UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

POTENTIAL APPLICATION OF MACHINE LEARNING IN OPTICAL COMMUNICATION SYSTEMS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirement for the

Degree of

MASTER OF SCIENCE

By

MOHAMED ABDALLA ELGREW
Norman, Oklahoma
2017

POTENTIAL APPLICATION OF MACHINE LEARNING IN OPTICAL COMMUNICATION SYSTEMS

A THESIS APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

____
Dr. Kam Wai Clifford Chan, Chair

____
Dr. Pramode Verma

____
Dr. Gregory Macdonald

Acknowledgement

First, I would like to praise almighty GOD for his supporting and strengthening specially during the period of my graduate studies and for my entire life.

I must express my very profound gratitude to my Mother and to my Wife for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis.

I would like to thank my thesis advisor Dr.Chan for his unlimited support and guidance. The door to Dr.Chan office was always open whenever I ran into a trouble spot or had a question about my research.

I will always be grateful to Dr. Verma' s support during the tough times. It would have been very difficult to carry on without his support. Thank you very much for your help.

Finally, I would like to acknowledge Dr. MacDonald for his critical advices during the research time, and I am gratefully indebted to him for very valuable guidance to finish this task.

TABLE OF CONTENTS

List of Tables

List of Figures

ABSTRACT

In this work, we will examine practically encoding information in the state of polarization using M- POLSK. We employ polarization-shift keying (POLSK) to generate an expected pattern of changing states of polarization (SOPs). We generate a random sequence of voltages that represents a random sequence of bits, then we encode this sequence in the state of polarization. we apply the sequence of voltages -using the DAQ assistant- to the polarization modulator to embed this random sequence of bits in the state of polarization. At the far end, we collect the Stokes parameters data of the encoded file using polarization detector. Then, we reduce the dimensions of the collected data to one dimension using a Matlab code. In the final stage of the data recovery, we process the data and discriminate bits using both averaging and machine learning techniques to recover the random sequence that has been sent. Finally, by comparing the sent data set with the recovered data set. We can calculate the efficiency of data recovery process or the bit error rate of the received file. These POLSK symbols will be encoded in a fully polarized light. We will encode binary POLSK, 4-POLSK and 8-POLSK symbols in the SOPs of light in different runs. Also, we will propose the use of averaging to process the Stokes parameters that result from encoding binary POLSK and machine learning techniques to analyze the process of the Stokes parameters that belong to 4-POLSK and 8-POLSK. The state of polarization is presented by five variables. Three of them are the Stokes parameters S1, S2, and S3. The other two are the angles 2γ, 2β. The final dimension is going to be the horizontal angle in the Poincare Sphere

representation. Then, we are going to predict the class (symbol) that belongs to each part of processed data. The length of the data points that represent each symbol is dominated by the sampling rate at the receiver. Symbol prediction will be accomplished using the classification learner's techniques called K-Nearest Neighbor and Support Vector Machine. Basically, these techniques predict the class of the data based on the model built using guided data (row data and its class). Finally, we match the predicted class with the original symbol file to measure the accuracy of the prediction models. In other words, the number of symbols that has been predicted successfully.

Chapter 1: Introduction

1.1 Background & Scope

There are many encoding schemes that have been used to send information over a single mode fiber. These schemes were based on altering amplitude, frequency, phase and the polarization of the light wave. As a potential replacement or dual use of the standard coherent modulation techniques like ASK, FSK, PSK and DPSK to coherent optical communications, modulation methods exploiting the vector characteristics of the propagating light radiation have been proposed. The use of the state of polarization (SOP) of a fully polarized light wave as the information carrier parameter exploits the two orthogonal channels available in single-mode fiber propagation. In single-mode fibers fed by a monochromatic light source, orthogonal SOP pairs at the input lead to orthogonal output SOP pairs although the input state of polarization is not maintained in general [1].

The modulation of the state-of-polarization (SOP) of a light wave has been attractive for digital optical transmission systems [1]. Polarization-shift keying (POLSK) is one of the proposed modulation schemes for practical transmission systems [2]. The use of POLSK have been theoretically analyzed and experimentally demonstrated for Coherent and direct detection systems [1]. Because of the previous extensive research work, many research outcomes have shown that depolarization phenomena or polarization dependent losses are of little importance even after relatively long fiber spans. The most effective phenomena is the rigid rotation of the constellation of signal points on the Poincare sphere. In the meanwhile, signal points will keep the spatial relationship between them [1]. The advantages of POLSK over other modulation schemes are: it is a constant-power modulation, it is less vulnerable to nonlinear fiber effects and multilevel POLSK shows a good sensitivity because the minimum distance between signal points of the constellation can be kept large by exploiting the three-dimensional Stokes space. For these reasons, POLSK seems to be promising for long-haul communications [3][4].

There are different approaches that have been proposed to apply POLSK. Some research has proposed the conversion of wavelengths of multichannel polarization shift keying by using four wave- mixing [5]. Others, proposed to use the modulate the state of polarization and the optical power to send more constellation [6]. Also, the combination of altering both polarizations and phase of light wave was one of the suggested way to apply the POLSK [7]. In the addition to the way that POLSK was applied, many works investigated the way how the constellation recovered. Most of them were focused on separating the Stokes parameters using complicated receivers [8]. The use of machine learning techniques was proposed for the first time to compensate the fiber nonlinearity and phase noise last year in the Journal of Light Wave Technology [9]. But it has never been proposed to handle the affection of birefringence and rigid rotation when we use POLSK.

1.2 Thesis Statement and Contribution

The research community has been talking about polarization shift keying for decades as the innovative approach of the encoding schemes in optical communications. There were many ways that have been used to recover information encoded using POLSK schemes. Since we need to track the polarization due to the rigid rotation, we propose to use machine learning techniques to learn the channel behavior as an alternative approach to the available depolarization methods. Because machine learning offers huge capabilities to learn and model the behavior of the fiber and the rigid rotation effect, we propose to use a machine learning classification learner to learn the channel behavior by creating a model that eventually will be used to classify the processed Stokes parameters to recover the bit sequence. This model can be continuously updated with the channel behavior by using supervised training periodically in the data stream. Updating the model periodically prevents the miss classification of the Stokes parameters after being processed.

> **Thesis Statement:** *it is possible to use machine learning models to recover data points that have been encoded in the state of polarization using (POLSK) for short distance communication.*

1.3Thesis Outline

The rest of the thesis is organized as follows: Chapter 2 addresses the theoretical foundation of light where the author introduces the nature of light, polarization, Poincare sphere and birefringence. In Chapter 3, the author discussed the relevant works in literature focused on polarization shift keying and the modulation scheme that has been used to encode information in the state of polarization. Also, the author discussed some machine learning techniques that has have been used in the data processing of the experimental work. In Chapter 4, the author illustrated in detail the experiment setup, experiment procedures, Stokes parameters processing, information recovery and the observations of using 4-POLSk and 8-POLSK modulation schemes.

Chapter 2: Theoretical Foundation

2.1 Nature of Light

Light is an electromagnetic radiation that consists of oscillating electric and magnetic fields. The main components of this wave are the electric field E and magnetic field H vectors which are mutually perpendicular and are both orthogonal to the direction of propagation S (also called the Poynting vector). When we talk about the polarization, we usually consider only the direction of the electric field vector E [10].

The figure below (Figure 1) shows the Cartesian coordinates. If we assumed that the direction of propagation S is along the positive z-axis (towards the reader), we can write the electric field E vector (for a quasi-monochromatic source) as the sum of two orthogonal components.

$$E0 = \hat{i}E0x + i\hat{j}E0y$$

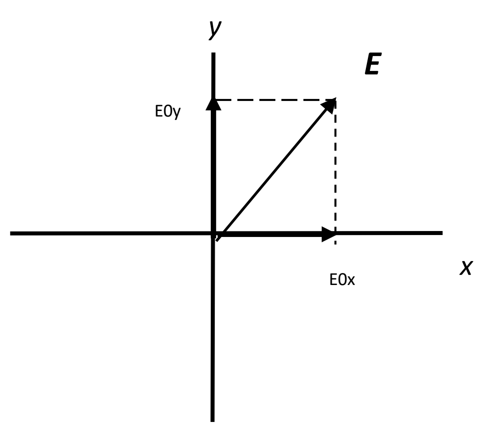

Figure 1 Cartesian presentation of Electric field.

The light wave is represented by the following function:

$$E = E_0 e^{i(kz-\omega t)}$$

where, k is the propagation constant, z is the distance in the direction of propagation, and ω is the angular frequency [11].

2.2 Polarization

Polarization is a parameter applying to transverse waves that specifies the geometrical orientation of the oscillations. In electromagnetic waves, polarization refers to the direction of the electric field. There are three polarization classifications: linear polarization, circular polarization and elliptical polarization. The light is linearly polarized, if E_0 is real. The light is elliptically polarized, if E_0 is complex. The light is circularly polarized if the real and imaginary parts of E_0 are equal. The result of the combination of both perpendicular components with different amplitudes and phases dominate the trace of the electric field E with respect to time which is called state of polarization. In the case of circular polarization and elliptical polarization, the direction of rotation of the electric field can be clockwise, the polarization state is called right-handed, or direction of rotation is counterclockwise, the polarization state is called left-handed [10].

2.3 Poincare Sphere

The Poincare sphere is a graphical tool in real, three-dimensional space that allows convenient description of polarized signals and polarization transformations caused by propagation through devices. Any state of polarization can be uniquely represented by a point on or within a unit sphere centered on a rectangular (x,y,z) coordinate system. The coordinates of the point are three normalized Stokes parameters describing the state of polarization [12].

Even though the polarization is usually described by the electric field of the light wave, there are other methods used to measure the polarization. It is common to use the Stokes parameters components as measurement to state polarization. Stokes parameters are determined by a set of intensity measurements taken when light is passed through several types of polarizers. Stokes parameters can be used mutually with Poincare sphere to represent the state of polarization [11].

The value of each parameter is a function of measured intensity levels. The four Stokes parameters are defined as follows [13]:

s_0= total power (polarized + un polarized)

$$s_1 = s_0 \cos 2\gamma \cdot \cos 2\beta$$

$$s_2 = s_0 \cos 2\gamma \cdot \sin 2\beta$$

$$s_3 = s_0 \sin 2\gamma$$

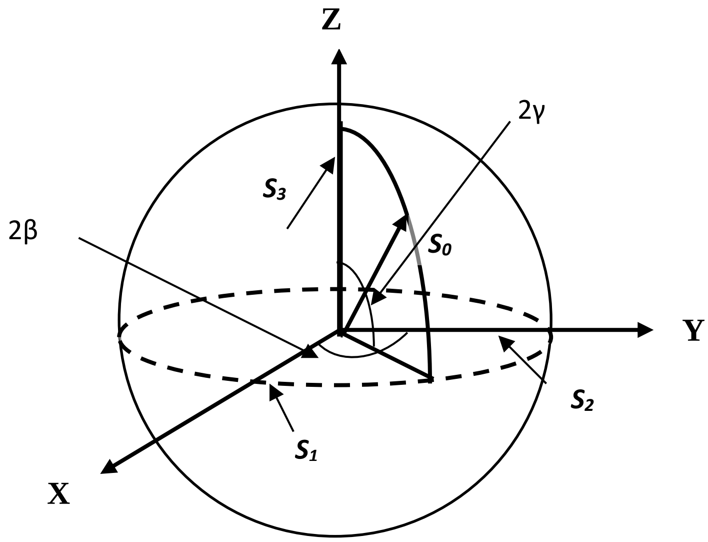

Figure 2 Stokes Parameters and the Poincare Sphere [11].

