KUNGL TEKNISKA HÖGSKOLAN

Communication and Error Correction via Polarisation of Single Photons and Time Ordering

Master's thesis in Engineering Physics (Quantum Technology)

Shek Lun Leung

DEPARTMENT OF APPLIED PHYSICS
ROYAL INSTITUTE OF TECHNOLOGY
Stockholm, Sweden 2024
www.kth.se

Master's thesis 2024

Communication and Error Correction via Polarisation of Single Photons and Time Ordering

Shek Lun Leung

Department of Applied Physics
Royal Institute of Technology
Stockholm, Sweden 2024

Communication and Error Correction via Polarisation of Single Photons and Time Ordering
Shek Lun Leung

Supervisor:
Dr. Jonas Almlöf, Quantum Tehcnologies, Ericsson AB
Dr. Richard Schatz, Department of Applied Physics, KTH
Dr. Oskars Ozolins, Department of Applied Physics, KTH
Examiner:
Dr. Sergei Popov, Department of Applied Physics, KTH

Master's Thesis 2024
Department of Applied Physics
Royal Institute of Technology
SE-100 44 Stockholm
Telephone +46 8 790 60 00

Cover:
An oil painting constructed in DALL.E 2 showing a single photon with polarization being sent by Alice and being corrected by Bob by distant transmission.

Typeset in LaTeX
Printed by KTH Reproservice
Stockholm, Sweden 2024

Abstract

This research study aims to investigate the capacity of single photons to carry information through polarization and time ordering and proposes a protocol called Beyond Pulse Position Modulation (BPPM) to improve photon-based communication reliability over longer distances with limited power. Such a protocol may be used in any communication scenario where energy efficiency is important, e.g., in satellite communication or where pulse position modulation (PPM) typically is used. The study compares various metrics such as information bits per symbol, photon, and time bin to evaluate the system's efficiency and conducts a comparative analysis of BPPM, Pulse Position Modulation (PPM), On-Off Keying (OOK), and General protocol's effectiveness. (The simulations were conducted using the Python programming language with Visual Studio Code IDE.)

Acknowledgements

I would like to express my deepest gratitude to Jonas Almlöf, my main supervisor, for his exceptional guidance and unwavering support throughout this project. Jonas has consistently shown a willingness to address any questions or concerns I had, making themselves available whenever possible. His insightful perspectives and collaborative approaches have greatly contributed to the success of our work. Together, we have brainstormed ideas, discussed concepts on the whiteboard, and built this project as a cohesive team. Even in remote work situations, Jonas maintained strong rapport and ensured effective communication through email and messaging apps.

I extend my sincere thanks to my academic supervisor, Richard Schatz, for his invaluable support and the knowledge he shared. Richard's guidance has made complex concepts easier to comprehend, and his willingness to address my inquiries has been instrumental in my learning journey. Richard's meticulous reviews of my mathematical writings and calculations have not only improved the quality of my work but also provided valuable comments and advice. Furthermore, I appreciate Richard's efforts in organizing several lectures and dedicating meeting time for us to delve into the intricacies of the desired protocol's work mechanism, principles of error correction, and potential metrics for comparative analysis. These discussions facilitated effective collaboration and progress in our project.

I am immensely grateful to Gemma Vall-llosera for her consistent efforts in reviewing my calculations and plots before our weekly sharing. Gemma's insightful feedback and guidance have continuously enhanced my understanding of the subject matter. Furthermore, I express my appreciation to Hanna Brännström for organizing meetings for the thesis projects and providing thought-provoking questions and invaluable insights on presenting data with mathematical rigor and meaning. Hanna's dedication to guiding me in programming and visualizing my work has been invaluable.

I would also like to extend my thanks to Danielle Gustafsson, Guilherme B Xavier, Joakim Argillander for their contributions in conducting the experimental work at Linköping University. Our weekly discussions have been instrumental in realizing how single-photon transmission via our desired protocol works in the experimental setup and understanding the limitations we need to consider.

To all of you who took the time to review the draft of my master's thesis, I am deeply grateful. Your input and suggestions have been invaluable in ensuring the quality and completeness of my work.

Beyond our academic endeavors, I have cherished the time during our lunches, engaging in conversations on various topics from culture, fascinating world of quantum computation to the latest advancements in generative AI and satellite communication, which have enriched my overall experience and broadened my horizons.

Once again, I want to express my profound gratitude for your unwavering support, guidance, and inspiring discussions. Working with such an exceptional team has made this journey incredible, and I am truly grateful for the opportunity.

Shek Lun Leung, Stockholm, June 2023

Contents

List of Figures

List of Tables

1

Introduction

1.1 Brief Overview of the Topic

Data transmission through various communication channels has become increasingly essential to our daily lives. The advancement of the Internet of Things (IoT) is driving the need for faster and more reliable telecommunication technologies to connect an increasing number of devices and sensors to the Internet [10]. As more devices connect, bandwidth and low-latency communication demand are expected to grow significantly. 5G wireless networks revolutionised communication technology by offering higher speeds, lower latency, and greater capacity to meet the demands of modern end devices. This transformation is already underway and will reshape how we communicate. [19].

Implementing 5G networks paves the way for innovative applications like autonomous vehicles, remote healthcare, and virtual reality experiences, ushering in a new era of technological possibilities. Besides meeting the demands of high-speed, low-latency communication, one of the challenges is the development of error-correcting codes due to the corruption caused by the noise or interference of data during transmission. Error-correction coding is a technique to detect and possibly correct damaged received data [33].

Satellite communication systems offer unique advantages for global communication, particularly in remote regions where traditional communication infrastructure is unavailable [14]. However, satellite communication also presents unique challenges, such as atmospheric attenuation and signal interference, that require the development of innovative and efficient communication protocols [12]. Optical fibre communication is pivotal in modern data centres due to its advantages, such as high bandwidth, low latency, and immunity to electromagnetic interference. As highlighted by Cholewa et al. [18], optical interconnects offer superior data transmission speeds, reduced signal delay, and enhanced reliability. However, as the signals travel through the fibre, they experience signal loss and dispersion. Signal loss occurs due to the absorption and scattering of light, leading to a decrease in signal strength.

One potential solution to overcome telecommunication challenges in an energy-starved scenario is to use the polarisation of photons and time ordering. By investigating the information capacity of a single photon through polarisation and time ordering, it may be possible to develop error correction techniques that can improve

the reliability of photon-based communication over longer distances [2]. Ultimately, this contributes to the ongoing pursuit of communication and error correction knowledge by investigating the polarisation of single photons and time ordering.

If energy is limited, one can use the vacuum state, i.e., an unexcited state as an empty Time Bin, to increase the information carried per photon. However, there is a price to pay: the transmission time goes up.

1.2 Research Question and Objectives

Over the last few decades, there has been a question about how much information a single photon can carry [8, 20, 21, 23], which drives the research question: What is the maximum amount of information that a single photon can carry through polarisation and time ordering?

The motivation for investigating the information capacity of a single photon is to facilitate energy conservation when available power is limited, which can be applied to satellite communication. In particular, it is shown that a single photon can be used to convey more than 10 bits (10.5 bits) of information per detected photon between the sender and the receiver by associating each photon with one of the pixels of a grid of (alphabet) detectors with a size of 9,072 spatial modes [31]. This technique works well over short distances but would only help a little in telecommunication over longer distances. Therefore, this master's thesis aims to investigate the information capacity of a single photon through polarisation and time ordering and to develop error correction techniques that can improve the reliability of photon-based communication over longer distances.

1.3 Significance of the Study

This master's thesis aims to demonstrate a new protocol called Beyond PPM (BPPM), using a vacuum or unexcited state in the encoding to send polarised single photons in the form of a polarised electromagnetic pulse. BPPM allows for message transmissions at a much lower total energy (significantly more than 1 bit per photon) than brute force approaches that use many photons per bit.

The approach also addresses the lack of error correction in pulse position modulation (PPM) and devises a code to correct lost pulses and new unwanted pulses appearing in a block due to reflections. To this end, a single photon or pulse is associated with an integer Number of Time Bins. While this approach requires more time for transmission than traditional block-coded communication protocols, it is energy-efficient. It avoids sending a wider spectrum that may cause potential health problems for the environment and interference with other telecommunications. Besides, the scheme resembles the permutation codes, work by Blake [6]. Therefore, information can be transmitted more reliably on BPPM compared to PPM due to the special type of error correction adopted for an asymmetric channel while still having the benefits

of PPM in that communication can be used in photon-starved scenarios where considerably less than one photon per information bit is being sent.

Overall, developing an efficient and reliable communication protocol for satellite communication can have significant implications for space applications.

2

Background

2.1 Communication Process in the Digital Transmission System

The encoded message is transmitted over the channel, where errors may occur, necessitating decoding the received codeword into the original message. The encoding and decoding processes are illustrated in Figures 2.1 and 2.2. The source message is transmitted to the source encoder of the communication system to convert it into a digital representation that reduces the redundancy of the source message, achieving a lower amount of data to be transmitted.

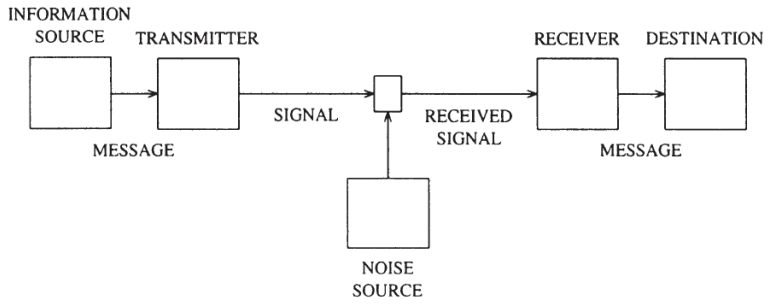

Figure 2.1: The simplified diagram depicts the digital transmission system. Image is extracted from [35].

The communication process involves the transmission of a message or information from a sender, Alice, to a receiver, Bob, over a channel. The source message is initially analogue and must be converted into a digital representation that can be transmitted. This conversion is performed by a source encoder, which uses techniques such as pulse code modulation or delta modulation to reduce the redundancy of the source message and represent it as a sequence of binary digits.

To ensure that the source codeword can be transmitted over the channel with minimal errors, an encoded source codeword by a channel encoder can solve this problem. The channel encoder adds redundant information to the source codeword, enabling the detection and correction of errors that may be introduced during transmission. The output of the channel encoder is the channel codeword, which contains both the source codeword and additional redundant bits that enable the detection and

correction of errors.

The channel codeword is then modulated into a signal that can be transmitted over the channel. Modulation techniques, such as amplitude modulation, frequency modulation, or phase modulation, convert the binary signal into a form that can be transmitted over the channel.

During transmission, the signal may be affected by noise and other sources of interference, which can cause errors in the received signal. The received signal is demodulated by a demodulator, which converts the modulated signal back into a digital representation that the channel decoder can process.

The channel decoder uses the redundant information in the channel codeword to detect and correct errors that may have been introduced during transmission. The output of the channel decoder is the source codeword, which is a digital representation of the source message closer to the original than the received codeword.

The source codeword is then decoded by a source decoder, which converts the digital representation back into the original analog form of the source message. The receiver is the destination of the transmitted message or information, and receives the decoded source message. The received information is the output of the source decoder, the original message transmitted.

In summary, the communication process involves several key components, including source encoding, channel encoding, modulation, transmission over a channel, demodulation, channel decoding, and source decoding. The use of error-correcting codes is a critical aspect of the process, enabling the reliable transmission of information even in the presence of noise and other sources of error.

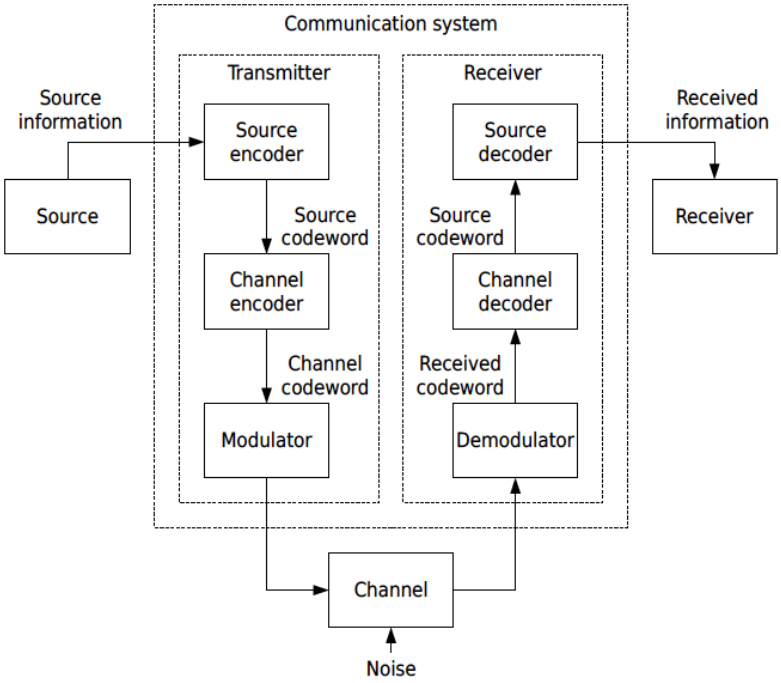

Figure 2.2: A detailed diagram depicts the digital communication system. Image is extracted from [5].

In an error correction framework, Alice transmits a message to Bob through a channel via air or data cable. Noise may be present in the channel, and that causes errors in the received message, as shown in 2.1. While high-power signal amplification and short channels help minimise errors, they may not be feasible. For instance, Global System for Mobile Communications (GSM) telephones [22] are designed to operate efficiently within the constraints of battery power, and Ethernet cables [4] are designed to wired data transmission over considerable distances in networking. However, GSM phones are often battery-powered and need to conserve energy for extended usage, and longer Ethernet cables may introduce signal degradation due to attenuation. Hence, codewords (which will be introduced in the Section 2.3) are encoded to reduce energy usage and enable long-distance transmission.

In order to enhance the efficacy of communication between Bob and Alice in the presence of noise, particularly in optical fibre or satellite communication systems, the utilisation of photons as the transmission medium presents a viable approach. Using polarization of a single photon for telecommunication will be introduced in the next section.

2.2 Using Single Photons in Telecommunication

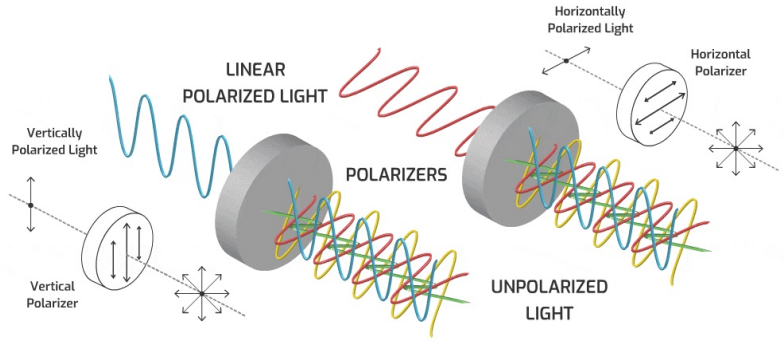

Figure 2.3: A illustration of light being polarized vertically and horizontally by polarizers. Image is extracted from [9].

In quantum communication and cryptography, single photons are used as carriers of quantum states. In this technique, information is encoded in the state of a single photon, which is then transmitted through a communication channel to a receiver. The information is then decoded by measuring the state of the photon at the receiver.

It is well known that one advantage of using single photons in telecommunication is that they allow for secure communication by employing Quantum Key Distribution (QKD) protocols such as BB84 [30], since any attempt to intercept or eavesdrop on the communication will disturb the state of the photon and thus be detectable by the receiver. According to the no-cloning theorem [34], it is impossible to create an exact, perfect copy of an arbitrary, unknown quantum state that does not have distinct measurement outcomes due to the fundamental principles of quantum mechanics. However, if the quantum state is known precisely or if the state is orthogonal, a quantum cloning machine can create approximate copies of the states with a limit on the accuracy of the clone.

Photons, being elementary particles of electromagnetic radiation, possess distinct advantages that are in line with the objectives of error correction capability and energy efficiency. Quantum error correction techniques, like as the surface code [25], are employed by these systems to effectively mitigate errors over extended distances, all while maintaining energy efficient. Also, single photon signals are less likely to interfere with other signals or experience cross-talk, leading to improved signal integrity and reduced noise. Besides, single photon-based system can operate at high data rates with low power consumption, enabling the potential for high data transmission rates for energy-efficient communication over long distances, which is demonstrated by implementing Low-density parity-check (LDPC) codes to QKD [36].

Single photons in telecommunication allow the photons to have non-orthogonal states, such as the basis states $\{|0\rangle, |1\rangle\}$, $\{\frac{|0\rangle+|1\rangle}{\sqrt{2}}, \frac{|0\rangle-|1\rangle}{\sqrt{2}}\}$ due to the polarization

of light. Polarisation is an inherent characteristic of light that pertains to the orientation of the electric field oscillations associated with it. The predominant portion of light sources, such as the sun, emit light that is unpolarized. As shown in Figure 2.3, unpolarized light exhibits vibrations that are randomly directed in directions perpendicular to the path of its propagation. In order to achieve polarisation, light undergoes a process where randomly directed vibrations are eliminated or converted into specific forms of electromagnetic waves, such as linear, circular, or elliptical. In the subsequent illustrations, our focus will solely be directed towards the process of linear polarisation of unpolarized light.

As shown in Figure 2.4, Poincare's sphere represents the quantum states of the polarised light, which provides a valuable means of visualising the state of a single photon and its operation on it. $|0\rangle$ represents the horizontal polarisation state $|H\rangle$, while $|1\rangle$ represents the vertical polarisation state $|V\rangle$.

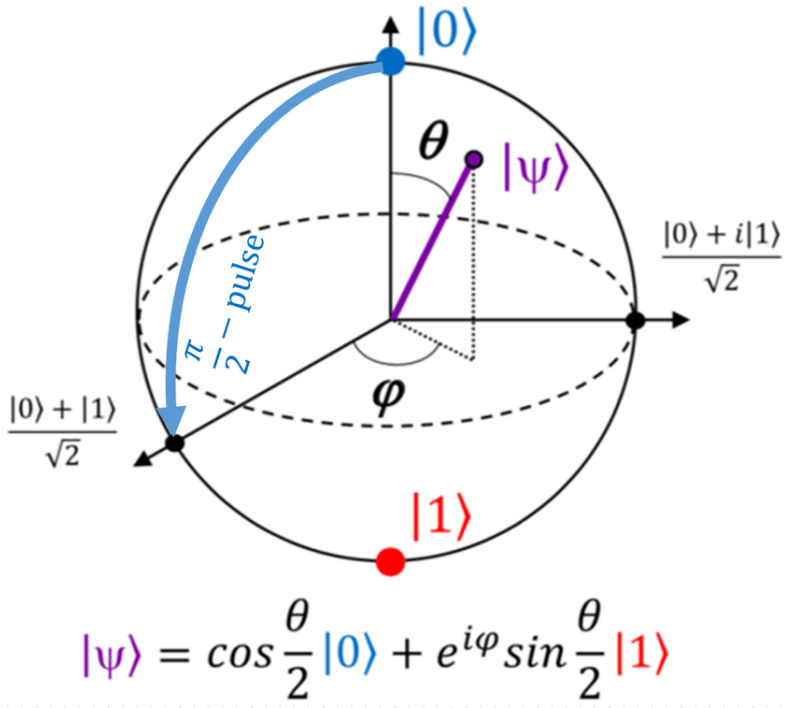

Figure 2.4: The diagram shows Poincare's sphere of various quantum states of a polarised single photon. Image is extracted from [17].

The above representation is a unit 1-sphere (i.e., the sphere's radius is 1), where two states are located at the north and south poles. The range of values for θ and orthogonals ϕ such that they cover the whole sphere is $\theta \in [0, \pi)$ and $\phi \in [0, 2\pi)$. Angle θ corresponds to latitude, and the angle ϕ corresponds to longitude.

A single photon state $|\phi\rangle$ can be written as:

$$|\psi\rangle = cos\frac{\theta}{2}|H\rangle + e^{i\phi}sin\frac{\theta}{2}|V\rangle = cos\frac{\theta}{2}|0\rangle + sin\frac{\theta}{2}e^{i\phi}|1\rangle \tag{2.1}$$

where $|H\rangle = |0\rangle$ and $|V\rangle = |1\rangle$. The parameters can be set to represent the 4 bases

$$\begin{aligned}
&\text{For} \quad \theta = 0 \quad \text{and} \quad \phi = 0 \\
&\quad |\psi\rangle = 1 \cdot |0\rangle + e^0 \cdot 0 \cdot |1\rangle = |0\rangle \\
&\text{For} \quad \theta = \frac{\pi}{2} \quad \text{and} \quad \phi = 0 \\
&\quad |\psi\rangle = \frac{1}{\sqrt{2}} \cdot |0\rangle + \frac{e^{i\cdot 0}}{\sqrt{2}} \cdot |1\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\
&\text{For} \quad \theta = \pi \quad \text{and} \quad \phi = 0 \\
&\quad |\psi\rangle = 0 \cdot |0\rangle + e^0 \cdot 1 \cdot |1\rangle = e^0 \cdot |1\rangle = |1\rangle \\
&\text{For} \quad \theta = \frac{3\pi}{2} \quad \text{and} \quad \phi = 0 \\
&\quad |\psi\rangle = \frac{1}{\sqrt{2}} \cdot |0\rangle - \frac{e^0}{\sqrt{2}} \cdot |1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{aligned} \tag{2.2}$$

The use of single photons in telecommunication is still an active area of research [26], with ongoing efforts to improve the efficiency and reliability of single-photon sources and detectors and develop new techniques for encoding and decoding information using single photons. One example of a quantum communication channel is an optical fibre designed to minimise losses and maintain polarization. Twin-field (TF) Quantum Key Distribution enables photons tolerating a maximum loss exceeding 100dB an across optical fibre to form a common secret string of bits between two remote users over 600km away [24]. Besides, Free Space Optical System (FSO) can also be used based on THz and quantum communication that offers a practical solution for data transfer in non-Line of Sight (NLOS) technology, which can offer inter-space link applications [16].

2.3 Transmission Concept

2.3.1 Time Bin

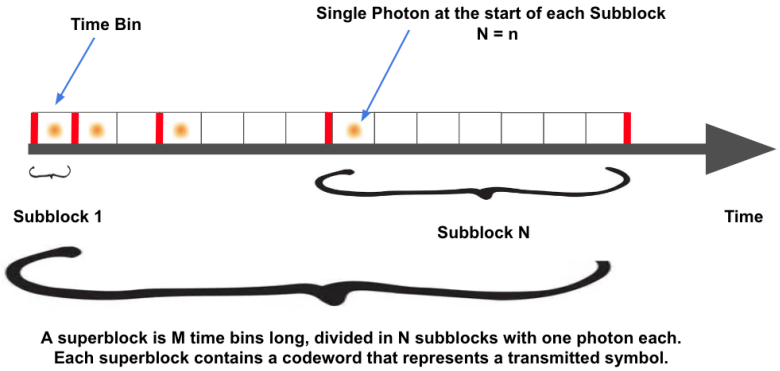

Figure 2.5: The diagram illustrates that a superblock is M time bins long, divided in N subblocks with one photon each. Each superblock contains a code word that represents a transmitted symbol.

