in the industry.

Furthermore, while the FBS was not able to identify the proper thickness as the power of noise related to scattering was too high, the FBSS was in fact able to. Thus, if

there is the ability to separate the source frequencies into individual signals, this should be implemented due to the inherent nature of the noise, as confirmed from the experimental results.

3.4 Conclusions

We introduced an SNR enhancement technique, Fractional BiSpectrum Separated, that would enable proper thickness measurement from the separation of respective frequencies. As was observed, for a particularly high level of noise power relative to the signal, the FT and FBS techniques failed to identify the proper thickness. When separating the interference patterns into respective images however, we were able to observe the proper thickness using the FBSS technique. Thus, our hypothesis was confirmed for the purpose of providing an SNR enhancement technique which relates inherent frequencies within separate signals.

REFERENCES

[1] Glindemann, A., Lane, R., Dainty, J.C.: ‘Estimation of binary star parameters by model fitting the bispectrum phase’Opt. Soc. Am., 1992, 9, (4), pp. 543–548.

[2] Totsky, A., Lukin, V., Zelensky, A., et al.: ‘Bispectrum-based methods and algorithms for radar, telecommunication signal processing and digital image reconstruction’ (Tampere International Center for Signal Processing, 2008)

[3] Babaie, J., Abolbashari, M., Farahi, N., Kim, S.M., Farahi, F.: ‘Optical Film Thickness Measurement of Turbid Materials using the Fractional BiSpectrum Noise-Reduction Technique’Opt. Commun., 2019, TBD, (TBD), p. TBD.

[4] Abolbashari, M., Kim, S.M., Babaie, G., Babaie, J., Farahi, F.: ‘Fractional bispectrum transform definition and properties’IET Signal Process., 2017, 11, (8), pp. 901–908.

[5] Bergauer, A., Eisenmenger-Sittner, C.: ‘Properties and characterization of thin films’, in ‘Physics of Thin Films’ (2017), pp. 75–143

[6] Pedersen, K.: ‘Ellipsometry’ (2004)

[7] Hlubina, P., Lunacek, J., Ciprian, D., Chlebus, R.: ‘Spectral interferometry and reflectometry used to measure thin films’Appl. Phys., 2008, 92, (2), pp. 203–207.

[8] Sheppard, C.: ‘Measurement of thin coatings in the confocal microscope’Micron, 2001, 32, pp. 701–705.

[9] Groot, P.J.D.: ‘Coherence Scanning Interferometry’, in Leach, R. (Ed.): ‘Optical measurement of surface topography’ (Springer, 2011), pp. 187–208

[10] Blu-ray Disc Association: ‘White paper Blu-ray Disc format’ (Burbank Blu-ray Disc Association, 2015, 4th edn.)

CHAPTER 4: PROFILOMETRY MEASUREMENT USING FRACTIONAL BISPECTRUM ON ROUGH SURFACES[3]

Abstract

In practice, the measurement of rough surfaces using laser interferometry is very difficult as the reflected light from the surface scatters and tends to dominant to the interference fringes. Using Wavelength Scanning Interferometry, phase unwrapping as a technique to determine the surface profile is not necessary as each pixel can independently extract the height information. Thus, the measurement is more resistant to spatial noise in the interference pattern generated from sources such as scattered light. Furthermore, with the use of the Fractional BiSpectrum, the signal corresponding to the temporal frequency inherent within each pixel intensity variation can be enhanced with respect to speckle noise.

4.1 Introduction

Laser Interferometry is a common technique in optical metrology for the measurement of physical characteristics of a sample such as refractive index, surface profile, etc. with very high resolution [1-4]. An example of such interferometer is a Michelson configuration where an incoming beam from a light source is divided to travel into two different paths, Figure 4-1 [5,6]. One is used as a reference and the second beam will travel toward a sample of interest. The reflection of the beam from the sample will form the optical path length of the cross-section of the beam point by point. The two beams will then be recombined, and a spatially varying interference pattern is then detected by the camera based on the differences in the phase of the beams from the two paths. There are

[3] This paper is submitted to the Applied Optics, OSA journal.

well known methods that can be used to measure the phase difference of the two beams from recorded intensity pattern [7].

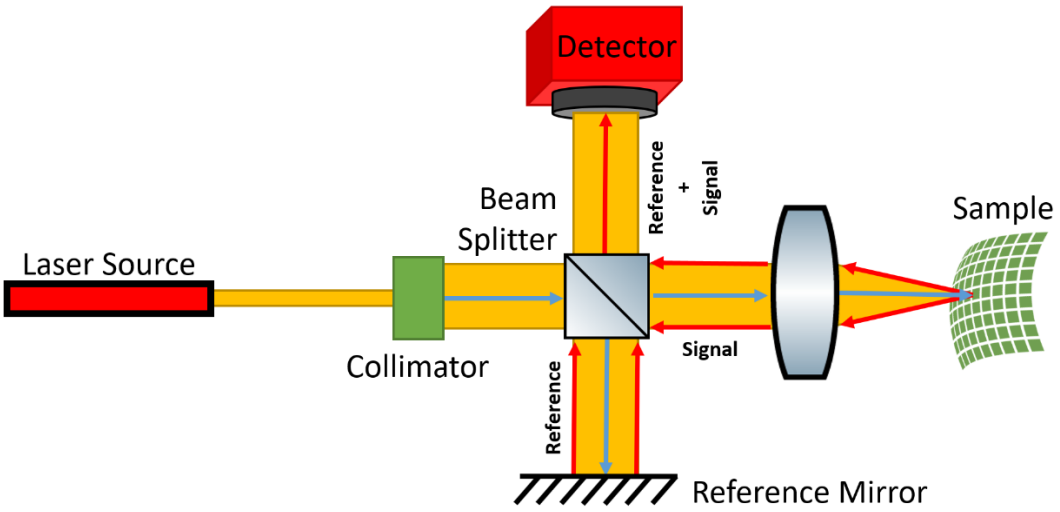

Figure 4-1: This shows the Michelson interferometer configuration to be used for measurement of a surface profile.

Since the map of the phase difference corresponds to surface height variation with respect to an optically flat surface, this interference pattern can be used to directly find the profile of the sample of interest. The equation defining the relation between the phase in the interference pattern and the respective height is formulated in Eqn. 4-1.

$$\phi(x,y) = \frac{4\pi}{\lambda} h(x,y) \tag{4-1}$$

where ϕ is the phase over the area of the interference pattern, λ is the wavelength of the laser source and h is the corresponding height over the illuminated area of the sample. While solving for the phase between 0 and 2π radians may be trivial, since the corresponding height is not necessarily restricted to this range, a phase-unwrapping technique must be used. To use a phase-unwrapping technique properly, however, there

must be minimal noise in the phase map as well as a continuity in phase from pixel to pixel, i.e. no phase jumps greater than π, to avoid ambiguity [7-9].

One of the advantages of a laser interferometer over other measurement techniques is the precision in resolution. Using the interferometry technique described above will give a resolution in height of $\lambda/500$ assuming 8-bit resolution in the intensity detection and minimal other noise sources [10]. In practice this is not achievable however, as detectors are not perfectly linear, they have an inherent noise level, and pixel bleeding exists [7,10,11]. Also, any presence of turbulence in the physical setup will affect the overall resolution. Thus, to enhance the resolution, techniques such as Phase-Scanning/Shifting Interferometry, or PSI, have been developed [12,13]. This uses a single wavelength source and shifts the sample offset to achieve exact phase steps for each pixel. Doing so up to 5 times or more dramatically reduces noise effects and can allow for sub-nm resolution [4,13].

The other major advantage – in particular over Coordinate Measuring Machines (CMMs) – is the ability to measure a surface area in a relatively short time without any contact. As there is no mechanical scanning involved, there are no errors related to mechanical drift and a high-density surface map can be generated within one image capture with minor processing.

Laser interferometry also has some disadvantages. Unwrapping techniques introduce a limitation as there should be no phase jumps between pixels in the interference pattern greater than π as unwrapping would not correctly predict the absolute phase [8]. This will limit the maximum slopes that may be present within the sample of interest which would also inherently limit the total measurement height range based on the pixel width of

the camera. Furthermore, since the spatial resolution is determined by the pixel size of the camera in combination with any optics placed in-front of the camera, sample features smaller than this resolution will get washed out due to pixelization of sensor [4]. Since this pixelization also impacts the phase-map, it could lead to height errors in neighboring pixels as well. Hence, while this technique works well for finding the form on optically smooth components, it is not recommended to use this interferometry technique for samples having a roughness, or S_a, greater than the order of $\lambda/4$. This is due to the fact that $\lambda/4$ corresponds to phase differences of π by Eqn. 4-1, and the roughness parameter, S_a is not defined as the maximum height variation present in the sample. Thus, there is potential for phase jumps between pixels greater than π especially where roughness features may be very dense as compared to the area of a pixel.

One laser interferometry technique developed to overcome some of these limitations is known as Wavelength Scanning Interferometry, or WSI [14-16]. Instead of using a singular wavelength source, a tunable laser is used which can scan over a large wavelength range. This technique does not use spatial information to determine the phase, but instead uses how each individual pixel intensity varies as a function of wavelength.

However, all laser interferometry techniques, including WSI, will suffer significantly from another phenomena. When a surface has a high level of roughness, due to the randomness of the surface structure, there will be many instances where rays of light propagating from different locations on the sample will interfere. When the phases from these scattered reflections are completely in phase or out of phase at the detector plane, they will be observed as localized high and low intensity points – known as a speckle

pattern - that tend to dominate the desired interference signal [17]. A typical speckle pattern is shown in Figure 4-2.

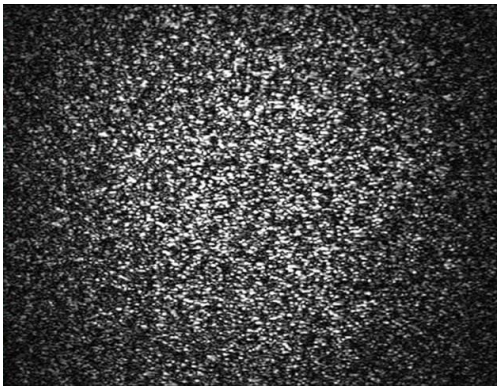

Figure 4-2: This shows the typical laser speckle pattern produced by rough surface.

When roughness of a surface reaches a particular roughness threshold, i.e. greater than $\lambda/4$ (see Appendix A), the power of the noise directly related to the speckle content will not be negligible as compared to the power related to the signal of the measurement. Furthermore, we can reason that when a sample containing a high degree of roughness produces a noise power within 3dB of the signal power or greater, a confident measurement can no longer be made. If the signal detection (or measurement analysis) can be performed in the Fourier domain, this has the advantage of tolerating a higher level of noise power. However, Fourier analysis, has its own limits as well. To extend this roughness threshold where a confident measurement can still be made, we apply Fractional BiSpectrum (FBS) to Wavelength Scanning Interferometry for the purpose of measuring the profile of a rough surface [18].

We have previously introduced Fractional BiSpectrum as a noise reduction technique and used this technique for film thickness measurement [18,19]. In this paper, we will briefly re-introduce this noise-reduction technique in section 4-2. Section 4-3 will

discuss the interferometry technique used for profile measurement and how the FBS algorithm can be applied. The experimental setup including the physical configuration as well as the automated software developed will be presented as well in section 4-3. In section 4-4, measurement of smooth and rough surfaces will be analyzed and Section 4-5 will summarize the results of this work.

4.2 Fractional BiSpectrum

In our previous work, we introduced the Fractional BiSpectrum, or FBS, which allows for enhancing the SNR as compared to that of the FT by using inherent relationships between frequencies present within the signal [18,19]. The FBS was developed as an extension of the BiSpectrum and maintains the property of suppressing the Gaussian noise inherent within the signal [20]. The magnitude component of the diagonal cross-section of the discrete FBS for a finite, real-valued discrete spatial or time-domain signal is shown in Eqn. 4-2.

$$|FBS[p;K]| = |X[p]|^2 \cdot |X[K \cdot p]| \tag{4-2}$$

where p is the frequency indices in FBS space that range from $-N/2$ to $N/2-1$, N is the finite data length of the discrete spatial or time-domain data, $x[.]$, $X[.]$ is the discrete FT of $x[.]$, and K is a constant value which relates the inherent frequencies present within $x[.]$. It is important to know the expected relationship between the inherent frequencies for determination of the proper K value to use this signal enhancement technique properly. For example, if there were two frequencies present within the spatial-domain data f_1 and f_2, the relationship $f_2/f_1 = K$ should be known such that $|X[f_1]|^2 \cdot |X[K \cdot f_1]|$ would result in an enhanced magnitude where $p = f_1$.

While this theory was experimentally verified in a film thickness measurement setup to produce better results than that of the FT, it was then demonstrated that if the two frequencies of interest were separated into separate signals, the relative noise level would be lower than if the two were combined within one signal [19]. This was due to the principle that each source producing each respective frequency also produced a noise power related to the amplitude of this frequency. When both inherent frequencies were present within the same signal, while inherent frequency powers are separated into respective frequencies, the combination of the noise powers were spread over all frequencies and thus creating a higher relative noise threshold in the frequency domain. Therefore, the Fractional BiSpectrum Separated, or FBSS, technique was developed to achieve the same Gaussian noise suppression qualities as the FBS while maintaining the natural advantage of the separated signals [19]. The magnitude component of the discrete FBSS of a set of real-valued discrete spatial or time-domain data is shown in Eqn. 4-3.

$$|FBSS[p;K]| = |X_1[p]|^2 \cdot |X_2[K \cdot p]| \tag{4-3}$$

where X_1 and X_2 are the discrete Fourier transforms of each respective signal x_1 and x_2, where x_1 and x_2 each contain one respective simple harmonic function. In this case, K is determined based on which signal is defined as X_1 and X_2.

Next, we apply this technique to a highly explored interferometry technique, Wavelength Scanning Interferometry.

4.3 Wavelength Scanning Interferometry

4.3.1 Wavelength Scanning Interferometry Theory

Wavelength Scanning Interferometry (WSI) is a technique that varies the input wavelength, λ, by shifting in precise steps, $\Delta\lambda$, over a range of wavelengths to create a

variation in the interference pattern [14-16,21-23]. WSI differs from Phase-Shifting Interferometry (PSI) as it does not use a set number or bucket of wavelengths that will alter the phase of the interference pattern spatially and use a 'phase-unwrapping' technique to recover the absolute height. Instead, WSI uses wavelength-dependent variation in intensity of each pixel to determine the height since the frequency of each pixel in the detector is directly related to the path imbalance associated with that pixel. In practice this wavelength-dependent frequency is a temporal frequency since the wavelength is scanned in time. The common WSI setup is shown in Figure 4-3.

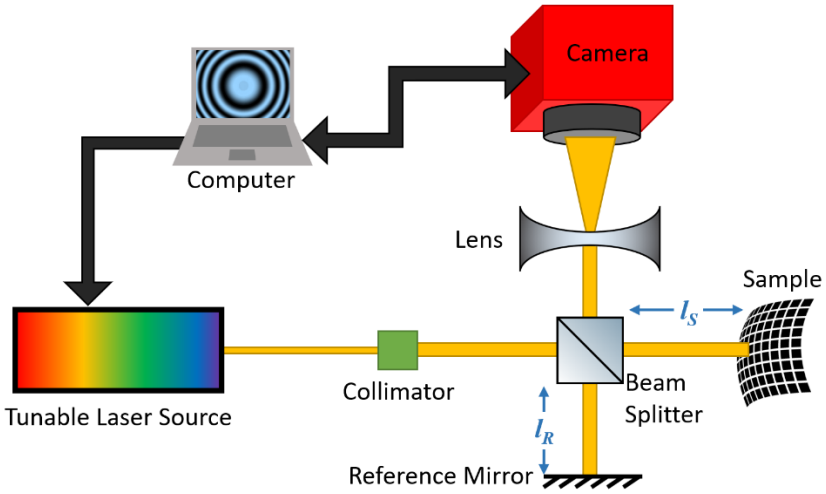

Figure 4-3: This shows the Wavelength Scanning Interferometry setup used in Michelson configuration.

As is seen, light from a tunable laser is incident onto a beam-splitter. The light travels down the two paths - to the sample of interest and a reference mirror - and is recollected by the beam-splitter. The interfered light is then detected by a camera with a lens used for setting the proper FOV on the sample. The phase with respect to each path are as follows in Eqn. 4-4.

$$\varphi_s(x,y) = 2\frac{2\pi}{\lambda_i} l_s(x,y), \qquad \varphi_R = 2\frac{2\pi}{\lambda_i} l_R \tag{4-4}$$

where φ_s and φ_R are the phase of each respective path, l_s and l_R are the path lengths from the beam-splitter to the test surface and the mirror surface, λ_i is the *i*th wavelength where i = 0,1,2,...,N-1 and there are N wavelengths in the scan, and x and y are the spatial parameters related to the sample of interest. As can be seen, the reference mirror is assumed to be optically flat and is perpendicular to the light path and thus, the phase associated with it should not vary in x and y. It can then be shown that based on the phase of each respective beam, the phase of the interference pattern is as follows,

$$\phi_i(x,y) = \frac{4\pi}{\lambda_i} h(x,y) \tag{4-5}$$

where ϕ_i is the interference phase at the *i*th wavelength, and h is the difference in respective beam paths, which corresponds to the surface height. As can be noted from this equation, there should be a substantial offset between reference and sample to create an observable variation in phase as the wavelength is varied. The variation in the interference phase with step in wavelength, $\Delta\lambda$, is as follows,

$$\Delta\phi_{kj}(x,y) = \phi_k(x,y) - \phi_j(x,y) = \frac{4\pi h(x,y)\Delta\lambda_{kj}}{\lambda_k\lambda_j}, \tag{4-6}$$

where $\Delta\phi_{kj}$ is the step in phase related to the step in wavelength, $\Delta\lambda_{kj} = \lambda_k - \lambda_j$. For $k = i + 1$ and $j = i$, it can be shown

$$\Delta\phi_{i(i+1)}(x,y) = \frac{4\pi h(x,y)\Delta\lambda_{i(i+1)}}{\lambda_i(\lambda_i + \Delta\lambda_{i(i+1)})}. \tag{4-7}$$

where $\Delta\phi_{i(i+1)}$ is the step in phase related to the step in wavelength, $\Delta\lambda_{i(i+1)}$, and both are in terms of the *i*th wavelength in the scan. It is noted that the number of phase steps and wavelength steps has a length of one less than the total number of wavelengths as expected.

As can be seen from Eqn. 4-7, $\Delta\phi_{i(i+1)}$ could be approximated to a constant shift in phase if $\Delta\lambda_{i(i+1)}$ was also constant, $\lambda_{i(i+1)}$ was set to the average wavelength over the scan, and the scanning range was much less than that of the average wavelength. However, for the purpose of our measurement and to achieve an adequate resolution in profile, this approximation breaks down as the wavelength scan range is not negligible as compared to the average wavelength. Since there is a desire to have uniform sampling with respect to the phase of the temporal signal of interest, $\Delta\phi_{i(i+1)}$ must be constant over the entire scan. Failure to do so would spread the power of the signal to a wider bandwidth of frequencies in the FT and FBSS domains or create additional complications when performing the respective transforms of the non-linear sampling in the temporal domain. Thus, to achieve constant phase steps, the wavelength must be stepped in a non-linear fashion. It can be shown that Eqn. 4-8 can be used for determination of all respective scan wavelengths. The derivation of this equation is shown in Appendix B.

$$\lambda_i = \frac{\lambda_{i-1} \cdot \lambda_{i-2}}{2 \cdot \lambda_{i-2} - \lambda_{i-1}}, \qquad for\ i = 2\ to\ N-1 \tag{4-8}$$

To use Eqn. 4-8, λ_0 and λ_1 must be known, so initial wavelength of scan and initial step of scan must be known. Thus, stepping in this non-linear way will allow the simplification of Eqn. 4-7 and the constant phase will be in relation to the initial wavelength and initial step as shown in Eqn. 4-9.

$$\Delta\phi(x,y) \cong \frac{4\pi h(x,y)\Delta\lambda_{01}}{{\lambda_0}^2} \tag{4-9}$$

For a variation in phase, $\Delta\phi_p$ of 2π, we can say the detected pixel intensity has completed one period. While this period is with respect to the change in wavelength of the

source, since the wavelength domain is not sampled uniformly, we will demonstrate the period as p samples in the index domain, i. This is depicted in Figure 4-4.