The physical interpretation of the above is as follows:

s_0 = Total power of the received signal (polarized + unpolarized)

s_1= Power received through a horizontal linear polarizer – Power received through a vertical linear polarizer

s_2= Power received through a 45-degree linear polarizer –Power receive through a – 45-degree linear polarizer

s_3 = Power received through right hand circular polarizer –Power received through a left hand circular polarizer

The "normalized" Stokes parameters are given by:

$$s_1 = s_1/s_0$$

$$s_2 = s_2/s_0$$

$$s_3 = s_3/s_0$$

where for fully polarized light,

$$s_0 = \sqrt{(s_1)^2 + (s_2)^2 + (s_3)^2}$$

2.4 Birefringence Phenomena

Birefringence is the optical phenomena which happens when a material having a refractive index that changes with respect to the polarization and propagation direction of light wave [14]. These optically anisotropic materials are said to be birefringent (or birefractive). The birefringence is often quantified as the maximum difference between refractive indices exhibited by the material.

There is also a slow time-varying birefringence component that is random and unpredictable. The result is that the original state of polarization at launch is transformed as it propagates through the fiber [11]. The good news is that if two mutually orthogonal light signals are launched into a fiber, they emerge at the receiver with their orthogonality preserved [15].

Chapter 3: Literature Review

Many approaches have been used to modulate and demodulate data in the SOPs of light using the POLSK scheme [16] [17] [18] [19] [24]. All these methods were aiming to achieve efficient reception of data with minimum cost and highest spectral efficiency [23]. Moreover, these demodulation techniques tended to save the power consumption as well as include as maximum as possible symbols in the polarization channel.

3.1 POLSK

The use of the state of polarization (SOP) of light wave as a carrier of the information, taking advantage of the two orthogonal components of the electrical field of light in both free space or in single-mode fiber is called polarization shift keying (POLSK). Orthogonal SOP pair in the input will result orthogonal SOP pair in the output, even though the out state is not the same as the input state of polarization. In other words, the spacing between the two SOP will be kept fixed [1]. The detection of the encoded information is done by the analysis of the state of polarization. The main representation of the SOP in the Poincare sphere is the Stokes parameters. By processing the Stokes parameters at the output side, the encoded information can be recovered. There are two ways of encoding information in the two electrical field components: one way is encoding the information dependently and this is what happened in the case of using the POLSK. The other way is to encode information independently and that occurred in the case of multiplexing.

3.1.1 2-POLSK

The binary POLSK is a modulation scheme that switches two linear orthogonal SOPs to represent the information. In the three-dimension space, these two SOPs are represented by two antipodal points. To detect the information of binary POLSK, a reference vector is needed to decide which SOP has been received. Based on the sign of the scalar product of the received SOP vector and the reference vector, a decision can be easily taken in the absence of the noise. In the case of the noise, the reference vector need to track to overcome the induced changes done by the fiber and this is the reason why SOPs of 2-POLSK systems need to be tracked [1].

3.1.2 M -POLSK

In M-POLSK modulation scheme, SOP of the input light wave is being changed in such a way that the corresponding SOP point in the Stokes space is moved onto one of the M points belonging to the signal constellation. The transmission of multilevel POLSK seems to be more efficient in the three-dimensional Stokes space. There are many ways to transmit multilevel in POLSK. But there is one common issue between all ways that the power is constant [1]. In case of 4-POLSK systems. schemes differ in signal geometry: in one case, we have four signals lying on a maximum circle over the Poincare sphere as in figure 3.

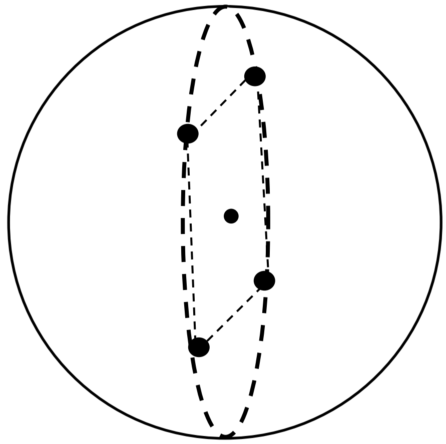

Figure 3 Four signal points on one circle.

Figure 4 shows signals are the vertices of a tetrahedron inscribed into the Poincare sphere

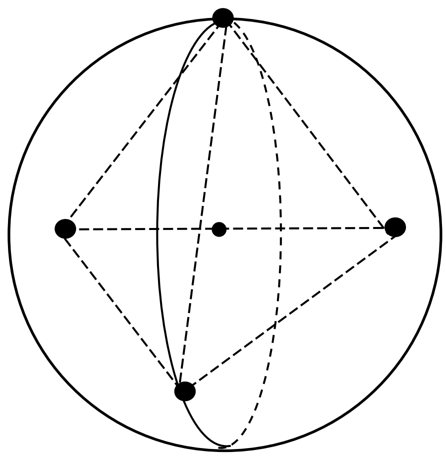

Figure 4 Signals lying on tetrahedron inside the Poincare sphere.

For 8-POLSK scheme, the signal set is made up by the vertices of a cube inscribed into the Poincare sphere [2] [22] as in Figure 5.

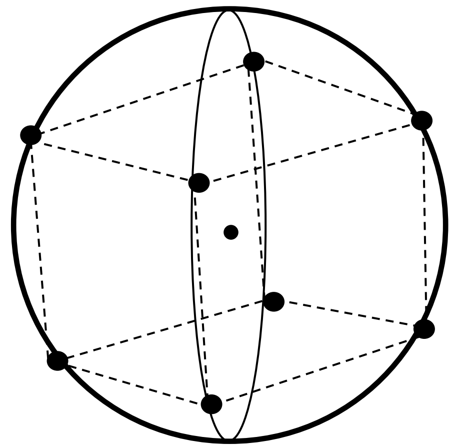

Figure 5 Signals on the cube inside the sphere.

Also, the signals can be on the great circle on the Poincare sphere as in Figure 6.

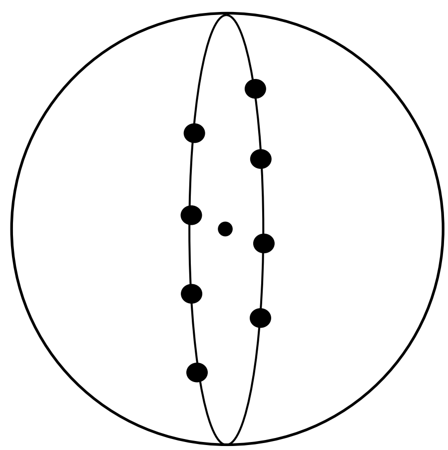

Figure 6 Eight signal points on one maximum circuit.

3.2 Classification Models

Classification models predict categorical class labels. Usually classification models are used to perform two types of predictions. These two prediction types are a continuous valued prediction, in this case the prediction is the probability of membership to a certain class and its value between 0 and 1. In addition to a continuous prediction, classification models provide a predicted class; in this case the prediction is one of discrete categories [20]. In this thesis, we are going to use a discrete category prediction to classify the data processed to one of the original categories (symbols). The two main classifiers that have been used are K-Nearest Neighbors and Support Vector Machines.

3.2.1 K-Nearest Neighbors

In the K-nearest neighbor classification model, a new sample is predicted based on the K-closest data points in the training set. The classification is performed by using a sample's geographic neighborhood to predict the sample's classification. KNN for classification predicts a new sample using the K-closest samples from the training set. "Closeness" is determined by a distance metric. For example, a 5-nearest neighbor model is shown in Figure 7. In this example, two new samples are going to be predicted. The first sample is indicated by the symbol (●) near a set of mixed classes. Since three out of the five neighbors are class one, then it is classified to class one. The other sample (▲) has all five points; neighbors are from the second class. So, it is predicted as second class [20]. The number of neighbor's K should be selected in a way that avoid ties. In other words, if the number of classes are even, then K should be an odd number as in the example below K = 5 for two classes case. If two or more classes are tied for the highest estimate, then the tie is broken at random or by looking ahead to the K + 1 closest neighbor [20] [21].

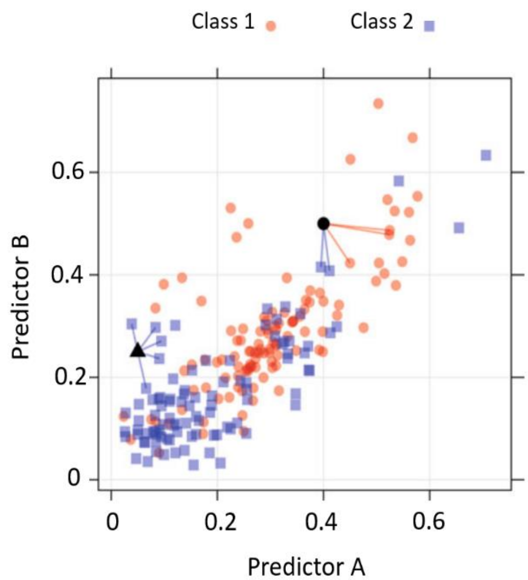

Figure 7 The K-nearest neighbor classification model. Two new points, symbolized by filled triangle and solid dot, are predicted using the training set [20].

3.2.2 Support Vector Machines

Support vector machines are a class of statistical models first developed in the mid-1960s by Vladimir Vapnik. In later years, the model has evolved considerably into one of the most flexible and effective machine learning tools available. Initially, the model was developed in the classification setting. In 2010, Vapnik developed the regression version of Support Vector Machine which was the extension of the classification model. The example of data set that has two classes is shown in the left side of Figure 8. It is clear that the two classes are separate. Also, there are an infinite number of boundaries that separate the two classes. The question here is what is the best optimum boundary between all these boundaries? And what are the measures that can be considered to evaluate these boundaries?

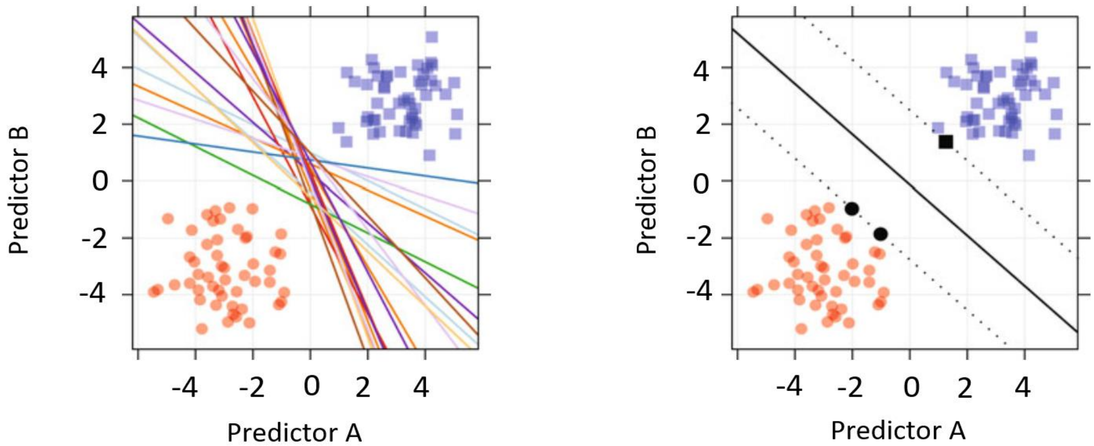

Figure 8 *Left*: A data set with completely separable classes. An infinite number of linear class boundaries would produce zero errors. *Right*: The class boundary associated with the linear maximum margin classifier. The solid black points indicate the support vectors [20].

A metric was defined called the margin. Basically, the margin is the distance between the classification boundary and the closest training set point. The right side of Figure 8 shows one possible classification boundary as a solid line. The distance between the dashed lines and the boundary is the margin in this example and it's the maximum between the training data set and the classification margin. In this example, the solid triangle and circle points in the right side of Figure 8 are equally closest to the classification boundary. The quantity of the margin can be used to evaluate the classification models. For SVM, the slope and the intercept that achieve the maximum distance between the classification boundary and the nearest training data set is known as the maximum classification margin [20].