A Time Bin typically refers to a specific time interval or window within which multiple events or signals can be detected or measured. For example, in single photon detectors used in quantum communication, a time bin is a specific duration during which a single photon at the start of each subblock can be detected and measured. In this context, the duration of the time bin is typically determined by the characteristics of the detector and the requirements of the communication system.

2.3.2 Subblock and Superblock

In digital communication, a superblock is a data structure that efficiently organises and transmits a large amount of data. The superblock comprises several subblocks, each containing number of time bins. The duration of each time bin is determined by the specific modulation scheme used in the communication system and can vary depending on the system design and requirements.

In Figure 2.5, there is a single photon at the start of each subblock, and there are 4 subblocks in a superblock in this figure. Therefore, the number of photons is equal to the number of subblocks in this case.

The purpose of using a superblock is to reduce the overhead associated with transmitting large amounts of data. By grouping multiple subblocks into a single symbol, the communication system can reduce the number of symbols that need to be transmitted, improving the data rate and reducing the time required to transmit the data. Additionally, using a superblock can improve the reliability of the transmission by allowing for error correction and detection at the subblock level, which can help to

ensure that the data is transmitted accurately and without errors.

2.3.3 The Number and the Length of Codewords

The number of codewords is an important aspect of coding theory, which is used in digital communication and data storage systems to improve the reliability and efficiency of data transmission and error correction [7]. The number of codewords determine the number of symbols. Each symbol is mapped to a digital codeword. Hence the number of codewords are hence equal to the number of symbols, K, to be encoded. In Figure 2.5, there are 4 subblocks, which can be arranged in ([1,2,4,7], [1,4,2,7], [4,1,2,7], ...), i.e., $K = 4! = 24$ different ways to create various codewords. Therefore, there can be 6 possible ways to transmit a symbol.

The length of each code word, M, is determined by the number of symbols and the modulation format used. There are 7 time bins in a superblock in Figure 2.5, which corresponds to the length of a codeword. In order to optimise data transmission and reduce the overhead associated with transmitting error correction and detection codes, the optimal length of codewords for a specific communication channel or storage system can be designed, such as variable-length codes [13]. This can improve the data rate and reduce the time required to transmit or store the data. Also, it is important for error correction and detection. In general, longer codewords can provide higher reliability and better error correction and detection capabilities. Encoding the original data with error correction codes that add redundancy can detect and correct errors when the data is received. One can also add extra parity bits at the end of a code word to provide redundancy for error detection and correction. But, it may come at the cost of a reduced code rate and increased overhead. The optimisation of error correction and detection processes, ensuring accurate and reliable data transmission or storage.

In summary, a superblock is M time bins long, divided in N subblocks with one photon each. Each superblock contains a codeword that represents a transmitted symbol.

2.3.4 Number of Information Bits

The number of information bits refers to the count of individual binary digits (bits) that carry meaningful data in a given data communication. These bits represent the information conveyed, excluding any additional bits used for error correction. In the designed protocol BPPM, the superblock can be arranged in n! ways since each subblock length occurs exactly once in a superblock. The conveyed information bit, B, in such a subblock is upper bound by

$$B = log_2(n!) \tag{2.3}$$

2.4 Transmission in Protocol

2.4.1 Transmission in Beyond Pulse Position Modulation (BPPM)

In a noise-free channel, Beyond PPM (BPPM) is built out of superblocks defined by n photons with M time bins that entail $n!$ combinations and $\frac{log_2 n!}{n}$ bits per photon and $\frac{log_2 n!}{M}$ bits per time bin. We define a subblock as an integral number of time bins (or time bins, or other orthogonal encoding sources, i.e., that can be perfectly discriminated). The number of codewords corresponds to the number of ways we can organise the subblocks in a superblock (permutation)

$$K = n! \tag{2.4}$$

where K is the number of the codewords and n is the Number of Photons in a superblock.

The information content per symbol of the superblock is

$$log_2(n!) \quad \text{bits/symbol} \tag{2.5}$$

The information content per photon of the superblock is

$$\frac{log_2(n!)}{n} \quad \text{bits/photon} \tag{2.6}$$

Moreover, the information content per time bin of the superblock is:

$$\frac{log_2(n!)}{M} \quad \text{bits/time bin} \tag{2.7}$$

The metrics will be illustrated with the example of the photon number n = 4. Let us represent the [1,2,4,7] superblock in terms of time bins and polarization, which comprises $1+2+4+7 = 14$ lengths of the subblock, equivalent to 14 time bins. For 4 photons, which is [1,2,4,7] in a superblock according to the sequence, then there are $4! = 24$ permutations of ways to organise the subblocks to generate different superblocks representing the corresponding symbols

$$[1,2,4,7] \rightarrow A, [1,2,7,4] \rightarrow B, [1,4,2,7] \rightarrow C,$$

$$[1,4,7,2] \rightarrow D, [1,7,2,4] \rightarrow E, [1,7,4,2] \rightarrow F$$

$$[2,1,4,7] \rightarrow G, [2,1,7,4] \rightarrow H, [2,4,1,7] \rightarrow I$$

$$[2,4,7,1] \rightarrow J, [2,7,1,4] \rightarrow K, [2,7,4,1] \rightarrow L$$

$$[4,1,2,7] \rightarrow M, [4,1,7,2] \rightarrow N, [4,2,1,7] \rightarrow O$$

$$[4,2,7,1] \rightarrow P, [4,7,1,2] \rightarrow Q, [4,7,2,1] \rightarrow R$$

$$[7,1,2,4] \rightarrow S, [7,1,4,2] \rightarrow T, [7,2,1,4] \rightarrow U$$

$$[7,2,4,1] \rightarrow V, [7,4,1,2] \rightarrow W, [7,4,2,1] \rightarrow X$$

Using n = 4 in BPPM protocol, we can cover nearly all the English alphabets (A-Z) for digital communication.

For 4 photons in a superblock, we have

$$log_2(4!) = 4.6 \quad \text{bits/symbol}$$

In each photon, we have

$$\frac{log_2(4!)}{4} = 1.15 \quad \text{bits/photon}$$

and for each time bin, it has

$$\frac{log_2(4!)}{1+2+4+7} = 0.33 \quad \text{bits/time bin}$$

For 4 photons and 14 time bins, BPPM provides:
1. $K = n! = 4! = 24$ ways to order codewords
2. 4.6 bits/symbol
3. 1.15 bits/photon
4. 0.33 bits/time bin

2.4.2 Transmission in Pulse Position Modulation (PPM)

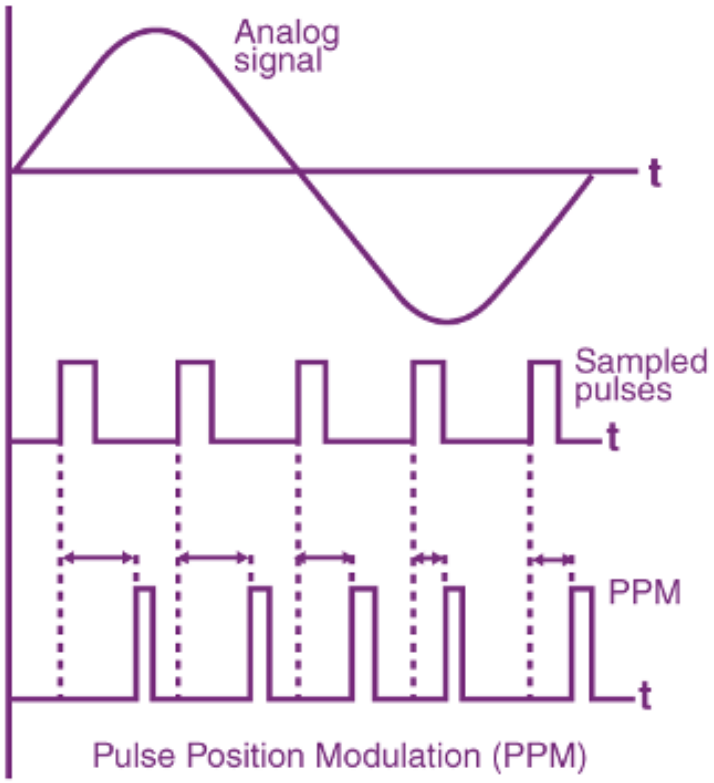

Figure 2.6: The diagram illustrates the waveform of PPM. Image is extracted from [1].

Pulse Position Modulation (PPM) is a type of signal modulation used for analogue and digital signal transmissions in various communication systems, including optical fibre and infrared remote controls. As shown in Figure 2.6, PPM is a form of pulse modulation where data is transmitted using short pulses of fixed width and amplitude, but the time delay between each pulse varies due to the position of pulse with repsect to that of a reference pulse differ [37]. In PPM, the time delay between each pulse represents the encoded data, with each delay corresponding to a different data symbol. The length of each delay is typically shorter than the pulse width, which allows for multiple symbols to be transmitted within a single pulse. The time delay between each pulse is typically measured from a fixed reference point, such as the beginning of the pulse, which allows for accurate data decoding.
PPM has several advantages over other modulation types, particularly in optical communication systems. Because PPM uses short pulses with a fixed amplitude, it is less susceptible to distortion from noise and other sources of interference for constant power transmission. Additionally, PPM can be used to transmit analogue and digital signals, allowing for greater flexibility in the types of data that can be transmitted. Also, PPM can be easily easily separated from a noisy signal [37].

However, PPM is highly complex, and requires more bandwidth for transmission [37]. PPM can be applied to air traffic control system and telecommunication system. In order to effectively use PPM in communication systems, we must carefully optimise the pulse width, pulse interval, and pulse position to ensure accurate and reliable transmission of data. Additionally, sophisticated signal processing techniques may be required to decode the data and compensate for noise and other sources of interference.

Figure 2.7: The diagram illustrates a single photon at the start of the superblock using PPM.

In this protocol, there is only one photon, as shown in Figure 2.7, no matter how many the time bins are in a superblock. We can freely choose how many the time bins are.

$$n = 1, \text{Time Bin} = M \tag{2.8}$$

The information content per symbol of the superblock

$$\log_2 \mathrm{K} = log_2 \mathrm{M} \quad \text{bits/symbol} \tag{2.9}$$

The information content per photon of the superblock

$$\frac{log_2 \mathrm{M}}{1} = log_2 M \quad \text{bits/photon} \tag{2.10}$$

Since we have 1 photon, the value of information bits per symbol is the same as that of information bits per photon.

The information content per time bin of the superblock

$$\frac{log_2 \mathrm{M}}{\mathrm{M}} \quad \text{bits/time bin} \tag{2.11}$$

In order to fairly compare PPM to BPPM, we choose 14 time bins in PPM, which corresponds to 4 photons in BPPM.

The representation of bits to symbol is as follow

$$[10000000000000] \rightarrow A, [01000000000000] \rightarrow B$$

$$[00100000000000] \rightarrow C, [00010000000000] \rightarrow D$$

$$[00001000000000] \rightarrow E, [00000100000000] \rightarrow F$$

$$[00000010000000] \rightarrow G, [00000001000000] \rightarrow H$$

$$[00000000100000] \rightarrow I, [00000000010000] \rightarrow J$$

$$[00000000001000] \rightarrow K, [00000000000100] \rightarrow L$$

$$[00000000000010] \rightarrow M, [00000000000001] \rightarrow N$$

For sending and receiving 4 photons in PPM protocol, we have 14 letters out of 26 of the alphabet (A-Z) for digital communication.

For 14 time bin in a superblock, we have

$$log_2 14 = 3.8 \quad \text{bits/symbol}$$

For each photon, it contains

$$\frac{log_2 14}{1} = 3.8 \quad \text{bits/photon}$$

For each time bin, it has

$$\frac{log_2 14}{14} = 0.27 \quad \text{bits/time bin}$$

For 1 photon and 14 time bins, PPM provides

1. M = K ways to order codewords
2. 3.8 bits/symbol
3. 3.8 bits/photon
4. 0.27 bits/time bin

2.4.3 Transmission in On-Off Keying (OOK)

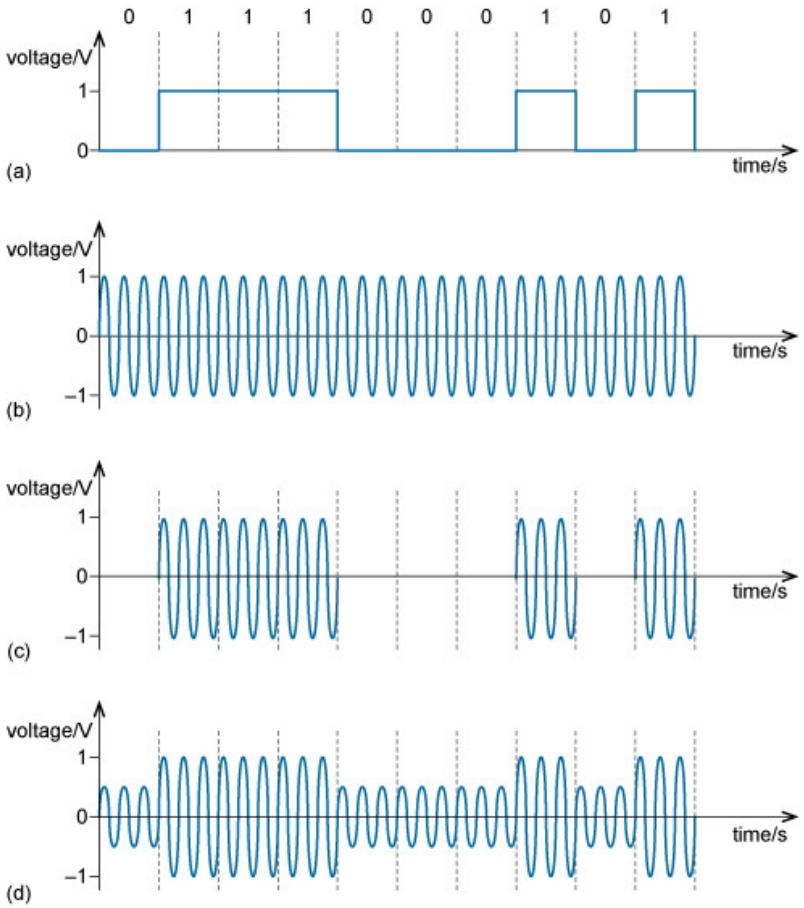

Figure 2.8: A digital communication signal using two voltage levels is depicted in figure (a). One level corresponds to 1 and the other to 0. Figure (b) shows the unmodulated carrier in more details. The modulated waveforms utilising the two ASK versions are shown in Figures (c) and (d). OOK is used in Figure (c), while binary ASK, or BASK, is used in Figure (d). Image is extracted from [32].

The most straightforward type of amplitude-shift keying is on-off keying (OOK). Depending on whether the input message is binary 1 or binary 0, OOK will either have bursts of carrier wave to send or nothing at all. OOK is more noise-sensitive when employing a regenerative receiver or a poorly designed superheterodyne receiver, although it is more spectrally efficient than frequency-shift keying [3]. The bandwidth of an OOK signal and a BPSK (Binary Phase Shift keying) signal are equal for a given data rate. OOK is applied to early telephone modem , it can also transmit digital data over optical fiber.
The number of time bins in OOK is double its number of photon in average.

$$M = 2n \quad \text{time bins} \tag{2.12}$$

, where n is the number of photons in the protocol.

The number of a codeword corresponds 2 to the power of the time bin to organize the subblocks in a superblock

$$K = 2^{2n} = 2^M \tag{2.13}$$

, where K is the number of the codewords, n is the number of photons, and M is the number of the time bins in a superblock.

The information content per symbol of the superblock is

$$log_2 K = M \tag{2.14}$$

The information content per photon of the superblock is

$$\frac{log_2 K}{n} = \frac{M}{n} \quad \text{bits/photon} \tag{2.15}$$

The information content per time bin of the superblock is

$$\frac{log_2 K}{M} = \frac{M}{M} = 1 \quad \text{bits/time bin} \tag{2.16}$$

Figure 2.9: The diagram illustrate 7 photons arbitrarily placed among 14 time bins in a superblock using OOK.

In order to fairly compare OOK with BPPM, we choose 14 time bins, which correspond to 7 photons in BPPM.

$$M = 2 * 7 = 14 \quad \text{time bins}$$

OOK uses 7 photons on average in 14 time bins.

$$K = 2^{2*7} = 2^{14} = 16,384$$

For sending and receiving 7 photons in OOK protocol, we have 16,384 codewords for digital communication.

For 14 time bin in a superblock, we have

$$log_2 16,384 = 14 \quad \text{bits/symbol}$$

For each photon, it contains

$$\frac{14}{7} = 2 \quad \text{bits/photon}$$

For each time bin, it has

$$\frac{14}{14} = 1 \quad \text{bits/time bin}$$

For 7 photons and 14 time bins, OOK provides
1. $K = 2^M = 2^{14} = 16,384$ ways to order codewords
2. 14 bits/symbol
3. 2 bits/photon
4. 1 bits/time bin

2.4.4 Transmission in a general protocol

We define a protocol where n photons can be arbitrarily placed among M time bins as a "General Protocol". In the general protocol, for n photons per superblock (noise free), the number of ways n photons are placed among M positions is the number of codewords that corresponds to the number of allowed binomial combinations to organise the subblocks in a superblock

$$K = \binom{M}{n} = \frac{M!}{(M-n)!n!} \tag{2.17}$$

where K is the number of codewords, n is the number of photons, and M is the number of time bins in a superblock.

The information content per symbol of the superblock is

$$log_2 K = log_2 \binom{M}{n} \quad \text{bits/symbol} \tag{2.18}$$

In order to fairly compare General with BPPM, we choose 4 photons and 14 time bins. The Information content per photon of the superblock is

$$\frac{log_2 \binom{M}{n}}{n} \quad \text{bits/photon} \tag{2.19}$$

The information content per time bin of the superblock is

$$\frac{log_2 \binom{M}{n}}{M} = \frac{log_2 \binom{M}{n}}{M} \quad \text{bits/time bin} \tag{2.20}$$

Figure 2.10: The diagram illustrate 4 photons arbitrarily placed among 14 time bins in a superblock using general protocol.

For sending and receiving 4 photons in the general protocol, we have

$$K = \binom{14}{4} = \frac{14!}{(14-4)!4!} = 1,001 \text{codewords}$$

For 14 time bin in a superblock, we have

$$log_2\binom{14}{4} = 9.97 \quad \text{bits/symbol}$$

Each photon contains

$$\frac{log_2\binom{14}{4}}{4} = 2.49 \quad \text{bits/photon}$$

For each time bin, it has

$$\frac{log_2\binom{14}{4}}{14} = 0.71 \quad \text{bits/time bin}$$

For 4 photons and 14 time bins, general protocol provides
1. $K = \binom{M}{n} = \binom{14}{4} = 1,001$ ways to order codewords
2. 10 bits/symbol
3. 2.5 bits/photon
4. 0.71 bits/time bin

Table 2.1: Table depicts BPPM, PPM, OOK and General protocol in terms of permutation, bits per symbol, bits per photon, bits per time bin when there are 4 photons and 14 time bins in BPPM, the corresponding parameters in other protocols are chosen for fair comparison. For 4 photon, BPPM does not exhibit an advantage among other 3 protocols, however, it could provide higher information content for higher number of photon, which will be shown in the section Results and Discussion.

Protocol	Number of Codeword, K	Bits/Symbol	Bits/Photon	Bits/Time Bin
BPPM	n! = 24	4.6	1.15	0.33
PPM	M = 14	3.8	**3.8**	0.27
OOK	2^{2M} = **16,384**	**14**	2	**1**
General	$\binom{M}{n}$ = 1,001	9.97	2.49	0.71

2.5 BPPM's Principle Error Correction Capabilities

2.5.1 Principle

2.5.1.1 A Subblock and a Superblock in BPPM

First, a subblock is defined as an integral number of time bins (or other orthogonal encoding resources, i.e., that can be perfectly discriminated). Then, a superblock of n smaller subblocks is formed, whose lengths exhaust the first n integers of the sequence

$$a(n) = (1, 2, 4, 7, 12, 20, 29, 38, 52, 73, 94, 127, 151, 181, 211, 257, 315, 373, 412, ...) \tag{2.21}$$

This sequence has the property that pairwise sums of distinct elements from the sequence are all distinct. Hoey first described it. See [15].

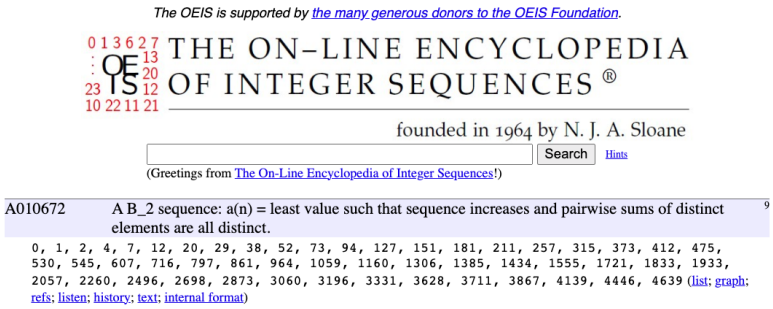

Figure 2.11: A diagram shows a sequence A010672 used in BPPM. Image is extracted from [15].

The sequence shown in Figure 2.11 states that a(n) is the smallest value such that sequence increases and pairwise sums of distinct elements are all distinct, which can be viewed as the number of time bins in the corresponding subblock. In Figure 2.5, there are four photons in the superblock, the sequence is $a(4)$, where the last subblock, subblock 4 contains 7 time bins. The length of the superblock containing n subblocks is the sum of the first nth sequence, $M(n)$, which is equivalent to the sum of time bins in a superblock. Putting 4 into the series $M(n)$ and get 14, which corresponds to $1+2+4+7 = 14$ time bins in that superblock in Figure 2.5.

$$\begin{aligned} M(n) &= (1, 3, 7, 14, 26, 46, 75, 113, 165, 238, 332, 459, 610, 791, 1002, 1259, 1574, 1947, 2359, ...) \\ &= \sum_{i=1}^{n} a(i) \end{aligned} \tag{2.22}$$

2.5.1.2 More than 1 Bit of Information pr Photon

Note that this results in more than 1 bit of information per photon, $\frac{B}{n}$, when n>3, i.e., when

$$\frac{B}{n} = \frac{log_2(n!)}{n} > 1 \tag{2.23}$$

Example: Consider the case with three subblocks of length 1, 2, 4 and 7. They can be ordered in 4! = 24 ways ([1,2,4,7],[1,4,2,7],[2,1,4,7],[2,4,1,7],[4,1,2,7],[4,2,1,7],...) and therefore transmit

$$log_2(24)/4 \approx 1.15 > 1 \quad \text{bits of information per superblock.} \tag{2.24}$$

2.5.1.3 Ways of encoding with Polarization

For each subblock, the first time bin contains a photon with one of two orthogonal polarisations, e.g., horizontal (H) or vertical (V) linear polarizations. In the remaining text, a single photon will be used. However, in principle, the scheme could also be implemented with any polarised light pulse.

Any given subblock i is tagged with an "H" state if the next subblock $i+1$ is longer than subblock i and with a "V" if the next subblock is shorter than the previous subblock. For the last subblock $i = n$, it is tagged with an "H" ("V") if the first subblock $i = 1$ is longer (shorter).(The string indicates how the symbols are ordered in the time bins, starting with the first bin and finishing with the last.)