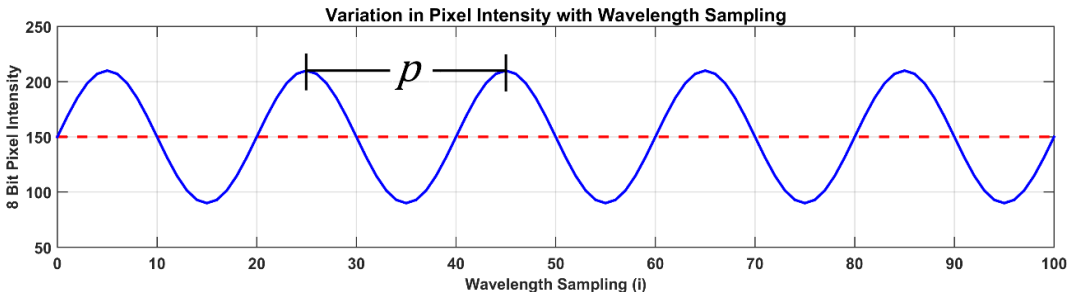

Figure 4-4: This shows the variation in pixel intensity with wavelength scanning indices. The number of sampling per period is denoted as p within the graph.

As can be seen, a width of p samples will generate one period of pixel intensity variation as corresponding to the wavelength scan. Thus, we can set $\Delta\phi_p$ equal p times $\Delta\phi$, or the phase step corresponding to a sampling width of 1. Substituting 2π in for $\Delta\phi_p$ and using Eqn. 4-9, we can solve for the height profile, h as follows.

$$h(x,y) = \frac{{\lambda_0}^2}{2p(x,y)\Delta\lambda_{01}} \tag{4-10}$$

Using properties of the discrete Fourier transform, it is known that the corresponding temporal frequency index, k_f, must be equal to N/p, where the frequency indices range from $k = -N/2$ to $N/2 - 1$ [24]. Thus, the height profile can be found with respect to the frequency index k_f, as follows.

$$h(x,y) = \frac{{\lambda_0}^2 k_f(x,y)}{2N\Delta\lambda_{01}} \tag{4-11}$$

Thus, since the index k_f can be found by using the magnitude of the Fourier transform and finding the index associated with the maximum amplitude, we can determine the height at each pixel location and hence the surface profile of the sample of interest. It

should be noted that for convenience, the Fourier domain can be mapped in terms of height by use of Eqn. 4-11. Thus, using discrete Fourier properties, all criteria such as minimum and maximum heights as well as resolution in height can be determined [24].

4.3.2 Applying Fractional BiSpectrum

While scanning through the set of wavelengths can achieve a height measurement, when the level of roughness on the surface is high, it could skew the interference images and produce pixel intensity patterns that contain a high level of noise. If the noise power level becomes too significant relative to the signal, the Fourier transform will break-down and the proper height determined will either be incorrect or there will be a high level of uncertainty in the result.

However, since we have control of how the tunable laser cycles through the wavelengths, we can adjust the initial wavelength step, $\Delta\lambda_{01}$ and thus the sampling over the wavelength range due to Eqn. 4-8. Hence, if we take two scans - each with different initial wavelength steps while holding all other scan parameters the same - we will in effect receive two scans with one sampling more densely over the wavelengths and one sampling more dispersed. As can be seen in Figure 4-4, this will impact the pixel temporal period. Furthermore, since the index k_f is directly proportional to $\Delta\lambda_{01}$ – as can be seen by rearranging terms in Eqn. 4-11 - if the initial wavelength step is altered, the frequency produced will be altered by the same factor. Thus, using the ratio between two scans with two initial steps, i.e. $K = \Delta\lambda_{01,1}/\Delta\lambda_{01,0}$, it will be known that the ratio in output frequencies should be the same, i.e. $k_{f,1}/k_{f,0} = K$. Since these are two separate signals with two respective inherent frequencies, this is the condition for the FBSS technique to be used as formulated in Eqn. 4-3. Thus, for a surface with a high level of roughness, this

technique can be used for enhancing the SNR in the FBSS domain and thus producing a height map with a higher degree of confidence in the measurement.

A simulation was performed to demonstrate how the FBSS can be implemented within WSI and to visually see the advantage of using this technique versus just using the Fourier transform. The simulation parameters included λ_0= 1470nm, $\Delta\lambda_{01,0}$ = 0.2nm, $\Delta\lambda_{01,1}$ = 0.17nm, K = 0.85, N = 500, spatial parameters cover -1.8mm to 1.8mm in a 100x100 grid of uniform spacing, l_s = 150.3mm, and l_R = 150.0mm giving an offset between the two arms equal to 300μm. A sample profile was randomly generated at the given offset location with a variation in height of ±10μm. Random speckle patterns were embedded into each set of interferograms for the initial wavelength with speckle diameters randomized from 3 to 9 pixels wide, speckle intensity amplitudes randomized with respect to the signal intensity from ±20, and speckle pixel locations randomized over the whole space of image. The respective speckle locations were then randomized for each discrete wavelength of the first scan. Since the speckle patterns generated in the second scan should be predictable based on the first scan speckles, an interpolation technique was used to determine what the second scan speckle patterns should look like. The roughness profile used and the second wavelength speckle patterns generated for each respective scan are shown in Figure 4-5.

As is seen, a typical roughness profile is generated that has a total variation of approximately 20μm. Figures 4-5b and 4-5c demonstrate typical randomized speckle patterns that contain speckles varying in size, amplitude, and location. It should be mentioned that as the first wavelength in each scan is the same, the speckle will look

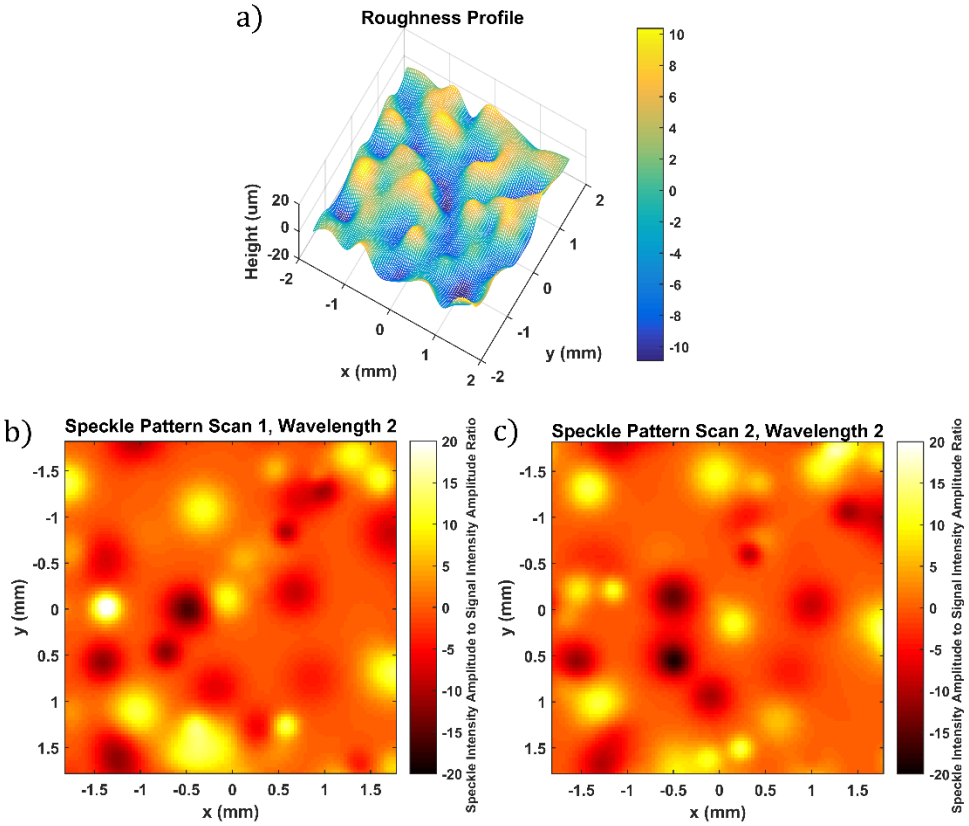

Figure 4-5: This shows the randomly generated functions from simulated MATLAB code over 100x100 grid. (a) depicts the surface profile to be used in the WSI simulation. (b) and (c) depict the randomly generated speckle patterns to introduce noise in the intensity of the interference patterns for the second wavelengths of each respective scan.

identical, thus the second wavelength speckles in each respective scan are displayed so speckle variation with wavelength can be visualized. Applying the scans defined above, we calculated three respective profiles – one using only the first scan and applying the Fourier transform to each pixel, one using only the second scan and applying the Fourier transform to each pixel, and the final profile was calculated by applying the FBSS technique between each respective pixel and both scans. These sample profiles are depicted in Figure 4-6.

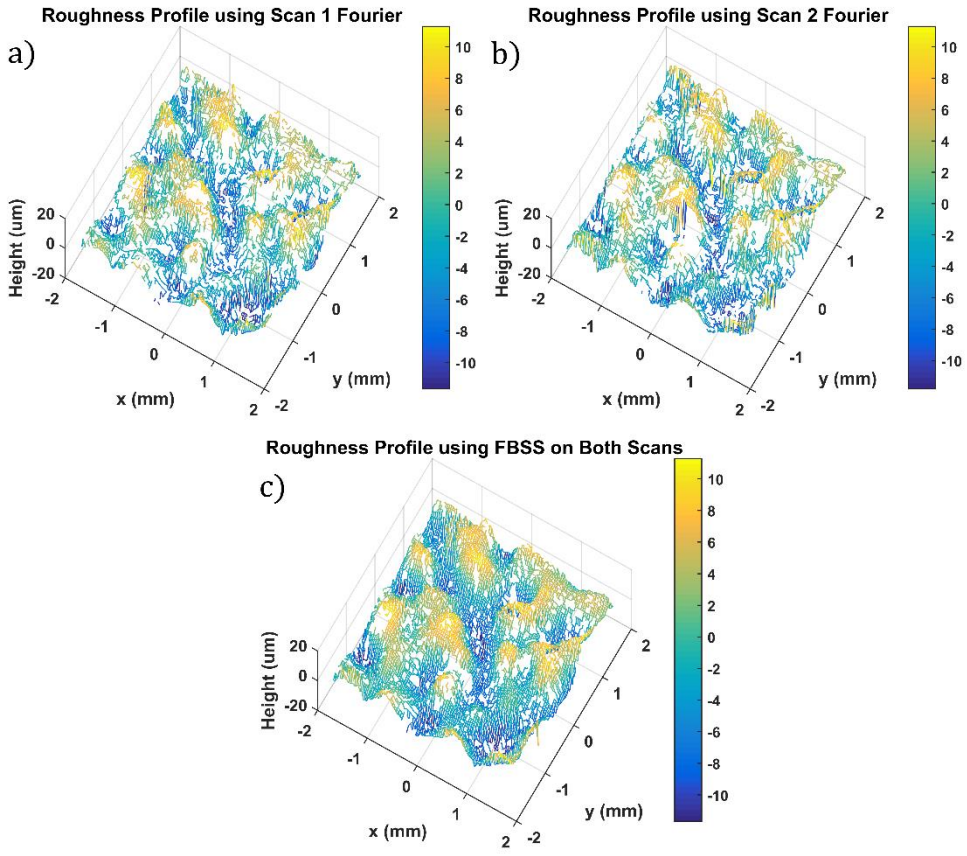

Figure 4-6: This shows the calculated profiles using Fourier and FBSS techniques. (a) and (b) depict calculated profiles for respective scans using the typical Fourier technique, while (c) depicts applying the FBSS technique to the combination of both scans.

As is seen, each calculated profile has been restricted to the expected range of the actual profile in Figure 4-5a, with the offset of 300μm removed from absolute height. As a lot of the data from Figures 4-6a and 4-6b is out of this range, it is shown as missing data on the profile plot. The data missing in each respective scan using the Fourier technique was 32% and 35% out of the 10,000 total pixels. For the FBSS technique, as is seen in Figure 4-6c, the profile more closely resembles that of the actual profile and only 12% of the respective pixels were missing from the spatial grid.

For the purpose of providing a comparison metric to be used when evaluating the actual roughness sample, it was desired to determine the average difference in SNR (in dB)

for each scan using the Fourier technique over all the pixels as compared to the FBSS results over all the pixels. The calculated average SNR difference was 0.96dB and 1.17dB for each respective individual scan as compared to the FBSS technique. An in-depth discussion on how these SNR differences were calculated are discussed in Appendix C.

4.3.3 Experimental Setup

A tunable laser source, Agilent 81680A, that could scan from 1460nm to 1580nm with 5pm accuracy in wavelength was used. The source was outputted through a single mode optical fiber to a collimator producing a beam with diameter of 3.6mm and rated for a 1550nm source. The collimated laser light was then projected to a 50:50 non-polarizing beam-splitter and output to both a reference flat and a sample of interest. The reference flat was the perpendicular side of an 18-degree wedge prism with a surface flatness of $\lambda/10$ rated for 633nm. The wedge prism was used to verify that only light reflected from the front surface would reflect back to the beam-splitter. Both the sample and reference were attached to a two-knob optical mount to adjust tilt in both x-z and y-z directions to align each respective plane proper with respect to the axis of the beam. The reference flat was also attached to a micrometer that could be varied in the direction of the beam to create the proper offset between sample and reference. After the beams were re-combined, light propagated to a double-concave lens with an effective focal length of -12mm which was placed 90mm in front of the camera to adjust for the proper FOV on the sample. A negative lens was used because the camera, SU320-1.7RT, had a sensor size of 12.8mm by 9.6mm which was much bigger than the beam size. The camera used was in the SWIR range with a light sensitivity of 900nm to 1700nm and a quantum efficiency greater than 70% from 1000nm to 1600nm. The schematic of the setup is the same as the depiction in Figure 4-3.

As seen in Figure 4-3, the computer has two-way communication as it receives the camera images, but also controls the camera settings, such as integration time. The computer also has control of the wavelength scanner with commands to vary the set wavelength as well as the power of the source. Using LabVIEW, we developed an interface for setting up the two sets of scans, running the scans, and processing the images from the scans similar to that of the simulation above to determine the height values for each respective pixel.

4.4 Experimental Results

For the purpose of testing the performance of the setup, the system was used to first measure a relatively smooth surface, then shift to surfaces with a greater roughness content. A GAR electroforming roughness set ranging from 2µin to 500µin was used to test the measured profile as seen in Figure 4-7.

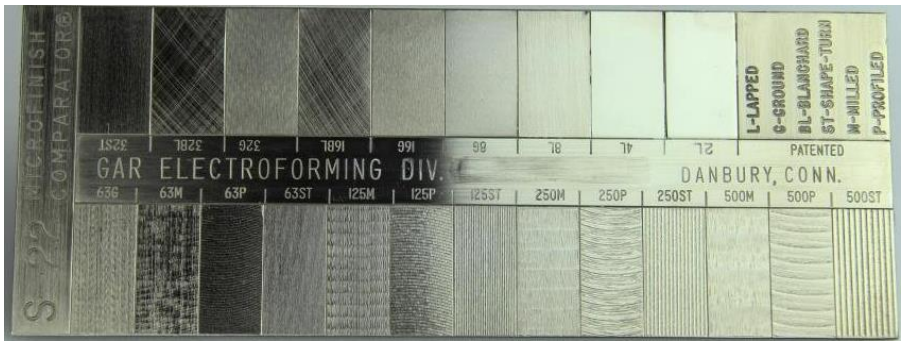

Figure 4-7: This shows the roughness sample set used for verifying roughness profile detected from WSI system. Sample set varies from an S_a of 2µin to 500µin.

4.4.1 Smooth Measurement

In order to characterize the system, a measurement on a smooth surface was taken to verify the resolution in height. For this the 2µin (0.051µm) sample was used. The resultant profile of the measurement is shown in Figure 4-8. For this measurement and all subsequent measurements, the scan parameters used were λ_0= 1475nm, $\Delta\lambda_{01,0}$ = 0.2nm,

$\Delta\lambda_{01,1}$ = 0.16nm, K = 0.80, N = 463, spatial parameters cover -1.2mm to 1.2mm in the x direction and -0.9mm to 0.9mm in the y direction on a 320x240 grid of uniform spacing, and an offset between the two arms of approximately 500µm.

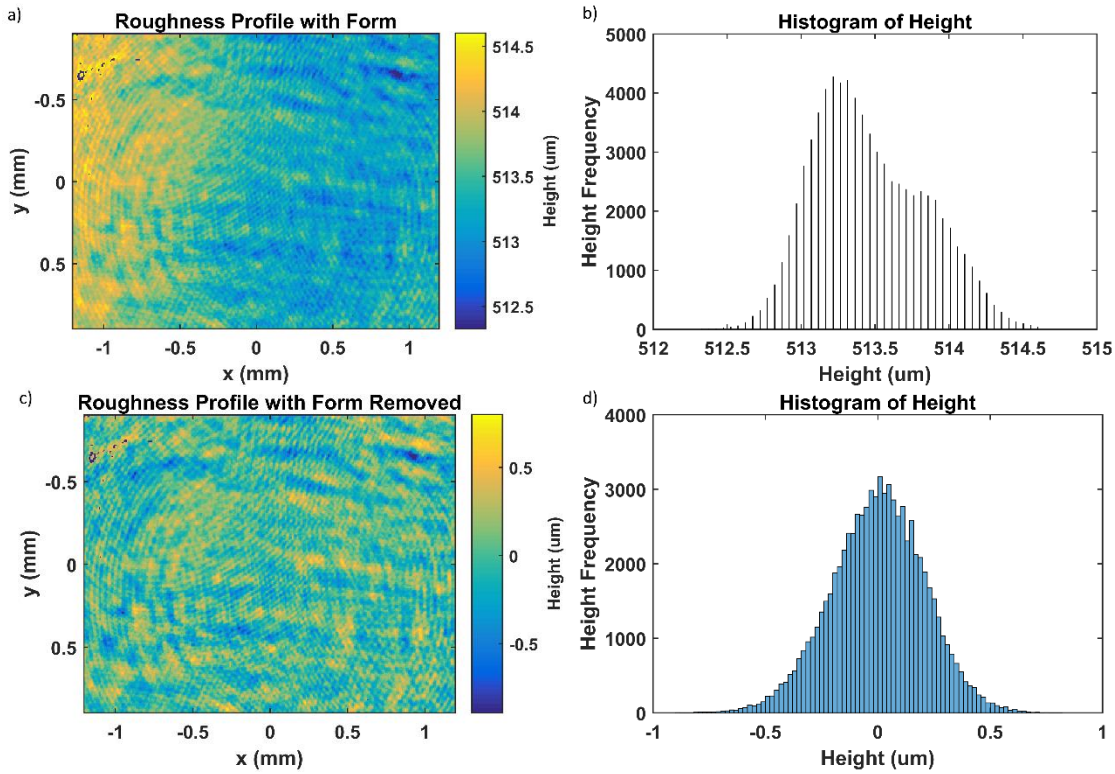

Figure 4-8: This shows the calculated profile for 2µin (~51nm) roughness sample using the common scan parameters. (a) shows the raw profile as detected directly from the frequency domain while (b) depicts the histogram of the heights presented in (a). Similarly (c) shows the profile after the low-order form has been removed while (d) depicts its respective histogram.

As is seen in Figure 4-8b, the histogram shows only 10 discrete heights within a 0.5µm window and thus gives a calculation resolution of approximately 50nm for these particular scan parameters. Since the roughness is on the order of this resolution, there would be no way to accurately obtain the roughness for this sample.