Chapter 4: Experimental Work

4.1 Experiment Setup

A simple block diagram of the experiment setup is shown in Figure 9. We have used the Agilent 8509c as a light source (Agilent 8509C is a polarization modulator, but we use it as a light source in this set up. The wavelength range is from 1310 nm to 1550 nm). The light source is connected to the input of Versawave - Versawave Technologies currently a division of Optelian - polarization modulator (40Gb/s bandwidth, Bias Voltage -12Vto 12V, Wavelength range 1530nm – 1610nm). The wavelength that was used in this experiment is 1550nm. When the light passes through the polarization modulator, the polarization changes as function of the applied DC voltage on the modulator. The output of the Versawave polarization modulator is connected to the POD-201 Polarimeter (manufactured by General Photonics- monitoring and logging SOPs up to 1 billion points, sampling rate 4Msamples/s). The POD-201 – Polarimeter works as a receiver in this diagram where it collects and stores the Stokes parameters that carries the encoded information. The light is sent from the light source to the POD through the polarization modulator. When the light passes through the polarization modulator, a piece of Lab View code at PC1 applies voltages at a certain speed. These voltages correspond to a sequence of random bits that was created using a math lab code. POD samples the Stokes parameters with a rate that is at least double of the speed of the applied voltages (Nyquist Rate).

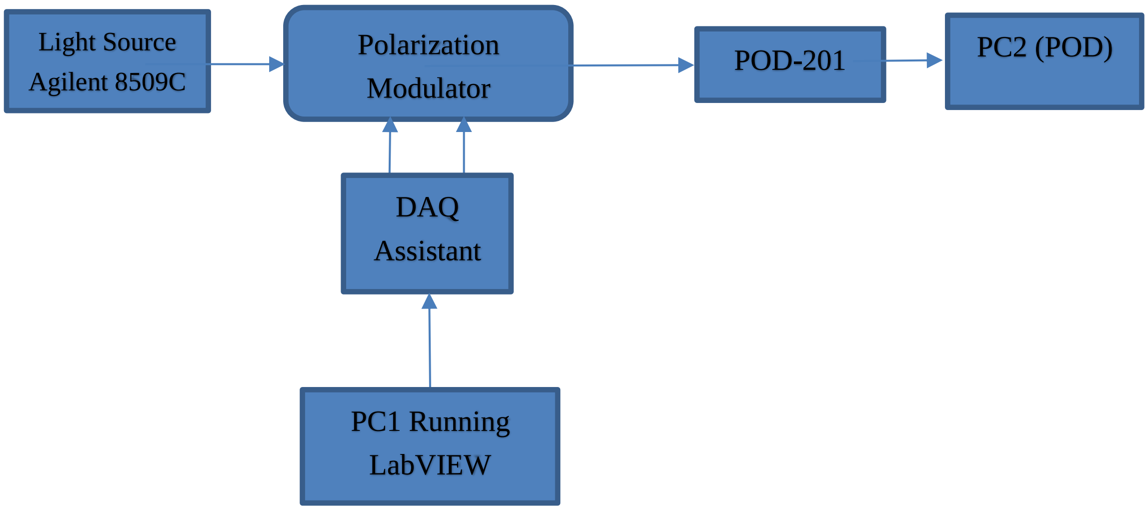

Figure 9 Experiment setup diagram.

4.2 Experiment work

This experiment aims to prove the possibility of encoding information using M-POLSK and recover the information at the far end using machine learning techniques. The distance between the transmitter and receiver is a short distance (2 Meters). We generated a sequence of bits using a piece of code in math lab (code#1-Appendix A). These bits are mapped to voltages based on the POLSK scheme used. 2-POLSK, 4-POLSK and 8- POLSK were examined during this experiment. In 2-POLSK scheme two voltages were used to represent one bit as in Table-1. In 4-POLSK, four voltages were mapped to two bits (each one voltage or symbol presents two bits at a time) as in Table-2. In 8-POLSK scheme each symbol presents three bits (see Table-3). The author applied all the voltages between (-10V,10V) with 0.5V increment to the polarization modulator to examine the behavior of each individual voltage. He selected the optimal voltages that achieves the best resolution between the state of polarization and voltages.

Voltage(V)	Bit
2.5	0
-2.5	1

Table-1 2-POLSK Voltage Mapping

Voltage (V)	Bits
4	00
6	01
8	10
10	11

Table-1. 4-POLSK Voltage Mapping

Voltage (V)	Bits
-4	000
-2	001
0	010
2	011
4	100
6	101
8	110
10	111

Table-3. 8-POLSK Voltage Mapping

4.3 Stokes Parameters Processing

The information is encoded in the Stokes parameters S1, S2 and S3. We used voltages to express this information. Each symbol (voltage) is represented by one state of polarization in the three-dimension space (Poincare sphere). This SOP can be represented by S1, S2, S3 and the angles 2γ, 2β. The polarization modulator has in the input side a linear polarizer. This linear polarizer results that all SOPs are on one great circle. To recover this information, we need to reduce the dimension of the SOP presentation to only one dimension. This dimension can be used to discriminate between the symbols that are encoded. In the POD software, a visual presentation shows the SOP track in a great circle. Since we are sampling five times of the encoding rate at the receiver (POD) (sampling rate = 5 * encoding speed), the receiver will detect some points in transit as in Figure 10.

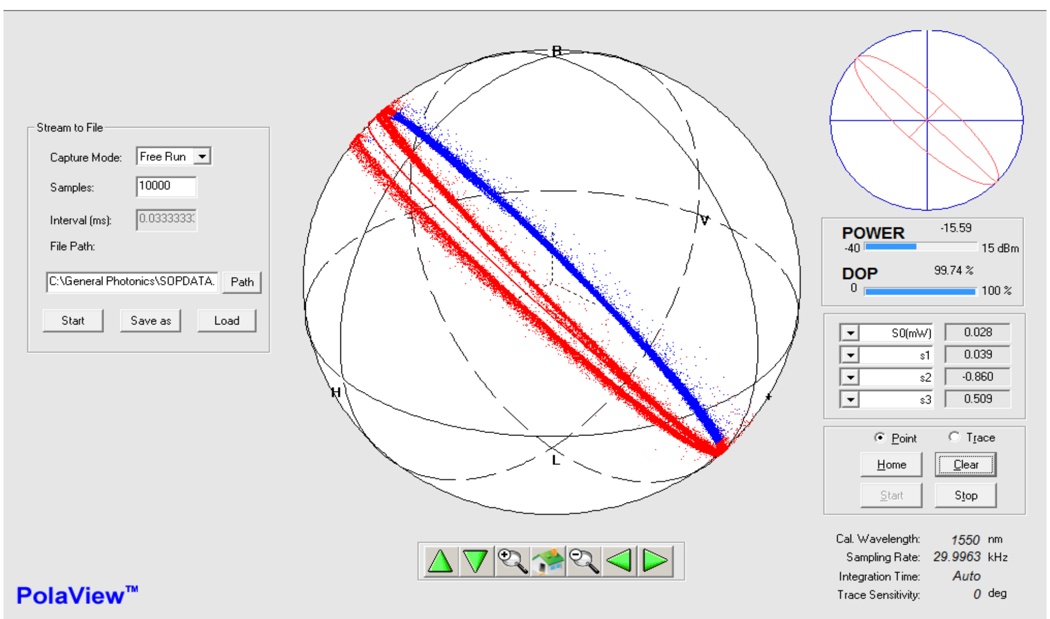

Figure 10 Pola View Screenshot shows the SOP great circle for 8-POLSK.

A Matlab code (code#2- Appendix A) was developed to project the great circle that has been performed by sampling different symbols (Voltages) at higher speed. That means every single point is going to be projected on X-Y plane. The benefit of this projection is to reduce the dimension of the SOP representation to one dimension which is the angle 2β in this case. In other words, based on the values of the angle 2β, we will be able to decide which symbol was sent. The mathematical explanation behind the projection process is as following:

i- we find three sets of points on the great circuit for each Stoke parameter and take the mean of each of the three sets. We make sure that these sets are apart by applying some conditions.

ii- We normalize the mean of the three reference points.

iii- We create vector using the three Stokes S1, S2 and S3. The Vectors are SV1, SV2 and SV3.

iv- Form a triangle using the vectors SV1, SV2, SV3. Since all the SV points lie on the plane of the triangle, and it is in general not necessary that the origin also lie on this plane, therefore, we can construct the normal vectors using the v1 and v2 defined below.

v1 = SV1 - SV3 v2 = SV2 - SV3

v- Then, we find out the perpendicular vectors to the vectors v1 and v2 as following:

$\overrightarrow{e_1} = \overrightarrow{v_1}/|v_1|$

$\overrightarrow{e_2} = \overrightarrow{v_2} - \overrightarrow{e_1}.(\overrightarrow{e_1}'.\overrightarrow{v_2})$

$\overrightarrow{e_2} = \overrightarrow{e_2}/|e1|$

$\overrightarrow{e_3} = \overrightarrow{e_1} \times \overrightarrow{e_2}$

$\overrightarrow{e_3} = \overrightarrow{e_3}\ /|e_3|$

vi- Now, we can find the transformation matrix for the projection.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = M.\overrightarrow{e_1}$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = M.\overrightarrow{e_2}$$

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = M.\overrightarrow{e_3}$$

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = M.(e1\ e2\ e3)$$

M = $(e1\ e2\ e3)$^-1

4.4 Information Recovery

The experiment was run for different POLSK schemes with different encoding speeds and information recovered by different information recovery methods. We applied number of random bit sequences of 2-POLSK, 4-POLSK and 8- POLSK to the polarization modulator. For 2-POLSK, we mainly used the averaging approach. Although averaging might be the easiest way to recover the information, it cannot be maintained for a long time since the mapping between the voltages and SOP is not fixed because of birefringence phenomena. First, we pick two voltages that corresponds to two SOPs separated with enough distance to distinguish in the Poincare sphere space. Second, we apply the sequence of voltages that represents a random sequence of bits one and zero. Simultaneously, we collect the Stokes parameters that presents the encoded sequence. Third, we take the average of the angle 2β points that belong to each symbol after SOP being projected to the X-Y plane. Finally, we run a code to decide which symbol was sent based of the SOP position in the sphere. The other methods that we used to recover information were machine learning technique. We used K – nearest neighbors and support vector machine algorithms to recover information that has been encoded using 4-POLSK and 8-POLSK. These algorithms were found in the Matlab application called Classification Learner. Initially, we apply the sequence of voltages that presents 2 or 3 bits (4 Voltages for 2 bits and 8 Voltages for 3 bits). At the same time, we collect the Stokes parameters that belong to this sequence. Then, we process the Stokes parameters by reducing the dimension to one dimension (symbol is function in 2β). Also, we find out the start point of samples that represent the symbols. By knowing the sampling rate at the POD, we can associate the samples points with the correspondent symbol. After we organize the angle and voltage, we build our model classifier by providing training data set to software. This data set is organized in columns and rows in a way that shows each voltage (bits) and the associated angle values. Finally, we use this model to classify raw data. The classifier will predict the voltage associated with each set of samples. Then, author calculate the accuracy by comparing the predicted voltages with initial sequence of voltages.

4.4.1 2-POLSK Observations

This section will describe the 2-POLSK results. For 2-POLSK, we used the averaging method to recover encoded data. After the generation of a random binary sequence of 1Kbit size, we used two voltages to express the bits 0 and 1. The bit 0 is mapped to 2.5V and the bit 1 is mapped to -2.5V. This voltage file was applied to the polarization modulator. At the same time, the Stokes parameters were collected at the far end (POD-201). Figure (11) a, b, c shows the Stokes parameters vs the time. The collected Stokes parameters in 3D space are shown in Figure 12a. These Stokes parameters were processed and projected to the X-Y plane as in Figure 12b to make the voltages (Symbols) function in one variable which is the angle as shown in Figure 12c. We averaged the five points that represent each voltage. Based on this value, the decision has been taken if it is 2.5V or -2.5V or not. Finally, we matched the recovered file with the initial file. The accuracy of data recovery was 100%. That means we could recover back all the encoded 1Kbits file.