The relation is summarized as follows:

$$\begin{aligned} l_i < l_{i+1} &\implies H_i \\ l_i > l_{i+1} &\implies V_i \\ l_1 > l_n &\implies H_n \\ l_1 < l_n &\implies V_n \end{aligned} \tag{2.25}$$

where l_i is the length of the subblock i, 1 is the first subblock, n is the last subblock, H_i is the ith subblock in H polarisation, and V_i is the ith subblock in V polarization. The rule is observed to be cyclical, i.e., $l_n < l_1 \implies H_n$ and so on (the previous subblock is shorter than the next subblock).

Example If the codeword [1, 2, 4, 7] in Figure 2.5 is sent, the transmitted encoded word will be

$$[1, 2, 4, 7] \rightarrow HH0H000V000000 \tag{2.26}$$

Since all the previous subblocks are shorter than the next ones except the last subblock is longer than the first one, then there are 3H 1 V in this encoding.

Example If the codeword [2, 1, 4] is sent, the transmitted encoded word will be

$$[2, 1, 4] \rightarrow V0HV000 \tag{2.27}$$

Following the rules in 2.25, the encoding can be easily obtained.

2.5.2 Error Correction Capabilities

The suggested encoding can protect the superblock from the loss and addition of a photon. The proposed error correction code does not rely on tables. Instead, the error can be identified and corrected based on simple rules.

2.5.2.1 Addition of photon errors

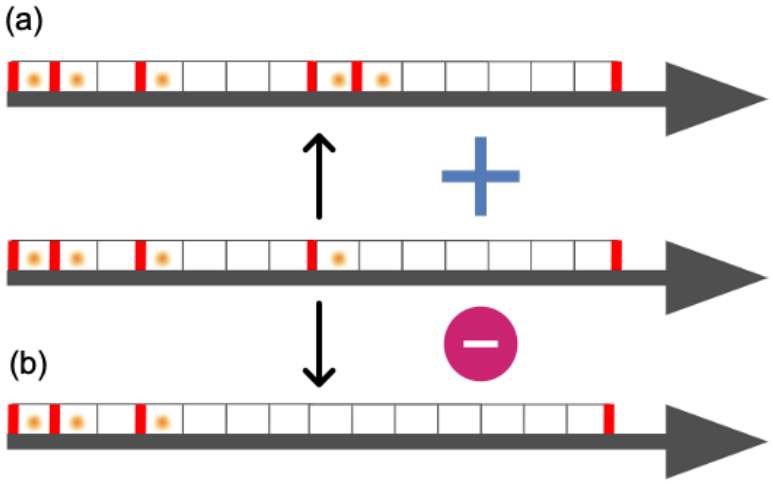

Figure 2.12: A diagram shows the detection of error correction of BPPM. Figure a illustrates the added error to the superblock cause an additional photon to the last subblock and change in the last subblock length. Figure b illustrates the added error to the superblock cause an lost photon to the last subblock and change in the last subblock length.

Suppose [1,2,4,7] was transmitted, and an additional photon appears inside the "7" subblock. The receiver would get one of [1,2,4,1,6], [1,2,4,2,5], [1,2,4,3,4], [1,2,4,4,3], [1,2,4,5,2], [1,2,4,6,1].

Let [1,2,4,1,6] as the situation, the additional photon is horizontally polarized.

$$[1, 2, 4, 1, 6] \rightarrow HH0H000VH00000 \tag{2.28}$$

In Figure 2.12a, the receiver would see more subblocks than expected, and one disallowed subblock length would indicate the position of the error. Next to the not allowed subblock length , i.e. 6, there will be a subblock whose length appears once in the sequence. Thus, the error can be corrected by identifying this subblock and merging the two subblocks. In one case, the situation is symmetric, i.e., for [1,2,4,3,4], and it is not immediately clear which subblock to merge with the not allowed subblock length. However, the error can be corrected using the polarisation information in this case.

2.5.2.2 Loss of Photon Error

When a photon is lost, a subblock with length l, not in the list of allowed lengths, will be present and signal an error. Polarisation is needed to decide how to proceed.

Example:
Suppose the codeword starts with [1,2,4,7]: After a loss, the sequences [3,4,7], [1,6,7],[1,2,11],[1,2,4] could be received. the first photon could also be lost, but this case is easy to correct since the start time of each superblock is assumed to be known. Each of these configurations has exactly one subblock with a disallowed length. If this subblock is initialised with an "H", it is composed of two lengths, the first one smaller than the second.
If the superblock becomes [1,2,11] as shown in Figure 2.12b,

$$[1, 2, 11] \rightarrow HH0H00000000000 \tag{2.29}$$

then subblock of length 11 must be composed of [4,7] and not [7,4] due to the original polarization since H but not V as shown in the following

$$[1, 2, 7, 4] \rightarrow HH0V000000V000 \tag{2.30}$$

2.5.2.3 Correctable Multiple Errors

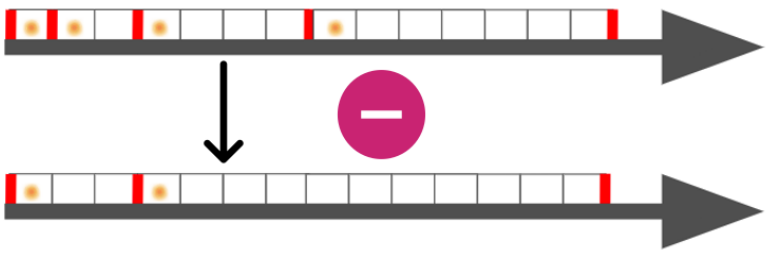

Figure 2.13: A diagram illustrates correctable multiple errors to the superblock by BPPM.

Consider the situation where multiple non-adjacent photons were lost. For this purpose, superblock [1,2,4,7] with equaiotn 2.26 is used as an example. Suppose that

after transmission, the receiver gets

$$[3, 11] \rightarrow H00H0000000000. \tag{2.31}$$

As shown in Figure 2.13, the receiver can conclude that likely 2 loss errors occurred (it could also be that 3 loss errors and one photon addition error occurred, but with a much lower probability). The 3 subblocks can be corrected since the initial photon is H, meaning that [2,1] is ruled out and, therefore, 3 → [1,2] is the correct mapping. After correcting the 3 subblocks, the 11 subblocks can be corrected similarly.

Suppose the codeword stars with [1, 2, 4]

$$[1, 2, 4] \rightarrow HH0V000 \tag{2.32}$$

$$HH0V000 \rightarrow H00V000 \rightarrow HV0V000 \tag{2.33}$$

If the second H is lost and a V is added in the same subblock, or it can be

$$HH0V000 \rightarrow HH00000 \rightarrow HH000V0 \tag{2.34}$$

If the last photon is lost and another one is added in the same subblock. However, in both examples, the resulting word is not a valid codeword or a valid syndrome. It can therefore be detected as an error without the prospect of being able to correct it. Typically, such errors are reconciled by asking the sender to resend the erroneous subblock.

Another example of considering the loss in different positions in [1,2,4,7] is if there is an added error in sending 4 photons from Alice to Bob.

$$[1, 2, 4, 7] \rightarrow [1, 2, 2, 2, 7] \tag{2.35}$$

In this case, it is known for sure that the [2,2,2] sequence must be either [4,2] or [2,4] due to the allowed subblock lengths. The polarisation of the first two subblocks can be used to decide which one it is. If the polarisation is H, then it is known that the sent sequence must have been [1,2,4,7] under 1 error occurrence.

2.5.3 Most Uncorrectable Errors can be Detected

The code can not always correct for two or more adjacent loss errors and some other multi-errors. However, most errors can be detected since the agreed protocol states that a fixed number n of distinct subblock lengths should be present. Errors from a combination of addition and deletion occur with low probability, which makes them undetectable. However, such errors would require two subblocks to switch places,

requiring at least three errors. An example of this would be

$$\begin{aligned}
&[1,2,4] = HH0V000 \\
&(1 \text{ loss}) \rightarrow 0H0V000 \\
&(1 \text{ loss}) \rightarrow 000V000 \\
&(1 \text{ add}) \rightarrow V00V000 \\
&(1 \text{ add}) \rightarrow V0HV000 = [2,1,4], \text{ which is a valid codeword}
\end{aligned} \tag{2.36}$$

However, the more data bits sent through the channel, the more likely the errors are to appear, which raises the question of when implementing an error correction code pays off. This research question motivates the investigation of Reed-Solomon Code implementation into the proposed protocol for error-free single-photon telecommunication with low energy consumption.

Reed-Solomon codes (RS codes) are a type of widely-used error-correcting code in telecommunication, known for efficiently correcting errors in data transmission. RS codes are based on finite field theory for efficient encoding and decoding algorithms. Block codes operate on fixed-size data blocks to allow comprehensive data to be divided into packets. The added redundancy to the original data takes the form of extra symbols to allow the receiver to recover from errors in the received data. Also, they are capable of both error detection and correction. The receiver can determine if errors can occur by examining the received symbols. The errors can be corrected to restore the original data if they are detected. RS Codes are versatile enough to handle various types of errors, such as noise and burst errors, for a wide range of communication channels. Therefore, RS codes suit satellite communication, wireless networks, and data transmission protocols. More details can be found in the Appendix. After explaining the error correction capabilities, binomial distribution of error will be introduced in the next section.

2.6 Binomial Distribution of Error

2.6.1 Combination

In mathematics, a combination is a fundamental concept used to describe the selection of objects from a collection where the selection order is irrelevant. The notation $\binom{n}{r}$ is known as "n choose r," which refers to the number of ways r objects can be selected from a set of n objects. The explicit formula for a combination without repetition (replacement) is expressed as

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{2.37}$$

For instance, a list of three photons with the length of seven-time bins, such as [1,2,4], can be organised in 3! ways. The number of ways to select these three photons from a collection of seven-time bins in which the selection order does not matter is given by $\binom{7}{3}$. This concept of combinations can be extended to evaluate the probability of an event in a communication system, such as the probability of an error in transmitting data.

2.6.2 Error Probability

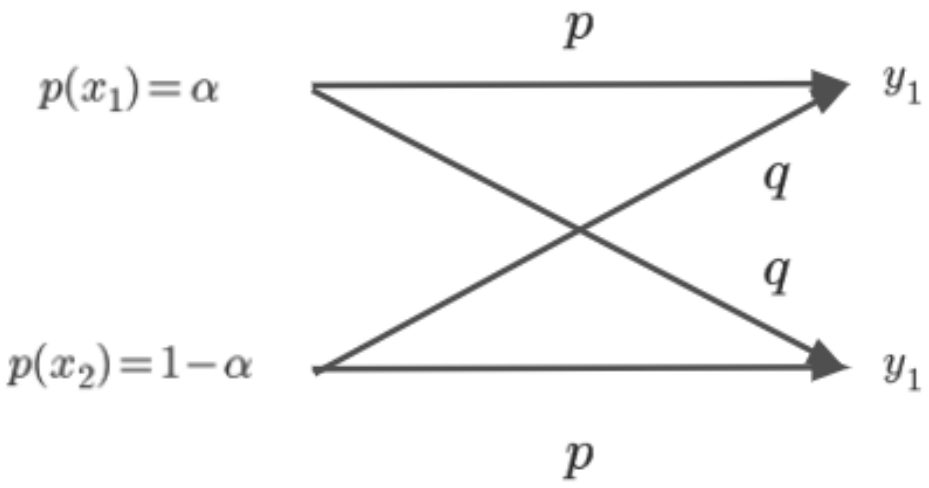

Figure 2.14: The diagram shows the binary symmetric channel. $p(x_i)$ input, where $p(x_1) = \alpha$, $p(x_2) = 1 - \alpha$. $p(y_i)$ is the output probability, and $p(e|xi)$ is the transition probability that the input is transferred to the output with error e, of which its probability is indicated as p and q in this communication channel. $i = 1, 2$

From Figure 2.14, the error probability P_E of a binary symmetric channel is computed as

$$\begin{aligned} P_E &= \sum_{n=1}^{2} p(e|x_i)p(x_i) \\ &= p(e|x_1)p(x_1) + p(e|x_2)p(x_2) \\ &= qp(x_1) + qp(x_2) \\ &= q\alpha + q(1-\alpha) \\ &= q \end{aligned} \tag{2.38}$$

where $p(e|x_i)$ is the error probability given input x_i, E is the overall error in the channel, and e is the error in a given input x_i. It states that the unconditional error probability P_E equals the conditional error probability $\sum_{i \neq j} p(y_j|x_i)$.

2.6.3 Probability Distribution

Suppose there are n photons in a superblock. The binomial loss error distribution is the probability of losing l and not losing $n - l$ photons.

$$P_{loss}(n, l, P_l) = \binom{n}{l} P_l^l (1 - P_l)^{n-l}, \tag{2.39}$$

n is the number of photons, l is the number of lost photons, and P_l is the independent probability of loss error for a single photon.

Similarly, the binomial distribution of added errors represents the probability of occurrence for the added errors. The number of errors remained from the total error needed to be considered after dealing with the loss error, which is written as

$$P_{add}(M, n, a, P_a) = \binom{M-n}{a} P_a^a (1 - P_a)^{M-n-a} \tag{2.40}$$

n is the number of photons, a is the number of added photons, M is the number of time bins in the coded subblock, and P_a is the probability of added error.

In general, the probability that exactly a added photon errors and l lost photon errors occur is

$$P(l, a, n, M, P_l, P_a) = \binom{n}{l} P_l^l (1 - P_l)^{n-l} \times \binom{M-n}{a} P_a^a (1 - P_a)^{M-n-a}, \tag{2.41}$$

Where n is the number of photons, M is the number of time bins in the coded subblock, l is the number of lost photons, a is the number of added photons, P_l and P_a are the independent probability of loss error and added photon error respectively. The (M-n) number of time bins in the coded subblock are empty, so only these are possible for added photon errors.

Let us take an example with consideration of the error probability distribution

$$[1, 2, 4] \rightarrow HH0V000 \tag{2.42}$$

Considering only loss error in the above codeword, the total probability distribution of loss error is

$$P(l, 0, 3, 7, P_l, 0) = (1 - P_l)^3 + 3(1 - P_l)^2 P_l + 3(1 - P_l) {P_l}^2 + {P_l}^3 \tag{2.43}$$

The probability for zero loss error is $(1-P_l)^3$, the probability for at most 1 loss error is $(1-P_l)^3+3(1-P_l)^2P_l$, and the probability for exactly 2 loss errors is $3{P_l}^2(1-P_l)$.

Coefficient 3 represents that two type of loss errors can occur in the subblocks [1,2], [1,4], and [2,4] for 1 loss error and [1], [2], and [4] for 2 loss errors. The 2 errors cannot occur in the same subblock since only one photon exists in each subblock. There is a possibility that the following scenario can occur

2.6.4 Binomial Distribution without Adjacent Errors

Since in a string of n elements, there exist n adjacent elements that cannot be corrected, the overall combination of errors detected in the distribution must be edited. The adjusted binomial distribution only applies to more than 1 error for having adjacent errors. Correcting 2 loss errors, except the adjacent ones, is

$$\text{Prob(2 non-adjacent loss errors)} = \{[\binom{n}{2} - n] \times P_l^2(1-P_l)^{n-2} \tag{2.44}$$

Since adjacent errors are counted cyclically, an error in the last and first subblocks would count as an adjacent error. Similarly, 3 adjacent errors can be obtained.

$$\text{Prob(3 non-adjacent loss errors)} = \{[\binom{n}{3} - n] \times P_l^3(1-P_l)^{n-3} \tag{2.45}$$

The formulas are valid for $n \leq 2$ and $n \leq 3$, respectively. For considering up to 2 errors of either loss or added errors, the probability of "Success" (correcting the received superblock)

$$\begin{aligned} &P(\leq 2 \text{ non adjacent errors of either loss or added type}) \quad \text{where} \\ &= P(0,0) + P(0,1) + P(1,0) \\ &\quad + [\binom{n}{2} - n][P_l^2(1-P_l)^{n-2} \times (1-P_a)^{M-n} \\ &\quad + (1-P_l)^n \times P_a^2(1-P_a)^{M-n-2} \\ &\quad + P_l(1-P_l)^{n-1} \times P_a(1-P_a)^{M-n-1}], \end{aligned} \tag{2.46}$$

The last term is one loss error and one added error that cannot be corrected. In general, the probability of correcting l loss errors and a added errors without the adjacent ones can be expressed as follow

$$\text{Prob}(l,a,n,M,P_l,P_a) = \{[\binom{n}{l} - n] \times P_l^l(1-P_l)^{n-l}\} \times \{[\binom{M-n}{a} - (M-n)] \times P_a^a(1-P_a)^{M-n-a}\} \tag{2.47}$$

Figure 2.15 plots the error probability distribution.

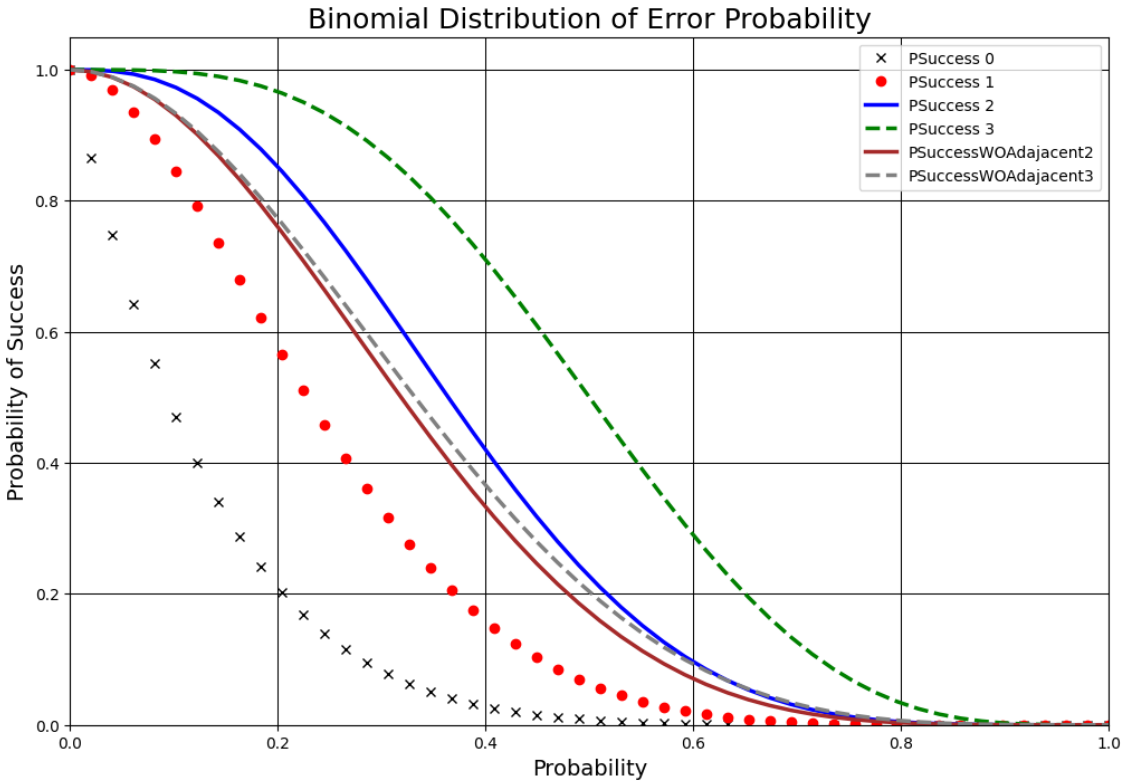

Figure 2.15: P(Success) is plotted vs the loss probability P_l for codes correcting various numbers of errors. The no-error correction case is also plotted. P(Success) is the probability of receiving the correct superblock.

2.6.5 Error Probability and Distance

Error probability can be indirectly related to distance or time. For example, in wireless communication systems, error probability is often used to quantify the likelihood that a transmitted signal will be received incorrectly due to interference or noise, and this can have implications for the quality of the communication link over distance or time. The error probability of a photon traveling a certain distance can be affected by various factors such as absorption, scattering, and noise.

For wired telecommunication, the attenuation of light in an optical fiber can be modeled using the following equation

$$I(d) = I_0 \cdot e^{-\mu d} \tag{2.48}$$

where $I(d)$ is the intensity of the light at a distance d from the source, I_0 is the initial intensity of the light, μ is the attenuation coefficient of the fiber. Thus, the probability that a photon was lost $P_{loss} = 1 - e^{-\mu d}$.

The attenuation coefficient β depends on the properties of the fiber, such as its material, geometry, and any impurities or defects. It also depends on the wavelength

of the light, with certain wavelengths being more strongly attenuated than others.

The error probability in an optical fiber communication system can be affected by signal-to-noise ratio, dispersion, and nonlinear effects. The relationship between error probability and distance traveled will depend on the specific system parameters, such as the signal power, modulation scheme, and fiber properties.

Signal power represents the strength of the signal being transmitted through a medium. It is often measured in watts or dBm (decibels referenced to 1 milliwatt). The initial intensity I_0 corresponds to the signal power at the source, and the intensity at a distance d is $I(d)$.

The signal-to-noise ratio (SNR) measures the strength or quality of a signal compared to the background noise present in a system. It is the ratio of the power of the desired signal to the power of background noise or interference. When a photon travels a certain distance in wired or wireless network, SNR should be considered, where a higher SNR indicates a better quality with loss or noise during transmission. Mathematically, SNR can be expressed as

$$\frac{\text{Signal Power}}{\text{Noise Power}} = \frac{I_0 \cdot e^{-\mu d}}{\text{Noise Power}} \tag{2.49}$$

It helps determine the error probability in the channel. The noise power depends on various factors, such as the characteristics of the communication channel and the sources of interference.

Dispersion refers to the spreading out of a signal as it propagates through a medium, typically due to variations in the propagation speeds of different signal components (wavelengths or frequencies). In telecommunication, dispersion can lead to signal distortion by spreading the signal to overlap with adjacent bits, that causes errors in data transmission. Chromatic dispersion (by wavelength) and modal dispersion (by propagation modes in multi mode fibers). Chromatic dispersion (D) can be expressed as

$$D = \beta \cdot (\lambda^2) \cdot \Delta L \tag{2.50}$$

where β is the group velocity dispersion parameter, λ is the wavelength of light, and ΔL is the distance over which dispersion is considered.

Modal dispersion occurs in multi-mode optical fibers, where different modes (paths) through the fiber and have slightly different propagation speeds. The modal dispersion can be expressed as

$$\Delta t = \frac{L \cdot \Delta \eta}{c \cdot \eta} \tag{2.51}$$

where Δt is the modal dispersion in seconds, L is the length of the optical fiber in meters, and $\Delta\eta$ is the refractive index difference between the fiber's core and the cladding. c is the speed of light in vacuum, η is the average refractive index of the fiber core.

Nonlinear effects occur when a system's response is not directly proportional to the input signal. In optical fibers, nonlinearities can arise due to factors like high signal power or specific fiber properties. It can introduce distortions, such as harmonics, and inter-modulation distortion, which can affect signal quality and error probability. Nonlinearity can be expressed as

$$\Delta P = \gamma \cdot P^2 \cdot L \tag{2.52}$$

where ΔP represents the change in optical power due to nonlinear effects, γ is the fiber's nonlinear coefficient, which depends on the material properties, and P is the optical power.

Since the change in optical power is proportional to the square of the optical power itself and the length of the fiber, nonlinear effects become more significant with higher power levels.

2.7 Information Theory

Information theory provides a measure of the information in message signals, which can help determine a system's capability to transfer this information from source to destination and how data is being compressed.The metrics employed in this study are entropy and Mutual Information (MI), both of which are mathematical functions derived from the probability distributions that underpin the act of communication [11].