Furthermore, it is seen from Figure 4-8a there was an offset of approximately 513µm with a slight tilt. For the purpose of calculating the actual variation in height, this was readily removed as demonstrated in Figure 4-8c giving an expected Gaussian

distribution in height (Figure 4-8d). In this case, since the surface is relatively flat, the Gaussian distribution in height is characterizing how much error is present in the measurement. Calculating the standard deviation of this variation roughly characterizes the resolution in height of this setup - using the respective scan parameters. Thus, it can be stated that the height resolution of the system is approximately 0.164µm. Hence, as long as the roughness is much greater than this threshold, it should provide an adequate roughness measurement.

4.4.2 Rough Surface Measurement

For the purpose of measuring rough surfaces, the same process as above would be carried out – i.e. using the same scan parameters, removing the offset and form from the absolute height measured, and finding the variation in height of the sample. However, for the purpose of measuring roughness, since the variation in height should be well above the resolution in height of the system, this distribution can be used to measure the roughness of the sample. To characterize the roughness of the sample based of the roughness standard used, we will be calculating the parameter, S_a. To determine S_a from a discrete area of roughness data, Eqn. 4-12 can be used [25].

$$S_a = \frac{1}{A}\sum_{A}|Z(x,y)| \tag{4-12}$$

where S_a is the surface roughness, A is the area of the measurement, and Z is the measured height after the offset, tilt, and form have been removed from the absolute height. For each of the rough samples, a measurement was taken in 4 locations over the sample to verify consistency in the roughness measurement. The first roughness sample measured

was 125P, or rated for a roughness of 3.2µm. The measured profile and histogram for one of the 4 measurements are shown in Figure 4-9.

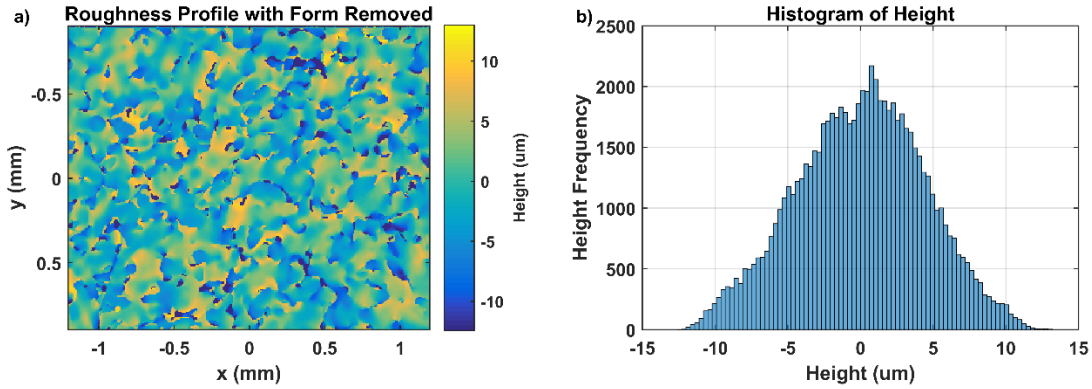

Figure 4-9: This shows the calculated profile for 3.2µm roughness sample using the common scan parameters. (a) shows the profile after the low-order form has been removed and (b) depicts its respective histogram.

As is seen in Figure 4-9, as expected, the relative measured roughness was much greater than that of the smooth sample. For the four measurements taken over different locations of the sample of interest, an average S_a of 3.59µm was observed with a standard deviation of 80nm. Thus, the roughness calculated was approximately 400nm off of the expected result. For all respective techniques (Fourier and FBSS), the roughness calculated was approximately the same with a minimal number of pixels (<5%) out of the expected measurement range. Thus, the absolute SNR for the measurement was relatively high. However, when using the criteria for finding the SNR differences for Fourier and FBSS techniques – as referenced in Section 4-3.2 - it was observed that for the first scan, the average SNR for the Fourier technique was 2.56dB as compared to the average FBSS SNR and for the second scan the average SNR for the Fourier technique was 1.83dB respectively. Thus, if the noise level in the measurement were increased to be comparable to that of the simulation, it can be reasoned that the Fourier measurements would break down to a much greater extent than that of the simulation.

Again, for roughness samples of 250ST and 500ST (6.35µm and 12.7µm roughness respectively), four measurements were taken over different locations on the samples. The following table contains the respective data for the measurements.

Again, for roughness samples of 250ST and 500ST (6.35µm and 12.7µm roughness respectively), four measurements were taken over different locations on the samples. The following table contains the respective data for the measurements.

Table 4-1: Roughness data from measurements on 250ST and 500ST roughness samples.

Sample		**250ST (6.35µm)**		**500ST (12.7µm)**	
		Average	**Standard Dev.**	**Average**	**Standard Dev.**
FBSS	**S_a (microns)**	7.30	0.105	13.40	0.934
	Pixels Missing	1.30%	0.13%	2.49%	0.71%
Scan 1 FT	**S_a (microns)**	7.31	0.087	13.50	0.427
	Pixels Missing	1.44%	0.24%	3.82%	2.29%
	SNR Diff (dB)	1.55	0.25	2.18	0.58
Scan 2 FT	**S_a (microns)**	7.58	0.294	14.71	1.16
	Pixels Missing	1.76%	0.84%	4.68%	4.26%
	SNR Diff (dB)	2.45	0.54	2.60	0.49

As is seen in Table 4-1, for the subsequent measurements, the average calculated roughness for each respective sample was 7.30µm and 13.40µm respectively. Both of these measurements were within one micron of the rated roughness. Thus, it can be concluded that this system can adequately measure surface roughness up to approximately 13µm. Using the raw data from each measurement, the standard deviation of each sample was calculated to be 105nm and 934nm respectively. It was also seen that the SNR must be relatively high as both Fourier calculated measurements over each scan were able to approximately measure the proper roughness without many pixel heights lost. However, although the SNR was high, it is seen that the number of pixels missing from the profile using the FT technique was always greater than the FBSS technique. Furthermore, if we

set a performance criterion as the comparison of missing pixels, we see that for the 6.35μm roughness sample, the FBSS has 11% less missing pixels as compared to scan 1 and 35% less for scan 2. For the 12.7μm roughness sample, the FBSS has 53% less missing pixels as compared to scan 1 and 88% less than the FT for scan 2. Thus, it is seen that the FBSS is providing a much more reliable measurement than the FT.

Observing the SNR difference criteria, it is seen that the FBSS technique maintains approximately a 2.2dB enhancement in SNR as compared to that of the FT technique. This is consistent with the measurements taken on the 125P sample. Thus, for measurements taken with this system, it can be concluded that using the FBSS technique provides a major advantage in suppressing the speckle noise that arises in the interference patterns of rough surfaces. If the surface roughness was more densely populated over the sample of interest, or a sample of greater roughness were measured, the FBSS technique would still be able to provide a proper height measurement on pixels where the typical Fourier technique would not.

4.5 Conclusion

The Fractional BiSpectrum technique was briefly reintroduced. It was shown that for multiple time-domain data containing inherent frequencies with a known relationship, this technique could be used to enhance the signal with respect to the inherent noise in the frequency domain. It was then demonstrated how this technique could be applied to a WSI system, as multiple scans would generate respective inherent frequencies across each of the pixels. The relationship between the inherent frequencies within these pixels could be directly determined by the initial step of the wavelength scan. A simulation was provided demonstrating the advantage of using the FBSS technique in this system.

An experimental setup was then developed for the purpose of testing this method. To verify the roughness calculated, a rated roughness standard with various roughness samples was used. After determining the resolution of the measurement system on a relatively smooth sample, we measured three of the roughest samples on the standard. The average of the measurement results for the three respective samples were all within one micron of the rated roughness. Furthermore, all the measurements demonstrated approximately a 2.2dB enhancement in the SNR for the FBSS technique as compared to that of the Fourier transform. Thus, it can be concluded that for samples with a higher level of roughness - which would produce more speckle and thus a greater level in noise power in the interference pattern – this FBSS technique would be able to recover the proper height information for many pixels over the measurement area where the common Fourier transform technique would not.

REFERENCES

[1] E. Hecht, "Interferometers," in *Optics,* 5th ed. (Pearson, n.d.), pp. 398–457.

[2] P. Hariharan, "Interference: A Primer," in *Basics of Interferometry*, 2nd ed. (Elsevier, 2007), pp. 3–12.

[3] P. de Groot, "Principles of interference microscopy for the measurement of surface topography," Adv. Opt. Photonics **7**, 1–65 (2015).

[4] S. Yang and G. Zhang, "A review of interferometry for geometric measurement," Meas. Sci. Technol. **29**, 1–29 (2018).

[5] E. Hecht, "Mirrored Interferomters," in *Optics*, 5th ed. (Pearson, 2016), pp. 424–428.

[6] P. Hariharan, "The Michelson Interferometer," in *Basics of Interferometry*, 2nd ed. (Elsevier, 2007), pp. 16–18.

[7] B. Dorrio and J. Fernandez, "Phase-evaluation methods in whole-field optical measurement techniques," Meas. Sci. Technol. **10**, R33–R55 (1999).

[8] I. Shevkunov, "A new phase unwrapping method," J. Phys. Conf. Ser. **737**, 1–5 (2016).

[9] L. Ying, *Phase Unwrapping* (2006).

[10] G. Jaeger, "Limitations of precision length measurement based on interferometers," in *Fundamental and Applied Metrology* (2009), pp. 1915–1919.

[11] N. Bobroff, "Residual errors in laser interferometry from air turbulence and nonlinearity," Appl. Opt. **26**, 2676–2682 (1987).

[12] P. de Groot, "Phase Shifting Interferometry," in *Optical Measurement of Surface Topography*, R. Leach, ed. (Springer, 2011), pp. 167–186.

[13] A. J. Lewis, "Fringe analysis & phase stepping interferometry," University of London (1993).

[14] A. Davila, J. M. Huntley, C. Pallikarakis, P. D. Ruiz, and J. M. Coupland, "Wavelength scanning interferometry using a Ti:Sapphire laser with wide tuning range," Opt. Lasers Eng. **50**, 1089–1096 (2012).

[15] I. Yamaguchi and A. Yamamoto, "Surface topography by wavelength scanning interferometry," Opt. Eng. **39**, 40–46 (2000).

[16] A. Yamamoto, C.-C. Kuo, K. Sunouchi, S. Wada, I. Yamaguchi, and H. Tashiro, "Surface shape measurement by wavelength scanning interferometry using an electronically tuned Ti:Sapphire laser," Opt. Rev. **8**, 59–63 (2001).

[17] E. Hecht, "The speckle effect," in *Optics*, 5th ed. (Pearson, 2017), pp. 634–635.

[18] J. Babaie, M. Abolbashari, N. Farahi, S. M. Kim, and F. Farahi, "Optical film thickness measurement of turbid materials using the Fractional BiSpectrum noise-reduction technique," Opt. Commun. **440**, 106–116 (2019).

[19] J. Babaie, M. Abolbashari, N. Farahi, and F. Farahi, "SNR enhancement technique: a new approach to Fractional BiSpectrum," IET Electron. Lett. **TBD**, TBD (2019).

[20] M. Abolbashari, S. M. Kim, G. Babaie, J. Babaie, and F. Farahi, "Fractional bispectrum transform definition and properties," IET Signal Process. **11**, 901–908 (2017).

[21] S. Kuwamura and I. Yamaguchi, "Wavelength scanning profilometry for real-time surface shape measurement," Appl. Opt. **36**, 4474–4482 (1997).

[22] D. O'Connor, X. Jiang, R. K. Leach, G. Moschetti, and A. Forbes, "Quadrature wavelength scanning interferometry," Appl. Opt. **55**, 5332–5340 (2016).

[23] P. D. Ruiz, Y. Zhou, J. M. Huntley, and R. D. Wildman, "Depth-resolved whole-field displacement measurement using wavelength scanning interferometry," J. Opt. A Pure Appl. Opt. **6**, 679–683 (2004).

[24] D. Voelz, "Sampled functions and the discrete Fourier transform," in *Computational Fourier Optics* (SPIE Press, 2011), p. 19.

[25] "Geometrical product specifications (GPS) — Surface texture: Areal — Part 2: Terms, definitions and surface texture parameters," U.S. patent ISO 25178-2:2012 (n.d.).

[26] B. Bhushan, "Surface roughness analysis and measurement techniques," in *Modern Tribology Handbook* (CRC Press, 2001), pp. 1–71.

[27] J. E. Harvey, S. Schröder, N. Choi, and A. Duparré, "Total integrated scatter from surfaces with arbitrary roughness, correlation widths, and incident angles," Opt. Eng. **51**, 013402-1-013402-11 (2012).

[28] E. M. Stein and R. Shakarchi, "Mean-square convergence of Fourier series," in *Fourier Analysis: An Introduction* (Princeton University Press, 2003), p. 79,80.

4.6 Appendix A

The total reflected light from a rough surface can be broken down into light from specular reflection and light from scatter. The following equations show the percentage of light from each respective reflection as a function of root-mean-square (rms) surface roughness, S_{rms} [26,27].

$$\frac{R_{SP}}{R_T} = e^{-\left(\frac{4\pi S_{rms}}{\lambda}\right)} \tag{4-13}$$

$$\frac{R_{SC}}{R_T} = 1 - e^{-\left(\frac{4\pi S_{rms}}{\lambda}\right)} \tag{4-14}$$

where R_{SP} is the specular reflectance coefficient, R_{SC} is the scattered reflectance coefficient, R_T is the total reflectance, and λ is the wavelength of propagated laser light. Assuming the surface mimics that of a Lambertian surface, the detectable scattered light by the aperture is as follows,

$$\frac{R_{SCD}}{R_T} = \frac{R_{SC}}{R_T}\left(\frac{\pi d^2}{2{z_0}^2}\right) = 2\pi(NA)^2\left(\frac{R_{SC}}{R_T}\right). \tag{4-15}$$

Where R_{SCD} is the reflectance corresponding to the scattered light detected by the aperture, d is the diameter of the aperture, z_0 is the distance from the rough surface to the aperture and NA is the numerical aperture. To determine when the power of scattered incoming light detected is comparable to that of the specularly reflected light, the following equation can be derived.

$$\left[1 - e^{-\left(\frac{4\pi S_{rms}}{\lambda}\right)}\right] 2\pi(NA)^2 \geq e^{-\left(\frac{4\pi S_{rms}}{\lambda}\right)} \tag{4-16}$$

Thus, it can be shown that

$$S_{rms} \geq \frac{\lambda}{4\pi}\sqrt{\ln\left(\frac{1}{2\pi}\right) - 2\ln(NA)}. \tag{4-17}$$

To better conceptualize this, an *NA* of 10^{-1} corresponds to $S_{rms} \geq \lambda/7.5$, an *NA* of 10^{-2} corresponds to $S_{rms} \geq \lambda/5$, an *NA* of 10^{-3} corresponds to $S_{rms} \geq \lambda/4$, and an *NA* of 10^{-6} corresponds to $S_{rms} \geq (4/10)\lambda$ approximately. For Michelson interferometry, the numerical aperture of the setup is typically on the order of 10^{-3} for a laboratory setup, and thus a surface roughness of approximately $\lambda/4$ will induce a significant amount of speckle – of equal power or greater than the specular light related to the signal – that arises from the scattered light. It should be noted that although S_{rms} is used for this derivation, the value is on the same order as that of S_a provided in the text.

4.6 Appendix B

When the range of the wavelength scan starts to become comparable to that of the initial wavelength, the non-linear effect in phase with the wavelength steps cannot be ignored if a singular temporal frequency is to be expected in the pixel intensity. Thus, non-consistent wavelength steps are used to compensate for this effect and keep the change in phase over the whole scan consistent. Using Eqn. 4-7 as a starting point, the first consideration for determination of the proper step is to set the first two phase steps equal to each other as shown in Eqn. 4-18.

$$\frac{4\pi h\Delta\lambda_{01}}{\lambda_0(\lambda_0+\Delta\lambda_{01})} = \frac{4\pi h\Delta\lambda_{12}}{\lambda_1(\lambda_1+\Delta\lambda_{12})} \tag{4-18}$$

Simplifying as necessary and utilizing $\lambda_{i+1} = \lambda_i + \Delta\lambda_{i(i+1)}$, we get the following.

$$\frac{\lambda_1-\lambda_0}{\lambda_0\lambda_1} = \frac{\lambda_2-\lambda_1}{\lambda_1\lambda_2} \tag{4-19}$$

Finally, we can solve for the next wavelength in the scan, λ_2.

$$\lambda_2 = \frac{\lambda_0\lambda_1}{2\lambda_0-\lambda_1} \tag{4-20}$$

With the next wavelength in the scan solved, we can reiterate this process to generate the rest of the wavelengths in the scan. Based on the first two wavelengths in the scan – which should be known based on initial wavelength and wavelength step – we can define the rest of the wavelengths in the scan as follows.

$$\lambda_i = \frac{\lambda_{i-1} \cdot \lambda_{i-2}}{2 \cdot \lambda_{i-2} - \lambda_{i-1}}, \qquad for\ i = 2\ to\ N-1 \tag{4-21}$$

4.6 Appendix C

The calculation for the average difference in SNR (in dB) of respective Fourier results as compared to the FBSS is demonstrated here. To determine the respective power for the signal and noise for each of the pixel data, Parseval's theorem was used. The equation for finding the total energy in the discrete frequency domain is as follows [28].

$$E_x = \frac{1}{N_f} \sum_{k=0}^{N_f-1} |X[k]|^2 \tag{4-22}$$

where Ex is the total energy over N_f samples in the frequency domain, k is the frequency index, and X[.] is the discrete Fourier data. To find average power based on the sampling in the time domain, the following expression can be used.

$$P_x = \frac{1}{N_t}\frac{1}{N_f} \sum_{k=0}^{N_F-1} |X[k]|^2 \tag{4-23}$$

where N_t is the total number of samples in the time domain. It can be noted that when analyzing the entire Fourier domain, $N_t = N_f = N$ and thus the first term would become $1/N^2$. It is then assumed that the discrete time domain signal, $x[n]$ can be broken down into the sum of signal (or harmonic function) and noise as shown below.

$$x[n] = s[n] + p[n] \tag{4-24}$$

where $s[n]$ is the signal, $p[n]$ is the noise, and n is the index in the time domain. Furthermore, since the Fourier transform is a linear system, Eqn. 4-25 can be stated.

$$X[k] = S[k] + P[k] \tag{4-25}$$

where $S[k]$ and $P[k]$ are the respective discrete Fourier transforms of the time domain functions. We then predict $S[k]$ to be centered at a particular frequency, f_s in the Fourier domain with a bandwidth, w_s in discrete Fourier index units. A critical assumption is made that the Fourier content within this bandwidth is mainly due to the signal and not the noise. Thus, for a more accurate calculation, the absolute SNR should be much greater than 1. Finally, we can find the power of the signal by analyzing only these frequency indices as shown in Eqn. 4-26.

$$P_S \approx \frac{1}{N_t}\frac{1}{w_s}\sum_{k=f_s-\frac{w_s}{2}}^{f_s+\frac{w_s}{2}} |X[k]|^2 \tag{4-26}$$

where P_S is the power of the signal. Using the same assumption, we can formulate an expression for power of noise with respect to the Fourier domain.

$$P_P \approx \frac{1}{N_t}\frac{1}{N_f - w_s}\left[\sum_{k=0}^{f_s-\frac{w_s}{2}} |X[k]|^2 + \sum_{k=f_s+\frac{w_s}{2}}^{N_F-1} |X[k]|^2\right] \tag{4-27}$$

where P_P is the power of the noise. Using this assumption, the SNR can be calculated for each pixel in the interference pattern as follows.

$$SNR_{x,y} = 10\log\left(\frac{P_{S_{x,y}}}{P_{P_{x,y}}}\right) \tag{4-28}$$

To find the average SNR over the whole interferogram, all pixel SNRs are summed up and divided by the total number of pixels. Furthermore, to obtain the differences

between each respective scan Fourier and the combination of scans using the FBSS, the following was used.

$$\Delta SNR_i = SNR_{FBSS} - SNR_i \quad (4\text{-}29)$$

where ΔSNR_i is the difference in SNR for the *i*th scan, SNR_i is the average SNR over all the pixels using the *i*th scan and applying the Fourier technique, and SNR_{FBSS} is the average SNR over all the pixels using both scans and applying the FBSS technique. It should be mentioned that to directly compare the SNRs of the FBSS and Fourier techniques, the FBSS data must be normalized properly such that magnitudes are directly comparable.