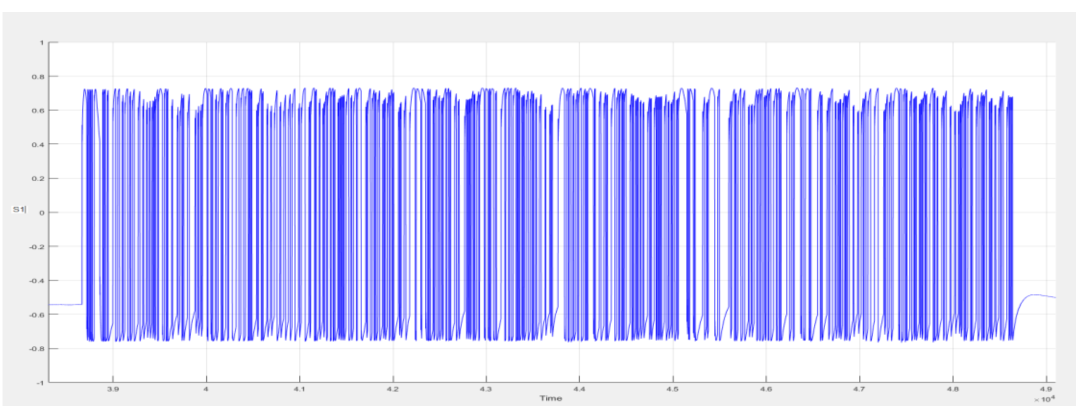

Figure 11a S1 vs Time for 2-POLSK.

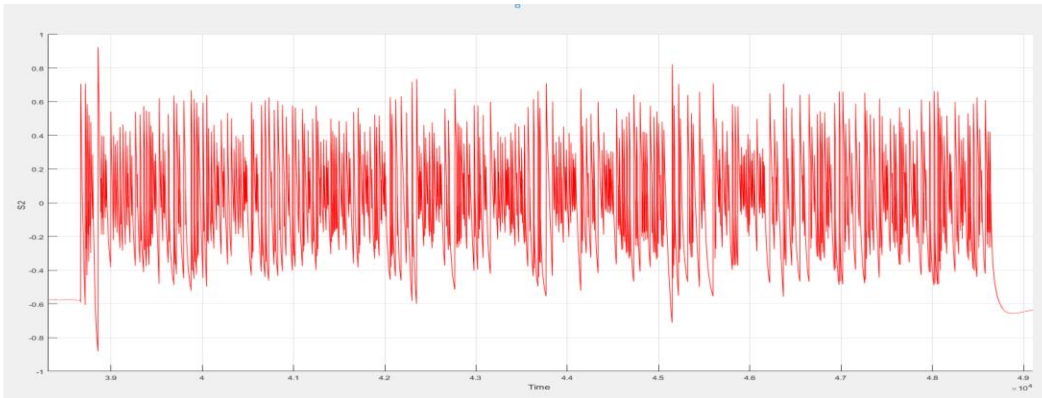

Figure 11b S2 vs Time for 2-POLSK.

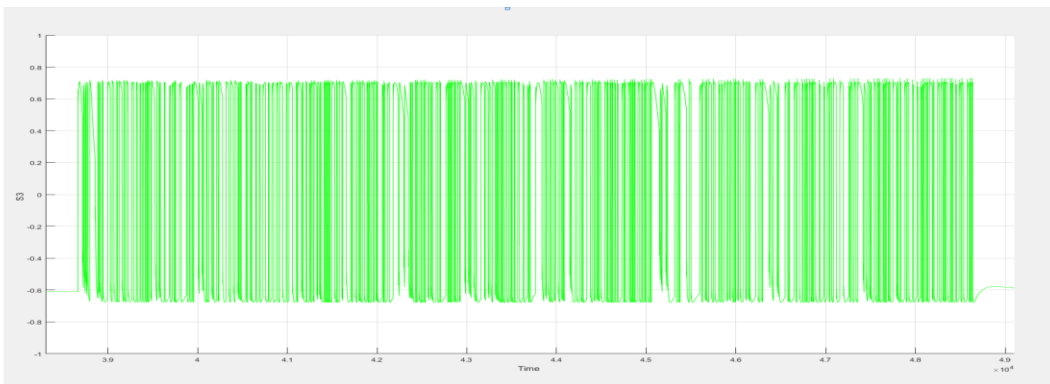

Figure 11c S3 vs Time for 2-POLSK.

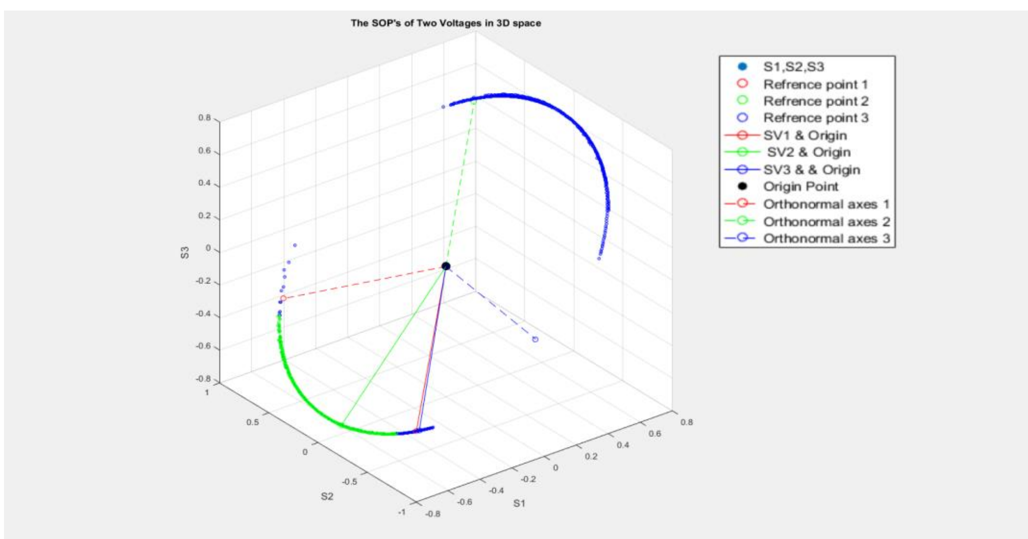

Figure 12a Stokes Parameters in 3D Space.

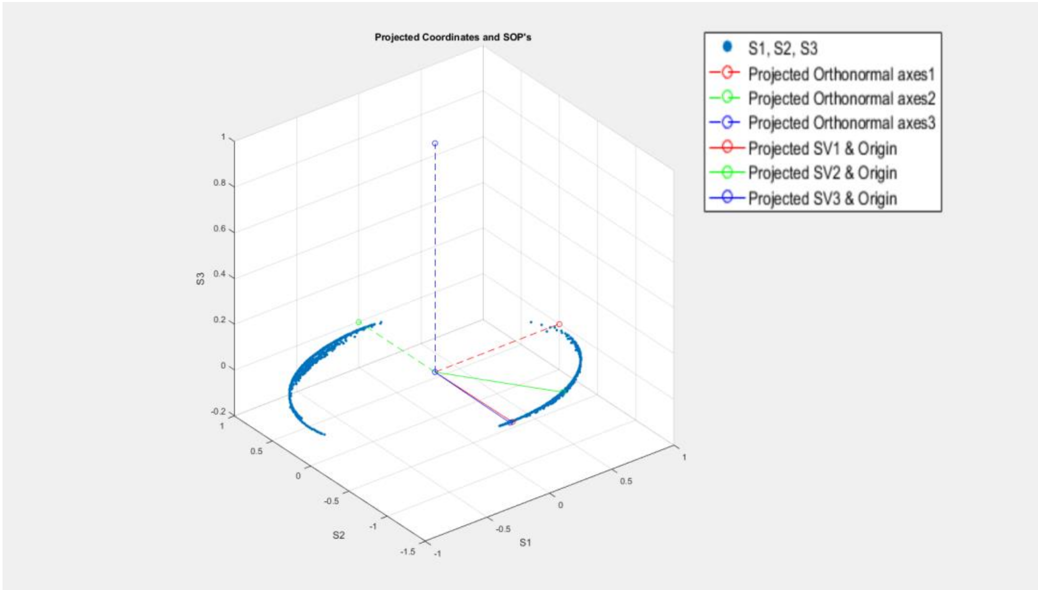

Figure 12b Projected Coordinates and SOP in 3D space.

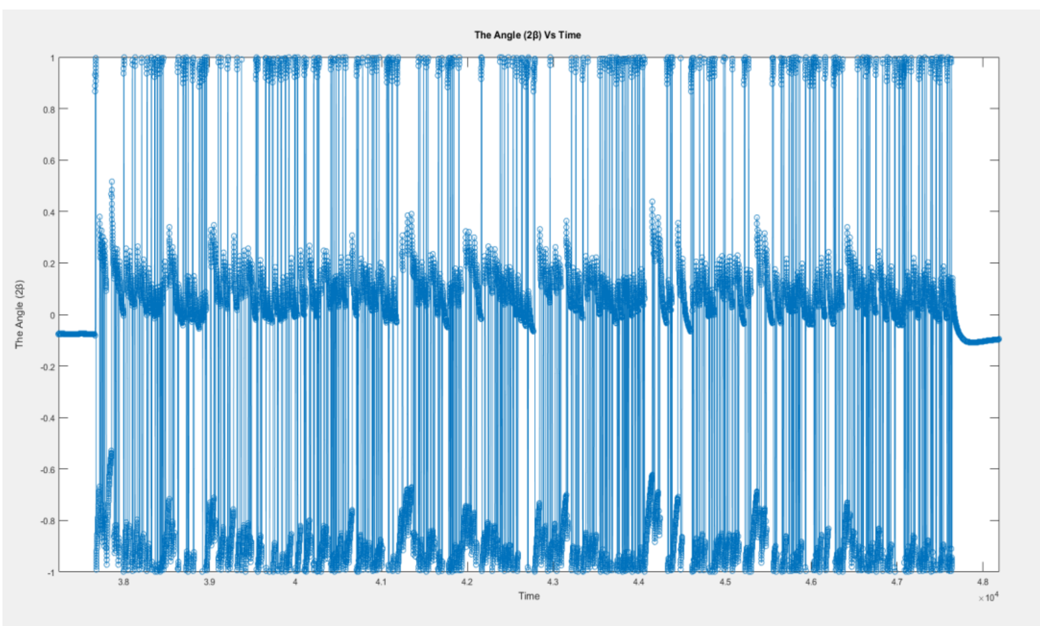

Figure 12c The Angle 2β vs Time for 2-POLSK.

4.4.2 4-POLSK Observations

First, we generated the file of random N Kbits. We mapped each two-different possible bit combination to one of the four voltage steps as in Table-1. The final size of voltage file is N/2 Kbits. We apply this file to the polarization modulator using the LabView code to encode the mapped bits in the polarization of the passing light. We collected the s Stokes parameters that belong to the encoded bits at (POD-201). Figure 13 a, b, c, d. shows the stokes parameters S1, S2, S3, the angle 2β vs time.