Entropy is a well-defined measure for every probability distribution, and it serves as a fundamental concept in information theory. Moreover, the concept of entropy is further utilised to establish the notion of mutual information, which quantifies the extent to which one random variable carries information about another. Entropy is then defined as the measure of self-information associated with a random variable. Mutual information may be classified as a specific instance of a broader concept known as relative entropy, which serves as a metric for quantifying the dissimilarity between two probability distributions. All of these numbers exhibit a strong correlation and possess a number of fundamental characteristics in common.

According to Shannon's noisy-channel coding theorem, digital source output can be transmitted over a channel with an arbitrarily small probability of error even in the presence of noise if the source's computable information rate is less than the channel capacity (or Shannon Capacity) [27]. The term "channel capacity" pertains to the theoretical upper limit of error-free data transfer that may be achieved via a channel

under the influence of random data transmission faults at a certain degree of noise. The initial description of this concept was provided by Shannon in 1948, and subsequently documented in a publication co-authored by Shannon and Warren Weaver titled "The Mathematical Theory of Communication" [28]. However, Shannon's theorem provides clues on implementing error correction to reach this capacity, which motivates the investigation of error-correcting codes in telecommunication protocols.

2.7.1 Information

Let x_j be an event with $p(x_j)$ probability. If that event x_j has occurred, then information units have been received.

$$I(x_j) = \mathbf{log}_a \frac{1}{p(x_j)} = -\mathbf{log}_a p(x_j) \tag{2.53}$$

The base of the logarithm determines the units by which information is measured, with the standard unit being the binary unit or "bit" when using logarithms to base 2. It is the most straightforward random experiment with equally likely outcomes, such as flipping a coin, which has one bit of information associated with it. The use of base 2 logarithms is also consistent with the binary nature of digital computers.

2.7.2 Discrete Channel Models

It is assumed that our communication channel is memoryless, meaning that the channel output at a given time depends on the channel input and is not influenced by previous inputs. Discrete memory-less channels are defined by the set of conditional probabilities that relate the probability of each output state to the input probabilities.

Considering the random variables on the transmitter and receiver side, let X be a discrete input random variable with alphabet $\mathcal{X}$ and probability mass function $p(x) = Pr\{X = x\}, x \in \mathcal{X}$, and Y is the output being defined in a similar procedure. The variables $p(x)$ and $p(y)$ are distinct random variables that correspond to two separate probability mass functions. The probability density functions of random variables X and Y are denoted as $p_X(x)$ and $p_Y(y)$, respectively.

For example, a channel with two inputs and three outputs, each input-to-output path indicating a conditional probability p_{ij}, a concise notation for $p(y_j|x_i)$. The conditional probability p_{ij} represents the probability of obtaining output y_j given that the input is x_i and is called a channel transition probability.

The set of transition probabilities completely specifies the channel. The matrix of transition probabilities $[P(Y|X)]$ specifies a discrete channel, determining the conditional distribution of the output given the input. Since each input to the channel

results in some output, each new row of the channel matrix must sum to unity. The probability matrix can be written as follow

$$P[(Y|X)] = \begin{bmatrix} p(y_1|x_1) & p(y_2|x_1) & p(y_3|x_1) \\ p(y_1|x_2) & p(y_2|x_2) & p(y_3|x_2) \end{bmatrix} \tag{2.54}$$

The channel matrix helps derive the output probabilities given the input probabilities. For example, if the input probabilities P(X) are represented by the row matrix

$$[P(X)] = \begin{bmatrix} p(x_1) & p(x_1) \end{bmatrix} \tag{2.55}$$

then

$$P(Y) = \begin{bmatrix} p(y_1) & p(y_2) & p(y_3) \end{bmatrix} \tag{2.56}$$

which is computed by

$$[P(Y)] = [P(X)][P(Y|X)] \tag{2.57}$$

If $[P(X)]$ is written as a diagonal matrix, the above equation yields a matrix $[P(X,Y)]$. Each element in the matrix has the form $p(x_i)p(y_j,x_i)$ or $p(x_i,y_j)$. This matrix is known as the joint probability matrix, and the term $p(x_i,y_j)$ is the joint probability of transmitting x_i and receiving y_j

2.7.3 Entropy

In general, it is more interesting to know the average information about an experiment's results than the specific information about each individual incident. The Entropy $H(X)$ is the measure of the average informational uncertainty connected to a discrete random variable X with a probability mass function $p(x_j)$.

$$H(X) = -\sum_{j=1}^{n} p(x_j) \log_2 p(x_j) \tag{2.58}$$

where $n \in \mathcal{X}$ is the total number of possible outcomes. There is a convention that $0 \log 0 = 0$ is easily justified by continuity since $x \log x \to 0$ is $x \to 0$. There is no change to the entropy when terms of zero probability are added to the system.

Given that entropy is a mathematical function associated with the probability distribution X, it is independent of the specific values assumed by the random variable X. Instead, it just relies on the probabilities $p(x_j)$ associated with each possible outcome x_j.

Example

Consider a random variable that follows a uniform distribution with 128 possible outcomes. In order to identify a result, it is necessary to employ a label that encompasses a total of 128 distinct values.

$$H(X) = -\sum_{j=1}^{128} p(x_j) \log_2 p(x_j) = -\sum_{j=1}^{128} \frac{1}{128} \log_2 \frac{1}{128} = log_2 128 = 7 \quad \text{bits} \tag{2.59}$$

, which corresponds to the amount of bits required to describe X.

Entropy can be regarded as the expected value of the random variable, i.e.,the average number of bits required to describe a random variable [11].

$$H(X) = E[I(x_j)] = -\sum_{x} p(x) \log_2 p(x) \tag{2.60}$$

Therefore, entropy should be maximum when each outcome is equally likely.

If one symbol is more likely to occur than the other, it is less uncertain about which symbol will appear at the source output.

Example
Suppose 8 symbols are being transmitted through a noiseless channel. Assume that probabilities of receiving them are $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256})$. The entropy can be calculated as follows

$$\begin{aligned} H(x) &= -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{16}\log_2\frac{1}{16} - -\frac{1}{32}\log_2\frac{1}{32} - \frac{1}{64}\log_2\frac{1}{64} - \frac{1}{128}\log_2\frac{1}{128} \\ &= \frac{1}{2}(1) + \frac{1}{4}(2) + \frac{1}{8}(3) + \frac{1}{16}(4) + \frac{1}{32}(5) + \frac{1}{64}(6) + \frac{1}{128}(7) + \frac{1}{256}(8) \\ &= 1.96 \approx 2 \quad \text{bits} \end{aligned} \tag{2.61}$$

A communication can be transmitted to indicate the specific symbol that is being requested. One such approach involves transmitting the index value corresponding to the symbol, with each symbol requiring a 3-bit description. However, the probabilities exhibit non-uniform distribution. Utilising shorter descriptions for symbols with higher probabilities and longer descriptions for symbols with lower probabilities is a logical approach to attaining a reduced average description length, in contrast to the uniform code's 3-bit length.

Conversely, if the probability of all events is zero, there is no uncertainty since which symbol will occur is precisely known.

Since

$$0 \leq p(x) \leq 1, \text{ implies that } \log \frac{1}{p(x)} \geq 0 \tag{2.62}$$

Then

$$H(x) \geq 0 \tag{2.63}$$

2.7.4 Joint Entropy

The entropy of a single random variable is defined in the previous subsection, and the definition of a pair of single vector-valued random variables (X, Y) is now extended. The joint entropy $H(X, Y)$ is the average uncertainty of the communication system.

If the input probabilities $p(x_i)$ are used, the output probabilities $p(y_j)$, the transition probabilities $p(y_j|x_i)$, and the joint probabilities $p(x_i, y_j)$, several different entropy functions for a channel with n inputs and m outputs can be defined. These are

$$H(X) = -\sum_{i=1}^{n} p(x_i)\mathbf{log}_2 p(x_i) \tag{2.64}$$

and

$$H(Y) = -\sum_{j=1}^{m} p(y_j)\mathbf{log}_2 p(y_j), \tag{2.65}$$

These entropies are easily interpreted. $H(X)$ is the average uncertainty of the source, whereas $H(Y)$ is the average uncertainty of the received symbol. The joint Shannon entropy of two discrete random variables X and Y with image $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$H(X, Y) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2 p(x, y), \tag{2.66}$$

Let x and y represent specific values of X and Y, respectively. The term p(x,y) denotes the joint probability of these values happening simultaneously, whereas $p(x, y)log2[p(x, y)]$ is assigned the value of 0 when the function $p(x, y)$ equals 0.

Joint entropy can also be expressed as

$$H(X, Y) = -E \log P(X, Y), \tag{2.67}$$

, where E is the expected value of the entropies.

2.7.5 Conditional Entropy

The concept of Conditional Entropy refers to the quantification of information required to characterise the result of a random variable Y, given the knowledge of another random variable X. It is denoted as $H(Y|X)$ and represents the conditional information entropy.

Given discrete random variables X with image $\mathcal{X}$ and Y with image $\mathcal{Y}$, if the random variable (X, Y) has the probability distribution $p(x, y)$, i.e., $(X, Y) \sim p(x, y)$, the conditional entropy of Y given X is defined as the weighted sum of $H(Y|X = x)$ for each possible value of x, using $p(x)$ as the weights, i.e.,

$$\begin{aligned} H(Y|X) &= \sum_{x \in \mathcal{X}} p(x) H(Y|X = x) \\ &= -\sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log_2 p(y|x) \\ &= -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x) p(y|x) \log_2 p(y|x) \\ &= -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, y) \log_2 p(y|x) \\ &= -E \log_2 p(Y|X). \end{aligned} \tag{2.68}$$

The function $H(Y|X)$ is the average uncertainty of the received symbol given that X was transmitted. Another entropy $H(X|Y)$, which is sometimes called equivocation, is defined as

$$H(X|Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \mathbf{log_2} p(x_i|y_j) \tag{2.69}$$

$H(X|Y)$ measures our average uncertainty of the transmitted symbol after a symbol has been received.

According to the chain rule in information theory, the entropy of a pair of random variables may be expressed as the sum of the entropy of one variable and the conditional entropy of the other variable.

The proof is as follows:

$$
\begin{aligned}
H(X,Y) &= -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i,y_j)\log_2 p(x_i,y_j) \\
&= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2 p(x,y) \\
&= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2 p(x)p(y|x) \\
&= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2 p(x) - \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2 p(y|x) \\
&= -\sum_{x\in\mathcal{X}} p(x)\log_2 p(x) - \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2 p(y|x) \\
&= H(X) + H(Y|X)
\end{aligned}
\tag{2.70}
$$

The above relation can be obtained by writing.

$$\log_2 p(X,Y) = \log_2 p(X) + \log_2 p(Y|X) \tag{2.71}$$

Two crucial and valuable relationships, which can be obtained directly from the definitions of the various entropies, are

$$H(X,Y) = H(X|Y) + H(Y) \tag{2.72}$$

$$H(X,Y) = H(Y|X) + H(X) \tag{2.73}$$

2.7.6 Mutual Information and Channel Capacity

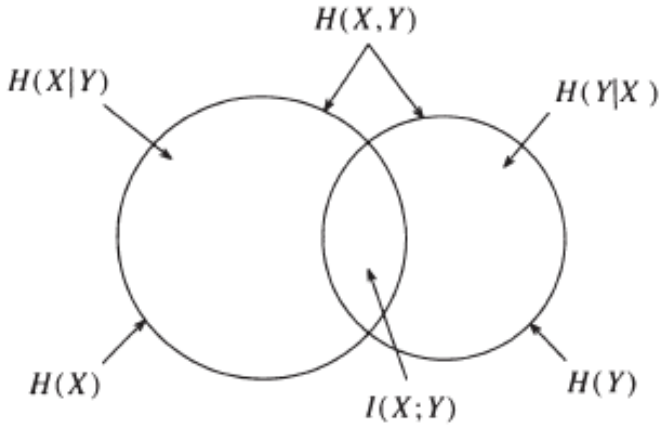

Figure 2.16: The Venn diagram depicts the additive and subtractive relationships among various information measures associated with correlated variables X and Y. Image is extracted from [35].

The concept of mutual information (MI) in probability theory and information theory is used to measure the level of dependency between two random variables. The concept under consideration measures the degree of information (expressed in units like Shannon's bits, nats, or hartleys) acquired regarding one random variable through the observation of another random variable. The notion of mutual information is inherently connected to the entropy of a random variable, a basic term within information theory that measures the anticipated quantity of information possessed by a random variable.

Claude Shannon introduced and examined the concept in question inside his influential publication titled "A Mathematical Theory of Communication" [29], but without explicitly referring to it as "mutual information." The word in issue was afterwards introduced by Robert Fano (Kreer, 1957). Mutual Information, sometimes referred to as information gain, is a concept in information theory.

The concept of mutual information (MI) pertains to quantifying the degree of information shared between variables X and Y. It captures the intrinsic interdependence present in the joint distribution of X and Y, in relation to the marginal distribution of X and Y assuming independence. This phenomenon refers to the situation when an increase in the specificity of one variable leads to a decrease in the uncertainty of the other variable, or alternatively, the uncertainty of the other variable remains unchanged. Assuming that the variables X and Y exhibit independence. In this scenario, the acquisition of knowledge pertaining to X would not yield any more insights into Y, and conversely, the mutual information between X and Y would exhibit symmetry, consistently maintaining a non-negative value of zero. Consequently, the mutual information between random variables X and Y equals zero when X and Y are independent.

On the contrary, if X or Y is deterministic to each other, then all information conveyed by one variable is shared with another one. In this case, MI is the same as the entropy of X and Y. At the channel output, an observer's average uncertainty concerning the channel input will have some value before the reception of output, and it will usually decrease when the output is received.

The relative entropy between two probability mass function $p(x)$ and $q(x)$ is defined as

$$D(p||q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)} \tag{2.74}$$

Let us consider two random variables, denoted as X and Y, which possess a joint probability mass function denoted as $P(x, y)$. Additionally, these random variables have marginal probability mass functions denoted as $P(x)$ and $P(y)$. The quantity denoted as $I(X; Y)$ represents the mutual information, which may be interpreted as the relative entropy between the joint distribution and the product distribution. The product of the functions P(x) and P(y).

$$
\begin{aligned}
I(X;Y) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)} \\
&= \sum_{x,y} p(x,y)\log_2\frac{p(x|y)}{p(x)} \\
&= -\sum_{x,y} p(x,y)\log_2 p(x) + \sum_{x,y} p(x,y)\log_2 p(x|y) \\
&= -\sum_{x} p(x)\log_2 p(x) - (-\sum_{x,y} p(x,y)\log_2 p(x|y))) \\
&= H(X) - H(X|Y)
\end{aligned}
\tag{2.75}
$$

Therefore, the decrease in the observer's average uncertainty of the transmitted signal X when the output Y is received is a measure of the average transmitted information, $I(X;Y)$.

By symmetry, it follows that

$$I(X;Y) = H(X) - H(X|Y) \quad \text{or} \quad I(X;Y) = H(Y) - H(Y|X)$$

It means the mutual information is a function of the source and channel transition probabilities.

Since $H(X,Y) = H(X) + H(Y|X)$, then

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{2.76}$$

The mutual information of a random variable X with itself is the entropy of the random variable.

$$I(X;X) = H(X) - H(X|X) = H(X) \tag{2.77}$$

Thus, entropy is sometimes referred to as self-information.

The Venn diagram presented in Figure 2.16 illustrates the additive and subtractive connections that exist between different information measures pertaining to correlated variables X and Y. The region enclosed by the two circles symbolises the collective entropy H(X, Y). The circle on the left depicts the entropy of the individual variable, denoted as H(X), while the size of the circle shows the conditional entropy H(X|Y). The circle on the right side of the diagram represents the entropy of the random variable Y, denoted as H(Y), while the size of the circle indicates the conditional entropy H(Y|X). The region of intersection symbolises the measure of mutual information between the random variables X and Y. The figure presented offers a graphical depiction of the interconnectedness of different measures of information and underscores their importance in the examination of variables that

exhibit correlation.I would like to kindly request that you rewrite my text in a more academic manner.

Mutual information can be equivalently expressed as

$$\begin{aligned} I(X;Y) &= H(X) - H(Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y) \\ &= H(X,Y) - H(X|Y) - H(Y|X) \end{aligned} \tag{2.78}$$

Imagine an observer stationed at the output of a communication channel. This observer initially holds some uncertainty regarding the information transmitted through the channel before receiving any output. Typically, this uncertainty decreases as the output is received. In mathematical terms, we express this as $H(X|Y) \leq H(X)$, indicating that the uncertainty about the input (X) diminishes as we obtain information (Y) from the channel.

We can establish this mathematically by demonstrating that

$$H(X) \geq H(X|Y) \tag{2.79}$$

This inequality holds because

$$H(X|Y) - H(X) = I(X;Y) \leq 0 \tag{2.80}$$

We can represent $-I(X;Y)$ as

$$-I(X;Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\mathbf{log}_2\frac{p(x_i)}{p(x_i|y_j)} \tag{2.81}$$

Utilizing the fact that

$$\mathbf{log}_2 x = \frac{\mathbf{ln}x}{\mathbf{ln}2} \tag{2.82}$$

and

$$\frac{p(x_i)}{p(x_i|y_j)} = \frac{p(x_i)p(y_j)}{p(x_i, y_j)} \tag{2.83}$$

We can simplify $-I(X;Y)$ to

$$-I(X;Y) = \frac{1}{\mathbf{ln}2}\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\mathbf{ln}\frac{p(x_i)p(y_j)}{p(x_i, y_j)} \tag{2.84}$$

To proceed with the derivation, we need the inequality

$$\mathbf{ln}x \leq x - 1 \tag{2.85}$$

This inequality can be proven by examining the function

$$f(x) = \mathbf{ln}(x) - (x - 1) \tag{2.86}$$

The derivative of f(x) is

$$\frac{df}{dx} = \frac{1}{x} - 1 \tag{2.87}$$

At $x = 1$, this derivative equals zero. Since $f(x)$ may be made arbitrarily tiny by selecting a high enough number for x, the maximum value of $f(x)$ is $f(1) = 0$. Then, by substituting into Eq. 2.84, we get

$$-I(X|Y) \leq \frac{1}{\mathbf{ln}2}\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)[\frac{p(x_i)p(y_j)}{p(x_i, y_j)} - 1] \tag{2.88}$$

This yields:

$$-I(X|Y) \leq \frac{1}{\mathbf{ln}2}[\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i)p(y_j) - \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)] \tag{2.89}$$

Since both double sums equal 1, we arrive at the desired result:

$$-I(X|Y) \leq 0 \tag{2.90}$$

Therefore, we have shown that mutual information is always positive, implying that $H(X) = H(X|Y)$.
The greatest mutual information, or the highest average information per symbol that can be conveyed efficiently across the channel, is symbolised by the symbol C, which stands for "channel capacity:

$$C = \mathbf{max}[I(X;Y)] \tag{2.91}$$

The capacity is the maximum rate at which information can be transmitted over the channel with an infinitesimally small probability of error at the output.
Example
A noiseless binary channel is one in which the output perfectly replicates the input binary value, allowing for the error-free reception of any sent bit. Thus, the capacity is 1 bit per transmission, and 1 bit may be conveyed successfully to the receiver if

the formula $C = \max \quad I(X;Y) = 1$ bit is used.

Figure 2.17: The diagram shows the noiseless binary channel that the binary input is reproduced exactly at the output. Therefore, C = 1 bit.

Example
Let us now consider that a channel of each input letter is received either as the same letter with probability $\frac{1}{2}$ or as the following letter with probability $\frac{1}{2}$. If all 4 input symbols are used, inspecting the output would not reveal which input symbol was sent with certainty. If, on the other hand, only inputs 1 and 3 are used, the input from the output can be determined immediately (if 1 or 2 is received, then the input must be 1). Since the channel is similar to the previous example, 1 bit per transmission over the channel without errors can be sent.

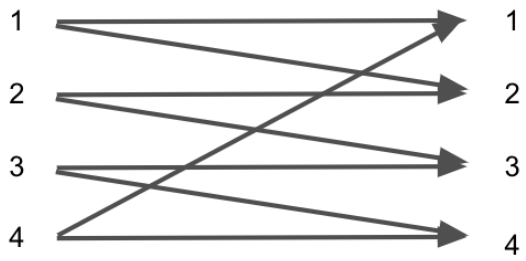

Figure 2.18: The diagram shows the noisy 4-symbol channel that acts like the previous example. Therefore, C = 1 bit.

A subset of the inputs can only sometimes be identified for a more complicated structure to send information without error. Instead, a subset of the input sequences (the codewords) can be identified to transmit information over the channel so that the sets of possible output sequences associated with each codeword are approximately disjoint.

Example

A binary symmetric channel has a binary input; its output equals the input with probability $1-p$. With probability p, on the other hand, a 0 is received as a 1, and vice versa. In this case, the channel's capacity can be $C = 1+p\log p+(1-p)\log(1-p)$ bits per transmission. Channel works similarly to the previous example, and then information can be sent at a rate of C bits per transmission with an arbitrarily low error probability.

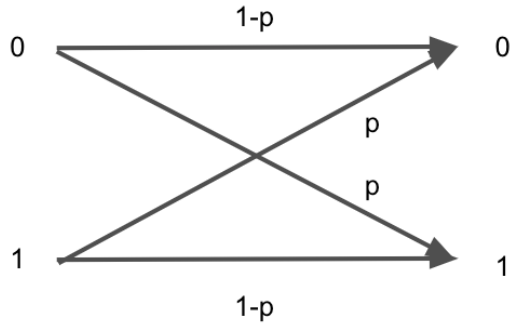

Figure 2.19: The diagram shows the binary symmetric channel that the input and output are the same with probability $1-p$. Otherwise, the probability would be p.

The maximization concerns the source probabilities since the channel fixes the transition probabilities. However, the channel capacity is a function of only the channel transition probabilities since the maximization process eliminates the dependence on the source probabilities.

The source probabilities are maximized, while the channel transition probabilities are fixed since they impact the channel capacity. However, the channel capacity only depends on the channel transition probabilities, as the maximization process removes the impact of the source probabilities.

Let us work on some examples to learn how to calculate relevant values of the channel information.