CHAPTER 5: ADDITIONAL CONSIDERATIONS WHEN USING THE FRACTIONAL BISPECTRUM

5.1 Applying FBS to more than 2 inherent frequencies

As was discussed in the previous chapters, the relationship between two inherent frequencies within the signal were used in order to enhance the SNR. It should be explored how this technique can be used to incorporate more inherent frequencies to extend the SNR further. Thus, as a starting point, we use the FBSS as formulated in Chapter 3 (See Eqn. 3-3, Pg. 72).

$$|FBSS[p;K]| = |X_1[p]|^2 \cdot |X_2[K \cdot p]| \tag{5-1}$$

where again $X_1[.]$ and $X_2[.]$ are the respective Fourier transforms of $x_1[.]$ and $x_2[.]$, p is the frequency index, and *K* is the variable relating inherent frequencies. It is seen that the first signal FT magnitude is squared as compared to the other. To simplify this a bit for the purpose of additional signals/inherent frequencies, we set each signal to the same order such that it takes on the feel of a 'scaled cross-spectrum.' Thus, for incorporating more inherent frequencies, the following form can be used.

$$|FSS[p;\boldsymbol{K},N]| = \prod_{i=1}^{N} |X_i[K_i \cdot p]| \tag{5-2}$$

where *FSS* is the Fractional Spectrum Separated which is incorporating *N* number of signals and inherent frequencies. K_i is the ratio between the *i*th inherent frequency and the first inherent frequency – i.e. f_i/f_1. Thus K_1 is equal to 1. For the purpose of comparison, we will explore how using three and four signals and inherent frequencies can extend the SNR in Section 5.2.2.

5.2 Applying the SNR enhancement techniques

When applying the FBS to a setup for the purpose of enhancing the SNR of the signal, it is important to note if and how this technique can be implemented. The methodology for using the FBS as described by the contents of the previous four chapters is to determine if a harmonic function, i.e. $\sin(a \cdot x)$, is present within the time-domain or spatial-domain measurement signal and whether the frequency content is the principle information used to find the characteristic of interest. If this is the case, and only a single harmonic function is present, it is then necessary to understand how the underlying physics produces this harmonic function and thus what variable can be used to alter the frequency. If another harmonic function is already present within the same signal, it is instead necessary to know the relationship between the frequency of the first and the second. If one of these two scenarios exist based on the setup, the FBS may be used to enhance the signal as compared to the noise present.

5.2.1 Conditions of experiment

As mentioned previously, there are different situations where using the FBS may be practical. The conditions for use of the FBS are laid out more explicitly below.

1) *No control over signal/source (i.e. astronomy)*

In this situation, one detects the information based on an assumption that multiple frequency data is inherent in the observable information and there is some intuition of the relationship between the multiple frequency data. There is no control over the source of the data. For this condition, the FBS technique can be used for a single time-domain / spatial domain signal. Also, the common Fourier transform can be used to detect one of the two inherent frequencies present.

2) Control over signal/source (i.e. metrology)

In this situation, some experimental parameters may be controlled. Thus, different potential setups and noise reduction techniques should be considered.

a. Single Data Set with Two or More Inherent Frequencies

In this situation, the experiment can be designed to include multiple sources to generate multiple frequency signal information. In the physics of this experiment, the amplitude of the noise generated from experiment is a function of each source amplitude as described in Chapter 2. This scenario is similar to that of condition *1* and thus the FBS technique and Fourier transform can be used.

b. Separate Data Sets each with One Associated Inherent Frequency

In this situation, the experiment can be designed such that two separate signals are analyzed each with a respective inherent frequency and similar to *2a*, the noise is associated with the amplitude of the source. This scenario is described in Chapters 3 and 4. Thus the FBSS technique must be used since there are multiple separate signals. Furthermore, if more inherent frequencies can be incorporated, Eqn. 5-2 can be used.

5.2.2 Comparison of SNR enhancement techniques

Table 5-1 shows the results of a 3000-iteration Monte-Carlo simulation of each noise reduction technique. White Gaussian noise was used where the noise power was varied with respect to the signal power of the respective inherent frequencies. The ratio of these powers is shown in the first row and is defined as *R*. The predicted frequency of each method was compared with the expected frequency in the time-domain signal to determine

the percentage of times each algorithm was correct. The results for each respective technique are depicted below.

Table 5-1: This shows the direct comparison of SNR enhancement techniques. The colored values shown are the percentage of correct frequencies predicted for each respective technique.

SNR Technique	*Number of Frequencies*	*Number of Signals*	*R = 1*	*R = 5*	*R = 7*	*R = 9*	*R = 12*	*R = 15*	*R = 18*
FT	**2**	**1**	100	83.13	70.17	53.37	22.33	9.67	4.03
FBS	**2**	**1**	100	100	98.97	88.77	45	19.83	7.4
FBSS	**2**	**2**	100	100	100	99.97	93.5	67.4	39.63
FSS_3	**3**	**3**	100	100	100	100	99.9	96.3	82.07
FSS_4	**4**	**4**	100	100	100	100	100	99.17	93.4

As can be seen in the Table 5-1, there is quite a variation in the performance of the different noise reduction techniques. When the technique predicted every result in the Monte-Carlo correctly, it is denoted with a green background, while greater than 90% correct is denoted with a yellow background. Anything worse is denoted with a red background symbolizing a break-down of the respective technique. To better characterize each respective technique, they will be compared based on the conditions of Section 5.2.1.

Condition 1 Analysis:

The first two rows in Table 5-1 show results for multiple inherent frequencies within one signal. As can be seen for this condition, as the noise power related to the additive Gaussian noise increases compared to the inherent frequency power, each technique starts to fail in correctly identifying the inherent frequency within the initial signal. However, it is obvious that the FBS technique outperforms that of a standard Fourier technique with two inherent frequencies as the percentage of failed predictions is much less than that of the Fourier technique. Thus, if there is no control of the source

signal – as condition one suggests, the FBS technique should be used to enhance the SNR as compared to the FT.

Condition 2a Analysis:

While Condition 2a uses the same SNR enhancement techniques as Condition 1, the distinction is that there is control of the signal information for Condition 2a. It is obvious, however, to note that the first two techniques in Table 5-1 underperform compared to that of the latter techniques in the table where the inherent frequencies are separated into multiple signals. This can easily be identified within the comparison of the FBSS for two separate signals as compared to that of the FBS which contains both inherent frequencies within one signal. This result is due to the fact that each inherent frequency introduces a respective noise power into the signal. Thus, it would not be advantageous if one had control over the source information to implement a system with multiple inherent frequencies within one signal. Therefore, Condition 2a is not preferred and should be avoided all-together if possible. This was demonstrated in the results of Chapter 3.

Condition 2b Analysis:

The remaining noise reduction techniques show interesting results. It is clear that if there is the ability to generate the multiple signals with multiple inherent frequencies that have a relation with each other, this is preferred for the enhancing the SNR. The FSS_3 technique seems to further enhance the SNR as well as for that of the FSS_4. Thus, incorporating more inherent frequencies into more respective signals is seen to be more advantageous. The results are depicted into graph form for a visual representation of the results for further understanding. This depiction, Figure 5-1, can be seen as an extension to Figure 2-3 used in Chapter 2.

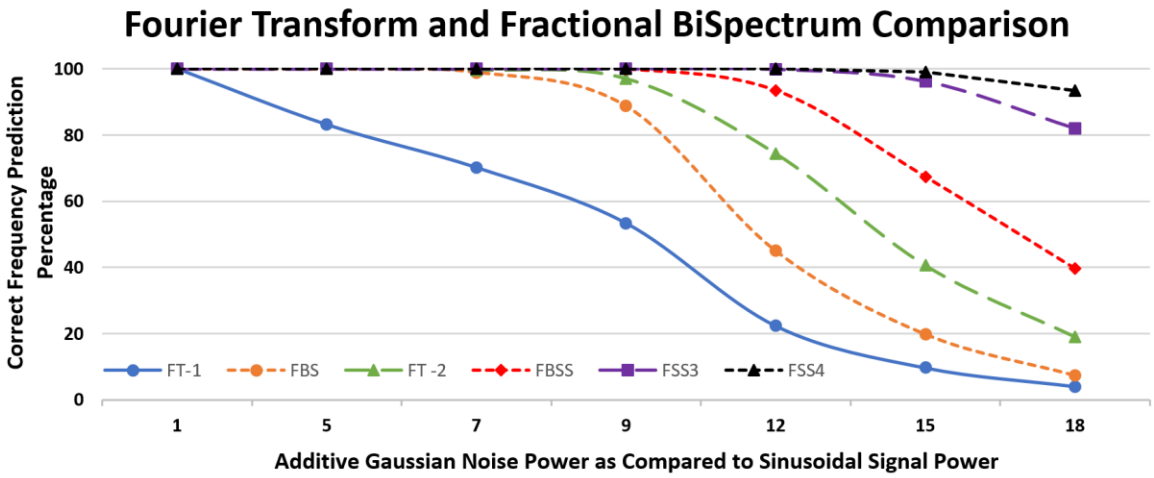

Figure 5-1: This shows a graphical representation of Table 5-1. Each respective technique is compared using a 3000 iteration Monte-Carlo simulation where the noise powers were varied as compared to the inherent frequencies present within the signal.

5.3 Error detection techniques

As shown in the previous section, all techniques will break down at a respective noise power. It is thus important to determine when the wrong inherent frequencies are predicted within the technique as it will lead to an incorrect measurement. In this section, we introduce a few methods for detecting incorrectly predicted inherent frequencies.

5.3.1 1D Error detection analysis

In this section, we demonstrate some techniques that can be applied using one-dimensional information present within the spatial / time domain signal.

Method 1) Multiple Lines of Data:

If for example a two-dimension spatial-domain signal is studied and the inherent frequency component within, a typical method to calculate the frequency is to look at a cross-section of the image. This is demonstrated in Figure 5-2.

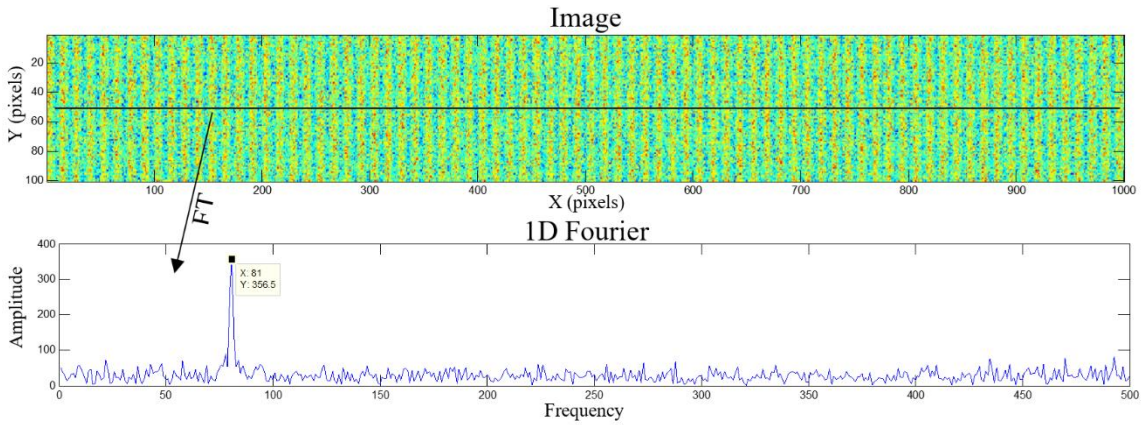

Figure 5-2: This depicts the FT of a horizontal cross-section of a 2D spatial-domain signal.

As can be seen, a singular peak frequency was extracted for the line Y = 50. For error detection, we can apply this several times over the image (multiple Y's). If for example 9 cross-sections are used and then our SNR enhancement technique for extracting the inherent frequency is used, then 9 respective frequencies will be detected. Then for example, if 4 of these 9 are required to match, this creates a check to make sure the result is valid. If this method for error detection is desired for a 1D time-domain or spatial-domain signal, the signal itself can be broken up into *N* number of equal length sections assuming there is a significant number of periods contained within each section.

Method 2) Extra Inherent Frequency and Signal:

With this method, an extra inherent frequency and signal will be used to detect whether a SNR enhancement technique using one less signal is incorrect. If for example the FBSS algorithm is used, with the addition of a third signal, the results of three different two-source methods can be compared to determine if a predicted inherent frequency is incorrect. This is depicted in the Figure 5-3.

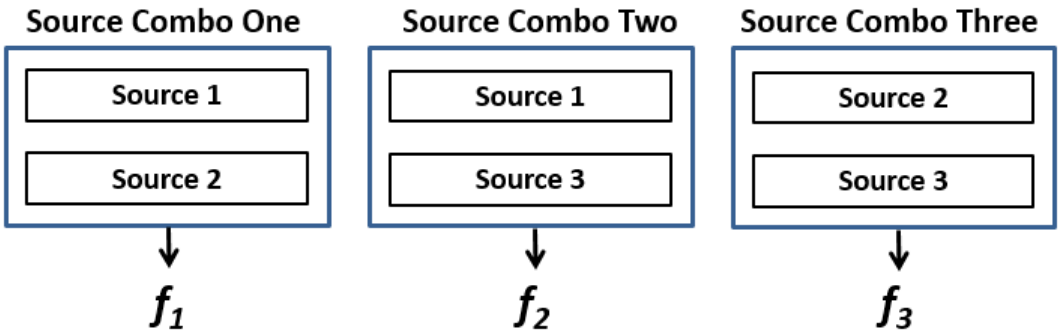

Figure 5-3: This depicts how Method 2 is implemented to determine if predicted inherent frequency is incorrect. For example, if two of the frequencies predicted agree, then one can assume a conclusive result.

In the case where $f_1 = f_2 = f_3$, we can be confident the predicted frequency must be correct. We can also make the assumption that if any combination of two signals match - i.e. $f_1 = f_2$, $f_1 = f_3$, or $f_2 = f_3$ – there is a correct result. However, if $f_1 \neq f_2$, $f_1 \neq f_3$, and $f_2 \neq f_3$, there is no repeating frequency and thus an 'error' is detected. This can obviously be applied to systems with more inherent frequencies than three.

Method 3) Combination of Methods 1 and 2:

As seen in the two previous methods, since the two are using completely different techniques which do not conflict with each other, the two methods can be used together. One can apply Method 1 on the first two signals of Method 2, first and third signals, and then finally for the second and third signals. If any one of these combinations satisfies the criteria for Method 1, it can be presumed that the correct inherent frequency has been identified. This will allow for more data to check than Method 1 and a better error detection technique than the Method 2.

Data Comparison:

The four following figures demonstrate how each error detection technique performs when an incorrect frequency has been determined. For each detection technique,

three 2-dimensional spatial-domain signals (100 pixels by 1000 pixels) each with inherent frequencies in the x direction were generated. White additive Gaussian noise was applied with a varying noise power compared to the inherent frequency's power for each signal. A 3000-iteration Monte-Carlo simulation was used to determine how often each method correctly determined errors. Also, for the two algorithms (FT and FBSS), the general percentage of correct results is shown to give a reference for how much each error detection technique is affecting the result of the noise reduction techniques as there is more data analyzed. For the first depiction, Figure 5-4, the FT technique is analyzed using error detection technique Method 1.

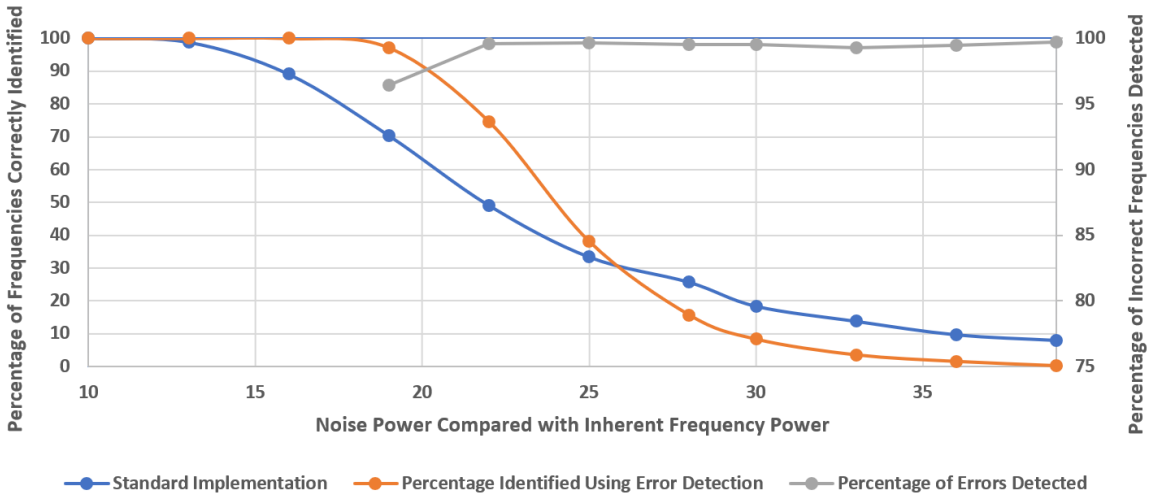

Figure 5-4: This depicts how Method 1 error detection technique performed using the FT as the SNR enhancement technique. It is seen that about 99% of the incorrect frequencies were detected.

The axes and plotted data on Figure 5-4 should firstly be explained properly as they are used in all the subsequent figures in this section. The left axis, which is used for the orange and blue curves, demonstrates the percentage of times - out of the 3000 random spatial-domain signals generated - that each technique produced the correct result or proper frequency. The blue curve is used as a reference for the overlying SNR enhancement technique used as shown in Figure 5-1 in section 5.2.2. The orange curve demonstrates

how often the error detection technique of interest identifies the proper result. The comparison between these two is important because we want to correctly identify errors, but also do not want to reject proper measurements. The rejection of proper measurements can easily be visualized. The axis on the right – which is only used for the gray curve – demonstrates the percentage of incorrectly identified frequencies that were properly rejected. More exactly, if the data in the orange curve was less than 100%, there must have been some incorrectly predicted results and thus of these, we must know what percentage did the error detection technique of interest tell us the result was indeed incorrect. Hence, since the gray curve data must have 'errors' or incorrectly identified frequencies, it could only be plotted for noise thresholds where the orange curve was below 100%.

Analyzing the data from this figure, the blue curve indicates the result of the standard FT technique. This curve was developed based on the result of the center line from the first image to detect a frequency. As can be seen from the orange curve, while the error detection technique is used as a method to detect incorrect frequencies identified, it has also increased the percentage of correct results when the noise power is relatively low. This is due to the fact that it is analyzing more data, 9 sets as compared to 1. There is a crossing point around thirty-three percent where the standard technique will outperform that of the error detection technique due the high level of noise power in the signal. The most important information in Figure 5-4 is the gray curve. As can be seen on the right Y-axis, this technique detects on average around 99% of incorrectly found frequencies for all the various levels of noise power. While not perfect, it provides a very good error indication as compared to 100% pass as with any standard technique. It should be mentioned that if one desired a better error detection technique using the same method, the

criteria of 4 matching inherent frequencies out of 9 could be increased to five or more. This would of course reduce the percentage of the orange curve and create a crossing with the blue curve at a lower noise power.