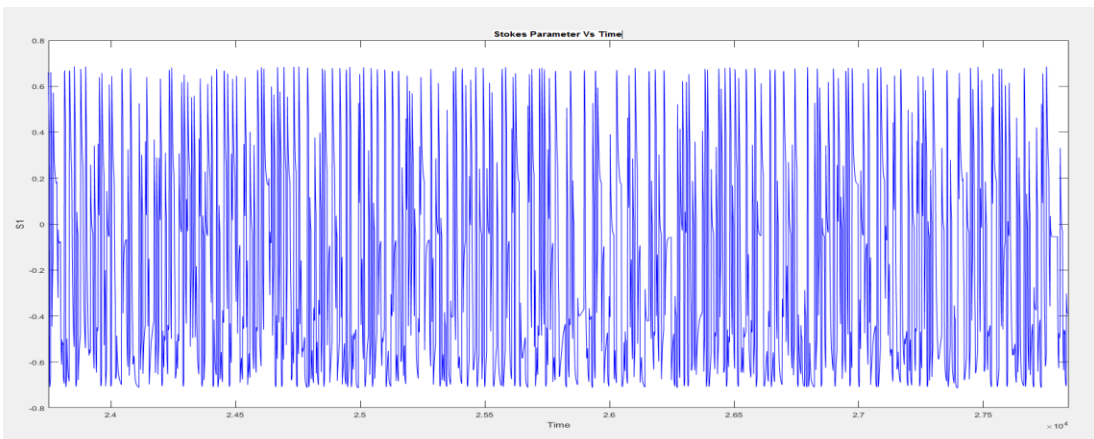

Figure 13a S1 vs Time for 4-POLSK.

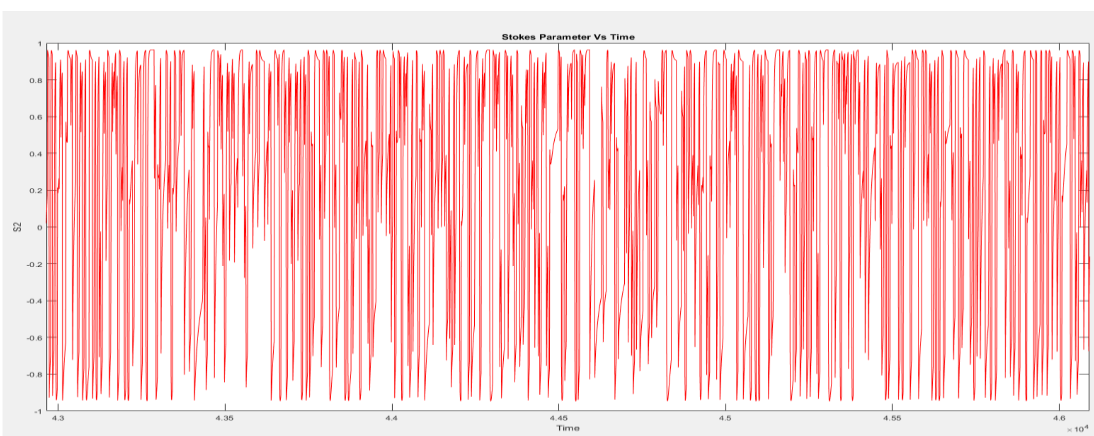

Figure 13b S2 vs Time for 4-POLSK.

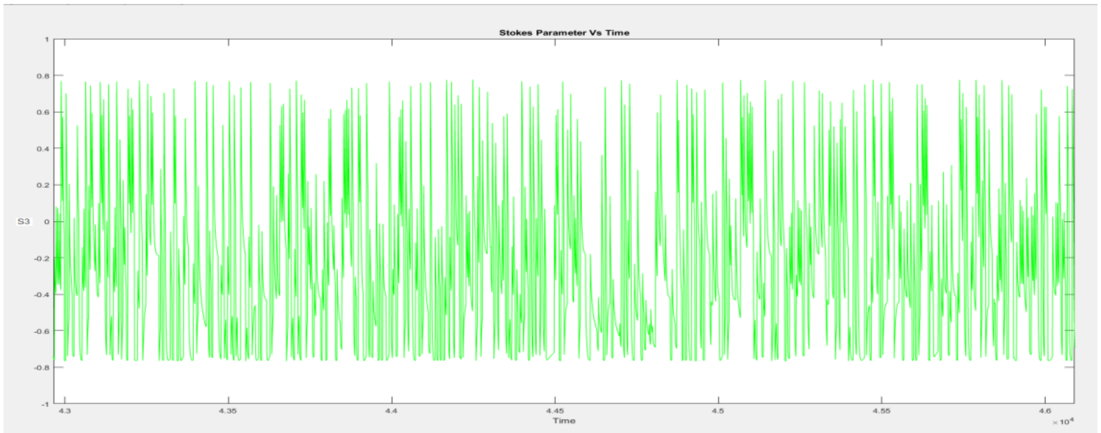

Figure 13c S3 vs Time for 4-POLSK.

The Stokes processing starts by projecting the Stokes parameters on the X-Y plane to reduce the dimension. Figure 14 a, b show the great circle that was created by encoding the voltages (2V,4V,6V,8V) and the projected coordinates and SOP respectively.

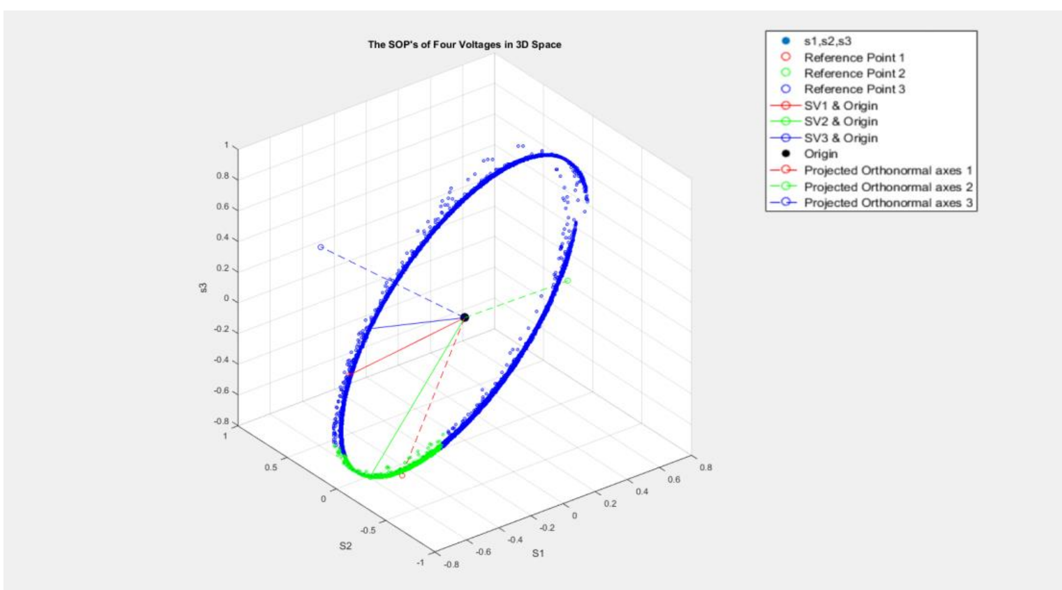

Figure 14a The SOP of Four Voltages in 3D Space.

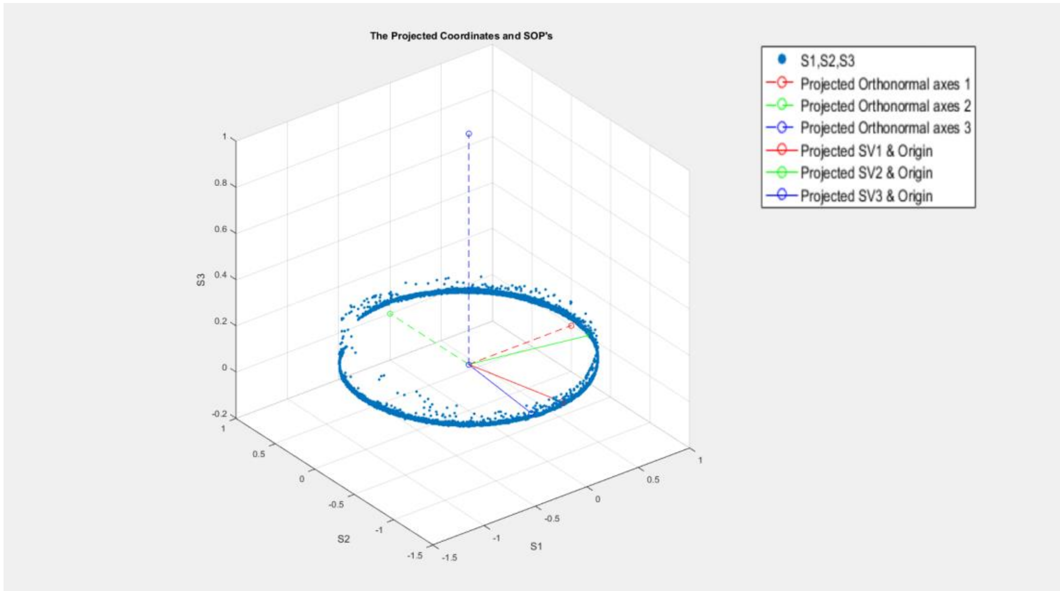

Figure 14b The Projected Coordinates and SOP.

A new attribute that presents the 4 voltage steps has been created. This new attribute is the angle 2β as shown in Figure (15). Since the sampling rate at the (POD-201) was five times of the encoding speed (sampling rate = 10KS/s for 2Kbit/s, sampling rate = 20Ks/s for 4Kbit/s, sampling rate = 30KS/s for 6Kbit/s), the angle data was reorganized in a way that each voltage chunked to have five points in each row. Author built a classification model using Matlab classification learner applications. Author tried multiple techniques for classification, the most effective two techniques were SVM and KNN. Each run, author used different data sequence, he used either SVM or KNN to classify data. Then based on model accuracy, author approved the results. The model was built using first 1Kbits where the angle points of the first 1Kbits were put in one table with respective voltage. The first five columns were the inputs to the classifier learner and the sixth one was the response. After building the classifier model, the model was used to classify the data points for the following 15Kbits. In other words, we input the angle data points that belong to the bits next to the first 1Kbits and the model classifies each five points to one of the initial voltage steps. This process has been applied to ten different random bits of 40Kbits length. The accuracy of our model- which represents the bit error rate - has been calculated to the ten different random sequences and plotted together in one graph. Figure (16) shows the mean and standard deviation of classification accuracy of the built models for each following 1Kbits of the following data in the sequence for different encoding speeds.

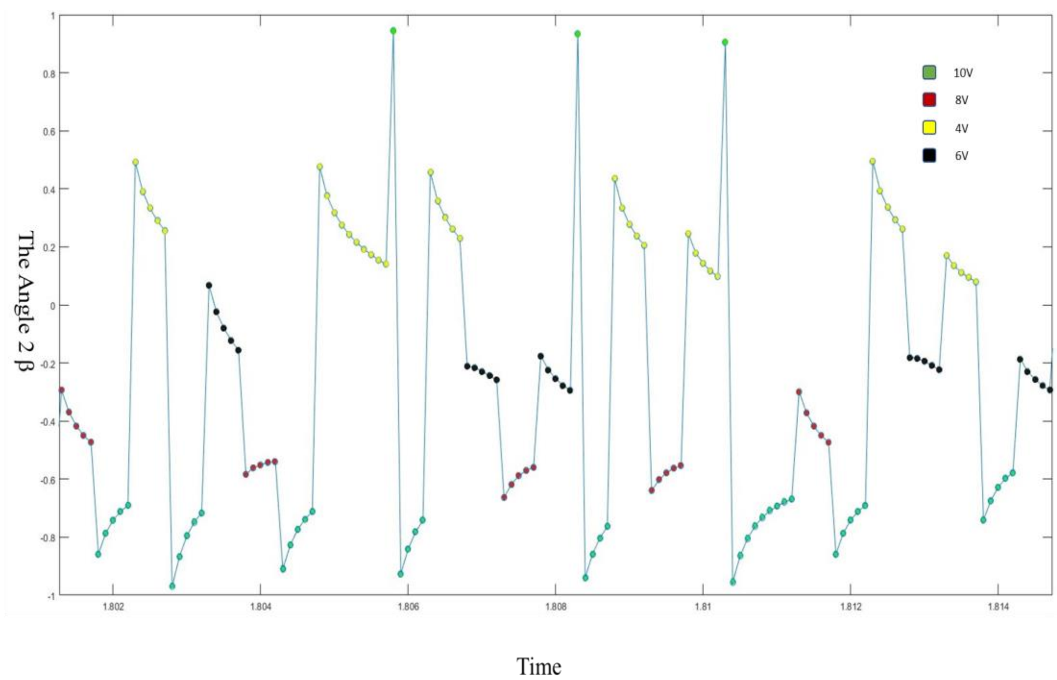

Figure 15 The Angle 2 β vs Time for 4-POLSK.