Example

Suppose the source probability $[P(X)]$ and the channel transition probability $[P(Y|X)]$, the output probability $[P(Y)]$ can be obtained

$$[P(Y)] = [P(X)][P(Y|X)] = \begin{bmatrix} 0.3 & 0.4 & 0.3 \end{bmatrix} \times \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 1 & 0 \\ 0 & 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.24 & 0.55 & 0.21 \end{bmatrix} \tag{2.92}$$

If $[P(X)]$ is written as a diagonal matrix, a joint probability matrix $[P(X,Y)]$ can be yielded.

$$P[(X,Y)] = [P(X)][P(Y|X)] = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.3 \end{bmatrix} \times \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 1 & 0 \\ 0 & 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.24 & 0.06 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0.09 & 0.21 \end{bmatrix} \quad (2.93)$$

Also, [P(Y|X)] can be calculated

$$P[(X|Y)] = \frac{[P(X,Y)]}{[P(Y)]} = \begin{bmatrix} \frac{0.24}{0.24} & \frac{0.06}{0.55} & 0 \\ 0 & \frac{0.4}{0.55} & 0 \\ 0 & \frac{0.09}{0.55} & \frac{0.21}{0.21} \end{bmatrix} = \begin{bmatrix} 1 & 0.109 & 0 \\ 0 & 0.727 & 0 \\ 0 & 0.164 & 1 \end{bmatrix} \quad (2.94)$$

After that, the corresponding $H(X)$, $H(Y)$ can be obtained, $H(X,Y)$, $H(X|Y)$, $H(Y|X)$ as follow

$$\begin{aligned} H(X) &= -\sum_{j=1}^{3} p(x_j) \log_2 p(x_j) \\ &= -0.3 \log_2 0.3 - 0.4 \log_2 0.4 - 0.3 \log_2 0.3 \\ &= 1.57 \end{aligned} \quad (2.95)$$

$$\begin{aligned} H(Y) &= -\sum_{k=1}^{3} p(y_k) \log_2 p(y_k) \\ &= -(0.24 \log_2 0.24 + 0.55 \log_2 0.55 + 0.21 \log_2 0.21) \\ &= 1.44 \end{aligned} \quad (2.96)$$

$$\begin{aligned} H(X,Y) &= -\sum_{j=1}^{3} \sum_{k=1}^{3} p(x_i, y_j) \log_2 p(x_i, y_j) \\ &= -(0.24 \log_2 0.24 + 0.06 \log_2 0.06 + 0.4 \log_2 0.4 + 0.09 \log_2 0.09 + 0.21 \log_2 0.21) \\ &= 2.05. \end{aligned} \quad (2.97)$$

$$H(X|Y) = -(0.24 \log_2 1 + 0.06 \log_2 0.109 + 0.4 \log_2 0.727 + 0.09 \log_2 0.164 + 0.21 \log_2 1) = 0.612 \quad (2.98)$$

$$H(Y|X) = -(0.24 \log_2 0.8 + 0.06 \log_2 0.2 + 0.4 \log_2 1 + 0.09 \log_2 0.3 + 0.21 \log_2 0.7) = 0.48 \quad (2.99)$$

Therefore, the mutual information would be

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= 1.57 - 0.612 = 0.96 \\ &= H(Y) - H(Y|X) \\ &= 1.44 - 0.48 = 0.96. \end{aligned} \tag{2.100}$$

The mutual information can be calculated from the joint probability matrix.

Example

Let (X, Y) have the following joint distribution

$$P[(X,Y)] = \begin{bmatrix} \frac{1}{8} & \frac{1}{16} & \frac{1}{32} & \frac{1}{32} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{32} & \frac{1}{32} \\ \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} \\ \frac{1}{4} & 0 & 0 & 0 \end{bmatrix}. \tag{2.101}$$

[P(X)] can be obtained by summing all the elements of $[P(X,Y)]$ in each column.

$$\begin{aligned} p(x_1) &= \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} = 0.5, \\ p(x_2) &= \frac{1}{16} + \frac{1}{8} + \frac{1}{16} + 0 = 0.25, \\ p(x_3) &= \frac{1}{32} + \frac{1}{32} + \frac{1}{16} + 0 = 0.125, \\ p(x_4) &= \frac{1}{32} + \frac{1}{32} + \frac{1}{16} + 0 = 0.125, \end{aligned} \tag{2.102}$$

Therefore,

$$[P(X)] = \begin{bmatrix} 0.5 & 0.25 & 0.125 & 0.125 \end{bmatrix}. \tag{2.103}$$

Similarly, [P(Y)] can be obtained as follow

$$\begin{aligned} p(y_1) &= \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{32} = 0.25 \\ p(y_2) &= \frac{1}{16} + \frac{1}{8} + \frac{1}{32} + \frac{1}{32} = 0.25 \\ p(y_3) &= \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = 0.25 \\ p(y_4) &= \frac{1}{4} = 0.25 \end{aligned} \tag{2.104}$$

$$[P(Y)] = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix} \tag{2.105}$$

Then, $[P(X|Y)]$ and $[P(Y|X)]$ can be obtained

$$P[(X|Y)] = \frac{P(X,Y)}{P(Y)} = \begin{bmatrix} \frac{0.125}{0.25} & \frac{0.0625}{0.25} & \frac{0.03125}{0.25} & \frac{0.03125}{0.25} \\ \frac{0.0625}{0.25} & \frac{0.125}{0.25} & \frac{0.03125}{0.25} & \frac{0.03125}{0.25} \\ \frac{0.0625}{0.25} & \frac{0.0625}{0.25} & \frac{0.0625}{0.25} & \frac{0.0625}{0.25} \\ \frac{0.25}{0.25} & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.25 & 0.125 & 0.125 \\ 0.25 & 0.5 & 0.125 & 0.125 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{2.106}$$

$$P[(Y|X)] = \frac{P(X,Y)}{P(X)} = \begin{bmatrix} \frac{0.125}{0.5} & \frac{0.0625}{0.25} & \frac{0.03125}{0.125} & \frac{0.03125}{0.125} \\ \frac{0.0625}{0.5} & \frac{0.125}{0.25} & \frac{0.03125}{0.125} & \frac{0.03125}{0.125} \\ \frac{0.0625}{0.5} & \frac{0.0625}{0.25} & \frac{0.0625}{0.125} & \frac{0.0625}{0.125} \\ \frac{0.25}{0.5} & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.125 & 1 & 0.25 & 0.25 \\ 0.125 & 0.25 & 0.25 & 0.25 \\ 0.5 & 0 & 0 & 0 \end{bmatrix}$$

(2.107)

After that, the corresponding $H(X)$, $H(Y)$, $H(X,Y)$, $H(X|Y)$, $H(Y|X)$ can be obtained as follow

$$H(X) = -(0.5\log_2 0.5 + 0.25\log_2 0.25 + 0.125\log_2 0.125 + 0.125\log_2 0.125) = 1.75 \tag{2.108}$$

$$H(Y) = -(0.25\log_2 0.25 + 0.25\log_2 0.25 + 0.25\log_2 0.25 + 0.25\log_2 0.25) = 2 \tag{2.109}$$

$$\begin{aligned} H(X|Y) = -\,(&2 \times 0.125\log_2 0.5 + 6 \times 0.0625\log_2 0.25 \\ &+ 4 \times 0.03125\log_2 0.125 + 0.25\log_2 1 + 3 \times 0) = 1.375 \end{aligned} \tag{2.110}$$

Therefore, the mutual information is

$$I(X;Y) = H(X) - H(X|Y) = 1.75 - 1.375 = 0.375 \tag{2.111}$$

Let us now consider bit flip during channel transmission by the on-off-keying (OOK) protocol. On-off keying (OOK) is a digital modulation technique widely used in optical communication systems. In OOK, digital data is transmitted by varying the intensity of an optical signal, with the signal being switched on and off to represent 1s and 0s, respectively. This digital representation of the presence or absence of carrier wave denotes the simplest form of amplitude-shift keying (ASK) modulation.

In OOK, the transmitter encodes the digital data into a series of electrical pulses, which are then used to control the on/off state of the optical signal. When the signal is on, it has a constant intensity representing a binary 1. When the signal is off, there is no detectable signal, which represents a binary 0. The receiver detects the optical signal and uses a threshold detector to determine whether the signal is on or off, and hence decodes the transmitted data.

OOK has several advantages over other types of modulation techniques. It is relatively simple and easy to implement, making it a popular choice for low-cost and low-power optical communication systems. Additionally, OOK is compatible with a wide range of optical sources and detectors, which makes it a versatile modulation technique.

However, OOK also has some limitations. It is less efficient than other modulation techniques, such as phase-shift keying (PSK), which can transmit multiple bits per symbol. Although OOK is more spectrally efficient than frequency-shift keying, it is more susceptible to noise and interference when using a regenerative receiver or a poorly implemented superheterodyne receiver [3], which can degrade the signal-to-noise ratio (SNR) and reduce the reliability of the transmission.

Sending a single photon with a 2-time bin and two photons with 4-time bins, respectively, which would display as follow:

Example

For $M = 1$, assume $P(X) = \frac{1}{2}$ and $H(X) = 1$

$$P[(Y|X)] = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \tag{2.112}$$

For $M = 2$, assume $P(X) = \frac{1}{4}$, $H(X) = 2$

$$P[(Y|X)] = \begin{bmatrix} 0.81 & 0.09 & 0.09 & 0.01 \\ 0.09 & 0.81 & 0.01 & 0.09 \\ 0.09 & 0.01 & 0.81 & 0.09 \\ 0.01 & 0.09 & 0.09 & 0.81 \end{bmatrix} \tag{2.113}$$

The probability matrix will look different if bit loss is considered during transmission. Assume that the loss probability is 0.1, and K is the number of codewords in the symbol using the OOK protocol.

Example

For $M = 1$, $K = 2^M = 2^1 = 2$, $p = 0.1$,

$$P[(X,Y)] \times \frac{1}{K} = \begin{bmatrix} 1 & 0 \\ p & 1-p \end{bmatrix} \times \frac{1}{2^M} = \begin{bmatrix} 1 & 0 \\ 0.9 & 1-0.9 \end{bmatrix} \times \frac{1}{2^1} = \begin{bmatrix} 0.5 & 0 \\ 0.05 & 0.45 \end{bmatrix} \tag{2.114}$$

Respectively, $[P(X)]$, $[P(Y)]$, $[P(X|Y)]$, $[P(Y|X)]$ can be obtained from [P(X,Y)]

$$P[(X)] = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \tag{2.115}$$

$$P[(Y)] = \begin{bmatrix} 0.55 & 0.45 \end{bmatrix} \tag{2.116}$$

$$P[(X|Y)] = \frac{P(X,Y)}{P(Y)} = \begin{bmatrix} \frac{0.5}{0.55} & \frac{0}{0.45} \\ \frac{0.05}{0.55} & \frac{0.45}{0.45} \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.09 & 1 \end{bmatrix} \tag{2.117}$$

$$P[(Y|X)] = \frac{P(X,Y)}{P(X)} = \begin{bmatrix} \frac{0.5}{0.5} & \frac{0}{0.5} \\ \frac{0.05}{0.5} & \frac{0.45}{0.5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.1 & 0.9 \end{bmatrix} \tag{2.118}$$

Then, $H(X)$, $H(Y)$, $H(X,Y)$, $H(X|Y)$ is moved on to be calculated

$$H(X) = -(0.5 \log_2 0.5 + 0.5 \log_2 0.5) = 1 \tag{2.119}$$

and

$$H(Y) = -(0.55 \log_2 0.55 + 0.45 \log_2 0.45 = 0.99 \tag{2.120}$$

$$H(X,Y) = -(0.5 \log_2 0.5 + 0 + 0.05 \log_2 0.05 + 0.45 \log_2 0.45) = 1.23 \tag{2.121}$$

$$H(X|Y) = -(0.5 \log_2 0.9 + 0 + 0.05 \log_2 0.09 + 0.45 \log_2 1) = 0.24 \tag{2.122}$$

$$H(Y|X) = -(0.5 \log_2 1 + 0 + 0.05 \log_2 0.1 + 0.45 \log_2 0.9 = 0.23 \tag{2.123}$$

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= 1 - 0.24 = 0.76 \\ &= H(Y) - H(Y|X) \\ &= 0.99 - 0.23 = 0.76 \end{aligned} \tag{2.124}$$

The mutual information per photon is

$$\frac{MI}{\frac{M}{2}} = \frac{0.76}{\frac{1}{2}} = 1.52. \tag{2.125}$$

Certainly, it is known that half-photons do not exist; the value means that there is a half photon on average within a 1-time bin of transmission. Similarly, the mutual information for 2 time bins can be calculated.

For $M = 2$, $K = 2^M = 4$, $p = 0.1$

$$\begin{aligned} P[(X,Y)] \times \frac{1}{2^M} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ P & (1-P)(1) & 0 & 0 \\ P & 0 & (1-P)(1) & 0 \\ P^2 & (1-P)P & (1-P) & (1-P)^2 \end{bmatrix} \times \frac{1}{2^2} \\ &= \begin{bmatrix} 0.25 & 0 & 0 & 0 \\ 0.025 & 0.225 & 0 & 0 \\ 0.025 & 0 & 0.225 & 0 \\ 0.0025 & 0.0225 & 0.0225 & 0.2025 \end{bmatrix} \end{aligned} \tag{2.126}$$

Respectively, $[P(X)]$, $[P(Y)]$, $[P(X|Y)]$, $[P(Y|X)]$ can be obtained from [P(X,Y)]

$$P[(X)] = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix} \tag{2.127}$$

$$P[(Y)] = \begin{bmatrix} 0.3025 & 0.2475 & 0.2475 & 0.2025 \end{bmatrix} \tag{2.128}$$

$$P[(X|Y)] = \frac{P(X,Y)}{P(Y)} = \begin{bmatrix} \frac{0.25}{0.3025} & \frac{0}{0.2475} & \frac{0}{0.2475} & \frac{0}{0.2025} \\ \frac{0.025}{0.3025} & \frac{0.225}{0.2475} & \frac{0}{0.2475} & \frac{0}{0.2025} \\ \frac{0.025}{0.3025} & \frac{0}{0.2475} & \frac{0.225}{0.2475} & \frac{0}{0.2025} \\ \frac{0.0025}{0.3025} & \frac{0.0225}{0.2475} & \frac{0.0225}{0.2475} & \frac{0.2025}{0.2025} \end{bmatrix} = \begin{bmatrix} 0.83 & 0 & 0 & 0 \\ 0.082 & 0.91 & 0 & 0 \\ 0.082 & 0 & 0 & 0 \\ 0.0082 & 0.091 & 0.091 & 1 \end{bmatrix} \tag{2.129}$$

$$P[(Y|X)] = \frac{P(X,Y)}{P(X)} = \begin{bmatrix} \frac{0.25}{0.25} & \frac{0}{0.25} & \frac{0}{0.25} & \frac{0}{0.25} \\ \frac{0.025}{0.25} & \frac{0.225}{0.25} & \frac{0}{0.25} & \frac{0}{0.25} \\ \frac{0.025}{0.25} & \frac{0}{0.25} & \frac{0.225}{0.25} & \frac{0}{0.25} \\ \frac{0.0025}{0.25} & \frac{0.0225}{0.25} & \frac{0.0225}{0.25} & \frac{0.2025}{0.25} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.1 & 0.9 & 0 & 0 \\ 0.1 & 0 & 0.9 & 0 \\ 0.1 & 0.09 & 0.09 & 0.81 \end{bmatrix} \tag{2.130}$$

$$H(X) = -4 \times (0.25 \times \log_2 0.25) = 2 \tag{2.131}$$

$$H(Y) = -(0.3025 \log_2 0.3025 + 2 \times 0.2475 \log_2 0.2475 + 0.2025 \log_2 0.2025 = 1.9855 \tag{2.132}$$

$$\begin{aligned} H(X,Y) = &- (0.25 \log_2 0.25 + 2 \times 0.025 \log_2 0.025 \\ &+ 2 \times 0.225 \log_2 0.225 + 0.0025 \log_2 0.0025) \\ &+ 2 \times 0.0225 \log_2 0.0225) + 0.2025 \log_2 0.2025) \\ =&2.469 \end{aligned} \tag{2.133}$$

$$\begin{aligned} H(X|Y) = &- (0.25\log_2 0.83 + 2\times 0.025\log_2 0.082 + 2\times 0.225\log_2 0.91 + \\ &+ 0.0025\log_2 0.0082 + 2\times 0.0225\log_2 0.091 + 0.2025\log_2 1) \\ =&0.48 \end{aligned} \tag{2.134}$$

$$\begin{aligned} H(Y|X) = &- (0.25\log_2 1 + 2\times 0.025\log_2 0.1 + 2\times 0.225\log_2 0.9 \\ &+ 0.0025\log_2 0.1 + 0.0225\log_2 0.09 + 0.0225\log_2 0.09 + 0.2025\log_2 0.81) \\ =&0.46 \end{aligned} \tag{2.135}$$

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= 2 - 0.48 = 1.52 \\ &= H(Y) - H(Y|X) \\ &= 1.9855 - 0.46 = 1.5255 \end{aligned} \tag{2.136}$$

The mutual information per photon is

$$\frac{MI}{\frac{M}{2}} = \frac{1.52}{\frac{2}{2}} = 1.52 \tag{2.137}$$

The above calculation shows that the mutual information per photon in the OOK protocol does not depend on the number of time bins, which will help the analysis in a later chapter.

The following are many characteristics pertaining to the joint entropy of two random variables, X and Y.

$$\begin{aligned} &H(X,Y) \geq 0 \\ &H(X,Y) \geq \max[H(X), H(Y)] \\ &H(X,Y) \leq H(X) + H(Y) \\ &H(X|Y) = H(X,Y) - H(Y) \\ &I(X;Y) = H(X) + H(Y) - H(X,Y) \end{aligned} \tag{2.138}$$

The Conditional Entropy and Mutual Information subsection will introduce the last two equations.

2.8 Comparative Metrics on Information Bits

Information Bit (IB), commonly referred to as IB, are a measure of the data carried by a particular entity, such as a symbol, photon, or time bin, during transmission

inside a communication system. Bits are units of measurement that quantify the amount of information present in a message, symbol, or signal. They serve the purpose of evaluating the effectiveness of data transfer and the capabilities of communication systems.

2.8.1 Information Bits per Symbol

Bits per symbol is an essential metric for measuring the capacity of a communication system in digital communication. It measures the information density of each symbol delivered across a given channel. This statistic gives a thorough assessment of the costs and benefits of varying data rate, error rate, and bandwidth utilisation.

various modulation techniques encode various numbers of bits per symbol, hence their use must be taken into account when calculating the number of bits per symbol. By using the representation of 16 unique signal states ($24 = 16$), a 16-QAM modulation method may encode up to 4 bits per symbol. In contrast, the 64-QAM modulation system uses 64 different signal states ($26 = 64$) to encode up to 6 bits per symbol, allowing for a greater possible data rate.

However, it is essential to understand that symbols can be encoded as error-correcting codewords. Mathematical methods are at the heart of these error-correcting codes, transforming symbols into redundantly longer codewords. Due to this redundancy, faults in transmission can be detected and, in certain cases, corrected at the receiving end.

When constructing communication systems, it is essential to take bits per symbol into account to ensure efficient data transfer. More bits per symbol may be sent with higher-order modulation methods, allowing for faster data rates. Although higher-order modulation systems may have better signal quality and more complex receiver designs, they may also be more vulnerable to noise and other impairments.

Evaluation and optimisation of communication system capacity are aided by familiarity of the symbol-bit relationship, the consequences of different modulation methods, and the possibility of mistake correction using error-correcting codes. The trade-offs between data rate, error rate, and bandwidth utilisation can help you choose the best modulation scheme for your application. This guarantees that digital communication systems are able to transmit data effectively and reliably.

2.8.2 Information Bits per Photon

In the realm of quantum communication systems, the concept of bits per photon serves as a fundamental metric for assessing the information capacity of these systems. It quantifies the amount of information a single photon conveys, thus playing a crucial role in evaluating trade-offs between data rate, error rate, and photon effi-

ciency. This concept is unique to quantum communication, where the distinct properties of individual photons are harnessed for information transmission. In contrast, in classical communication systems designed for energy-efficient communication, the primary focus revolves around data rate, often measured in bits per second. Classical systems typically involve the collective behavior of many photons or electrical signals to transmit information efficiently. Therefore, while bits per photon are a critical parameter in quantum communication, it does not have a direct counterpart in classical energy-efficient communication systems.

The determination of bits per photon hinges upon the specific quantum state encoding scheme utilized, as different schemes can encode varying numbers of bits per photon. For instance, a polarization encoding scheme can encode up to 1 bit of information per photon, utilizing any 2 orthogonal (distinct) polarization states available, e.g., the states $\{ |H\rangle, |V\rangle\}$. On the other hand, a time-bin encoding scheme can achieve a higher information capacity of up to 2 bits per photon, utilizing the distinguishable time-bin states (time bins can be empty).

In Figure 2.4, there are four degenerate polarization states on the Poincare's Sphere. From Equation 2.2, the states can be expressed as

$$\begin{aligned} |H\rangle &= |0\rangle \text{ for } \quad \theta = 0 \quad \text{and} \quad \phi = 0 \\ |V\rangle &= |1\rangle \text{ for } \quad \theta = \pi \quad \text{and} \quad \phi = 0 \\ |A\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \text{for } \quad \theta = \frac{\pi}{2} \quad \text{and} \quad \phi = 0 \\ |D\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}} \text{for } \quad \theta = \frac{3\pi}{2} \quad \text{and} \quad \phi = 0 \end{aligned} \tag{2.139}$$

Specific quantum states are experimentally obtained using polarizers by employing a setup that typically includes a source of polarized photons and one or more polarizing elements. A commonly used source is a laser that emits polarized light, often linearly polarized. The polarizing elements, such as polarizer sheets or beam splitters, are strategically placed within the experimental apparatus.

A linear polarizer aligned horizontally can be used to create the state $|H\rangle$, corresponding to horizontal polarization. This polarizer allows photons with horizontal polarization (aligned with its transmission axis) to pass through while blocking vertically polarized photons.

For the state $|V\rangle$, representing vertical polarization, a similar approach is applied with a vertically aligned linear polarizer.

States like $|A\rangle$ (right-circular polarization) and $|D\rangle$ (left-circular polarization) can be achieved by using quarter-wave plates and polarizers. A quarter-wave plate can transform linear polarization into circular polarization. For example, to obtain $|A\rangle$, a quarter-wave plate followed by a linear polarizer oriented at 45 degrees can be

used. This setup transforms linearly polarized light into right-circular polarization.

Bits per photon is a significant metric in assessing the capacity of quantum communication systems and understanding the efficiency of individual photon utilization for transmitting information. Unlike traditional communication systems that rely on the intensity or amplitude of electromagnetic waves, quantum communication leverages the principles of quantum mechanics to encode and transmit information using individual photons. Information is encoded in various quantum states of a single photon, such as Polarization, phase, or path, where the number of distinguishable states determines the information capacity per photon.

The efficiency and performance of quantum communication systems depend on the specific quantum encoding and detection techniques and the noise and loss characteristics inherent in the system. Informed decisions regarding quantum state encoding schemes and optimizing of information capacity per photon in quantum communication systems can be achieved by evaluating the trade-offs between data rate, error rate, and photon efficiency.

2.8.3 Information Bits per Time Bin

In communication systems, information bits per time bin is a critical metric for quantifying the amount of information conveyed within a specific time duration, known as a time bin. A time bin represents a discrete interval of time utilized for signal processing and transmission. This metric plays a pivotal role in assessing a system's data rate and processing capacity within the confines of the designated time interval while also factoring in power efficiency considerations.

The value of information bits per time bin depends on various factors, including the employed modulation scheme, coding techniques, channel conditions, and system constraints. Higher-order modulation schemes and advanced coding techniques can potentially augment the number of information bits effectively transmitted or processed within a time bin, thus yielding a higher data rate and optimizing power usage.

Comprehending information bits per time bin is instrumental in understanding the capabilities and limitations of a communication system, including its power efficiency. Informed decisions regarding system design, aiming to maximize the utilization of each time bin and enhance overall data transmission efficiency while minimizing power consumption, can be made by analyzing and optimizing the modulation scheme, coding methods, and transmission parameters. Power-efficient design is critical to ensuring the sustainability and effectiveness of modern communication systems.

2.8.4 Information Bits per Photon Versus Number of Time Bin

Information bits per photon versus the number of time bins metric measures the information efficiency of a communication system, defined as the relationship between the number of information bits that can be transmitted per photon and the temporal resolution of the communication system. A higher information bit per photon ratio indicates more efficient utilization of individual photons for carrying information that enables higher data rates and more efficient utilization of individual photons, while increasing the number of time bins allows for finer temporal resolution and more precise timing control in the system, providing advantages such as improved timing precision, reduced interference, or enhanced synchronization.