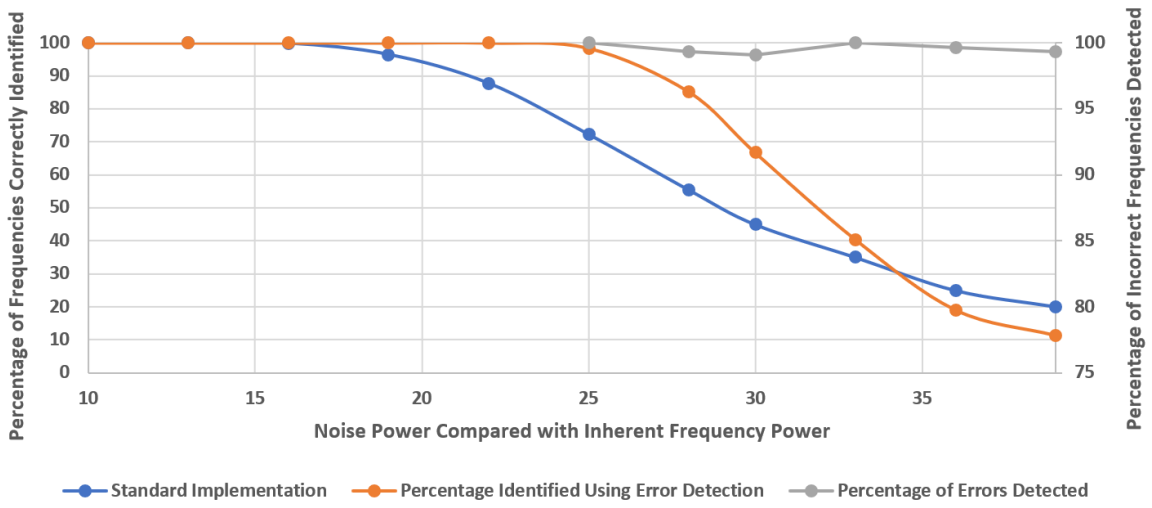

Figure 5-5: This depicts how Method 1 error detection technique performed using the FBSS as the SNR enhancement technique. It is seen that about 99% of the incorrect frequencies were detected.

As seen in Figure 5-5, using Method 1 with the implementation of the FBSS technique, there are a few differences. The main difference is that both the standard FBSS technique (blue line) and FBSS with Method 1 applied (orange line) have a higher percentage of detecting the correct frequencies as compared with the FT technique for the same noise power. This should be the case as it is known that the FBSS technique enhances the SNR to greater extent than that of the FT. The important information to take from this is that the percentage of correctly identified incorrect frequencies is again about 99% at all levels of noise power. Thus, it can be stated conclusively that Method 1 provides a good method for detecting errors in the respective frequency domains.

Next we apply the Method 2 error detection technique for the same spatial domain data in Figure 5-6.

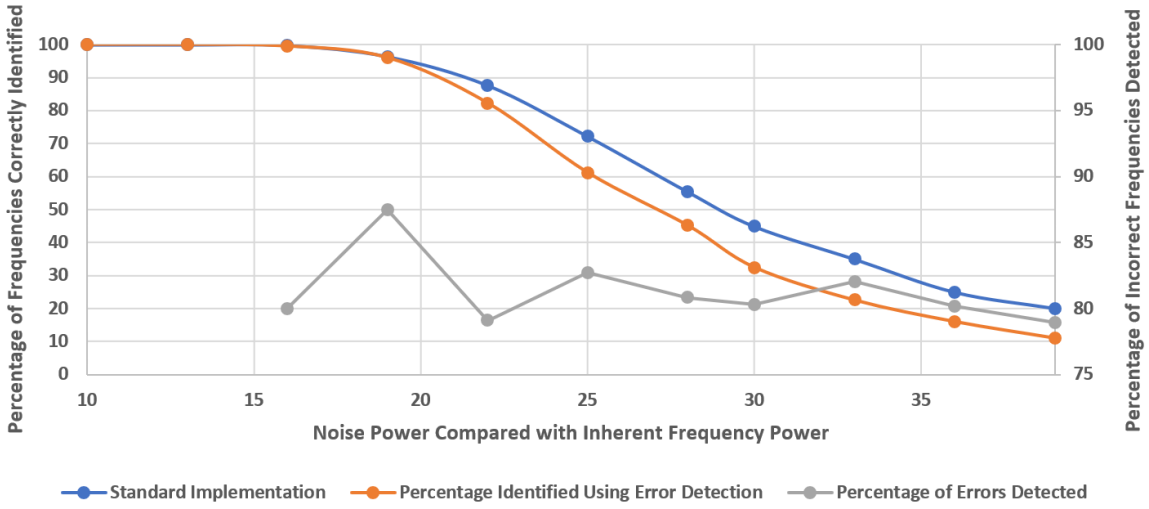

Figure 5-6: This depicts how Method 2 error detection technique performed using the FBSS as the SNR enhancement technique. It is seen that about 80% of the incorrect frequencies were detected.

There are many take-aways from Figure 5-6. Firstly, while the blue curve remains the same as the Figure 5-5 (same noise-reduction technique), the orange curve hovers slightly below it at all levels of noise power. Thus, there will be a slight loss in overall inherent frequencies identified. This is due to the fact that if one of the three signals provide noise content at a particular frequency much higher than that of the inherent frequency, it could impact two of the three combinations generated from the FBSS. The more important information seen from this plot is that the error-detection technique hovers around 80%. Thus, this technique does not perform as well as that of Method 1 which rejects around 99% of the incorrectly identified inherent frequencies. This lower rejection rate could also be due to the hypothesis that one signal could impact two of the three combinations. This could of course be improved if the criteria were to be that all three combinations of inherent frequencies must match, but it can also be assumed that the orange curve would dramatically drop to the point where using a noise-reduction technique is questioned.

There are a few points to note for Method 3 depicted in Figure 5-7. Firstly, there is seen a dramatic improvement in the overall correct frequency percentage as shown in the orange line. It is seen that an average of 98.5% of the incorrect frequencies are correctly identified. While this value is lower by one percent than Method 1, it is much better than Method 2.

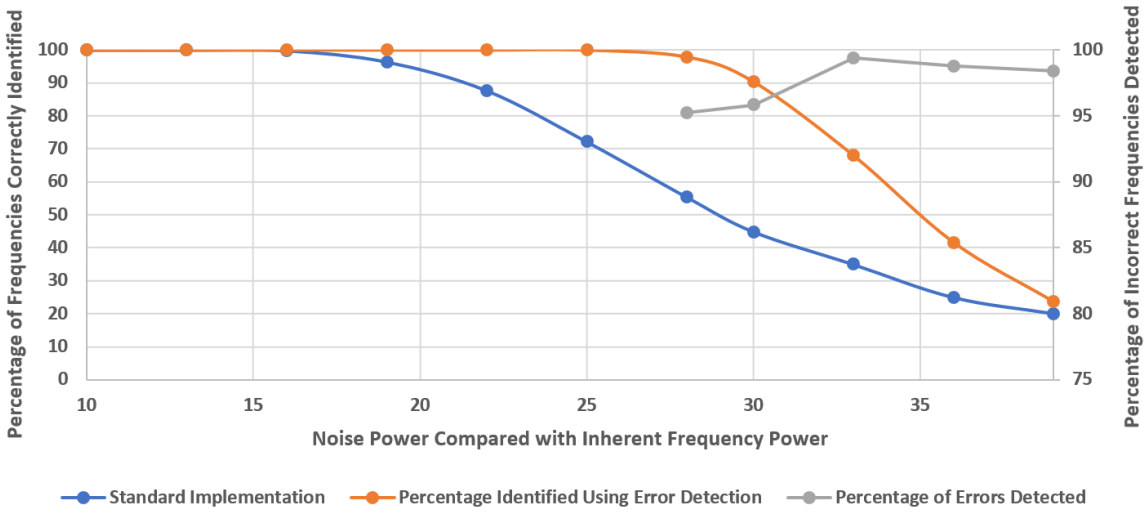

Figure 5-7: This depicts how the Method 3 error detection technique performed using the FBSS as the SNR enhancement technique. It is seen that about 98% of the incorrect frequencies were detected.

Thus, Methods 1 and 3 should be considered the best methods. There is a trade-off between the two error-detection techniques in that one will have less final correct inherent frequency identified (Method 1), while one will reject a slightly smaller number of incorrect frequencies (Method 3). Depending on the application of where the techniques will be used, one of these two methods can be chosen.

5.3.2 2D Error detection analysis

It should be noted in the previous section that since data analyzed was from a two-dimension spatial-domain signal, 2-D analysis can be used as well. If one were instead to

perform a 2-D FT and only look at the results in frequency space on the horizontal axis, one would include the whole dataset for determining the peak horizontal frequency. Thus, it could be assumed that a dramatic enhancement in SNR would occur as this would be analyzing all of the 2D data. However, if one were to apply an error detection technique, the only method possibility would be that of Method 2 of the previous section. Figure 5-8 shows the results of the 2-D analysis on the same 2-D data analyzed in the previous section using instead a Monte-Carlo simulation with 1000 iterations due to the greater number of computations needed.

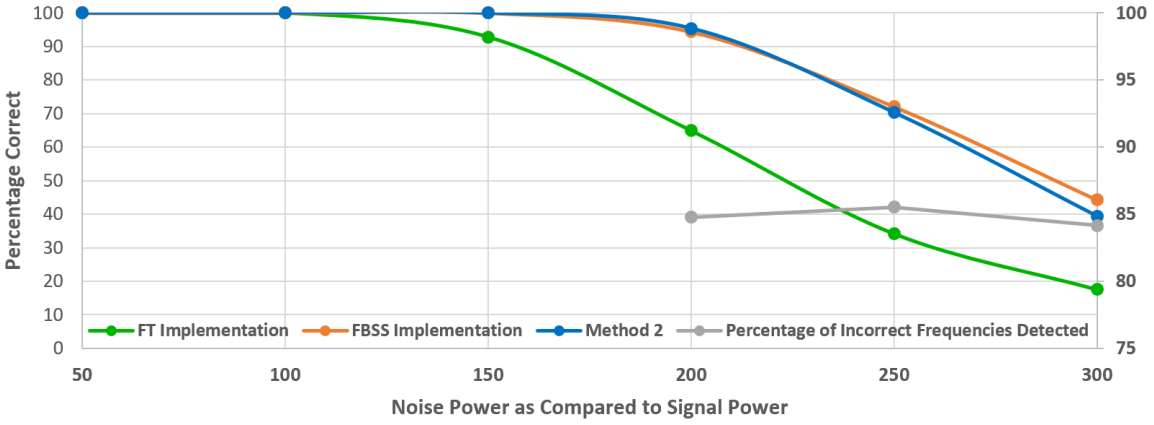

Figure 5-8: This depicts how using the Method 2 error detection technique can be applied to the entire 2-D spatial domain signal. The FBSS was used for the SNR enhancement technique with FT as a reference. It is seen that about 85% of the incorrect frequencies were detected.

It is important to note from Figure 5-8 that firstly the noise power axis ranges from 50 to 300 times that of the inherent frequency power for the two-dimensional spatial-domain signal. Thus, all levels of noise power analyzed in Figure 5-8 are above that of the previous section's highest analyzed noise power. Therefore, it can be quickly stated that analyzing the whole two-dimensional signal and looking at the horizontal component for its spatial frequency is the ideal choice as one should incorporate as much data as possible.

Of course, if there is only a one-dimensional time-domain or spatial-domain signal, one must use the analysis from the previous section, as would be the case for the Chapter 4. Furthermore, as expected, the FBSS technique outperformed the FT technique as can be seen in the blue curve as compared to the green curve. Similar to the previous section, it is seen that the Method 2 error detection technique does little to vary the overall percentage of correct frequencies detected, but maintains the advantage over the FT technique. Finally, it is seen that the Method 2 is rejecting approximately 85% of the incorrectly detected frequencies. This is similar to the result of the previous section for Method 2. While this is of course not as good as Methods 1 or 3, it is the only applicable error detection technique for the 2D implementation of the noise-reduction technique. If a greater reliability were desired, it is possible to use a greater number of signals and inherent frequencies to create more robust combinations than that of Figure 5-3.

CHAPTER 6: CONCLUSIONS

In this dissertation we have introduced an SNR enhancement technique, Fractional BiSpecctrum, extensively. We first introduced many of the metrology techniques for measuring film thickness of turbid mediums and surface profile of relatively rough surfaces. We showed that for such samples the speckle noise may dominate the signal information.

In Chapter 2 we introduced how the Fractional BiSpectrum was developed as an extension of the BiSpecturm, but without the frequency restriction as the relationship between frequencies in the FBS domain is a variable. Furthermore, we applied the FBS technique to a film thickness measurement system. This measurement technique propagated laser light at an incident angle to generate linearly spaced fringes in the x-direction of the sensor. After demonstrating the accuracy of the developed system, we showed how the SNR of the measurement technique was improved for rough/turbid samples by using the FBS which allowed us to successfully perform measurements where the common technique – the FT – failed.

In Chapter 3, we demonstrated how the FBS technique can be extended to measure separate signals, each containing one respective inherent frequency. This extension was defined as Fractional BiSpectrum Separated, or FBSS. It was shown through experimental results that it was advantageous for the purpose of suppressing inherent noise to apply the FBSS instead of using the FBS if there is the control over the source of the signal.

In Chapter 4, we applied the FBSS algorithm from the previous chapter to a highly explored laser interferometry technique, Wavelength Scanning Interferometry, or WSI.

We demonstrated in the theory that each pixel of information will generate two separate frequencies over two wavelength scans related to the different wavelength stepping increments. Using a roughness standard, it was then shown how surface roughness could be determined from the measurement of the surface using the FBSS technique and how the FBSS technique was able to enhance the SNR to a greater extent than that of the FT.

Chapter 5 explored some of the considerations for how to implement the FBS technique properly. Based on different conditions – control of the source of inherent frequency or not – it was discussed which SNR enhancement techniques were applicable and preferred. Furthermore, the implementation of more than two inherent frequencies was explored by expanding the FBSS technique into *N* inherent frequencies. A comparison of all SNR-enhancement techniques was presented. It was then discussed how the FBS technique could be utilized properly while detecting incorrectly predicted frequencies and thus incorrect measurements. Multiple error-detection strategies were presented with a comparison of how well they performed.

An additional application for the FBS noise reduction technique is in the measurement of sub-surface topography. If the topography of top surface is known, then with the method described in Chapter 4 one can measure the thickness map that is directly related to sub-surface structure by removing the reference top surface. If the topography of the top surface is not known, this method should be used in combination with a surface measurement technique [60,61].

Another future research direction would be to use the FBS technique as a general noise-reduction technique – similar to that of the BiSpectrum – by considering the phase information as well as intensity [44 – 47,52 – 57]. Using Appendix A as a starting point,

phase and magnitude reconstruction of the signal can be developed based on novel recursive algorithms based on FBS space [42]. This may be an area of study with significant number of applications.

REFERENCES

[1] Hecht, E.: ‘The speckle effect’, in ‘Optics’ (Pearson, 2017, 5th edn.), pp. 634–635

[2] Dainty, J.C.: ‘Laser speckle and related phenomena’ (Springer, 1975, 2nd edn.)

[3] Heavens, O.S.: ‘Optical properties of thin films’*Reports Prog. Phys.*, 1960, **23**, (1), p. 1.

[4] Bergauer, A., Eisenmenger-Sittner, C.: ‘Properties and characterization of thin films’, in ‘Physics of Thin Films’ (2017), pp. 75–143

[5] Webb, R.H.: ‘Confocal optical microscopy’*Reports Prog. Phys.*, 1996, **59**, (3), pp. 427–471.

[6] Gray, W., Price, R.: ‘Basic Confocal Microscopy’ (Springer, 2018, 2nd edn.)

[7] Sheppard, C.: ‘Measurement of thin coatings in the confocal microscope’*Micron*, 2001, **32**, pp. 701–705.

[8] Pedersen, K.: ‘Ellipsometry’ (2004)

[9] Tompkins, H., Irene, E.: ‘Handbook of Ellipsometry’ (William Andrew Publishing, 2005, 1st edn.)

[10] Miller, D.A.: ‘Optical properties of solid thin films by spectroscopic reflectometry and spectroscopic ellipsometry’. The City University of New York, 2008

[11] Heavens, O.S.: ‘Thin Film Optics’, in ‘Oprical Properties of Thin Solid Films’ (Dover Publications, 2011, 2nd edn.), pp. 46–59

[12] Hlubina, P., Lunacek, J., Ciprian, D., Chlebus, R.: ‘Spectral interferometry and reflectometry used to measure thin films’*Appl. Phys.*, 2008, **92**, (2), pp. 203–207.

[13] Olsen, T.: ‘Reflectometry of the precorneal film’*Acta Ophthalmol.*, 1985, **63**, (4), pp. 432–438.

[14] Groot, P.J.D.: ‘Coherence Scanning Interferometry’, in Leach, R. (Ed.): ‘Optical measurement of surface topography’ (Springer, 2011), pp. 187–208

[15] Fay, M.F., Dresel, T.: ‘Applications of model-based transparent surface films analysis using coherence-scanning interferometry’*Opt. Eng.*, 2017, **56**, (11).

[16] Montgomery, P.C., Montaner, D., Salzenstein, F.: ‘Tomographic analysis of medium thickness transparent layers using white light scanning interferometry and XZ fringe image processing’*Proc. SPIE*, 2012, **8430**, pp. 843014-1-843014–9.

[17] Vorburger, T., Raja, J.: ‘Surface finish metrology tutorial’ (1990)

[18] Bhushan, B.: ‘Surface roughness analysis and measurement techniques’, in ‘Modern tribology handbook’ (CRC Press, 2001, 1st edn.), pp. 49–120

[19] Bhushan, B.: ‘Specular reflection methods’, in ‘Modern tribology handbook’

(CRC Press, 2001, 1st edn.), pp. 87–89

[20] Rao, C., Raj, B.: 'Study of engineering surfaces using laser-scattering techniques'*Sadhana*, 2003, **28**, (3–4), pp. 739–761.

[21] Bennett, H., Porteus, J.: 'Relation between surface roughness and specular reflectance at normal incidence'*J. Opt. Soc. Am.*, 1960, **51**, (2), pp. 123–129.

[22] Kiuchi, T., Ikeuchi, K.: 'Determining surface roughness and shape of specular diffues lobe objects using photometric sampling device', in 'IAPR Workshop on Machine Vision Applications' (1992), pp. 175–178

[23] Golnabi, H.: 'Diffuse reflectance measurements from different surfaces'*J. Sci. I.R. Iran*, 2001, **12**, (4), pp. 359–364.

[24] Bhusan, B.: 'Diffuse reflection methods', in 'Modern tribology handbook' (CRC Press, 2001, 1st edn.), pp. 89–93

[25] Goodman, J.: 'Staistical properties of laser speckle patterns' (1963)

[26] Ohtsubo, J., Fuji, H., Asakura, T.: 'Surface roughness measurement by using speckle pattern'*Jpn. J. Appl. Phys.*, 1975, **14**, (Supplement 14-1), pp. 293–298.

[27] Asakura, T.: 'Surface roughness measurement', in 'Speckle Metrology' (Academic Press, 1978, 1st edn.), pp. 11–50

[28] Green, E.: 'Review of surface texture measurement and the associated metrological problems', in 'Proceedings of the Institution of Mechanical Engineers' (1968), pp. 330–343

[29] Whitehouse, D.: 'Surface metrology'*Meas. Sci. Technol.*, 1997, **8**, (9), pp. 955–972.

[30] Bhusan, B.: 'Mechanical stylus method', in 'Modern tribology handbook' (CRC Press, 2001, 1st edn.), pp. 81–85

[31] Bhusan, B.: 'Scanning tunneling microscope', in 'Modern tribology handbook' (CRC Press, 2001, 1st edn.), pp. 100–102

[32] Güntherodt, H.-J., Wiesendanger, R.: 'Scanning tunneling microscopy I' (Springer, 1994, 2nd edn.)

[33] Ruan, J.-A., Bhusan, B.: 'Atomic-scale friction measurement using friction force microscopy: part 1 - general principles and new measurement techniques'*Trans. ASME*, 1994, **116**, (2), pp. 378–388.

[34] Bhusan, B.: 'Atomic force microscopy', in 'Modern tribology handbook' (CRC Press, 2001, 1st edn.), pp. 102–106

[35] Bhusan, B.: 'Vertical scanning coherence peak sensing', in 'Modern tribology handbook2' (CRC Press, 2001, 1st edn.), pp. 96–97

[36] Gao, F., Leach, R., Petzing, J., Coupland, J.: 'Surface measurement errors using commercial scanning white light interferometers'*Meas. Sci. Technol.*, 2008, **19**,

(1), pp. 1–13.