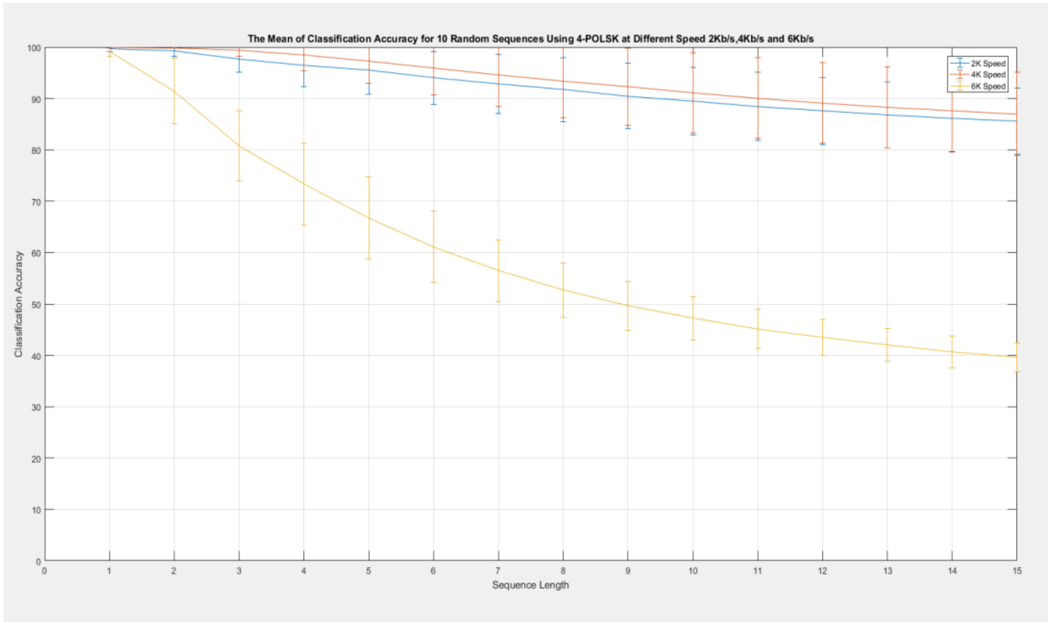

Figure 16 The Mean and Standard Deviation of Classification Accuracy for 10 Random Sequences using 4-POLSK at Different Speeds 2Kb/s,4Kb/s and 6Kb/s.

.

4.4.3 8-POLSK Observations

The procedures for the 8-POLSK is very similar to that for the 4-POLSK. We generated the file of random N Kbits. We mapped each three-different bit combination to one of the eight step voltages as in Table-2. The final size of voltage file is N/3 Kbits. We apply this file to the polarization modulator using the LabView code. That is to encode the mapped bits in the polarization of the passing light. As usual, we collect the Stokes parameters that belong to the encoded bits at (POD-201). Figure (17 a, b, c, d) shows the stokes parameters and angle 2β vs time.

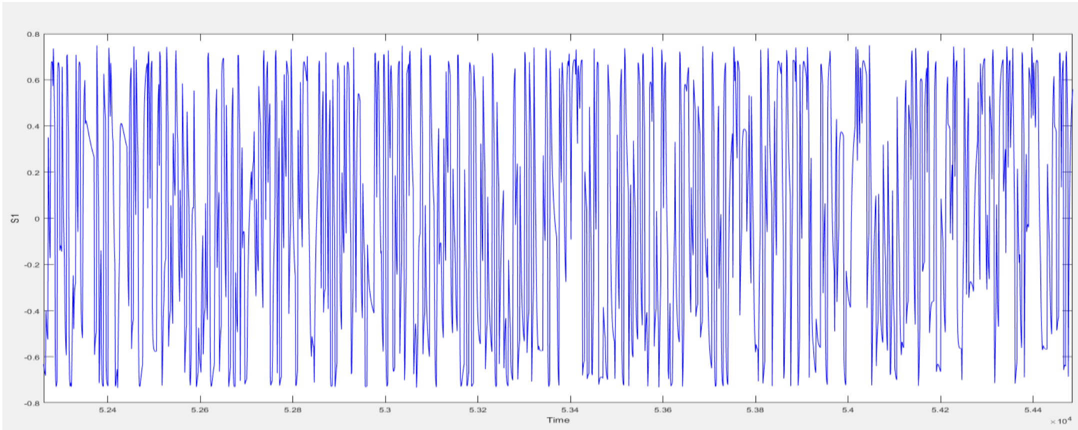

Figure 17a S1 vs Time for 8-POLSK.

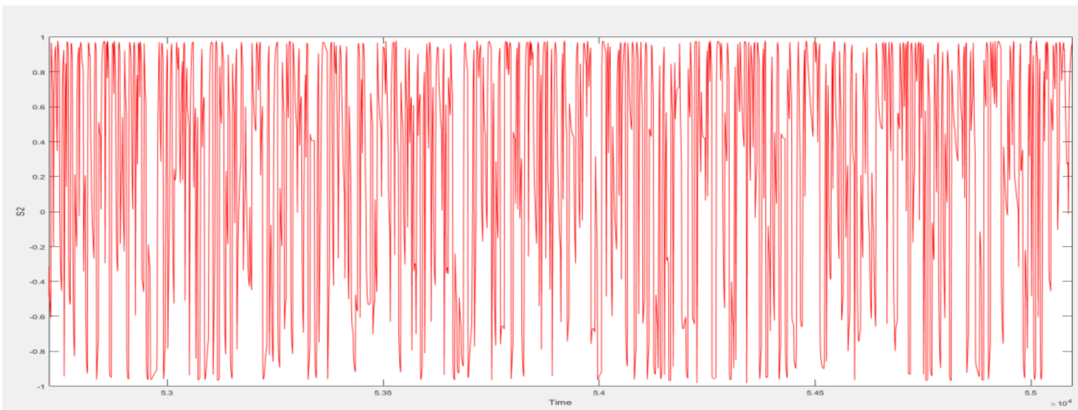

Figure 17b S2 vs Time for 8-POLSK.

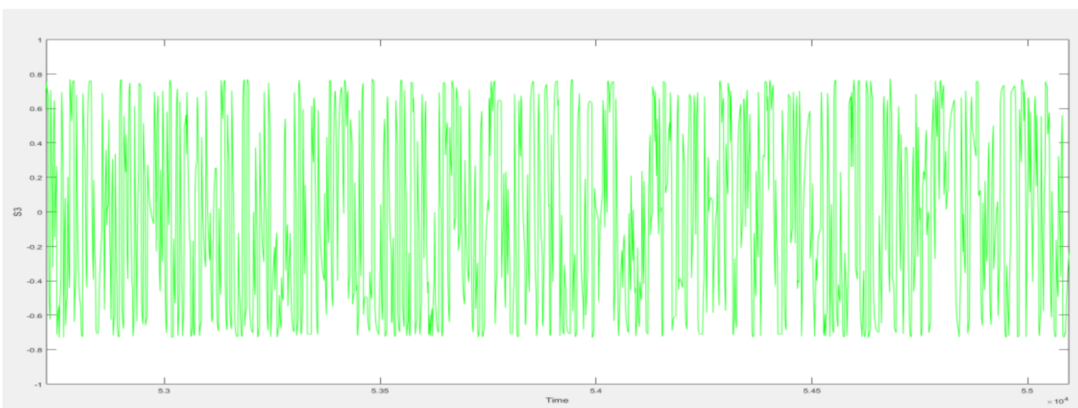

Figure 17c S3 vs Time for 8-POLSK.

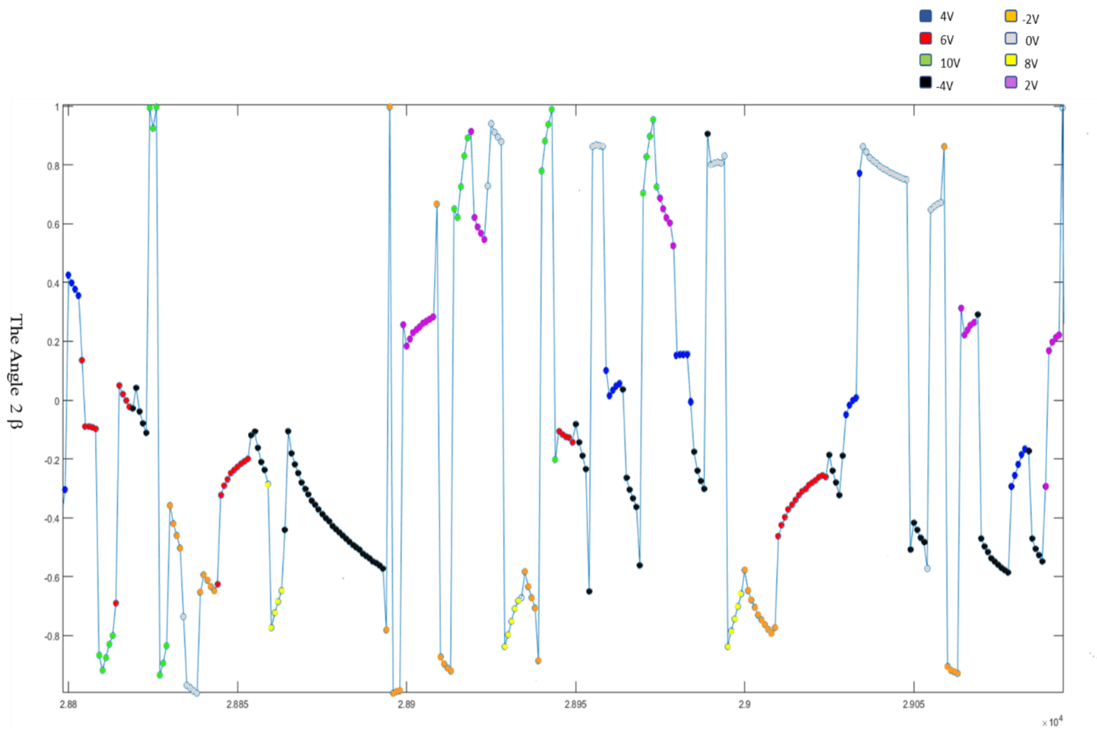

Figure 17d The Angle 2 β vs Time for 8-POLSK.

The Stokes processing starts by projecting the Stokes parameters on the X-Y plane to reduce the dimension. Figure (18 a, b) shows the great circle that was created by encoding the voltages (-4V, -2V, 0V, 2V,4V,6V,8V,10V) and the projected coordinates and SOP respectively.