For PPM, the metric of bits per time bin versus the number of photons can be used to evaluate the efficiency of the modulation scheme by calculating the number of bits that can be transmitted per pulse position or per unit of time, which would depend on the pulse width, pulse position resolution, and the available photon flux. In PPM, information is encoded in the position of the pulse within a time bin. The metric of bits per time bin versus photon can be utilized to assess the efficiency of the modulation scheme by determining the number of bits that can be transmitted per pulse position or per unit of time. The efficiency depends on parameters such as the pulse width (duration of each pulse), pulse position resolution (number of distinguishable positions within a time bin), and the available photon flux (number of photons available for transmission). A higher number of bits per time bin or per photon indicates higher efficiency in utilizing the modulation scheme.

Similarly, for OOK, the metric of bits per time bin versus the number of photons can be used to evaluate the efficiency of the modulation scheme by calculating the number of bits that can be transmitted per unit of time or photon. In OOK, information is encoded by turning the carrier signal on or off. The metric of bits per time bin versus photon can be employed to evaluate the efficiency of the modulation scheme by calculating the number of bits that can be transmitted per unit of time or photon. The efficiency is influenced by factors such as the modulation index (ratio of the signal's peak amplitude to the noise level), the bandwidth of the signal, and the available photon flux. A higher number of bits per time bin or per photon indicates greater efficiency in utilizing the modulation scheme.

2.8.5 Information Bits per Time Bin Versus Number of Time Bin

Information bits per time bin versus the number of time bins is a measure of the information efficiency of a communication system, defined as the number of information bits that can be transmitted using a fixed amount of time bins over a given period duration, known as the trade-off between the data rate and the temporal resolution of the system. The relationship between bits per time bin and the num-

ber of time bins is influenced by several factors, including the modulation scheme, coding techniques, channel conditions, and system constraints. Increasing the number of time bins can provide advantages such as improved timing precision, reduced interference, or enhanced synchronization. It allows for more precise control over the temporal aspects of signal processing or data transmission. On the other hand, As the number of time bins increases, the time available for encoding or transmitting information within each bin decreases, potentially reducing the data rate or processing capacity.

2.8.6 Information Bits per Time Bin Versus Number of Photon

Information bits per time bin versus the number of photons depicts the comparison between the number of information bits transmitted within a time bin and the total number of photons used. The relationship between information bits per time bin and the number of photons is influenced by several factors, including the quantum encoding and detection techniques employed, noise characteristics, system constraints, and the desired error rate. Increasing the number of photons can provide advantages such as improved signal-to-noise ratio, enhanced detection sensitivity, or higher transmission reliability. It can result in a more robust communication system with better performance in the presence of noise or channel impairments. On the other hand, increasing the number of photons may linearly increase the information bits per time bin. There can be diminishing returns or practical limits to the achievable information bits per time bin. This limitation arises due to factors such as the quantum encoding scheme, the signal-to-noise ratio, and the error rate requirements.

2.9 Comparative Metrics on Mutual Information

Mutual information, denoted as MI, measures the information shared between two random variables in a probabilistic system. It quantifies how much knowing one variable reduces uncertainty about the other. MI is concerned with the statistical relationship in bits between two random variables and how their information content is related. It is used to analyze and optimize communication channels, assess the performance of error-correcting codes, and understand the capacity of noisy channels.

2.9.1 MI per Photon Versus Number of Time Bin

MI per photon versus the number of time bins measures the mutual information per photon as a function of the number of time bins used in the simulation. This metric enables us to investigate how the information content of the photon changes as a function of the transmission time and determine the optimal number of time bins to use for maximum information transfer. MI per photon quantifies the shared

information between the input (transmitted) and output (received) signals, considering the noise and uncertainty present in the communication channel. However, the mutual information per photon may not necessarily increase with an increase in the number of time bins. The achievable mutual information depends on the system design, coding scheme, and the characteristics of the transmission medium.

2.9.2 MI per Time Bin Versus Number of Time Bin

Another important metric is the MI Per Time Bin Versus the Number of Time Bins for evaluating a single-photon's performance within a time bin and the temporal resolution of the telecommunication systems. The mutual information per time bin may not necessarily increase with an increase in the number of time bins. The achievable mutual information depends on the system design, coding scheme, noise characteristics, and channel capacity.

2.9.3 MI per Photon times Time Bin Versus Number of Time Bin

Mutual Information Per Photon times Time Bin Versus Number of Time Bin is a powerful metric for evaluating the performance of single-photon telecommunication systems in simulation. This metric measures the mutual information per photon times time bin as a function of the number of time bins used in the simulation.

3

Research Methodology

3.1 Introduction

The master's thesis aims to investigate the information capacity of a single photon through polarization and time ordering that facilitates energy savings when available power is limited and to develop error correction techniques that can improve the reliability of photon-based communication over longer distances.

3.2 Research Design

The research design of the patented protocol - Beyond Pulse Position Modulation (BPPM) needs to be demonstrated to allow for message transmission at a much lower total energy by comparing the mutual information quality based on various metrics as follows:

1. Information Bits per Symbol
2. Information Bits per Photon
3. Information Bits per Time Bin
4. Information Bits per Photon Versus Number of Photons
5. Information Bits per Photon Versus Number of Time Bin
6. Information Bits per Time Bin Versus Number of Time Bin
7. Information Bits per Time Bin Versus Number of Photon

Apart from the metrics mentioned above, Mutual Information (MI) for constant power and constant energy per information bit are measured to compare the effectiveness of BPPM, PPM, and a general protocol, i.e., a protocol where a fixed number of photons is arbitrarily placed among M time bins.

1. MI per Photon Versus Number of Time Bin
2. MI per Time Bin Versus Number of Time Bin

In the metrics mentioned earlier, the mutual information is normalized to measure communication efficiency and compare different protocols based on this measure.

3.3 Description of the Simulation

Our simulations use Python programming language, specifically version 3.8.5, in conjunction with Visual Studio Code IDE, version 1.76.2. The simulation software was developed using the NumPy library of version 1.23.5 for numerical computation and the matplotlib library of version 3.7.1 for visualization. Several assumptions are made in single-photon telecommunication simulation:

1. Ideal sources: The photons are emitted by ideal sources that have a perfect emission rate and polarization.

2. Perfect detectors: The detectors used to measure the photons have perfect efficiency and can detect every photon that arrives at them.

3. Lossless transmission medium: The photons travel through a near-lossless medium with minimal attenuation or distortion. This assumption considers that real-world transmission media may have some imperfections, but they are negligible for this simulation.

4. Single-mode propagation: The photons propagate in a single mode, meaning they do not interact with other modes or experience modal dispersion.

5. Distance travel and error probability: The error probability in the simulation is modeled as a function of the distance traveled by the photon using an attenuation model that considers the absorption, scattering, and other factors contributing to the reduction in signal strength.

6. Non-interacting photons: The photons do not interact with each other during transmission or detection.

7. Low photon flux: The photon flux is low enough that multiple-photon events can be neglected.

3.4 Conclusion

This research methodology section outlines the planned approach to simulate single-photon telecommunications. Firstly, a theoretical model of the system will be developed, taking into account the assumptions made about the sources, detectors, transmission medium, and other factors that affect the behavior of the photons. Python programming language will be used to implement the simulation, using libraries such as NumPy and Matplotlib for numerical computation and data visualization.

In the following research stage, a series of simulations to evaluate the performance of

the single-photon telecommunication system under different conditions. Parameters will be varied such as the error probability, the number of photons, and the number of time bins, and measure the resulting mutual information in the above relevant metrics for comparison of the effectiveness over BPPM, PPM, General Protocol, and OOK.

Finally, the results of our simulations will be analyzed and conclude the feasibility and potential applications of single photon telecommunications. Our research is expected to contribute to developing more efficient and secure communication systems, with important implications for fields such as satellite telecommunication, quantum computing, cryptography, and telecommunications engineering.

4

Results and Discussion

In this section, various metrics are measured and compared to evaluate the performance of different transmission protocols. These metrics include the number of permutations, information content per symbol, information content per photon, and information content per time bin. The transmission protocols under consideration are BPPM, PPM, OOK, and a general protocol (as described in Section 2.4). Our study will involve conducting comparisons between the quantities of photons and time bins. Specifically, we will analyse the relationship between the number of photons and the number of photons, the number of photons and the number of time bins, the number of time bins and the number of photons, and the number of time bins and the number of time bins. These studies will provide a comprehensive evaluation of the performance of these treatments.

According to Table 4.1,a notable relationship emerges within BPPM protocol between the number of photons (column 1) and the number of time bins (column 2). As the number of photons increases, there is a progressive growth in the number of time bins, which, in turn, leads to exponential permutations (column 3), This observation underscores the remarkable scalability and potential inherent in BPPM protocol.

Moreover, the investigation into other metrics, such as bits per symbol (column 4), bits per photon (column 5), and bits per time bin (column 6), reveals a consistent and proportional increment as the number of photons increases. These findings underscore BPPM protocol's remarkable consistency in transmitting information efficiently and effectively.

Table 4.1: Table depicts BPPM in terms of the metrics permutation, bits per symbol, bits per photon, bits per time bin up to photon number n = 20.

Number of Photon	Time Bin	Permutation	Bits/Symbol	Bits/Photon	Bits/Time Bin
1	1	1	0.000000	0.000000	0.000000
2	3	2	1.000000	0.500000	0.333333
3	7	6	2.584963	0.861654	0.369280
4	14	24	4.584963	1.146241	0.327497
5	26	120	6.906891	1.381378	0.265650
6	46	720	9.491853	1.581976	0.206345
7	75	5,400	12.299208	1.757030	0.163989
8	113	40,320	15.299208	1.912401	0.135391
9	165	362,880	18.469133	2.052126	0.111934
10	238	3,628,800	21.791061	2.179106	0.091559
11	332	3.991680e+07	25.250493	2.295499	0.076056
12	459	4.790016e+08	28.835455	2.402955	0.062822
13	610	6.227021e+09	32.535895	2.502761	0.053338
14	791	8.717829e+10	36.343250	2.595946	0.045946
15	1002	1.307674e+12	40.250140	2.683343	0.040170
16	1259	2.092279e+13	44.250140	2.765634	0.035147
17	1574	3.556874e+14	48.337603	2.843388	0.030710
18	1947	6.402374e+15	52.507528	2.917085	0.026968
19	2359	1.216451e+17	56.755456	2.987129	0.024059
20	2834	2.432902e+18	61.077384	3.053869	0.021552

4.1 Beyond Pulse Position Modulation (BPPM)

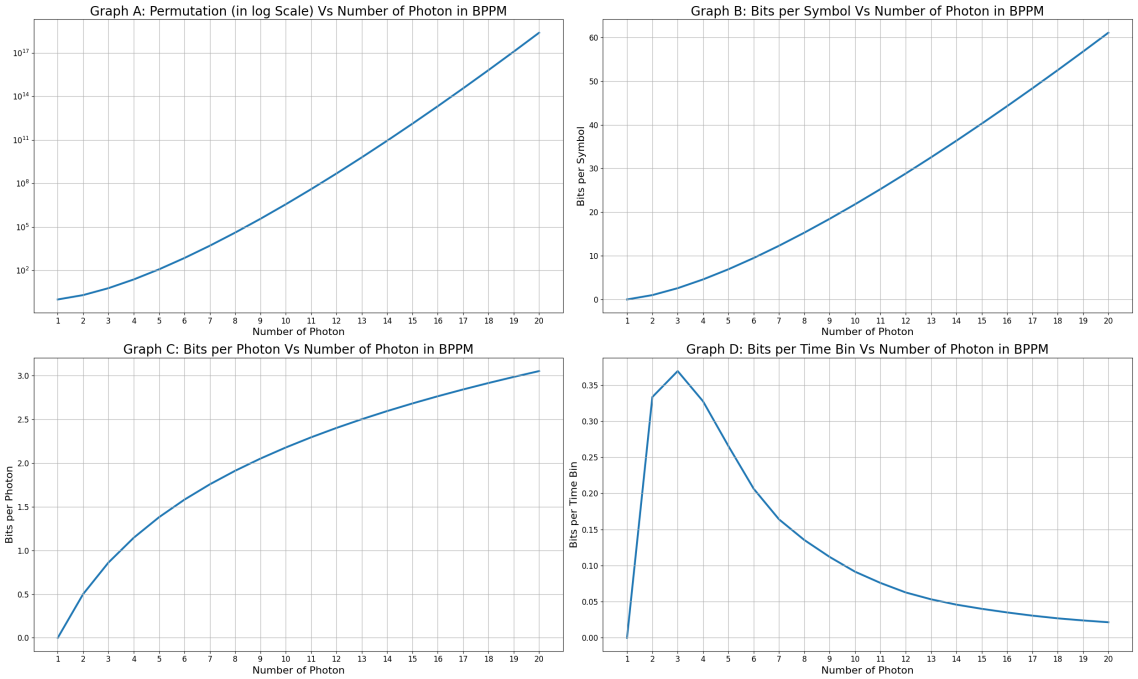

Figure 4.1: The 4 plots in a graph show the corresponding 4 metrics on BPPM versus the photon number.

Figure 4.1 presents the graphical representation of the permutation, bits per symbol, bits per photon, and bits per time metrics as a function of the number of photons within the BPPM protocol. These plots provide a visual depiction of the data presented in Table 4.1, shedding light on the performance characteristics of the BPPM protocol about the number of photons.

The relationship between permutations and bits per symbol concerning the number of photons shows a nonlinear increase, displaying a rapid growth for values up to 10. However, beyond this threshold, the increase becomes more gradual, indicating a more stable performance of the BPPM protocol in terms of symbol transmission efficiency.

In contrast, the relationship between bits per photon and the number of photons follows a different trend. Initially, there is a decelerated rate of increase in bits per photon as the number of photons increases. However, once the threshold of 10 photons is reached, the metric steadily rises, suggesting a more consistent utilization of photon resources for information transmission.

The relationship between bits per time bin and the number of photons is a remarkable plot. The data exhibits a notable surge followed by a subsequent exponential decline. This discovery underscores the intricate balance between time bin allocation and information transmission efficiency in the BPPM protocol.

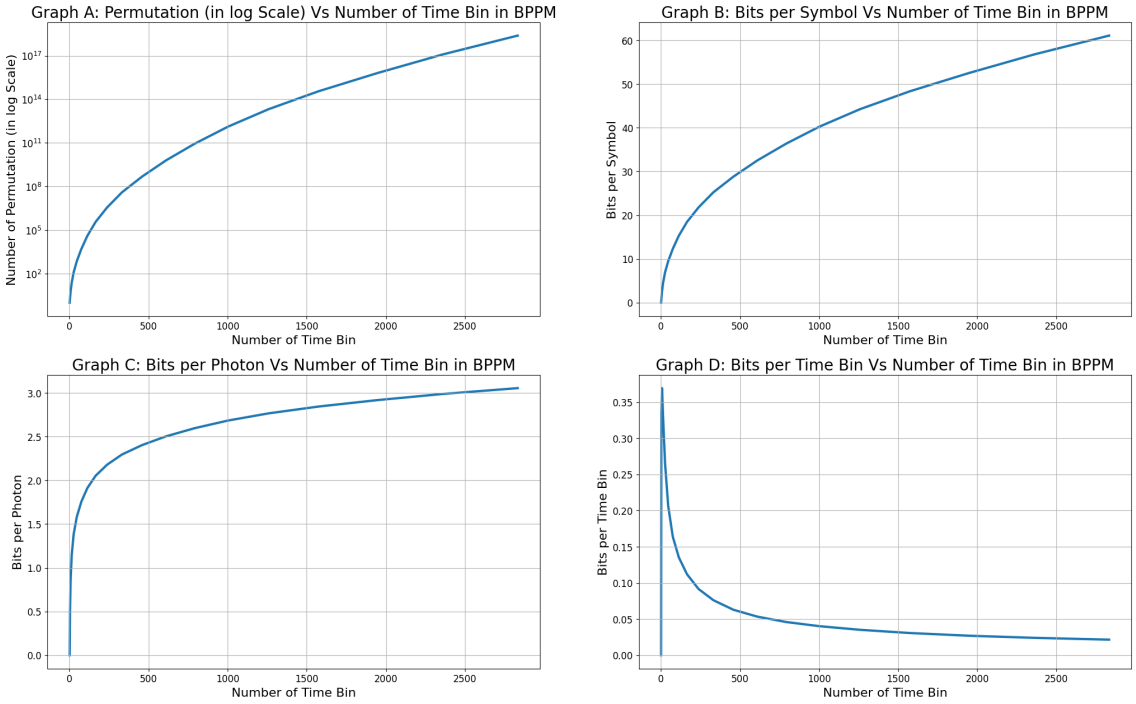

Figure 4.2: The 4 plots in a graph shows the corresponding 4 metrics on BPPM versus the number of time bin.

Figure 4.2 shows the permutation, bits per symbol, bits per photon, and bits per time bin versus the number of time bins. The first three metrics exhibit a similar increment pattern with varying rates. All three metrics exhibit a sublinear growth. On the contrary, there exists an inverse proportionality between the quantity of bits per time bin and the quantity of time bins.

4.2 Pulse Position Modulation (PPM)

Table 4.2 shows that while the number of photons remains constant, the number of time bins steadily increases over multiple iterations. This increase leads to exponential growth in the corresponding permutations, ranging from 1 to 2,834 within 20 iterations.

Furthermore, investigating bits per symbol and photon demonstrates a consistent and proportional increase. As the number of time bins progressively expands, both metrics display a corresponding growth pattern. However, it is worth noting that the relationship between bits per time bin and the number of time bins exhibits a distinct behavior. Initially, bits per time bin rises from 0 to 0.12 when the number of time bins reaches 6. Remarkably, the metric gradually decreases beyond this point, ultimately converging to 0 when the number of time bins reaches 14.

Table 4.2: Table depicts PPM in terms of the metrics permutation, bits per symbol, bits per photon, bits per time bin up to photon number n = 20.

Number of Photon	Time Bin	Permutation	Bits/Symbol	Bits/Photon	Bits/Time Bin
1	1	1	0.000000	0.000000	0.000000
1	3	3	1.584963	1.584963	0.528321
1	7	7	2.807355	2.807355	0.401051
1	14	14	3.807355	3.807355	0.271954
1	26	26	4.700440	4.700440	0.180786
1	46	46	5.523562	5.523562	0.120077
1	75	75	6.228819	6.228819	0.083051
1	113	113	6.820179	6.820179	0.060356
1	165	165	7.366322	7.366322	0.044644
1	238	238	7.894818	7.894818	0.033172
1	332	332	8.375039	8.375039	0.025226
1	459	459	8.842350	8.842350	0.019264
1	610	610	9.252665	9.252665	0.015168
1	791	791	9.627534	9.627534	0.012171
1	1002	1002	9.968667	9.968667	0.009949
1	1259	1259	10.298063	10.298063	0.008180
1	1574	1574	10.620220	10.620220	0.006747
1	1947	1947	10.927037	10.927037	0.005612
1	2359	2359	11.203960	11.203960	0.004749
1	2834	2834	11.468624	11.468624	0.004047

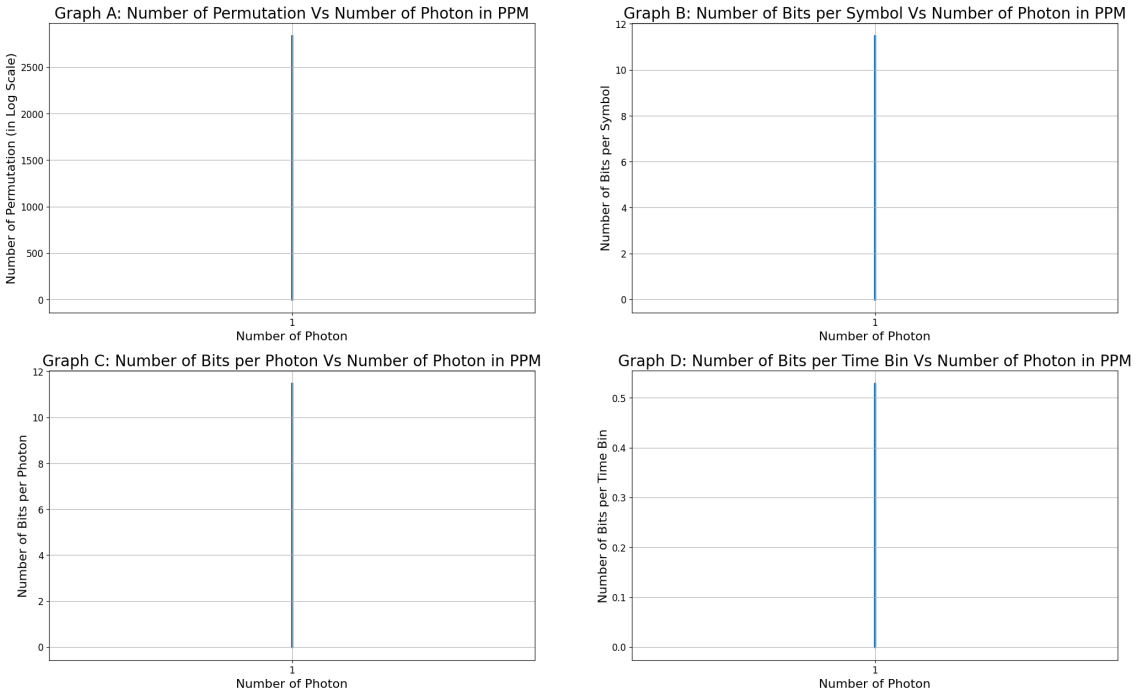

Figure 4.3: The 4 plots in a graph show the corresponding 4 metrics on PPM versus the photon number.

Figure 4.3 presents the graphical representation of the permutation, bits per symbol, bits per photon, and bits per time metrics as a function of the number of photons. These plots visualize the data in Table 4.2.

As expected, due to the fixed number of photons (a single photon), all the metrics demonstrate a vertical line relationship with the number of photons. This outcome arises from the inherent nature of the protocol, where the number of photons remains constant throughout the measurements.

This visualization provides valuable information about the observed metrics' behavior within the PPM method. The vertical lines indicate a consistent value for each metric as the number of photons remains unchanged. While this observation may seem trivial, it is essential to consider in analyzing and interpreting the subsequent findings.

The significance of these results lies in their ability to reinforce the understanding of the impact of the number of photons on the metrics of interest. By fixing the number of photons, we can effectively isolate and investigate the influence of other factors, such as time bins or information content, on the performance of the PPM method.