[37] Badami, V.G., Liesener, J., Evans, C.J., Groot, P. de: ‘Evaluation of the measurement performance of a coherence scanning microscope using roughness specimens’, in ‘Proceedings for the ASPE Annual Meeting’ (2011), p. 4

[38] Hariharan, P.: ‘Interference: A Primer’, in ‘Basics of Interferometry’ (Elsevier, 2007, 2nd edn.), pp. 3–12

[39] Jaeger, G.: ‘Limitations of precision length measurement based on interferometers’, in ‘Fundamental and Applied Metrology’ (2009), pp. 1915–1919

[40] Bobroff, N.: ‘Residual errors in laser interferometry from air turbulence and nonlinearity’*Appl. Opt.*, 1987, **26**, (13), pp. 2676–2682.

[41] Lewis, A.J.: ‘Fringe anlysis & phase stepping interferometry’. University of London, 1993

[42] Ying, L.: ‘Phase unwrapping’ (2006)

[43] Abolbashari, M., Kim, S.M., Babaie, G., Babaie, J., Farahi, F.: ‘Fractional bispectrum transform definition and properties’*IET Signal Process.*, 2017, **11**, (8), pp. 901–908.

[44] Babaie, J., Abolbashari, M., Farahi, N., Kim, S.M., Farahi, F.: ‘Optical Film Thickness Measurement of Turbid Materials using the Fractional BiSpectrum Noise-Reduction Technique’*Opt. Commun.*, 2019, **440**, p. 106-116.

[45] George, S. V., Ambika, G.: ‘Detecting dynamical states from noisy time series using bicoherence’*Nonlinear Dyn.*, 2017, **89**, (1), pp. 465–479.

[46] Totsky, A., Lukin, V., Zelensky, A., *et al.*: ‘Bispectrum-based methods and algorithms for radar, telecommunication signal processing and digital image reconstruction’ (Tampere International Center for Signal Processing, 2008)

[47] Sundaramoorthy, G., Raghuveer, M.R.: ‘Bispectral reconstruction of signals in noise Amplitude reconstruction issues’*IEEE Trans. Acoust.*, 1990, **38**, (7), pp. 1297–1306.

[48] Cetin, A.E.: ‘An iterative algorithm for signal reconstruction from Bispectrum’*IEEE Trans. Signal Process.*, 1991, **39**, (12), pp. 2621–2628.

[49] Matsuoka, T., Ulrych, T.J.: ‘Phase estimation using the Bispectrum’*Proc. IEEE*, 1984, **72**, (10), pp. 1403–1411.

[50] Nikias, C., Petropulu, A.: ‘Signal reconstruction from the phase of the Bispectrum’*IEEE Trans. Signal Process.*, 1992, **40**, (3), pp. 601–610.

[51] Cetin, A.E.: ‘An iterative algorithm for signal reconstruction from bispectrum’*IEEE Trans. Signal Process.*, 1991, **39**, (12), pp. 2621–2628.

[52] Glindemann, A., Lane, R., Dainty, J.C.: ‘Estimation of binary star parameters by model fitting the bispectrum phase’*Opt. Soc. Am.*, 1992, **9**, (4), pp. 543–548.

[53] Lukin, V., Totsky, A., Kurekin, A., Kurbatov, I., J. Astola, Egiazarian, K.: ‘Signal waveform reconstruction from noisy bispectrum estimations pre-processed by vector filters’, in ‘Seventh International Symposium on Signal Processing and Its Applications’ (2003), pp. 169–172

[54] Sadler, B., Giannakis, G.: ‘Shift and rotation-invariant object reconstruction using the bispectrum’*J. Opt. Soc. Am. A*, 1992, **9**, (1), pp. 57–69.

[55] Jin, S., Wear, S., Raghuveer, M.: ‘Reconstruction of speckled images using the bispectra’*J. Opt. Soc. Am. A*, 1992, **9**, (3), pp. 371–376.

[56] Dianat, S.A., Raghuveer, M.R.: ‘Fast algorithms for bispectral reconstruction of two-dimensional signals’, in ‘International Conference on Acoustics, Speech, and Signal Processing’ (1990), pp. 2377–2379

[57] Raghuveer, M.R., Nikias, C.: ‘Bispectrum estimation: A parametric approach’*IEEE Trans. Acoust.*, 1985, **33**, (4), pp. 1213–1230.

[58] Marron, J.C., Sanchez, P.P., Sullivan, R.C.: ‘Unwrapping algorithm for least-squares phase recovery from the modulo 2pi bispectrum phase’*J. Opt. Soc. Am.*, 1989, **7**, (1), pp. 14–20.

[59] Matson, C.L.: ‘Weighted-least-squares phase reconstruction from the bispectrum’*Opt. Soc. Am.*, 1991, **8**, (12), pp. 1905–1913.

[60] Edward, K., Farahi, F., Hocken, R.: ‘Hybrid shear force feedback/scanning quantitative phase microscopy applied to subsurface imaging’*Opt. Express*, 2009, **17**, (21), pp. 18408–18418.

[61] Edward, K., Farahi, F.: ‘Quantitative non-invasive intracellular imaging of Plasmodium falciparum infected human erythrocytes’*Laser Phys. Lett.*, 2014, **11**, (5), pp. 1–6.

APPENDIX A: FRACTIONAL BISPECTRUM TRANSFORM: DEFINITION AND PROPERTIES[4]

Abstract

A signal with discrete frequency components has a zero BiSpectrum if no addition or subtraction of any of the frequencies equals one of the frequency components. The authors introduce the Fractional BiSpectrum (FBS) transform in which for signals with zero BiSpectrum the FBS could be non-zero. It is shown that FBS has the same property as the Bispectrum for signals with a Gaussian probability density function (PDF). The FBS of a zero mean signal with a Gaussian PDF is zero. Therefore, it can be used to significantly reduce the Gaussian noise.

A.1 Introduction

The BiSpectrum is defined as the Fourier transform (FT) of the third order cumulant or moment of a stationary signal and can be used as a noise reduction technique in some circumstances [1, 2]. The definition of BiSpectrum of a 'stationary signal' is given as both the FT of the moment [3] or the cumulant [4]. For a stationary stochastic signal with zero mean, the moment and the cumulant are equal. The BiSpectrum has found applications in various fields including astronomy and cosmology [5–9], signal and image processing [10–13], interferometry applications [14], and mechanical [15–20], electrical [21–25], medical, and biomedical engineerings [19, 26–29].

The BiSpectrum transform has some useful properties. One of these properties is that the BiSpectrum of a stationary signal with a Gaussian probability density function (PDF) with zero mean is zero. Therefore, if a signal is contaminated with Gaussian noise,

[4] This paper is published in the IET Signal Processing journal.

the BiSpectrum of the signal removes or reduces the effect of such noise. More precisely, if the signal is zero mean and is independent of zero mean Gaussian noise, then the effect of noise will be eliminated. Interestingly, a signal can be reconstructed uniquely from its BiSpectrum [30] although the BiSpectrum is insensitive to shift in the domain in which the signal is recorded or sampled (e.g. time or space) [30], i.e. the BiSpectrum of a signal is the same as the BiSpectrum of the shifted version of the same signal.

Another property of the BiSpectrum transform is that it retains the phase information. The BiSpectrum can be used to extract the phase information [31, 32] which is lost in the power spectrum. The exception is the linear phase (shifted signal), which will be lost by a BiSpectrum transform [30]. In addition, the BiSpectrum can be used to reconstruct the amplitude of the FT of the signal [1]. Therefore, the signal can be reconstructed from its BiSpectrum.

As was mentioned, the BiSpectrum is the FT of the third order cumulant or moment of a stationary signal. In the case of a stationary signal with zero mean, the third order moment and cumulant are the same and are defined as

$$C(\rho,\tau) = E\{\mathbf{x}(t+\rho)\mathbf{x}(t+\tau)\mathbf{x}(t)\}, \tag{A-1}$$

where $E\{.\}$ represents the expectation value and x(t) is a stochastic process. The FT of the cumulant is defined as the BiSpectrum

$$S(\nu_1,\nu_2) = \mathcal{F}\{C(\rho,\tau)\} = \mathcal{F}\{E\{\mathbf{x}(t+\rho)\mathbf{x}(t+\tau)\mathbf{x}(t)\}\}, \tag{A-2}$$

where $\mathcal{F}\{.\}$ represents the FT. The definition of Lohmann and Wirnitzer [30] for the cumulant is adopted as follows

$$R(\rho,\tau) = \int_{-\infty}^{\infty} x\,(t+\rho)x(t+\tau)x(t)dt, \tag{A-3}$$

where $x(t)$ is an instance of the stochastic process $\mathbf{x}(t)$ and $R(\rho,\tau)$ is an estimator of the cumulant of $\mathbf{x}(t)$.

The above definition is replacing an ensemble average with a time average which is correct when the stochastic process is triplecorrelation-ergodic. Therefore, the BiSpectrum can be calculated as

$$\begin{aligned} S(\nu_1,\nu_2) &= \mathcal{F}\{R(\rho,\tau)\} \\ &= \int\int_{-\infty}^{\infty} R(\rho,\tau)e^{-i2\pi(\nu_1\rho+\nu_2\tau)}d\rho d\tau \\ &= \int\int\int_{-\infty}^{\infty} x(t+\rho)x(t+\tau)x(t)e^{-i2\pi(\nu_1\rho+\nu_2\tau)}dtd\rho d\tau \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} x(t+\rho)e^{-i2\pi\nu_1(\rho+\tau)}d\rho\right] \\ &\qquad \times\left[\int_{-\infty}^{\infty} x(t+\tau)e^{-i2\pi\nu_2(\rho+\tau)}d\tau\right]x(t)e^{-i2\pi(-\nu_1-\nu_2)}\,dt \\ &= X(\nu_1)X(\nu_2)X(-\nu_1-\nu_2), \end{aligned} \tag{A-4}$$

where $X(.)$ is the FT of the signal $x(.)$.

$$X(\nu)=\int_{-\infty}^{\infty} x\,(t)e^{-i2\pi\nu t}dt. \tag{A-5}$$

For a real signal, $X(-\nu)=X^*(\nu)$ and thus the BiSpectrum can be written as

$$S(\nu_1,\nu_2)=X(\nu_1)X(\nu_2)X(-\nu_1-\nu_2)=X(\nu_1)X(\nu_2)X^*(\nu_1+\nu_2). \tag{A-6}$$

As it is seen in Eqn. A-6, in order for the BiSpectrum to be non-zero there should exist non-zero components ν_1, ν_2, *and* $\nu_1+\nu_2$ for some frequency sets. Here, ν_1 *and* ν_2 can be any of f_1 *and* f_2. As examples, $\nu_1+\nu_2=f_1+f_2$ for $\nu_1=f_1$ *and* $\nu_2=f_2$ which needs three non-zero frequency components (f_1,f_2,f_1+f_2) and $\nu_1+\nu_2=2f_1$ for $\nu_1=f_1$ and $\nu_2=f_1$ which requires only two non- zero frequency components $(f_1,2f_1)$. If such combination does not exist, the BiSpectrum of the signal will be zero. For example, the BiSpectrum of a pure sinusoidal signal is zero. It is emphasised that the BiSpectrum of the

combination of two sinusoidal signals is zero if the frequency of one signal is not twice as the frequency of the other one.

Obviously, in some cases the signals have zero BiSpectrum and the benefits of this noise reduction method cannot be sought. It is in fact desirable to have a non-zero BiSpectrum even if the frequencies do not satisfy the required relation. This happens, for example, where the signal is the combination of two laser sources in which the wavelength of one of the laser sources is not exactly twice as the other one (which is very probable). In this work, a technique that expands the application of this noise reduction technique is described.

We introduce fractional BiSpectrum (FBS) and show that although it does not hold some properties of the BiSpectrum such as shift invariant, it holds the very attractive property of BiSpectrum; the FBS of a stationary signal with a Gaussian PDF with zero mean is zero. Therefore, this paper focuses on the following property of the FBS; for a zero mean signal with a Gaussian PDF, the FBS is zero. This is important for the initial target application for the FBS which is in the field of interferometry. By using the BiSpectrum one can improve the signal to noise ratio (SNR) of the output of an interferometer; however, for the BiSpectrum to be non-zero one must need at least two wavelengths (lasers) with one being the integer multiple of another. However, the problem is that the laser wavelengths with this property are not available. FBS solves this problem and allow the use of two arbitrary laser wavelengths while improving the SNR by removing the Gaussian noise. An example of interferometry system and the performance enhancement using FBS is given in Section A.4. The rest of the paper is organized as follows. In Section A.2, the FBS is defined and the above property is proven for the continuous time signal. The above

property for FBS of the discrete time signals is given in Section A.3. A simple example and an interferometry system simulation to measure film thickness are given to show how the FBS can be applied in Section A.4. This paper is concluded in Section A.5.

A.2 Fractional BiSpectrum: Definition

We define the FBS as

$$F(\nu_1, \nu_2; k) \equiv X(\nu_1)X(\nu_2)X(-\nu_1 - k\nu_2) = X(\nu_1)X(\nu_2)X^*(\nu_1 + k\nu_2), \quad \text{(A-7)}$$

where $k \in \boldsymbol{Q}^+$. As it is seen, the BiSpectrum transform is a special case of the FBS of $k = 1$. It is worth mentioning that in a series of papers [33–35] Ming Li *et al.* reviewed fractal time series and derived power spectrum for some fractal time series namely fractional Gaussian noise and generalised fractional Gaussian noise. Fractal time series can be described as a solution to the fractional differential equation (filter of fractional order) as explained in [33]. The dimension of output of such systems might be fractional and possess some properties such as not having an integrable autocorrelation function. On the other hand, FBS is derived from a signal with integer dimension and integrable power spectrum. Moreover, if we consider FBS transformation as a system, the system is not linear since it can generate frequencies that do not exist in the frequency domain of signal. The fractional part of FBS comes from the fact that another degree of freedom was added to the BiSpectrum by introducing the factor k which can be non-integer. The factor k enables the FBS to have flexibility to be used in scenarios that the BiSpectrum is not applicable (e.g. for signals with BiSpectrum equals zero) while benefiting from the properties of BiSpectrum. As it was discussed earlier, the BiSpectrum of a signal with a Gaussian PDF is zero. Therefore, if a signal has an additive Gaussian noise, the effect of the Gaussian noise will be removed or reduced by using the BiSpectrum. It will be shown

further on that this is valid for FBS as well. To show that the FBS of a zero mean signal with a Gaussian PDF is zero, it suffices to show that the inverse FT of the FBS is a third order statistics. Since the third and higher odd order statistics of a signal with a Gaussian PDF is zero [36], the FBS will be zero for a zero mean signal with a Gaussian PDF.

In general, the inverse FT of the FBS is calculated as

$$\begin{aligned}
R_F(\rho,\tau;k) &= \mathcal{F}^{-1}\{F(\nu_1,\nu_2;k)\} \\
&= \iint_{-\infty}^{\infty} F(\nu_1,\nu_2;k) e^{i2\pi(\nu_1\rho+\nu_2\tau)}\, d\nu_1 d\nu_2 \\
&= \iint_{-\infty}^{\infty} X(\nu_1) X(\nu_2) X(-\nu_1-k\nu_2) e^{i2\pi(\nu_1\rho+\nu_2\tau)} d\nu_1 d\nu_2 \\
&= \iint_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(t_1) e^{-i2\pi\nu_1 t_1} dt_1\right]\left[\int_{-\infty}^{\infty} x(t_1) e^{-i2\pi\nu_2 t_2} dt_2\right] \\
&\qquad \times \left[\int_{-\infty}^{\infty} x(t_3) e^{-i2\pi(-\nu_1-k\nu_2)t_3} dt_3\right] e^{i2\pi(\nu_1\rho+\nu_2\tau)} d\nu_1 d\nu_2 \\
&= \iiint_{-\infty}^{\infty} x(t_1) x(t_2) x(t_3) \left[\int_{-\infty}^{\infty} e^{i2\pi\nu_1(\rho+t_3-t_1)}\, d\nu_1\right] \\
&\qquad \times \left[\int_{-\infty}^{\infty} e^{i2\pi\nu_2(\tau+kt_3-t_2)}\, d\nu_2\right] dt_1 dt_2 dt_3 \\
&= \iiint_{-\infty}^{\infty} x(t_1) x(t_2) x(t_3) \delta(\rho+t_3-t_1)\delta(\tau+kt_3 \\
&\qquad\qquad - t_2) dt_1 dt_2 dt_3 \\
&= \int_{-\infty}^{\infty} x(\rho+t_3) x(\tau+kt_3) x(t_3) dt_3.
\end{aligned} \tag{A-8}$$

Therefore,

$$R_F(\rho,\tau;k) = \int_{-\infty}^{\infty} x(\rho+t) x(\tau+kt) x(t) dt = \tilde{E}\{x(\rho+t)x(\tau+kt)x(t)\}, \tag{A-9}$$

where $\tilde{E}\{.\}$ is the estimator of the expectation value.

Assume that $x(t)$ is an instance of the stochastic process $\boldsymbol{x}$ with a Gaussian PDF. Since ρ, τ, and t are independent variables and $x(\rho+t)$, $x(\rho+t)$, *and* $x(t)$ are Gaussian, the expectation value of the triple correlation in Eqn. A-7 equals zero according to the Isserlis theorem [36], i.e. $R_F(\rho,\tau;k) = 0$. As a result, it can be concluded that $F(\nu_1,\nu_2;k) = \mathcal{F}\{R_F(\rho,\tau;k)\} = 0$.

It should be noted that in practice, the signal $x(t)$is one instance of a zero mean Gaussian stochastic process and $R(\rho, \tau)$ is an estimator of the cumulant. Therefore, the FBS of a signal $x(t)$ might not be zero (it is zero with the assumption of $x(t)$ being triplecorrelation- ergodic); however, would be a small value. As a result, by using the FBS transform, the zero mean Gaussian noise would be reduced instead of being eliminated.

A.3 Fractional BiSpectrum of a Discrete Time Signal

For discrete time signals, the cumulant of a signal with zero mean is defined as

$$C[m,n] = E\{x[p+m]x[p+n]x[p]\}, \tag{A-10}$$

where m, n, and p are integers. Equivalent to the analogue signals, the following is defined as an estimator of the cumulant of a discrete time signal with zero mean

$$R[m,n] \equiv \sum_{p=-\infty}^{\infty} x[p+m]x[p+n]x[p]. \tag{A-11}$$

A.3.1 Continuous FBS of a discrete time signal

The continuous BiSpectrum of a discrete time signal is defined as

$$S_D(\omega, v) = \mathcal{F}\{R[m,n]\} = X\langle\omega\rangle X\langle v\rangle X\langle -\omega - v\rangle, \tag{A-12}$$

where $X\langle\omega\rangle$ is the continuous FT of the discrete time signal $x[n]$ [37]

$$X\langle\omega\rangle = \sum_{n=-\infty}^{\infty} x\,[n]e^{-i\omega n}, \tag{A-13}$$

and $\mathcal{F}\{.\}$ is the continuous FT of the discrete time signal as stated above. Note the different notation of $X\langle\omega\rangle$ compared with [37].

Therefore, analogous to the definition of the FBS for a continuous time signal, we define the continuous FBS of a discrete time signal as

$$F_D(\omega_1, \omega_2; k) \equiv X\langle\omega_1\rangle X\langle\omega_2\rangle X\langle -\omega_1 - k\omega_2\rangle. \tag{A-14}$$

It is assumed that k is a positive rational number, $k \in Q+$ and is represented as $k = k_1/k_2$ where k_1 and k_2 are relatively prime, $gcd(k_1, k_2) = 1$.

As it is known, the continuous FT of a discrete signal is periodic with a period of 2π. Hence, to calculate the continuous FBS of a discrete signal, we only consider one period of the continuous FT and thus we assume that the FT is zero for the frequencies outside that period. This 2π support can be considered as the interval $(-\pi, \pi)$ or $(0, 2\pi)$ and we consider it to be $(-\pi, \pi)$.