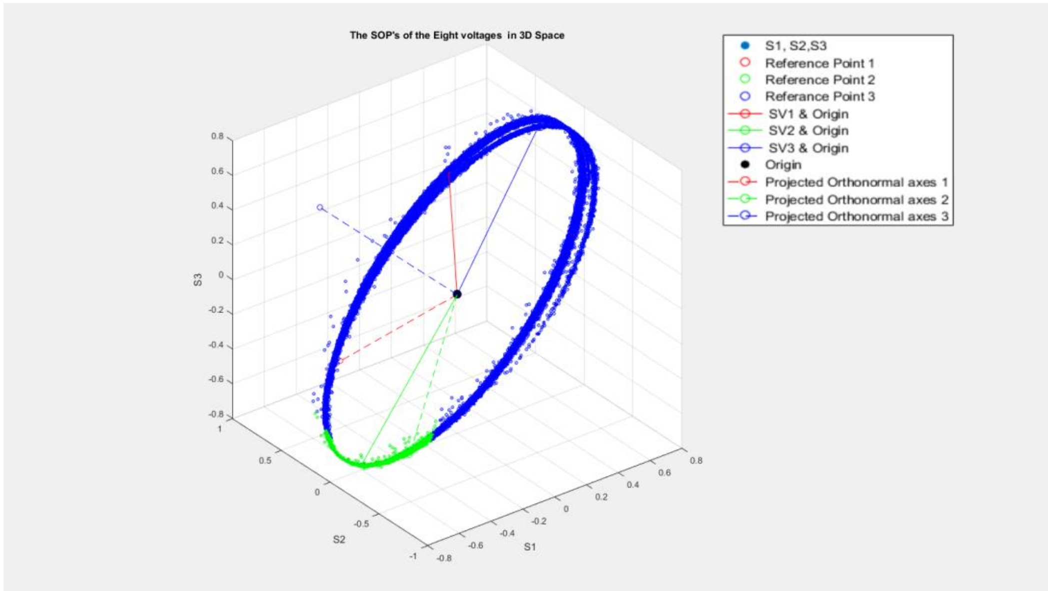

Figure 18a. The SOP of Eight Voltages in 3D Space.

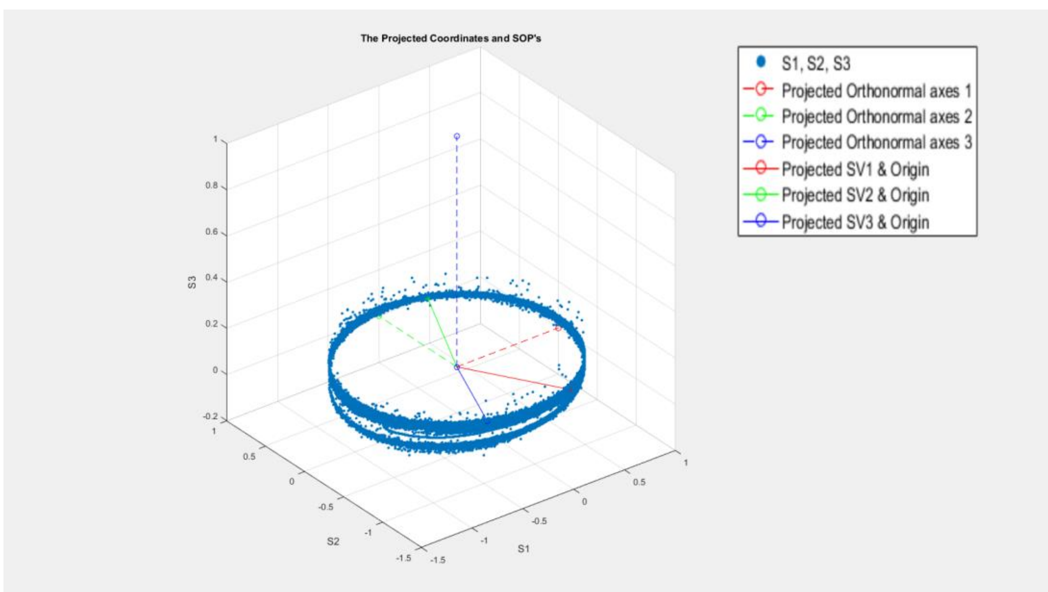

Figure 18b The Projected Coordinates and SOP for 8-POLSK.

The new attribute that presents the 8 voltage steps that we have is the angle 2β. Since the sampling rate at the (POD-201) was five times of the encoding speed (sampling rate = 10KS/s for 2Kbit/s, sampling rate = 20Ks/s for 4Kbit/s, sampling rate = 30KS/s for 6Kbit/s), the angle data was reorganized in a way that each voltage will be divided to have five points in each row. We built a classification model using math lab applications. The model was built using first 1Kbits where the angle points of the first 1Kbits were put in one table with respective voltage. The first five columns were the inputs to the classifier learner and the sixth one was the response. After building the classifier model, the model was used to classify the data points for the following 15Kbits. In other words, we input the angle data points that belong to the bits next to the first 1Kbits and the model classifies each five points to one of the initial voltage steps.

This process has been applied to ten different random bits of 60Kbits length. The accuracy of the model- which represents the bit error rate - has been calculated to the ten different random sequences and plotted together in one graph. Figure (19) shows the mean and standard deviation of classification accuracy of the built models for each following 1Kbits of the following data in the sequence for different encoding speeds:

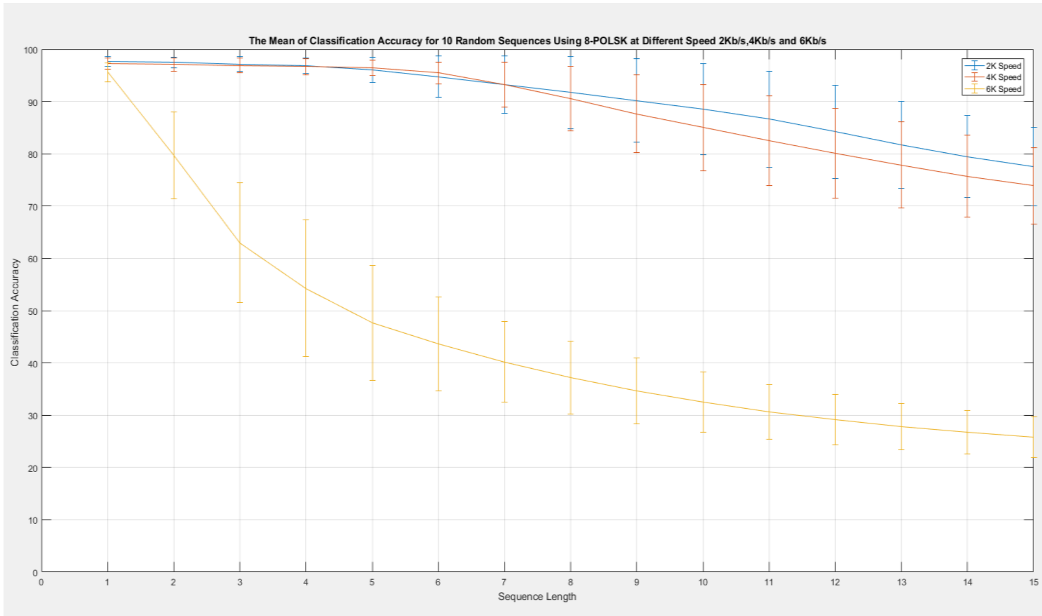

Figure 19 The Mean and Standard Deviation of Classification Accuracy for 10 Random Sequences using 8-POLSK at Different Speeds 2Kb/s,4Kb/s and 6Kb/s.

4.5 Observations Discussion

The results of both 4-POLSK and 8-POLSK have similar patterns. The modulator could map voltages to the state of polarizations with acceptable resolution at speed 2Kb/s and 4Kb/s. When voltages applied at 6Kb/s speed in both cases 4-POLSK and 8-POLSK, the modulator map the voltages to the state of polarizations with less resolution quality. Because of this low resolution, two or more voltages mapped to the same state of polarization. That lead to build models with over lapped boundaries and low accuracy predictions. While at 2Kb/s and 4Kb/s speeds, the prediction accuracy of predicting the 5Kbits following Kbits of the 1Kbits training set was very good in both 4-POLSK and 8-POLSK.

Chapter 5: Conclusion and Future Work

5.1 Conclusion

In this work, we have suggested a novel frame work to extract data that have been modulated in the polarization of light. We proposed to use the machine learning algorithms to model the behavior of the polarization channel. First, we applied a sequence of voltages to the polarization modulator using DAQ assistant via LabView code. These sequences were encoded in the polarization channel of a fully linearly polarized light. At the receiver, we collected the Stokes parameters using a POD (Polarimeter). The Stokes parameters were processed to reduce their dimension to one dimension. This dimension presented the encoded symbols and it was successfully used to discriminate between symbols and recover original sequences. The initial results show that we can use machine learning algorithms like K- Nearest Neighbors and Support Vector Machines to recover data from the Stokes parameters. Also, the results show that models were able to recover more than %97 of data in the 5Kbits following the 1Kbit data set.These models must be updated regularly to handle the changes that results from birefringence and rigid rotation. Even though the distance between the transmitter and receiver was very short and with low encoding speed, we believe that we have approved the possibility of exploiting the huge capabilities that have been offered by machine learning in classifying data in optical communication applications. To perform a communication protocol that use the machine learning algorithms to recover data at the reception side, we need to add some overhead as supervised bits to refresh the model each period. Based on the results, up to 20% overhead should be added to the payload. We return this to the limitation that we faced in the lab because of the polarization modulator that we used. It was not mapping voltages with the same state of polarizations after a certain encoding speed. i.e. there is no clear transfer function between the applied voltages and the correspondent state of polarization.

5.2 Future Work

The proposed work is expandable towards encoding and recovering data in the state of polarization in the real time. A complete modulation scheme needs to be developed. In other words, raw data should be mapped to voltages in the real time. Then, the mapped voltages can be applied to a polarization modulator in real time, too. This scheme should be able to update the classification models at the receiver periodically. That means supervised learning date sets should be included in the payload. A precise polarization modulator must be used to achieve better results. Some specifications need to be met in the polarization modulator like high resolution mapping between applied voltages and the state of polarization. Moreover, the polarization modulator should be able to switch the state of polarization with a higher speed with limited transitions between the states of polarization. Finally, we advise to include more symbols like 16-POLSK, 32-POLSK and 64-POLSK. That might be done using different degrees of polarization in the addition to different states of polarization.

REFERENCES

[1] S. Benedetto, and P. Poggiolini. "Theory of Polarization Shift Keying Modulation." IEEE Transactions on Communications. vol. 40, 4. April 4, 1992.

[2] Md. Nur-Al-Safa Bhuiyan , Motoharu Matsuura , Hung Nguyen Tan , Naoto Kishi, 2011. Simultaneous multichannel wavelength conversion of polarization shift keying signal with different channel group-delay and state of polarization. Optics Communications 284 (2011) 665–669.

[3] Benedetto S., Gaudino, R., Poggiolini, P., 1997. Polarization Recovery in Optical Polarization Shift-Keying Systems. IEEE Transactions on Communications. Vol. 45, No. 10, Oct. 1997.

[4] A. Carena, V. Curri, R. Gaudino, N. Greco, P. Poggiolini and S. Benedetto, POLARIZATION MODULATION IN ULTRALONG HAUL TRANSMISSION SYSTEMS: A PROMISING ALTERNATIVE TO INTENSITY MODULATION, Politecnico di Torino - Dipartimento di Elettronica Corso Duca degli Abruzzi, 24 10129 Torino Italy, ECOC'98, 20-24 September 1998, Madrid, Spain.

[5] Md. Nur-Al-Safa Bhuiyan, Motoharu Matsuura, Hung Nguyen Tan, Naoto Kishi, Simultaneous multichannel wavelength conversion of polarization shift keying signal with different channel group-delay and state of polarization, Optics Communications 284 (2011) 665–669.

[6] S. Benedetto, Senior Member, IEEE, R. Gaudino, and P. Poggiolini, Member, IEEE, Direct Detection of Optical Digital Transmission Based on Polarization Shift Keying Modulation, IEEE JOUKNAL Oh SELECTED AREAS IN COMMUNICATIONS, VOL. 13. NO 3, APRIL 1995.

[7] Xinquan Zhang, Qi Yang, Zhu Yang, D-mPSK-PoISK toward cost-effective application of polarization, State Key Laboratory of Optical Communication Technologies and Networks, Wuhan Research Institute of Posts and Telecommunications, Wuhan, Hubei, China, 430074.

[8] S. Benedetto, Senior Member, IEEE and P. T. Poggiolini, Member, IEEE, multilevel Polarization Shift Keying: Optimum Receiver Structure and Performance Evaluation, IEEE TRANSACTIONS ON COMMUNICATIONS, VOL 42, NO 21314. APRLL 1994.