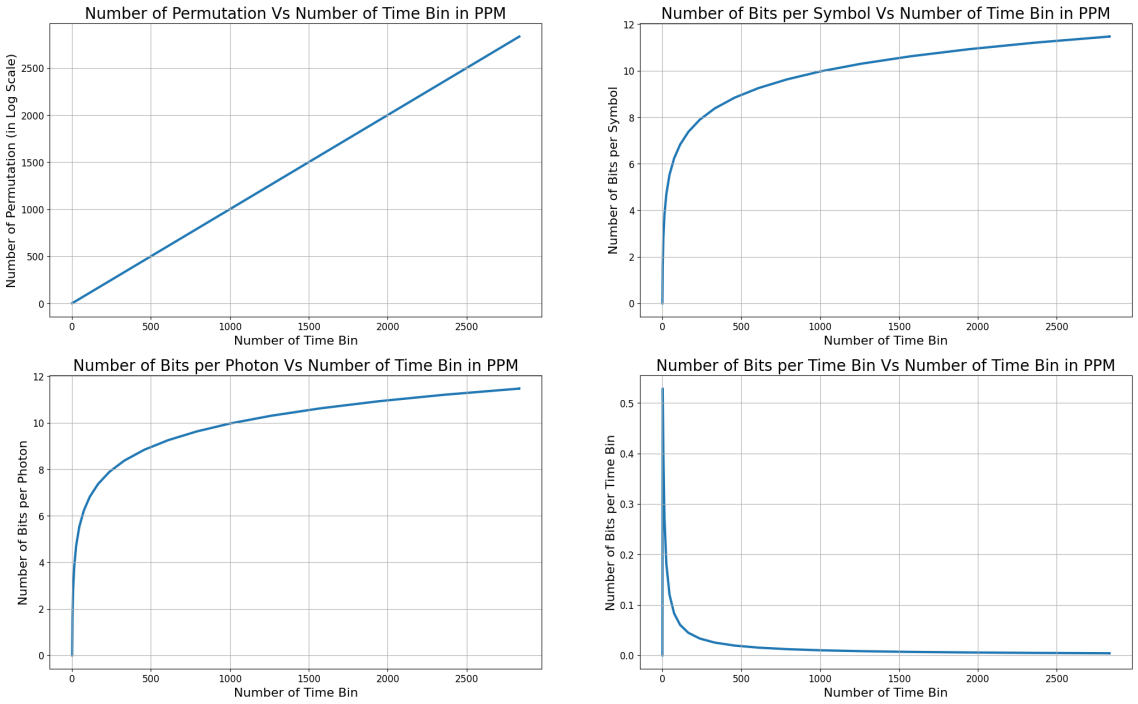

Figure 4.4: The 4 plots in a graph shows the corresponding 4 metrics on PPM versus the number of time bins.

Figure 4.4 depicts the permutation, bits per symbol, bits per photon, and bits per time metrics as functions of the number of time bins. These visualizations represent the data outlined in Table 4.2.Analyzing the first three metrics reveals a shared increment pattern as time bins increase. However, each metric exhibits a distinct growth rate, displaying a decelerated progression. Notably, the relationship between bits per time bin and the number of time bins exhibits a distinct behavior. The rate of decline in the number of bits per time bin follows an exponential pattern as the number of time bins increases.

4.3 On-Off Keying (OOK)

The analysis of the results presented in Table 4.3 shows that the number of time bins increases twofold compared to the number of photons across multiple iterations. This remarkable observation leads to exponential growth in the corresponding permutations, ranging from 4 to a value with a power of 12 within 20 iterations. Moreover, investigating the bits per symbol metric reveals a consistent and proportional increase of 2 with each iteration. As the number of time bins progressively expands, this metric consistently displays a corresponding growth pattern, reflecting the influence of the increased temporal resolution on the information content. However, it is noteworthy that the relationship between bits per photon and bits per time bin and the number of time bins exhibits a distinct behavior. Despite the continuous expansion of the time bins, both metrics exhibit a constant value unaffected by the time bin size. This observation suggests that allocating additional time bins does

Table 4.3: Table depicts OOK in terms of the metrics permutation, bits per symbol, bits per photon, bits per time bin up to photon number n = 20.

Number of Photon	Time Bin	Permutation	Bits/Symbol	Bits/Photon	Bits/Time Bin
1	2	4	2	2	1
2	4	16	4	2	1
3	6	64	6	2	1
4	8	256	8	2	1
5	10	1,024	10	2	1
6	12	4,096	12	2	1
7	14	16,384	14	2	1
8	16	65,536	16	2	1
9	18	262,144	18	2	1
10	20	1,048,576	20	2	1
11	22	4,194,304	22	2	1
12	24	1.677216e+07	24	2	1
13	26	6.7108864e+07	26	2	1
14	28	2.684355e+08	28	2	1
15	30	1.073742e+09	30	2	1
16	32	4.294967e+09	32	2	1
17	34	1.717987e+10	34	2	1
18	36	6.871948e+10	36	2	1
19	38	2.748779e+11	38	2	1
20	40	1.099512e+12	40	2	1

not impact the information efficiency on a per-photon or per-time-bin basis.

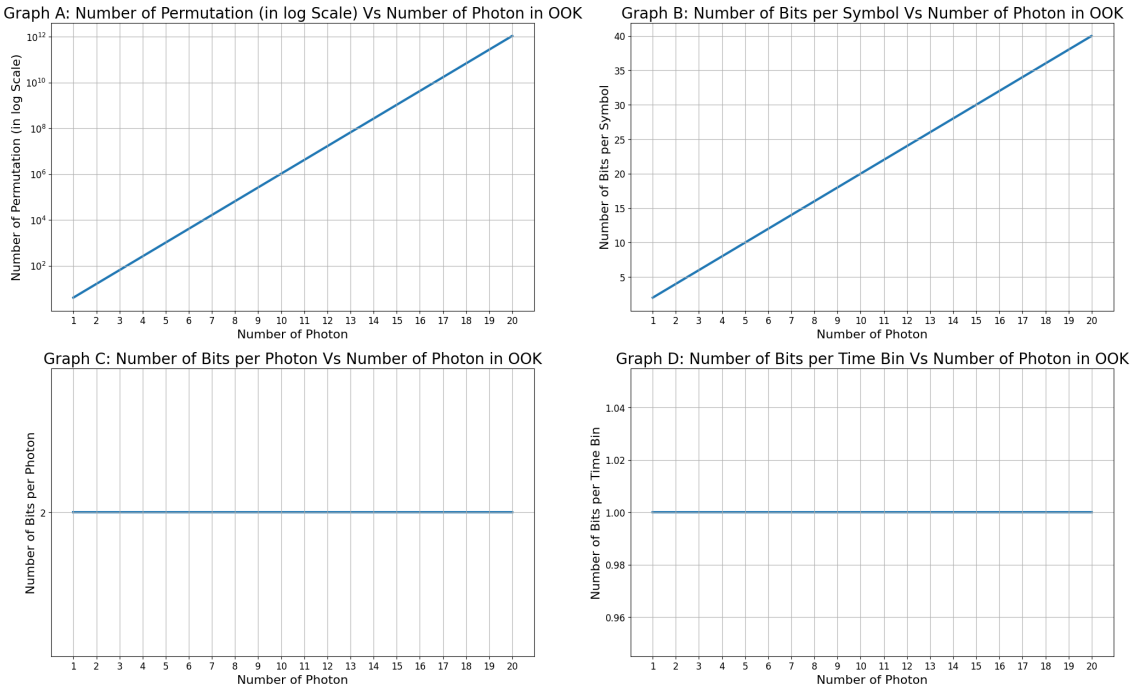

Figure 4.5: The 4 plots in a graph show the corresponding 4 metrics on OOK versus the photon number.

Figure 4.5 visually represents the permutation, bits per symbol, bits per photon, and bits per time metrics related to the number of photons. These plots complement the findings presented in Table 4.3 and offer additional information into the behavior of these metrics within the context of the OOK protocol. The permutation metric shows a pronounced increase from 4 to a value with a power 12 throughout 20 iterations. This exponential permutation growth directly results from the unique characteristics inherent to the OOK protocol. In contrast, the bits per symbol metric exhibits a consistent and linear increment in response to the number of photons. This observation indicates that as the number of photons increases, the information content per symbol grows proportionally. This finding highlights the potential for enhanced data transmission capabilities with higher photon counts. Interestingly, both the bits per photon and bits per time bin metrics, plotted against the number of photons, present horizontal lines that remain constant throughout the range of photon values considered. These metrics appear unaffected by the number of photons, suggesting that allocating additional photons does not increase information efficiency on a per-photon or per-time-bin basis.

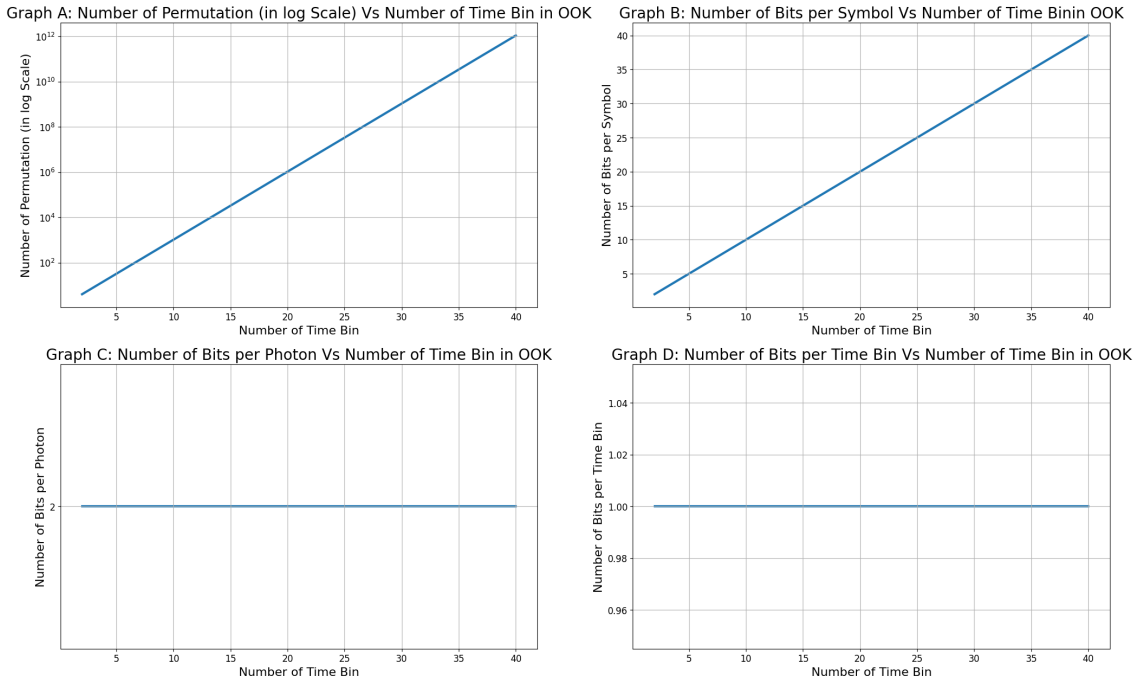

Figure 4.6: The 4 plots in a graph shows the corresponding 4 metrics on PPM versus the number of time bins.

Figure 4.6 shows the permutation, bits per symbol, bits per photon, and bits per time metrics versus the number of time bins. Notably, all four plots exhibit striking similarities to their counterparts in Figure 4.5, demonstrating consistent trends with the number of photons but at distinct rates of change with the number of time bins. This observation implies that the impact of the number of time bins on these metrics is analogous to the effect of the number of photons, albeit with varying degrees of sensitivity.

4.4 "general" protocol

The analysis of the results presented in Table 4.4 shows that the number of time bins progressively increases as the number of photons grows, leading to an exponential growth in the corresponding permutations. This exponential increment ranges from 1 to a staggering number with a power of 50 throughout 20 iterations. Furthermore, investigating the bits per symbol, bits per photon, and bits per time bin metrics demonstrates relatively steady increments concerning the number of photons. These metrics exhibit a consistent and proportional increase, reflecting the underlying principles of the General protocol. However, it is noteworthy that the relationship between the number of photons and the bits per time bin exhibits a distinct pattern. This metric exhibits a nonlinear increase in bits per time bin followed by a gradual decrease, indicating a nonlinear behaviour.

Table 4.4: Table depicts general protocol in terms of the metrics permutation, bits per symbol, bits per photon, bits per time bin up to photon number n = 20.

Number of Photon	Time Bin	Permutation	Bits/Symbol	Bits/Photon	Bits/Time Bin
1	1	1	0	0	0
2	3	3	1.584963	0.792481	0.528321
3	7	35	5.129283	1.709761	0.732755
4	14	1,001	9.967226	2.491807	0.711945
5	26	65,780	16.005361	3.201072	0.615591
6	46	9,366,819	23.159128	3.859855	0.503459
7	75	1.984830e+09	30.886368	4.412338	0.411818
8	113	5.117190e+11	38.896561	4.862070	0.344217
9	165	2.000631e+14	47.507449	5.278605	0.287924
10	238	1.326714e+17	56.880635	5.688063	0.238994
11	332	1.144580e+20	66.633380	6.057580	0.200703
12	459	1.579209e+23	77.063548	6.421962	0.167894
13	610	2.286081e+26	87.563007	6.735616	0.143546
14	791	3.835418e+29	98.275298	7.019664	0.124242
15	1,002	7.092269e+32	109.127946	7.275196	0.108910
16	1,259	1.730720e+36	120.380784	7.523799	0.095616
17	1,574	5.759108e+39	132.081041	7.769473	0.083914
18	1,947	2.333541e+43	144.065429	8.003635	0.073994
19	2,359	9.229376e+46	156.014926	8.211312	0.066136
20	2,834	4.292288e+50	168.198152	8.409908	0.059350

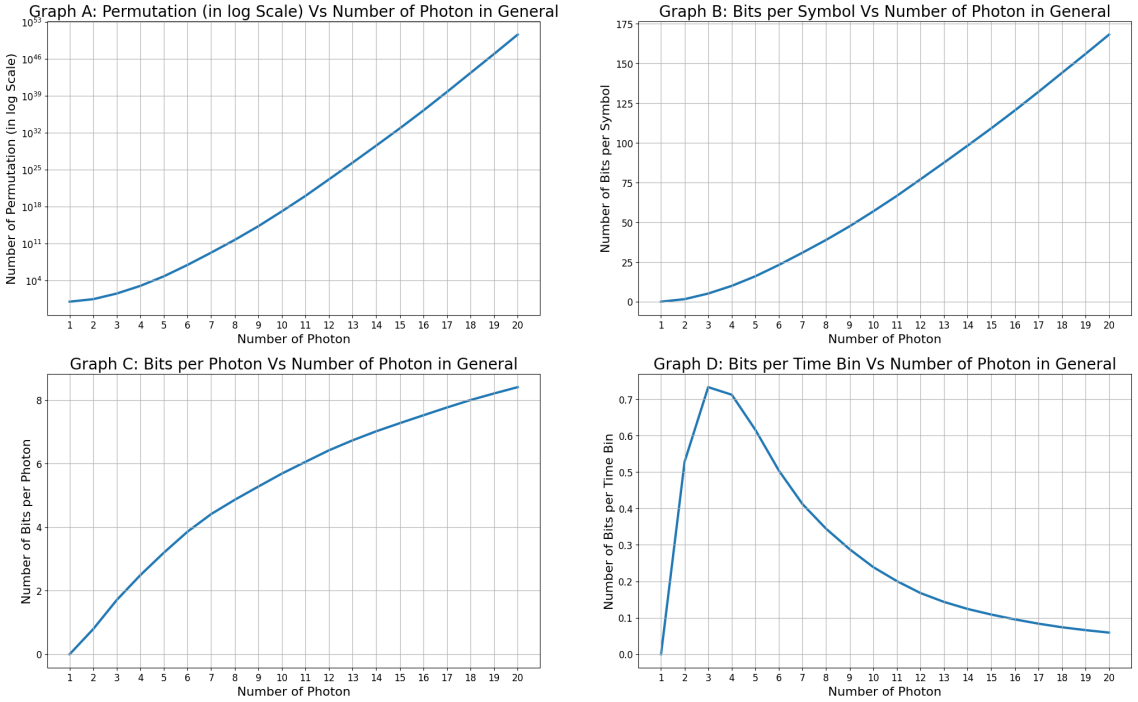

Figure 4.7: The 4 plots in a graph show the corresponding 4 metrics on General versus the photon number.

Figure 4.7 presents the graphical representation of key metrics within the General Protocol. These metrics, including the permutation, bits per symbol, bits per photon, and bits per time metrics, are analyzed as a function of the number of photons. These plots visually depict the data in Table 4.4. At the onset, there is a noticeable non-linear elevation in the number of permutations and the amount of information conveyed each symbol. Subsequently, the rate of growth exhibits a more gradual progression, suggesting a heightened level of stability in the performance of the General protocol with regards to the effectiveness of symbol transmission. In contrast, the relationship between bits per photon and the number of photons follows a different trend. Initially, there is a decelerated rate of increase in bits per photon as the number of photons increases. Moreover, the behaviour of bits per photon in response to varying photon numbers follows a distinct trajectory. As the number of photons increases, the rate of increase in bits per photon is initially moderate. However, subsequent to a specific increase in the quantity of photons, this growth experiences a further deceleration. The link between the quantity of photons and the bits per time bin is a notable plot of interest. The data exhibits a pronounced upward trend followed by a subsequent exponential decline in the measurements. This remark draws attention to the intricate balance between time bin allocation and information transmission efficiency within the General protocol.

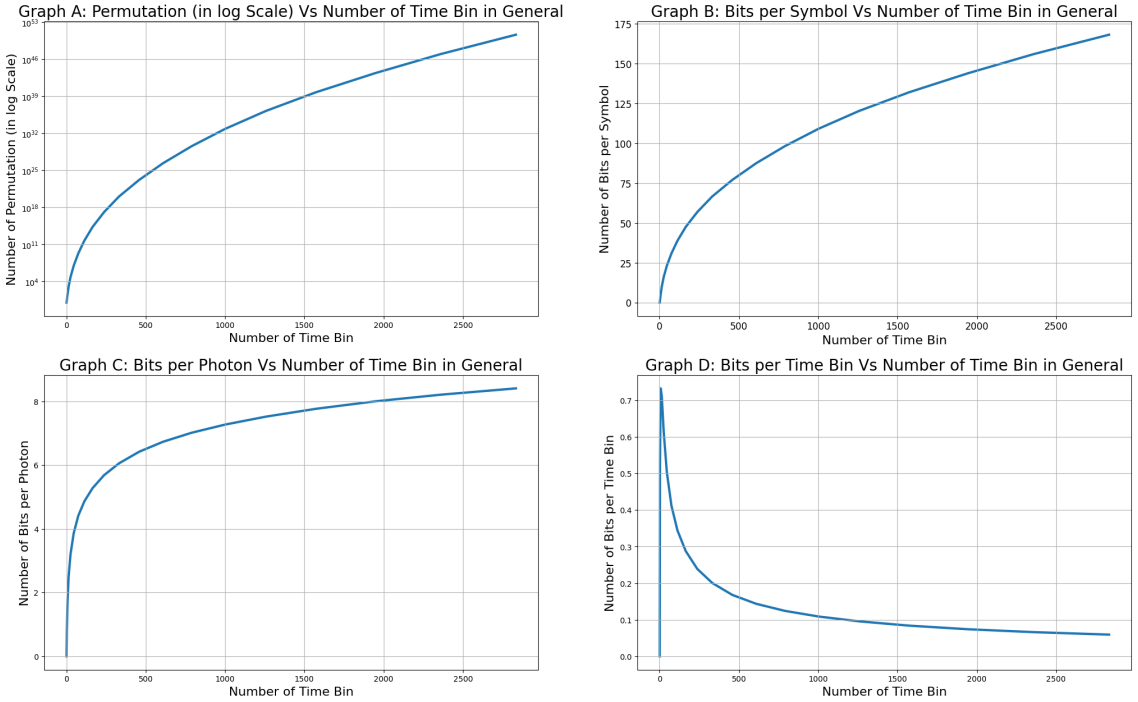

Figure 4.8: The 4 plots in a graph shows the corresponding 4 metrics on General versus the number of time bins.

Figure 4.8 displays the variations in permutation, bits per symbol, bits per photon, and bits per time as a function of the number of time bins. The three metrics exhibit an identical sublinear incremental pattern, albeit at different rates of change. Remarkably, the relationship between bits per time bin and the number of time bins mirrors the patterns observed in the preceding three metrics. Significantly, the plot exhibits a prominent peak that experiences an exponential decrease. These findings provide a comprehensive view of the dynamics surrounding time bin allocation and their effect on the performance metrics. The trends that have been noticed shed light on significant points of change, offering essential information into the complex tradeoffs that define the behaviour of the system.

4.5 Protocols Comparison in terms of Information Bits

4.5.1 Information Bits per Photon Versus Number of Photon

The bits per photon versus photon metric compares the number of information bits that can be transmitted using a fixed number of photons. This metric helps compare the efficiency of different modulation and encoding schemes and evaluate the impact of noise and interference on the performance of a communication system.

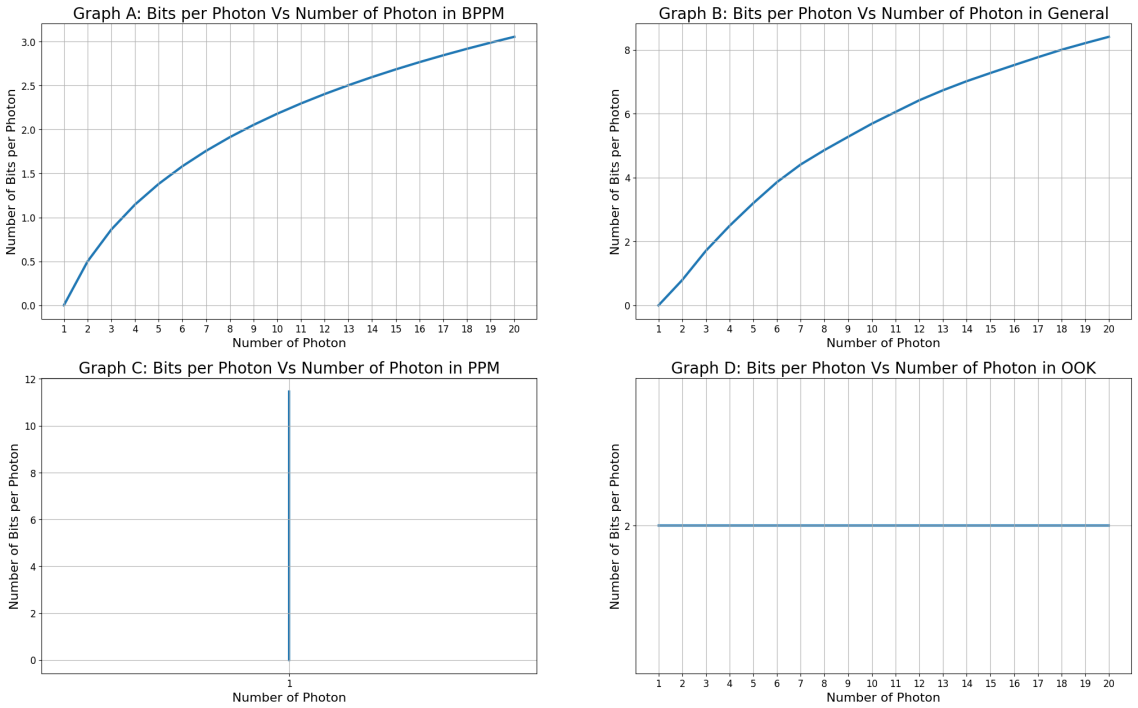

Figure 4.9: The figure displays four plots representing different protocols, showcasing the relationship between bits per photon and the number of photons.

Figure 4.9 illustrates the plots depicting the relationship between bits per photon and the number of photons for four different protocols: Binary Pulse Position Modulation (BPPM), General, Pulse Position Modulation (PPM), and On-Off Keying (OOK). BPPM and General exhibit a similar trend, with the bits per photon increasing at a decelerated rate as the number of photons increases. In contrast, PPM, utilizing only a single photon, demonstrates a vertical line. At the same time, OOK consistently achieves the same amount of bits per photon regardless of the number of photons in the protocol. However, the general protocol exhibits a higher number of bits per symbol as the number of photons increases compared to the other 3 protocols. It implies that general protocol is potentially a more efficient utilization of the available photon resources for encoding information. This observation suggests that as the number of photons employed for communication grows, the protocol can convey a relatively more significant amount of information within each symbol.