For the proof, in this section, we consider the condition that the down-sampling of $x[n]$ does not impact the spectrum content of it. This means that we apply the condition that for the discrete signal $x[n]$ the bandwidth is less than or equal to π/k_2. Therefore, the down-sampled signal $x_{k_2}^d[n]$ *of* $x[n]$ with the factor k_2 and its FT are related to $x[n]$ and its FT as follows

$$x_{k_2}^d[n] = x[k_2 n], \tag{A-15}$$

$$X_{k_2}^d\langle\omega\rangle = X\langle\omega/k_2\rangle, \tag{A-16}$$

where

$$X\langle\omega\rangle = \begin{cases} X_{k_2}^d\langle k_2\omega\rangle & |\omega| \le \dfrac{\pi}{k_2} \\ 0 & \dfrac{\pi}{k_2} < |\omega| < \pi \end{cases}, \tag{A-17}$$

$$X_{k_2}^d\langle k_2\omega\rangle = \sum_{n=-\infty}^{\infty} x_{k_2}^d[n] e^{-ik_2\omega n}.$$

On the basis of the above condition, the continuous signal $x(t)$, where $x[n]$ represents its samples ($x[n] = x(nT)$ with the sampling period T), can be recovered fully from both $x[n]$ and $x_{k_2}^d[n]$. Accordingly, the PDFs of both signals have to have the same shape. In other words, if PDF of a signal is Gaussian, its downsampled signal would have a Gaussian PDF, too.

On the basis of the above definitions and condition, it will be shown that the continuous FBS of a discrete time signal with a Gaussian PDF is zero. To show this, it is sufficient to show that the inverse FT of the discrete FBS is a triple correlation and thus it is zero.

$$
\begin{aligned}
R_F^C[m,n;k] &= \mathcal{F}^{-1}\{F_D(\omega_1,\omega_2;k)\} \\
&= \left(\frac{1}{2\pi}\right)^2 \oint\oint_{2\pi} X\langle\omega_1\rangle X\langle\omega_2\rangle X\langle-\omega_1-k\omega_2\rangle \\
&\qquad \times e^{i(\omega_1 m+\omega_2 n)} d\omega_1 d\omega_2 \\
&= \left(\frac{1}{2\pi}\right)^2 \oint\oint_{2\pi} X_{k_2}^d\langle k_2\omega_1\rangle X_{k_2}^d\langle k_2\omega_2\rangle X_{k_2}^d\langle-k_2\omega_1-kk_2\omega_2\rangle \\
&\qquad \times e^{i(\omega_1 m+\omega_2 n)} d\omega_1 d\omega_2 \\
&= \left(\frac{1}{2\pi}\right)^2 \oint\oint_{2\pi} \sum_{p_1=-\infty}^{\infty} x_{k_2}^d[p_1] e^{-ik_2\omega_1 p_1} \sum_{p_2=-\infty}^{\infty} x_{k_2}^d[p_2] e^{-ik_2\omega_2 p_2} \\
&\qquad \times \sum_{p_3=-\infty}^{\infty} x_{k_2}^d[p_3] e^{-i(-k_2\omega_1-\frac{k_1}{k_2}k_2\omega_2)p_3} e^{i(\omega_1 m+\omega_2 n)} d\omega_1 d\omega_2 \\
&= \sum_{p_1,p_2,p_3=-\infty}^{\infty} x_{k_2}^d[p_1] x_{k_2}^d[p_2] x_{k_2}^d[p_3] \\
&\qquad \times \frac{1}{2\pi}\oint_{2\pi} e^{i\omega_1(m+k_2p_3-k_2p_1)}\, d\omega_1 \frac{1}{2\pi}\oint_{2\pi} e^{i\omega_2(n+k_1p_3-k_2p_2)}\, d\omega_2 \\
&= \sum_{p_1,p_2,p_3=-\infty}^{\infty} x_{k_2}^d[p_1] x_{k_2}^d[p_2] x_{k_2}^d[p_3] \\
&\qquad \times \delta[m+k_2p_3-k_2p_1]\delta[n+k_1p_3-k_2p_2] \\
&= \sum_{p_1,p_2,p_3=-\infty}^{\infty} x[k_2p_1] x[k_2p_2] x[k_2p_3] \\
&\qquad \times \delta[m+k_2p_3-k_2p_1]\delta[n+k_1p_3-k_2p_2] \\
&= \sum_{p_3=-\infty}^{\infty} x[m+k_2p_3] x[n+k_1p_3] x[k_2p_3] \\
&= \tilde{E}\{x[m+k_2p_3]x[n+k_1p_3]x[k_2p_3]\} \\
&= \tilde{E}\{x((m+k_2p_3)T)x((n+k_1p_3)T)x((k_2p_3)T)\}.
\end{aligned}
\tag{A-18}
$$

As m, n, and p_3 are independent variables, $R_F^C[m,n;k]$ is the sampled version of the triple correlation of joint Gaussian signals. Since the triple correlation of the joint Gaussian signal is zero [36], $R_F^C[m,n;k]$ which is the samples of the triple correlation will be zero as well. Therefore, for a zero mean signal with a Gaussian PDF, $R_F^C[m,n;k] = 0$. As a result, it can be concluded that $F_D(\omega_1,\omega_2;k) = \mathcal{F}\{R_F^C[m,n;k]\} = 0$.

A.3.2 Discrete FBS of a discrete time signal

For a discrete signal with length *N*, the discrete FT (DFT) is defined as

$$\begin{aligned} X[r] &= \sum_{n=\left\lfloor -\frac{N}{2}\right\rfloor}^{\left\lfloor \frac{N}{2}-1\right\rfloor} x[n]e^{-i\frac{2\pi r}{N}n} = \sum_{n=-\infty}^{\infty} x[n]e^{-i\frac{2\pi r}{N}n} \\ &= \sum_{n=-\infty}^{\infty} x[n]e^{-i\omega n}\big|_{\omega=\frac{2\pi r}{N}} = X\langle\omega\rangle\big|_{\omega=\frac{2\pi r}{N}}, \end{aligned} \tag{A-19}$$

where $\lfloor . \rfloor$ and $\lceil . \rceil$ are floor and ceiling which map a real number to the largest previous integer and to the smallest following integer, respectively.

It is assumed that the support of the signal $x[n]$ *is* $\left[\left\lceil -\frac{N}{2}\right\rceil, \left\lceil \frac{N}{2}-1\right\rceil\right]$. Again, the condition that the bandwidth of the discrete signal *x*[*n*] (the support of $X\langle\omega\rangle$) is less than or equal to π/k_2 is applied. Therefore, the down-sampled signal $x_{k_2}^d[n]$ *of* $x[n]$ with the factor k_2 and its DFT are related to *x*[*n*] and its FT as follows

$$x_{k_2}^d[n] = x[k_2 n], \tag{A-20}$$

$$X[r] = \begin{cases} X_{k_2}^d[k_2 r] & |r| \le \left\lfloor \frac{N}{2k_2}\right\rfloor \\ 0 & \left\lfloor \frac{N}{2k_2}\right\rfloor < |r| \end{cases}, \tag{A-21}$$

where the down-sampled FT is

$$\begin{aligned} X_{k_2}^d[k_2 r] &= X_{k_2}^d\langle\omega\rangle\big|_{\omega=\frac{2\pi k_2 r}{N}} = X\langle\frac{\omega}{k_2}\rangle\big|_{\omega=\frac{2\pi k_2 r}{N}} \\ &= X\langle\omega\rangle\big|_{\omega=\frac{2\pi r}{N}} = X[r]. \end{aligned} \tag{A-22}$$

For a discrete signal, the discrete FBS is defined as

$$F[r,s;k] \equiv X[r]X[s]X[-r-ks]. \tag{A-23}$$

As it is seen for non-integer *k*, the index $-r-ks$ is not always an integer. $X[-r-ks]$ in this case is undefined and thus we define the discrete FBS as

$$\begin{aligned} F^D[r,s;k] &\equiv F_D(\omega_1,\omega_2;k)\big|_{\omega_1=\frac{2\pi r}{N},\omega_2=\frac{2\pi s}{N}} \\ &= X\langle\omega_1\rangle X\langle\omega_2\rangle X\langle-\omega_1-k\omega_2\rangle\big|_{\omega_1=\frac{2\pi r}{N},\omega_2=\frac{2\pi s}{N}}. \end{aligned} \tag{A-24}$$

It is again assumed that the support for $X\langle\omega\rangle$ is $(-\pi, \pi)$ as discussed before. The inverse DFT of the FBS is calculated as (see Eqn. A-25). As it is seen, $R_F^D[m,n;k]$ is the sampled version of the triple correlation of joint Gaussian signals, where *m*, *n*, and *p*3 are independent variables. According to the Isserlis theorem, the triple correlation of a joint Gaussian signal is zero [36]. Therefore, $R_F^D[m,n;k]$, which is the samples of the triple correlation, will be zero as well. Thus, we can conclude that $F^D[r,s;k] = \mathcal{F}\{R_F^D[m,n;k]\} = 0$.

As mentioned in the previous section, $x[n]$ is one instance of a zero mean Gaussian stochastic process and $R[m,n]$ is a cumulant estimator. Hence, in practice, the FBS of $x[n]$ might not be zero; however, it would be a very small value and, as a result, by using the FBS transform, the zero mean Gaussian noise in $x[n]$ would be significantly reduced.

A.3.2.1 Determination of FBS with non-integer *k* values

It should be noted that due to non-integer values of *k* in Eqn. A-23, a problem exists with the calculation of the third term of the discrete FBS of a discrete time signal. Since the index value is now noninteger, generating the FBS space is not as straight forward as

$$
\begin{aligned}
R_F^D[m,n;k] &\equiv \mathcal{F}^{-1}\{F^D[r,s;k]\} \\
&= \sum_{r,s=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} X[r]X[s]X[-r-ks]e^{i\frac{2\pi}{N}(rm+sn)} \\
&= \sum_{r,s=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} X\langle\omega_1\rangle X\langle\omega_2\rangle X\langle-\omega_1-k\omega_2\rangle\big|_{\omega_1=\frac{2\pi r}{N},\omega_2=\frac{2\pi s}{N}} e^{i\frac{2\pi}{N}(rm+sn)} \\
&= \sum_{r,s=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} X_{k_2}^d\langle\omega_1\rangle X_{k_2}^d\langle\omega_2\rangle X_{k_2}^d\langle-\omega_1-k\omega_2\rangle\big|_{\omega_1=\frac{2\pi k_2 r}{N},\omega_2=\frac{2\pi k_2 s}{N}} \\
&\qquad \times e^{i\frac{2\pi}{N}(rm+sn)} \\
&= \sum_{r,s=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} X_{k_2}^d[k_2 r]X_{k_2}^d[k_2 s]X_{k_2}^d[-k_2 r-k_2 ks]e^{i\frac{2\pi}{N}(rm+sn)} \\
&= \sum_{r,s=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} \sum_{p_1,p_2,p_3=-\infty}^{\infty} x_{k_2}^d[p_1]x_{k_2}^d[p_2]x_{k_2}^d[p_3]e^{-i\frac{2\pi k_2}{N}rp_1} \\
&\qquad \times e^{-i\frac{2\pi k_2}{N}sp_2}e^{-i\frac{2\pi k_2}{N}(-r-ks)p_3}e^{i\frac{2\pi}{N}(rm+sn)} \\
&= \sum_{p_1,p_2,p_3=-\infty}^{\infty} x[k_2p_1]x[k_2p_2]x[k_2p_3] \\
&\qquad \times \sum_{r=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} e^{-i\frac{2\pi}{N}r(k_2p_1-k_2p_3-m)} \sum_{s=\left\lceil -\frac{N}{2}\right\rceil}^{\left\lceil \frac{N}{2}-1\right\rceil} e^{-i\frac{2\pi}{N}s(k_2p_2-k_1p_3-n)} \\
&= \sum_{p_1,p_2,p_3=-\infty}^{\infty} x[k_2p_1]x[k_2p_2]x[k_2p_3] \\
&\qquad \times \delta[k_2p_1-k_2p_3-m]\delta[k_2p_2-k_1p_3-n] \\
&= \sum_{p_3=-\infty}^{\infty} x[k_2p_3+m]x[k_1p_3+n]x[k_2p_3] \\
&= \tilde{E}\{x[k_2p_3+m]x[k_1p_3+n]x[k_2p_3]\} \\
&= \tilde{E}\{x((k_2p_3+m)T)x((k_1p_3+n)T)x((k_2p_3)T)\}.
\end{aligned}
\tag{A-25}
$$

determining the FT for each integer once and re-using the calculated value through multiplication of the three terms such as is typically done in the calculation of the BiSpectrum. To demonstrate the proper way of generating the correct FBS space, the typical DFT of a discrete signal is shown again

$$X[r] = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} x[n]\, e^{-i\frac{2\pi n}{N}r}, \qquad r = -\frac{N}{2}\ to\ \frac{N}{2} - 1, \tag{A-26}$$

where the time-domain signal is $x[n]$, N is the number of samples, n is the index values of x, and r is the discrete independent values of X taken at integers. It is noted that choosing r to be integer values is the standard technique. From this, other techniques were developed such as the fast-FT to simplify the number of computations to significantly reduce the time of calculating the DFT. However, r does not have to be chosen to be integer values if there is significance to choosing the contrary. In the case of the FBS, it becomes necessary to compute the actual fractional values that pertain to the third term in Eqn. A-23. If one lets $t = -r - ks$, one can predetermine the set of values for t necessary to calculate for the DFT such that the calculation of the FBS can be readily available. It is then of consequence that an N by N matrix of DFT values be calculated prior to the FBS (since r and s that make up t are both of length N). Thus, if the standard computation time to calculate the DFT generates a Fourier space of length N, it would take on the order of N times the computation time to calculate the correct Fourier space necessary for generating the FBS.

A.4 Examples

A.4.1 Two sinusoidal signals

For the first example, a signal with two sinusoidal components were simulated. The frequencies of sinusoidal components are f_1 = 200 Hz and f_2 = 300 Hz. The BiSpectrum of such signal is zero. On the other hand, by selecting a proper k based on known frequencies the FBS would be non-zero. Namely, by choosing k = 0.5, $v_1 = f_1 = 200$, $v_2 = f_1 = 200$, and $v_1 + kv_2 = 200 + 0.5 \times 200 = 300 = f_2$ the FBS is non-zero. To compare the performance of Fourier analysis and FBS analysis, a Gaussian noise is added to the signal and then we tried to detect the frequency of sinusoidal components in the noisy signal. The frequency is detected if the amplitude of the signal in the Fourier or the FBS domain is at least twice as the amplitude of neighbouring noise components in that domain.

The range of the signal SNR is changed from −15 to −30 dB, and for each signal SNR, the detection ratio (the percentage of the correct frequency detection) and the detected SNR (SNR in the Fourier and the FBS domain) are calculated. Figure A-1 shows the detection ratio versus SNRs from Fourier and FBS analyses. The detected SNR that is calculated as the ratio of the power of signal (main lobe and main peak) to the power of the noise (side lobes and side peaks) in the corresponding domain (frequency domain or FBS domain). So, for example, in the frequency domain the signal is shown by a peak or a main lobe (depending on the bandwidth of the signal) and the noise is considered as the side lobe or peak next to the main lobe or peak.

As it is seen in the graphs, the FBS shows better performance both in the detection ratio and the detected SNR. Therefore, by using FBS analysis, one can achieve a better performance over the conventional Fourier analysis techniques.

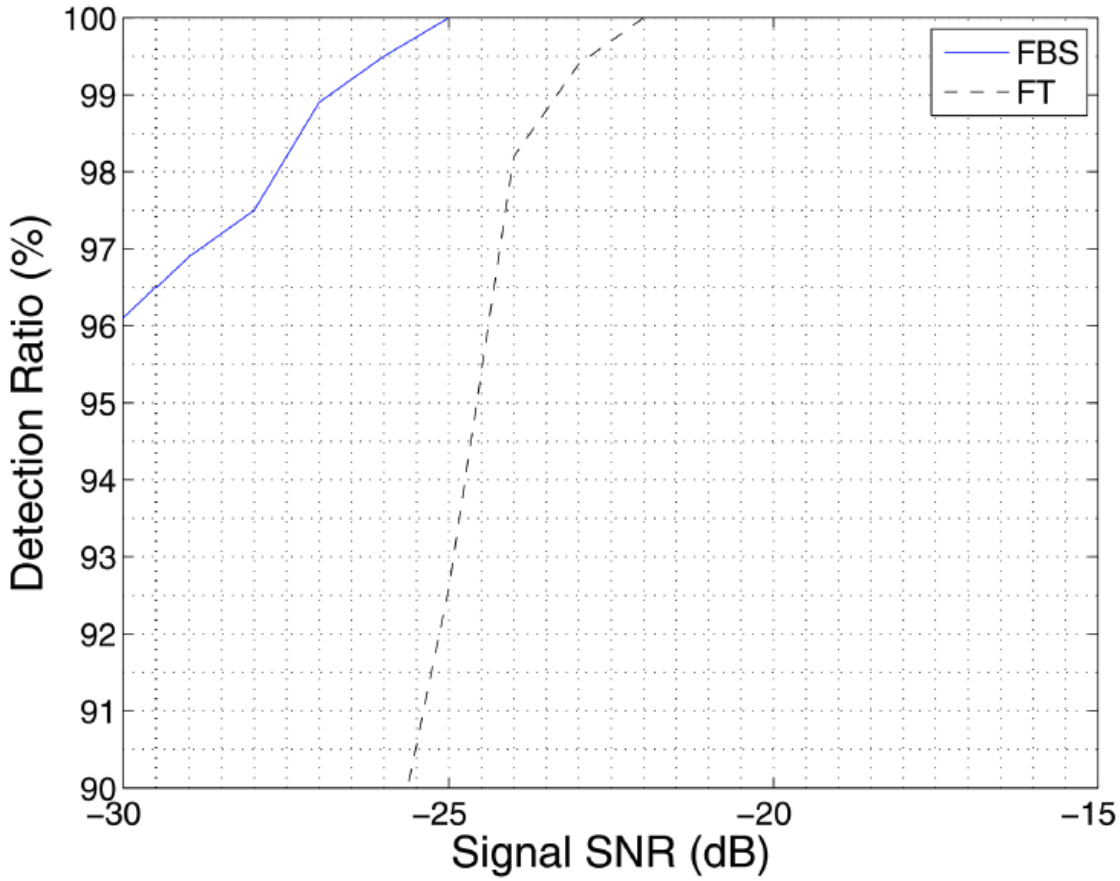

Figure A-1: This show how the Detection Ratio varies with the SNR for FT and FBS.

A.4.2 Interferometry system

In the field of metrology, optical techniques can be used to give information pertaining to the properties of a sample or dimensional information such as the thickness or profile of a sample. One of the common techniques for measuring the thickness of a sample on the range of tens of microns is to project laser light at a particular angle to generate an interference pattern. The setup for this technique is depicted in Figure A-2. As Figure A-3 shows, the interference pattern generates fringes in the *x*-direction. These fringes can be approximated to a singular spatial frequency which is related to the thickness of the sample and the wavelength of the source.

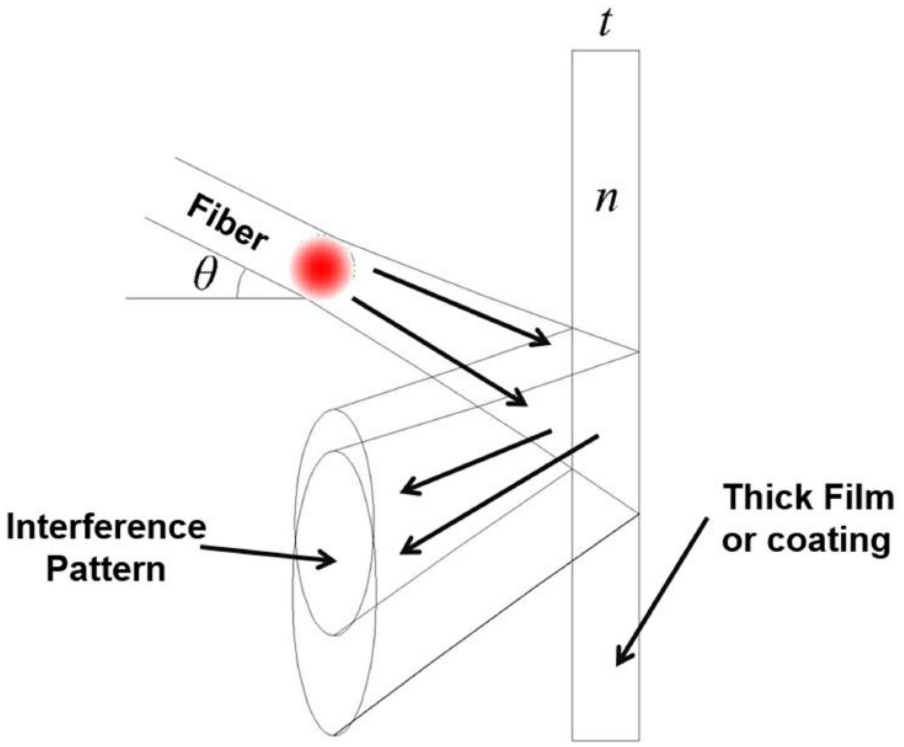

Figure A-2: This show the typical setup for the thickness measurement of a thick film or coating using interferometry technique.