[9] Darko Zibar, Member, IEEE, Molly Piels, Rasmus Jones, and Christian G. Schaeffer, Member, IEEE, Machine Learning Techniques in Optical Communication, JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 34, NO. 6, MARCH 15, 2016.

[10] Grant Fowles. Introduction to Modern Optics. Location of publisher: Dover Publications, Inc., 2nd Ed. 1975.

[11] Gregory G. MacDonald. A Polarization Based Intrusion Monitoring System for Optical Networks. Norman, Oklahoma: University of Oklahoma, 2003.

[12] Hewlett Packard, “The Poincare Sphere” White paper, February 1995, Hewlett Packard Journal, [Online], http://www.hpl.hp.com/hpjournal/95feb/feb95a4b.pdf

[13] Paul Green. Fiber Optic Networks. Location: Princeton Hall PTR.
[14] "Olympus Microscopy Resource Center".’Olympus America Inc. Retrieved 2011-11-13,[Online], http://www.olympusmicro.com/primer/lightandcolor/birefringence.html.

[15] L. Cimini, L. J., I. M. .Habbab, R. K. John A. A. Saleh On the Preservation of Polarization Orthogonality, a Through Linear System. Electronic Letters. Vol. 23, No. 23, pp. 1365-1366, December 3, 1987.

[16] Richard J. Blaikie, Member, IEEE, Desmond P. Taylor, Fellow, IEEE, and Peter T. Gough, Multilevel Differential Polarization Shift Keying, IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 45, NO. 1, JANUARY 1997.

[17] Zhenxing Chen, Jin Sub Bae, Se-Kyo Chung, Jin-Whan Koh, and Seog Geun Kang, Multi-Envelope 3-Dimensional Constellations for Polarization Shift Keying Modulation, School of Electrical and Electronic Engineering, Gyeongsang National University Jinju, Gyeongnam 660-701, Republic of Korea.

[18] Xinquan Zhang, Qi Yang, Zhu Yang, D-mPSK-PoISK toward cost-effective application of polarization, State Key Laboratory of Optical Communication Technologies and Networks, Wuhan Research Institute of Posts and Telecommunications, Wuhan, Hubei, China, 430074.

[19] Sergio Benedetto, Senior Member, IEEE, and Pierluigi Poggiolini, Theory of Polarization Shift Keying Modulation, IEEE TRANSACTIONS ON COMMUNICATIONS. VOL. 40, NO. 4, APRIL 1992.

[20] Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. New York, NY: Springer.

[21] Nicolás García-Pedrajas *, Domingo Ortiz-Boyer, Boosting k-nearest neighbor classifier by means of input space projection, Department of Computing and Numerical Analysis, University of Córdoba, Campus de Rabanales, 14071 Córdoba, Spain

[22] Zhenxing Chen, Jin Sub Bae, Se-Kyo Chung, Jin-Whan Koh, and Seog Geun Kang, Multi-Envelope 3-Dimensional Constellations for Polarization Shift Keying Modulation. School of Electrical and Electronic Engineering, Gyeongsang National University Jinju, Gyeongnam 660-701, Republic of Korea.

[23] Moshe Nazarathy, Senior Member, IEEE, Member, OSA, and Erez Simony, Stokes Space Optimal Detection of Multidifferential Phase and Polarization Shift Keying Modulation. JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 24, NO. 5, MAY 2006.

[24] Nan Chi, Siyuan Yu, Lin Xu, Palle Jappesen, Generation of a Polarization Shift Keying Signal and its Application in Optical Labeling. University of Bristol, BSB HR, United Kingdom. ECOC 2005 Proceedings - Vol.1

APPENDIX: Math Lab Codes

Code#1: Generating random bits

```
rand_bits= round(rand(1,60000));
y = [];
for i = 1:3:length(rand_bits)
    if rand_bits(i) == 0 && rand_bits(i+1)== 0 &&
rand_bits(i+2)== 0
        y = [y,-4];
    elseif rand_bits(i) == 0 && rand_bits(i+1)== 0 &&
rand_bits(i+2)== 1
        y= [y,-2];
    elseif rand_bits(i) == 0 && rand_bits(i+1)== 1 &&
rand_bits(i+2)== 0
        y = [y,0];
    elseif  rand_bits(i) == 0 && rand_bits(i+1)== 1 &&
rand_bits(i+2)== 1
        y = [y,2];
    elseif  rand_bits(i) == 1 && rand_bits(i+1)== 0 &&
rand_bits(i+2)== 0
        y = [y,4];
    elseif  rand_bits(i) == 1 && rand_bits(i+1)== 0 &&
rand_bits(i+2)== 1
        y = [y,6];
    elseif  rand_bits(i) == 1 && rand_bits(i+1)== 1 &&
rand_bits(i+2)== 0
        y = [y,8];
    elseif  rand_bits(i) == 1 && rand_bits(i+1)== 1 &&
rand_bits(i+2)== 1
        y = [y,10];

    end
end
y = y';
```

Code#2: SOP points projection

```
% set(0,'DefaultFigureWindowStyle','docked');
%set(0,'DefaultFigureWindowStyle','normal');
start_index = 1000;
stop_index = 200000;
t = VarName1(start_index:stop_index);
S0 = VarName2(start_index:stop_index);
S1 = VarName3(start_index:stop_index);
S2 = VarName4(start_index:stop_index);
S3 = VarName5(start_index:stop_index);

orig = [0;0;0]; %%% Default origin

figure(1)
plot(1:size(t,1), S1, 'b', 1:size(t,1), S2, 'r', 1:size(t,1), S3,
'g');
```

```
SV = [S1, S2, S3]';
figure(2)
scatter3(S1, S2, S3, 1, 'filled');
hold on
%%%%% Reference mean data point 1 %%%%%
S1_m1 = mean(S1(S2<0.6));
S2_m1 = mean(S2(S2<0.6));
S3_m1 = mean(S3(S2<0.6));
norm_m1 = sqrt(S1_m1^2 + S2_m1^2 + S3_m1^2);
S1_m1 = S1_m1 / norm_m1;
S2_m1 = S2_m1 / norm_m1;
S3_m1 = S3_m1 / norm_m1;

S1_temp = S1(S2> -0.2);
S2_temp = S2(S2> -0.2);
S3_temp = S3(S2> -0.2);
scatter3(S1_temp,S2_temp,S3_temp,10,'r');

%%%%% Reference mean data point 2 %%%%%
S1_m2 = mean(S1(S1<0.2));
S2_m2 = mean(S2(S1<0.2));
S3_m2 = mean(S3(S1<0.2));
norm_m2 = sqrt(S1_m2^2 + S2_m2^2 + S3_m2^2);
S1_m2 = S1_m2 / norm_m2;
S2_m2 = S2_m2 / norm_m2;
S3_m2 = S3_m2 / norm_m2;

S1_temp = S1(S1<0.2);
S2_temp = S2(S1<0.2);
S3_temp = S3(S1<0.2);
scatter3(S1_temp,S2_temp,S3_temp,10,'g');

%%%%% Reference mean data point 3 %%%%%
S1_m3 = mean(S1((S1<0.8)&(S1>0.2)));
S2_m3 = mean(S2((S1<0.8)&(S1>0.2)));
S3_m3 = mean(S3((S1<0.8)&(S1>0.2)));
norm_m3 = sqrt(S1_m3^2 + S2_m3^2 + S3_m3^2);
S1_m3 = S1_m3 / norm_m3;
S2_m3 = S2_m3 / norm_m3;
S3_m3 = S3_m3 / norm_m3;

S1_temp = S1((S1<0.8)&(S1>0.2));
S2_temp = S2((S1<0.8)&(S1>0.2));
S3_temp = S3((S1<0.8)&(S1>0.2));
scatter3(S1_temp,S2_temp,S3_temp,10,'b');

%%%%% Mean data points in vector form %%%%%
SV1 = [S1_m1; S2_m1; S3_m1];
SV2 = [S1_m2; S2_m2; S3_m2];
SV3 = [S1_m3; S2_m3; S3_m3];

%%%%% Plot lines from origin to the mean data points %%%%%
```

```
pt = [orig, SV1];
plot3(pt(1,:), pt(2,:), pt(3,:), 'ro-');
pt = [orig, SV2];
plot3(pt(1,:), pt(2,:), pt(3,:), 'go-');
pt = [orig, SV3];
plot3(pt(1,:), pt(2,:), pt(3,:), 'bo-');

% Form a triange using the vectors SV1, SV2, SV3. Since all the
SV points lie on the plane of the triangle, and %it is in general
not necessary that the origin also lie on this plane, therefore,
we can construct the normal vectors using the v1 and v2 defined
below.
v1 = SV1 - SV3;
v2 = SV2 - SV3;
e1 = v1 / norm(v1);
e2 = v2 - e1 * (e1' * v2);
e2 = e2 / norm(e2);
e3 = cross(e1, e2);
e3 = e3 / norm(e3);

%%%%% O is the circumcenter of the triangle formed by SV1, SV2,
SV3.
O1 = (norm(SV1)^2 - norm(SV3)^2) / 2 / norm(v1);
O2 = ((norm(SV2)^2 - norm(SV3)^2) - (norm(SV1)^2 -
norm(SV3)^2)*(v2' * v1)) / 2 / norm(v2 - v1 * (v1' * v2));
O = O1 * e1 + O2 * e2;

%%%%% If O = orig = [0;0;0], the SV data points center at the
origin, as it
%%%%% turns out to be.
scatter3(O(1), O(2), O(3), 100, 'k', 'filled');

%%%%% Plot the three orthonormal axes.
pt = [O, e1];
plot3(pt(1,:), pt(2,:), pt(3,:), 'ro--');
pt = [O, e2];
plot3(pt(1,:), pt(2,:), pt(3,:), 'go--');
pt = [O, e3];
plot3(pt(1,:), pt(2,:), pt(3,:), 'bo--');
axis square
hold off

%%%%% Find the rotation matrix.
G = inv([e1,e2,e3]);

%%%%% Find the SV in the rotated coordinate system.
SV_new = G * SV;
%%%%% Plot the SV in the rotated coordinate.
figure(3)
S1_new = SV_new(1,:);
S2_new = SV_new(2,:);
S3_new = SV_new(3,:);
scatter3(S1_new, S2_new, S3_new, 10, 'filled');
```

```
axis square;
hold on

tmp = G * e1;
pt = [O, tmp];
plot3(pt(1,:), pt(2,:), pt(3,:), 'ro--');
tmp = G * e2;
pt = [O, tmp];
plot3(pt(1,:), pt(2,:), pt(3,:), 'go--');
tmp = G * e3;
pt = [O, tmp];
plot3(pt(1,:), pt(2,:), pt(3,:), 'bo--');
tmp = G * SV1;
pt = [O, tmp];
plot3(pt(1,:), pt(2,:), pt(3,:), 'ro-');
tmp = G * SV2;
pt = [O, tmp];
plot3(pt(1,:), pt(2,:), pt(3,:), 'go-');
tmp = G * SV3;
pt = [O, tmp];
plot3(pt(1,:), pt(2,:), pt(3,:), 'bo-');
hold off

figure(4)
plot(S1_new);
hold on
plot(S2_new);
plot(S3_new);
hold off

figure(5)
SV_angle = angle(S1_new + 1i * S2_new);
subplot(2,1,1)
plot(SV_angle/pi, 'o-');
```

Code #3: Averaging and decision making.

```
x = SV_angle;
ss = [];
a = [];
ss = x(startindex:endindex);
    for i = 1:10:length(ss)
        a = [a, mean(ss(i:i+9))];
    end
aa = [];
for i = 1:length(a)
    if a(i)< -0.3 || a(i)<-0.9
        aa = [aa,2.5];
    else
        aa = [aa,-2.5];
    end
end
```