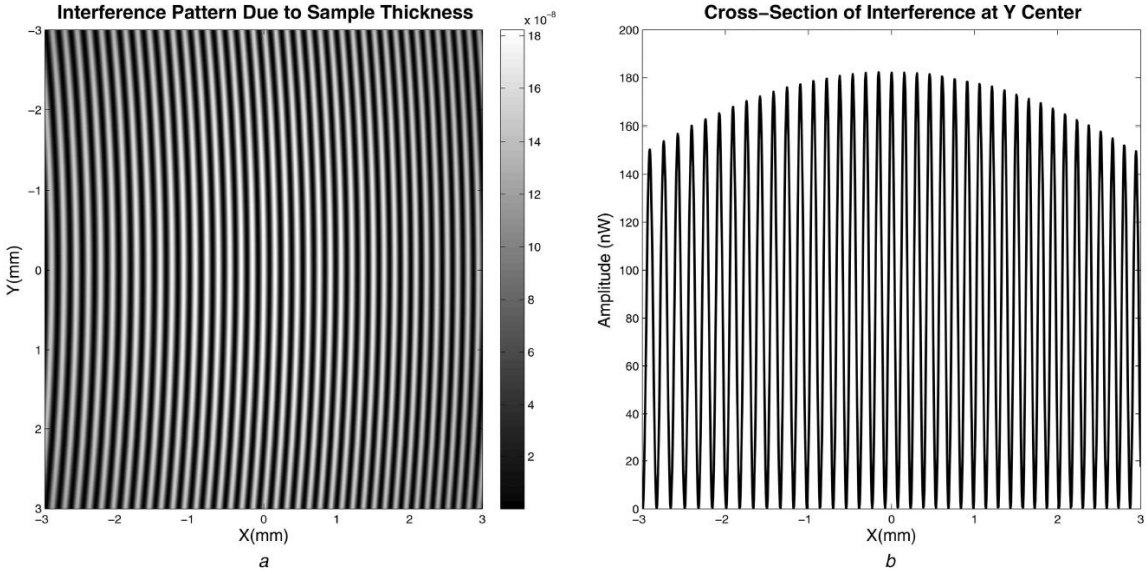

Figure A-3: This show the interference pattern generated from the thickness measurement: (a) over 6 mm by 6 mm area, (b) in the x-direction.

One can derive this frequency as a function of the setup variables using Gaussian beam theory as shown in Eqn. A-27 below.

$$P \simeq \frac{\lambda l_2 \sqrt{n^2 - \sin^2 \theta_i}}{2t \sin \theta_i \cos \theta_i} \tag{A-27}$$

In the expression above, P is the constant spatial period which is a function of n, the index of refraction, θ_i, the incident angle of the beam, l_2, the distance from the fibre to the film, t, the thickness of the film, and λ, the wavelength of the source. While all other components must be held fixed for the purpose of measurement, the experimenter does have the control over changing the wavelength for multiple measurement results. While this technique is good for smooth surface films/coatings, this technique suffers when the sample has a rough surface or fluctuations within the sample causing scatter. The effects of surface roughness can be further simulated to see how the fringe pattern is affected.

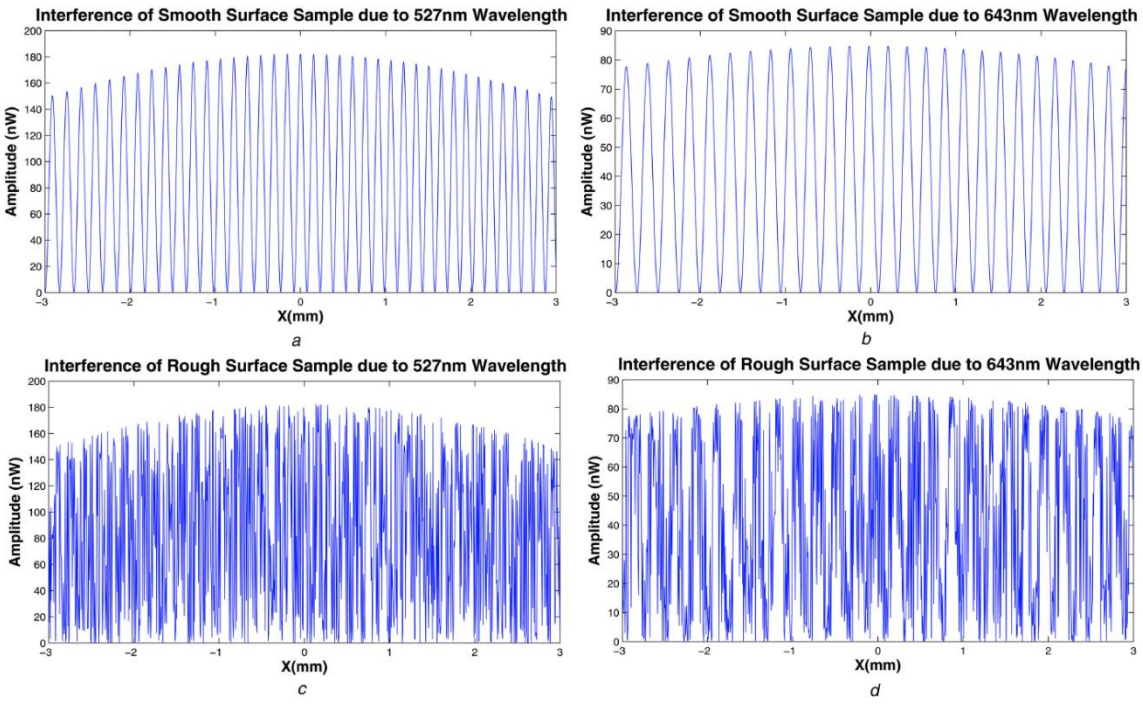

Figure A-4: This shows interference signal cross-sections for:
(a) 527 nm wavelength and clean surface, (b) 643 nm wavelength and rough surface,
(c) 527 nm wavelength and rough surface, (d) 643 nm wavelength and rough surface

Figure A-4 shows the comparison between a clean surface interference pattern and a rough surface interference pattern (50 nm roughness) cross section. As can be seen above, when comparing the fringe patterns of wavelengths 527 nm and 643 nm, it is obvious to see that the spatial period has increased for the 643 nm pattern proportional to the increase in wavelength when compared with the 527 nm wavelength. Since the rough surface simulation results do not show very clear fringe patterns, typically this optical technique does not allow for measuring thickness for these types of surfaces.

However, since the relationship between these two spatial frequencies are well known, the proposed FBS technique can be used to determine the thickness measurement by enhancing the SNR as claimed above. It should be noted that the relationship between the two spatial frequencies is a fraction which is directly related to the wavelength of the sources. Thus, the standard BiSpectrum technique would be incapable of reducing the noise in a system such as depicted in Figure A-4.

A.4.2.1 Interferometry system

To demonstrate the improvement using the FBS technique, a Monte Carlo simulation was generated comparing the typical FT and FBS techniques. The surface roughness of the thickness sample was varied from zero to 150 nm in 5 nm steps. Each roughness was generated by using an additive ‘white Gaussian noise’ to a smooth surface. The simulation was randomized 1000 times per surface roughness to obtain an average detection. The observation of a peak at the spatial frequency which relates to the thickness of the sample was defined as the detection criteria measured as shown in Figure A-5. For a peak to be considered detected, the amplitude at the spatial frequency of interest had to

be at least a factor of 2 greater than the noise surrounding it. As can be seen in Figure A-5, the FBS technique can clearly identify the thickness at levels of higher roughness on the sample. If given a 90% success rate cut-off, the FBS can measure the thickness for 30% rougher surfaces. It should be noted that although the roughness introduced on the system was Gaussian in nature, since the fringe generation is due to the interference between the two beams, the noise in the simulated resultant pattern is not purely Gaussian. Thus, although the FBS technique has over-performed the FT in this setup, the degree in which it has reduced the noise in the system is less than if the noise was purely Gaussian. There are other applications not to be mentioned in this document where the noise embedded in the system can be better modelled as purely Gaussian. Thus, one would expect better performance of this technique in those situations as demonstrated in example 1 (previous subsection).

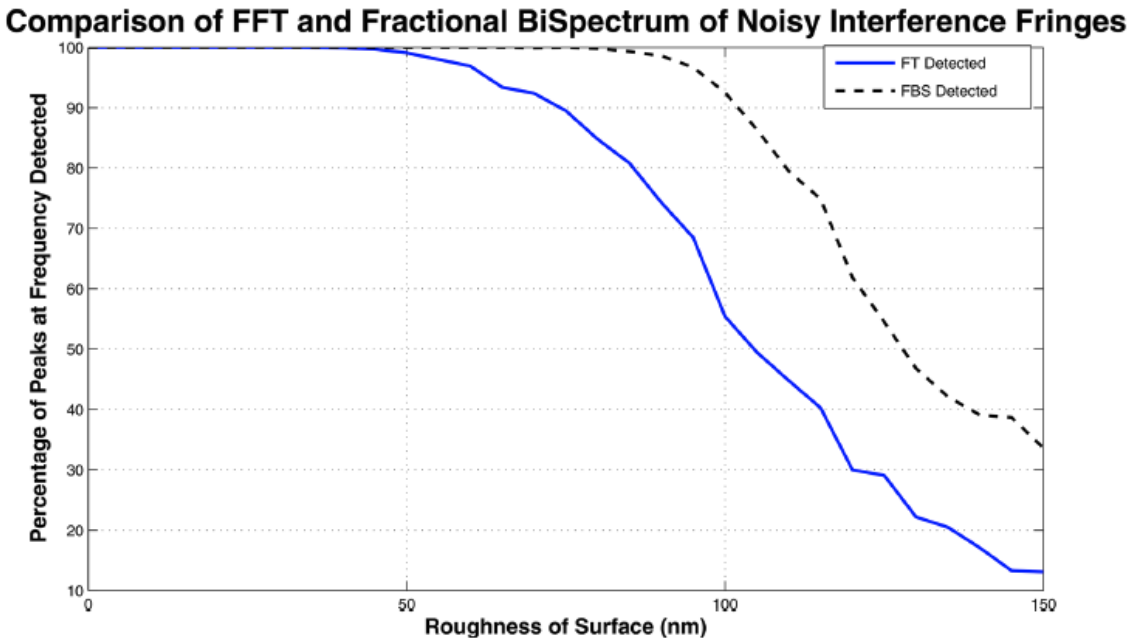

Figure A-5: This shows the Comparison of noise reduction techniques for the thickness measurement setup.

A.5 Conclusion

We introduced the FBS transform as a generalisation of the BiSpectrum transform. The FBS is useful in situations that the BiSpectrum of a signal is zero but in the frequency content of signal, there are some sets of non-zero frequency components of $\{v_1, v_2, v_1 + kv_2\}$ for some $k \in \boldsymbol{Q}+$. This situation happens, for example, when the signal comprised of two sinusoidal signals where the frequency of one sinusoidal signal is not twice of the other. Another example is where the signal is a bandpass signal and the highest non-zero frequency component is smaller than the twice of the lowest non-zero frequency component. In both cases, the BiSpectrum of the signal is zero while the FBS would be non-zero by choosing the proper value for k. The FBS has the advantages offered by the BiSpectrum transform without the limitation related to the frequency content of the signal. For example, for a signal with only two non-zero frequency components in which one frequency component is not twice of the other one, the BiSpectrum is zero while the FBS would be non-zero by choosing the proper value of the parameter k.

A.6 Acknowledgements

This work was supported in part by the National Science Foundation under grant no. 1520151.

REFERENCES

[1] Sundaramoorthy, G., Raghuveer, M.R., Dianat, S.A.S.: 'Bispectral reconstruction of signals in noise: amplitude reconstruction issues', *Inst. Electr. Electron. Eng.*, 1990, **38**, pp. 1297–1306

[2] Jin, S., Wear, S., Raghuveer, M.R.: 'Reconstruction of speckled images using bispectra', *J. Opt. Soc. Am. A*, 1992, **9**, (3), pp. 371–376

[3] Rosenblatt, M., Van Ness, J.W.: 'Estimation of the bispectrum', *Ann. Math. Stat.*, 1965, **36**, (4), pp. 1120–1136

[4] Brillinger, D.R.: 'An introduction to polyspectra', *Ann. Math. Stat.*, 1965, **36**, (4), pp. 1351–1374

[5] Fergusson, J.R., Gruetjen, H.F., Shellard, E.P.S., *et al.*: 'Combining power spectrum and bispectrum measurements to detect oscillatory features', *Phys. Rev. D*, 2015, **91**, (2), p. 023502

[6] Schmittfull, M., Baldauf, T., Seljak, U.: 'Near optimal bispectrum estimators for large-scale structure', *Phys. Rev. D*, 2015, **91**, (4), p. 043530

[7] Biern, S.G., Gong, J.-O., Jeong, D.: 'Nonlinear matter bispectrum in general relativity', *Phys. Rev. D*, 2015, **89**, (10), p. 103523

[8] Pollack, J.E., Smith, R.E., Porciani, C.: 'Modelling large-scale halo bias using the bispectrum', *Mon. Not. R. Astron. Soc.*, 2012, **420**, (4), pp. 3469–3489

[9] Bartolo, N., Fasiello, M., Matarrese, S., *et al.*: 'Large non-Gaussianities in the effective field theory approach to single-field inflation: the bispectrum', *J. Cosmol. Astropart. Phys.*, 2010, **2010**, (8), pp. 1–25

[10] Kakarala, R.: 'The bispectrum as a source of phase-sensitive invariants for Fourier descriptors: A group-theoretic approach', *J. Math. Imaging Vis.*, 2012, **44**, (3), pp. 341–353

[11] Dou, H., Wu, Z., Feng, Y. Feng, *et al.*: 'Voice activity detection based on the bispectrum'. IEEE 10th Int. Conf. on Signal Processing Proc., 2010, pp. 502–505

[12] Li, S., Liu, Y.: 'Feature Extraction of Lung Sounds Based on Bispectrum Analysis'. Third Int. Symp. on Information Processing, 2010, pp. 393–397

[13] Negrinho, R.M.P., Aguiar, P.M.Q.: 'Shape representation via elementary symmetric polynomials: A complete invariant inspired by the bispectrum'. IEEE Int. Conf. on Image Processing, 2013, pp. 3518–3522

[14] Gordon, J.A., Buscher, D.F.: 'Detection noise bias and variance in the power spectrum and bispectrum in optical interferometry', *Astron. Astrophys.*, 2012, **541**, p. A46

[15] Dong, G., Chen, J., Zhao, F.: 'A frequency-shifted bispectrum for rolling element bearing diagnosis', *J. Sound Vib.*, 2015, **339**, pp. 396–418

[16] Tian, X., Abdallaa, G.M., Rehab, I., *et al.*: 'Diagnosis of combination faults in a planetary gearbox using a modulation signal bispectrum based sideband estimator'. 21st Int. Conf. on Automation and Computing (ICAC), 2015, pp. 1–6

[17] Gu, F., Wang, T., Alwodai, A., *et al.*: 'A new method of accurate broken rotor bar diagnosis based on modulation signal bispectrum analysis of motor current signals', *Mech. Syst. Signal Process.*, 2015, **50-51**, pp. 400–413

[18] Guoji, S., McLaughlin, S., Yongcheng, X., *et al.*: 'Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears', *Mech. Syst. Signal Process.*, 2014, **43**, (1-2), pp. 76–89

[19] Peng, Z.K., Zhang, W.M., Yang, B.T., *et al.*: 'The parametric characteristic of bispectrum for nonlinear systems subjected to Gaussian input', *Mech. Syst. Signal Process.*, 2013, **36**, pp. 456–470

[20] Zhou, Y., Chen, J., Dong, G.M., *et al.*: 'Application of the horizontal slice of cyclic bispectrum in rolling element bearings diagnosis', *Mech. Syst. Signal Process.*, 2012, **26**, pp. 229–243

[21] Naumenko, V.V., Solodovnik, V.F., Totsky, A.V., *et al.*: 'Triplet signal generation for bispectrum-based digital wireless communication systems', *Telecommun. Radio Eng.*, 2015, **74**, pp. 695–707

[22] Naumenko, V.V., Solodovnik, V.F., Totsky, A.V., *et al.*: 'Experimental study of bispectrum-based encoding in radio communication system'. Int. Conf. on Antenna Theory and Techniques (ICATT), 2015, pp. 1–3

[23] Molchanov, P.O., Astola, J.T., Egiazarian, K.O., *et al.*: 'Target classification by using pattern features extracted from bispectrum-based radar Doppler signatures'. Radar Symposium (IRS), 2011 Proc. Int., 2011, pp. 791–796

[24] Chen, Y., Liu, J., Lv, S.: 'Modulation classification based on bispectrum and sparse representation in cognitive radio'. IEEE 13th Int. Conf. on Communication Technology, 2011, pp. 250–253

[25] Molchanov, P.A., Astola, J.T., Egiazarian, K.O., et al.: 'Object recognition in ground surveillance Doppler radar by using bispectrum-based time-frequency distributions'. Radar Symposium (IRS), 2010 11th Int., 2010, pp. 1–4

[26] Venugopal, G., Ramakrishnan, S.: ‘Analysis of progressive changes associated with muscle fatigue in dynamic contraction of biceps bacchii muscle using surface EMG signals and bispectrum features’, Biomed. Eng. Lett., 2014, 4, (3), pp. 269–276

[27] Orosco, E.C., Lopez, N.M., di Sciascio, F.: ‘Bispectrum-based features classification for myoelectric control’, Biomed. Signal Process. Control, 2013, 8, pp. 153–168

[28] Chen, X., Zhu, X, Zhang, D.: ‘A discriminant bispectrum feature for surface electromyogram signal classification’, Med. Eng. Phys., 2010, 32, (2), pp. 126–135

[29] Wang, R., Wang, J., Li, S., et al.: ‘Multiple feature extraction and classification of electroencephalograph signal for Alzheimer's’ with spectrum and bispectrum’, Chaos, 2015, 25, (1), p. 013110

[30] Lohmann, A.W., Wirnitzer, B.: ‘Triple correlations’, Proc. IEEE, 1984, 72, (7), pp. 889–901

[31] Matsuoka, T., Ulrych, T.J.: ‘Phase estimation using the bispectrum’, Proc. IEEE, 1984, 72, (10), pp. 1403–1411

[32] Nikias, C.L., Raghuveer, M.R.: ‘Bispectrum estimation: A digital signal processing framework’, Proc. IEEE, 1987, 75, (7), pp. 869–891

[33] Li M., Ming: ‘Fractal time series-A tutorial review’, Math. Probl. Eng., 2010, 2010, pp. 1–26

[34] Li M., Ming: ‘Power spectrum of generalized fractional Gaussian noise’, Adv. Math. Phys., 2013, 2013, pp. 1–3

[35] Li, M., Lim, S.C.: ‘A rigorous derivation of power spectrum of fractional Gaussian noise’, Fluct. Noise Lett., 2006, 6, (4), pp. C33–C36

[36] Isserlis, L.: ‘On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables’, Biometrika, 1918, 12, (1/2), pp. 134–139

[37] Oppenheim, A.V.: ‘Signals and systems’ (Prentice Hall, 1997)