4.5.2 Information Bits per Photon Versus Number of Time Bin

The bits per photon versus time bin metric measures the information efficiency of a communication system, defined as the number of information bits that can be transmitted per photon per time bin.

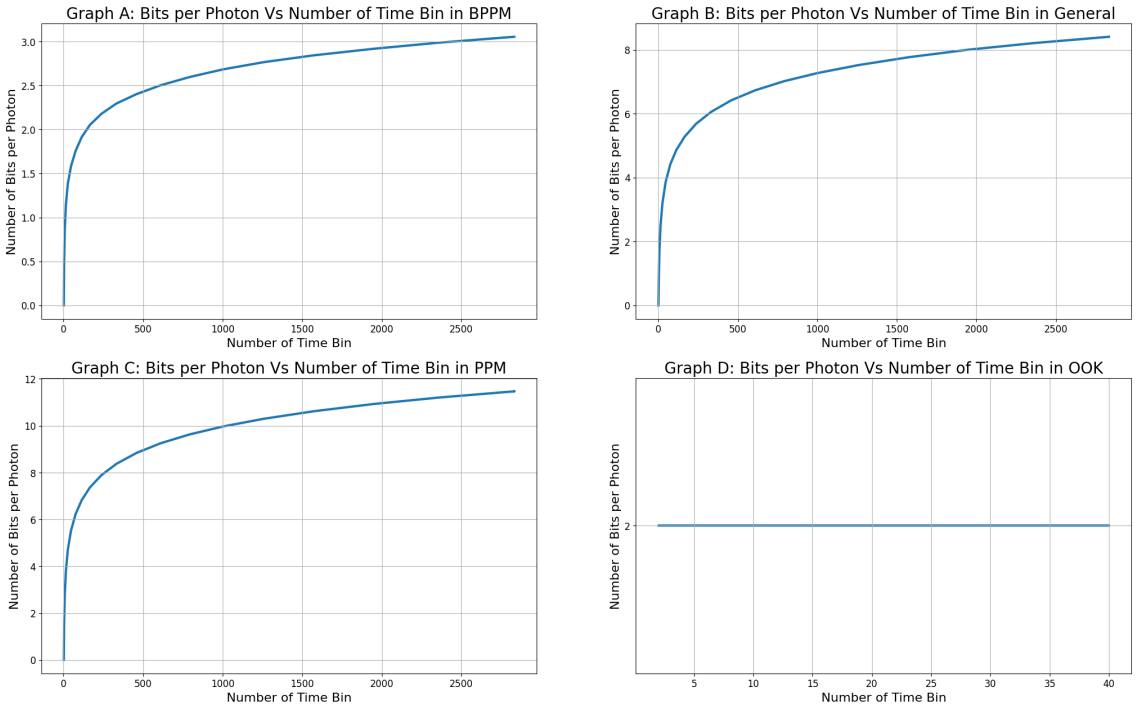

Figure 4.10: The figure displays four plots representing different protocols, showcasing the relationship between bits per photon and the number of time bins.

Figure 4.10 shows the relationship between bits per photon and the number of time bins for four different protocols: Binary Pulse Position Modulation (BPPM), General, Pulse Position Modulation (PPM), and On-Off Keying (OOK). The distinctive trends highlight the varying efficiencies and characteristics of the four protocols, providing a nuanced understanding of their performance in terms of bits per photon concerning the number of time bins. Increasing the number of time bins can increase the number of signals that can be transmitted, but it may come at the expense of other metrics, such as the number of photons per time bin. Therefore, the bits per photon versus time bin metric can be used to evaluate the efficiency and to determine the optimal number of time bins for a given system. In BPPM, General, and PPM, initially, as the number of time bins increases, there is a sharp surge in bits per photon. This surge signifies a rapid accumulation of information within a compact timeframe. However, this surge is followed by a gradual deceleration in the increase of bits per photon as more time bins are added. This pattern suggests that BPPM, General, and PPM protocols are particularly adept at capitalizing on the initial temporal expansion, enabling efficient encoding of information in a condensed span. As the number of time bins grows, their capacity to further enhance bits per photon becomes progressively limited, resulting in the observed deceleration. Their behaviors could be advantageous when the emphasis is on transmitting data rapidly within a constrained time window. However, as the number of time bins increases, their ability to further enhance bits per photon becomes limited. These protocols might be better suited for scenarios where rapid data transmission bursts are required. In stark contrast, the OOK protocol exhibits a consistent bits-per-photon value across time bins. Regardless of the number of time bins in the

protocol, the bits per photon remain relatively stable. This consistent performance can be attributed to the inherent modulation scheme of OOK, where each photon represents a discrete data bit. As a result, the protocol appears less sensitive to variations in time bin configurations. The steady bits-per-photon trend observed in OOK underscores its reliability in maintaining a consistent information-carrying capacity irrespective of temporal adjustments. The consistent bits-per-photon value across different time bin scenarios indicates that the OOK protocol maintains a steady information-carrying capacity. This stability is attributed to its modulation scheme, where each photon represents a single data bit.

4.5.3 Information Bits per Time Bin Versus Number of Time Bin

The 'bits per time bin versus time bin' metric is used to measure the information efficiency of a communication system. It is defined as the number of information bits that can be transmitted using a fixed number of time bins over a given period. This metric helps evaluate the overall efficiency of a communication system in terms of the amount of data that can be transmitted within a fixed time interval.

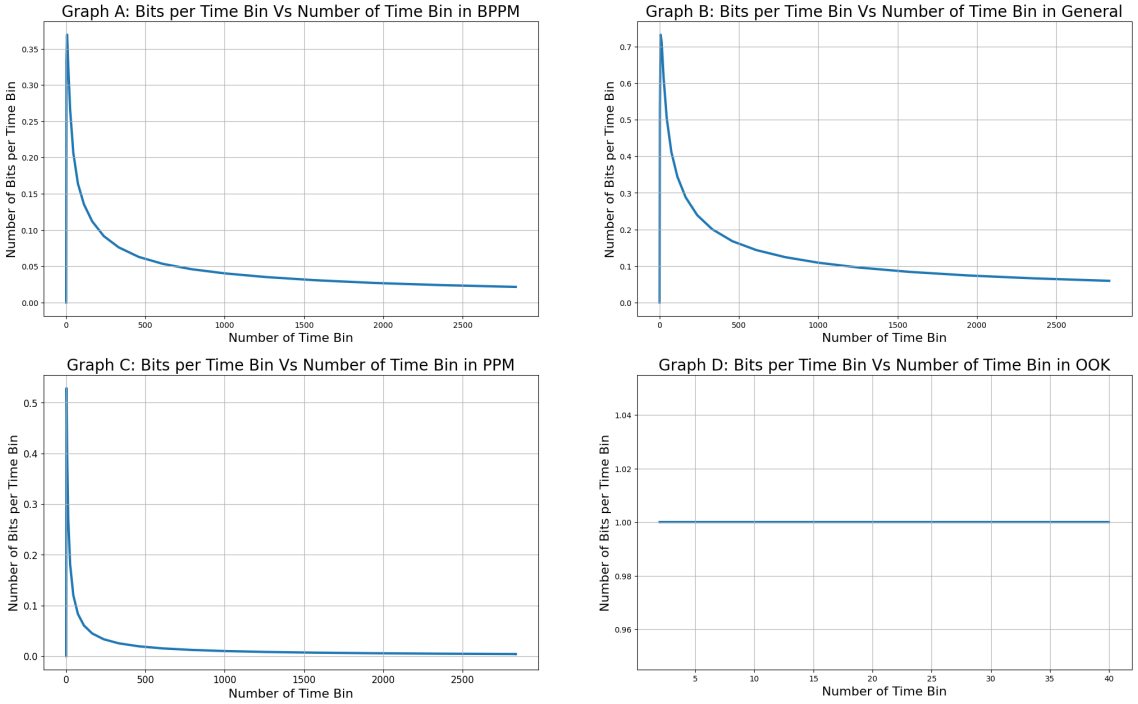

Figure 4.11: The figure displays four plots representing different protocols, showcasing the relationship between bits per time bin and the number of time bins.

Figure 4.11 presents the plots depicting the relationship between bits per time bin and the number of time bins for four different protocols: Binary Pulse Position Modulation (BPPM), General, Pulse Position Modulation (PPM), and On-Off Keying (OOK). BPPM, General, and PPM exhibit a similar pattern, where the bits per time bin undergo an abrupt increment with a slight increase in the number of time

bins, followed by an inverse decrease towards nearly zero. This distinct behavior indicates a unique optimization strategy utilized by these protocols, resulting in a concentrated distribution of bits within a specific range of time bins. This pattern indicates a strategy that capitalizes on a concentrated burst of information. These protocols might be well-suited for scenarios where burst data transmission is required within a defined temporal window. In contrast, OOK consistently maintains the same amount of bits per time bin, irrespective of the number of time bins in the protocol. This observation highlights the inherent characteristics of OOK, where the information encoding is independent of the temporal extent. The consistent maintenance of bits per time bin, regardless of the number of time bins, underscores OOK's independent encoding strategy. OOK could be a favorable choice for scenarios prioritizing consistent data transmission rates over different temporal configurations. These diverse trends shed light on the differing efficiencies and encoding strategies employed by the four protocols, offering a nuanced understanding of their performance in terms of bits per time bin concerning the number of time bins. For scenarios where bursts of data within specific time ranges are essential, BPPM, General, and PPM protocols could be more suitable due to their optimization for the concentrated distribution of bits per time bin. On the other hand, if maintaining a stable data rate across varying time bin scenarios is critical, the OOK protocol might be a better choice due to its consistent bits per time bin behavior.

4.5.4 Information Bits per Time Bin Versus Number of Photons

The bits per time bin versus photon metric compares the number of information bits that can be transmitted using a fixed number of photons over a given period, which helps evaluate the efficiency of a communication system in terms of the number of photons required to transmit a given amount of information.

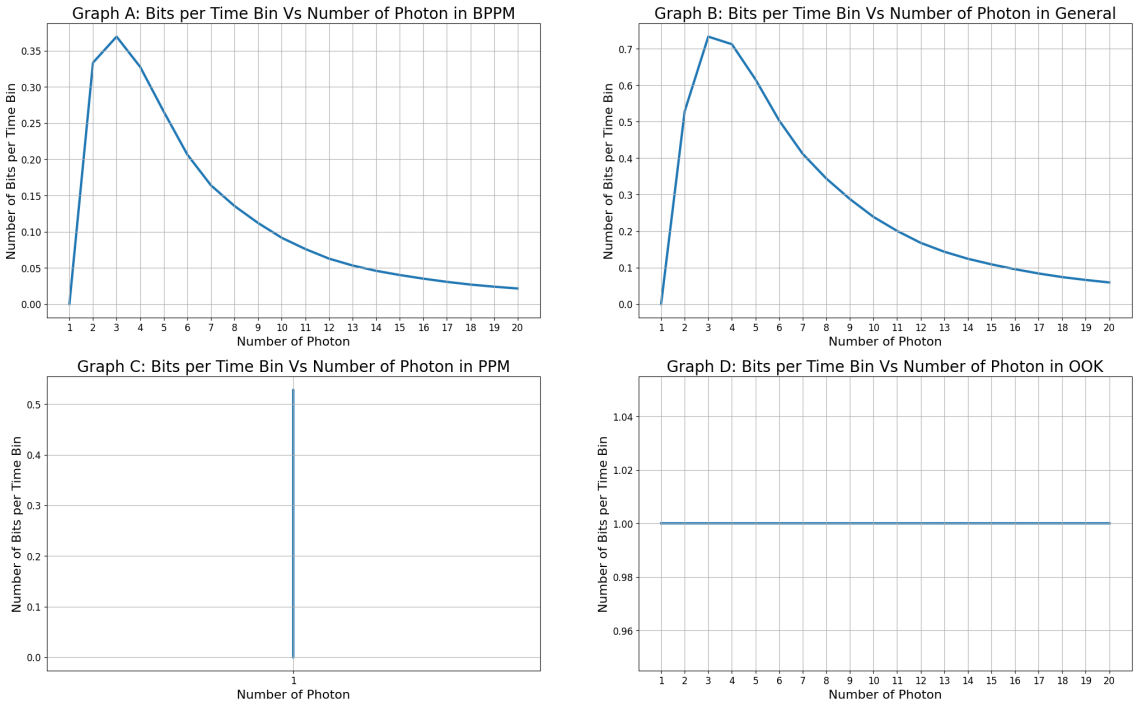

Figure 4.12: The figure displays four plots representing different protocols, showcasing the relationship between bits per time bin and the number of photons.

Figure 4.12 showcases the plots illustrating the relationship between bits per time bin and the number of time bins for four different protocols: Binary Pulse Position Modulation (BPPM), General, Pulse Position Modulation (PPM), and On-Off Keying (OOK). Both the BPPM and General protocol demonstrate a comparable trend, wherein the bits per time bin rise as the number of photons increases, reaching a maximum value and subsequently decreasing inversely towards zero. This observation indicates the existence of a specific range of photons that achieves the highest level of efficiency in terms of information encoding, measured in bits per time bin. These protocols demonstrate an effective utilization of photons within a specific range for maximizing bits per time bin. They could be well-suited for scenarios where the photon count can be controlled to achieve optimal encoding efficiency. In contrast, PPM demonstrates a vertical line, indicating that it utilizes only a single photon for encoding information. This unique behavior implies a specific mode of photon utilization, possibly making it suitable for situations with limited photon availability or constraints. On the other hand, OOK consistently achieves the same amount of bits per photon, irrespective of the number of photons present. This plot shows that OOK remains unaffected by changes in the photon count. This phenomenon indicates that OOK maintains a steady information-carrying capacity, making it favorable for scenarios requiring stable data transmission rates. These distinct patterns shed light on the encoding characteristics and limitations of the four protocols, providing valuable information about their performance in terms of bits per time bin concerning the number of photons. Overall, the choice of protocol depends on the specific needs of the communication system: BPPM and General protocols might be preferred when optimizing bits per time bin within an optimal

photon range is essential. PPM's characteristic of using a single photon could be advantageous in constrained photon resources. OOK's consistent performance is suitable when maintaining a stable data rate is a priority.

4.5.5 Protocols Comparison in terms of Mutual Information

The study investigates the efficiency of error correction techniques and the quality of information transmission across various communication protocols. To achieve this goal, we introduce the concept of Raw Information Rate. Essentially, it signifies the theoretical upper limit of data transfer without restrictions. In order to assess the efficiency of error correction techniques and evaluate the information quality of different communication protocols, our study employed graphical representations to depict the raw information rate for three specific protocols: Beyond Pulse Position Modulation (BPPM), Pulse Position Modulation (PPM), and a generalized protocol. These plots explored the survival rate of information bits per transmitted data bit, a crucial metric for assessing communication system performance. Moreover, we delved into the influence of error correction on the information quality of these protocols, explicitly focusing on scenarios involving low error probabilities. By visually illustrating this impact, we aimed to ascertain the tradeoff between error correction and the raw information rate, thereby determining the optimal level of error correction suited for varying error probabilities. The outcomes of our analyses provide valuable information into communication system performance and the significant role played by error correction in enhancing information quality.

4.5.5.1 Mutual Information (MI) for Constant Power

Mutual information (MI) for constant power explores how much information can be reliably transmitted over a channel while maintaining a fixed power ratio. It helps assess the channel's information capacity under power constraints and provides information about achievable data rates. The power ratio P means that the ratio of number of photons to the number of time bins should be constant among all protocols for comparison. Bin rate and photon energy play roles in conversion, assumes that time bins have the same length.

$$P \propto \frac{n}{M} \tag{4.1}$$

, where n and M are the number of photons and time bins, respectively. In order to assess the mutual information of three protocols, namely BPPM, PPM, and the General Protocol, for a comparative analysis while maintaining constant power transmission ratio. The results obtained from this analysis are shown in Table 4.5. By studying the power consumption ratio across the three examined protocols, our study pinpointed three noteworthy power ratios for investigating the mutual information quality. Specifically, we focused on 3 distinct ratio values, namely 0.2 and

0.05 and 0.01.

Protocol	P = 0.2	P = 0.05	P = 0.01
BPPM	n = 5, M = 26	n = 9, M = 165	n = 17, M = 1,574
PPM	n = 1, M = 5	n = 1, M = 18	n = 1, M = 100
General	n = 5, M = 26	n = 9, M = 165	n = 17, M = 1,574

Table 4.5: The table shows the corresponding number of photons and time bins for the 3 values of constant power among BPPM, PPM, and General protocol.

We used the data in Table 4.5 to plot the mutual information for the BPPM, PPM, and General protocols. Specifically, we determined the corresponding number of photons and time bins for each value of power consumed and used these values to plot the mutual information.

Figure 4.13 shows the comparative analysis of mutual information (MI) and the normalized mutual information (NMI) versus error probability for $0 \leq P_E \leq 0.1$ among the 3 protocols of BPPM, PPM, and General protocol based on the consistent power ratio. The number of photons and the time bin in BPPM are adjusted proportionally to maintain uniform power ratio across protocols. The analysis explicitly examines MI and NMI trends for power ratio $P = 0.2$, $P = 0.05$, and $P = 0.01$ indicated by graph A and B, C, D, and E and F, respectively. Initially, at $P_E = 0.01$, regarding mutual information (left-hand-side of the plot), the general protocol demonstrates the highest value for three cases. Although the value of mutual information in BPPM improves, it keeps as the same rank for 3 scenarios. Although the General protocol contains more information than other protocols initially, it exhibited the steepest linear slope, indicating that it lost information the fastest linearly. In addition, PPM exhibits the lowest value of mutual information, and the steepness of its linear curve insignificantly increase as the number of photons grows in other protocol. However, when the effect of increasing error probability P_E on various protocols, we observe the non-linearity pattern of both BPPM and General protocol that their mutual information experience a notable decline. The trend becomes more significant for the case in higher photon number like Graph C and Graph E.

The normalized mutual information is introduced to comprehensively assess the response of MI to varying P_E. The normalized plots investigate the duration for which information persists in the protocols when they have similar amounts of initial information. Displayed on the right-hand side of the plot, this metric offers a way to observe how mutual information changes with error probability P_E, which is calculated as the mutual information at specific P_E to that at $P_E = 0$, such that all the normalized values start from 1 for equitable comparison. Within $P_E = 0.02$, it is noteworthy that BPPM obtains the highest MI among the other 2 protocols in the zoomed subplots in Graph B and Graph D. However, as P_E increases till $P_E = 0.03$ and $P_E = 0.01$ as evidence by Graph D and Graph F respectively, the MI value for BPPM experience a sublinear decline, which leads to a convergence of BPPM's MI value with those of PPM. As expected, the normalized mutual information for

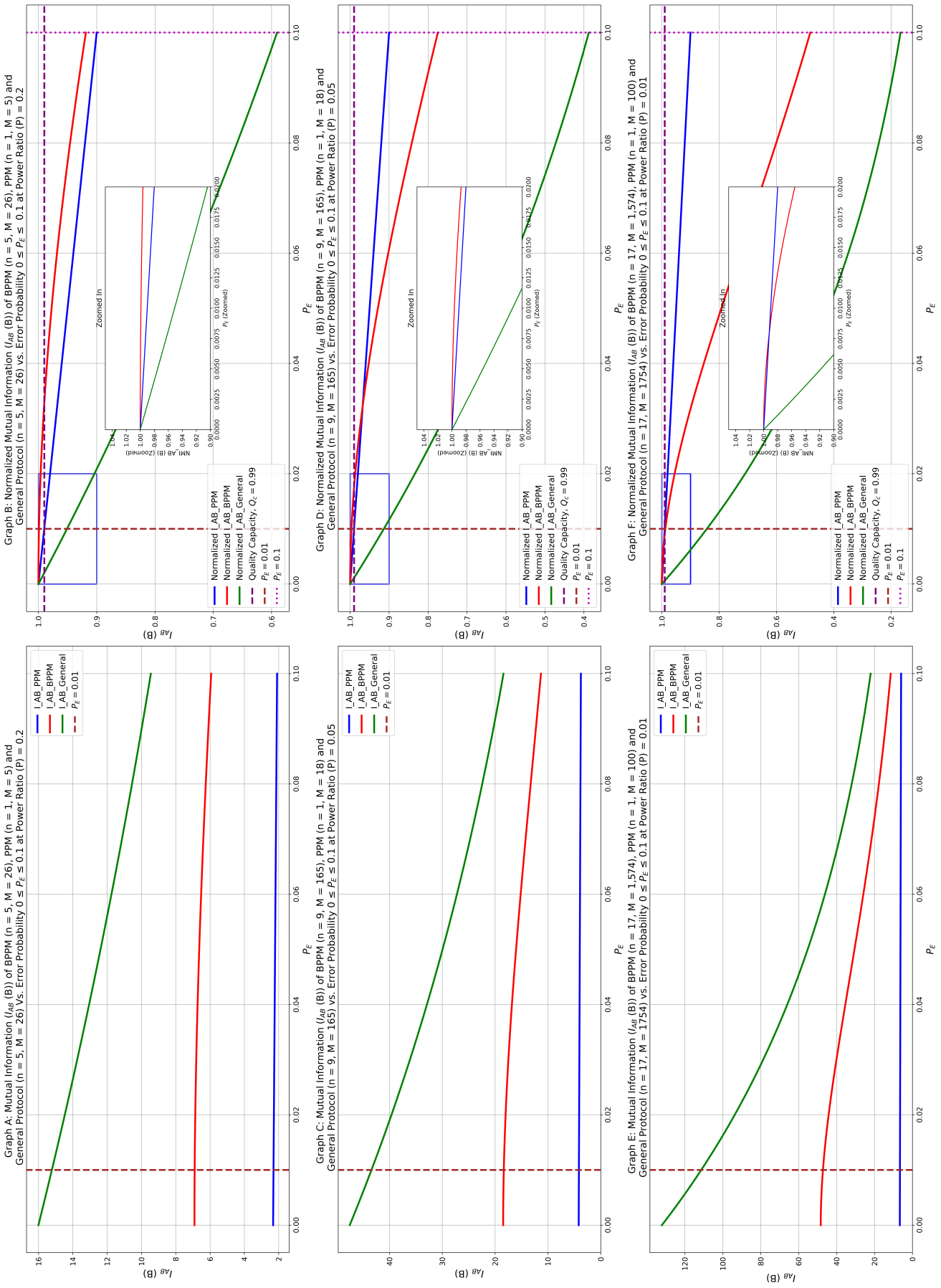

Figure 4.13: The diagram depicts the mutual information and normalized mutual information versus error probability for $0 \leq P_E \leq 0.1$ among the 3 protocols of BPPM, PPM, and General protocol based on the same ratio of power transmission. The value of P = 0.2, 0.05, 0.01 were selected.

all protocols except BPPM displayed a linear decrease, with PPM exhibiting the flattest curve at the start of the transmission. The curve's flatness for BPPM over three right-hand graphs indicates that the protocol can detect and correct errors, resulting in slower deterioration of information so that the information persisted longer than the other protocols. At specific error probability values, particularly within the range of $P_E \leq 0.01$, the NMI value of BPPM achieved superior to the established Quality Criterion (Q_c), denoted as $Q_c = 0.99$, represented by the purple horizontal dashed line. The benchmark delineates a threshold for effective communication across the protocols, signifying BPPM's advantageous position in conveying information over a noisy channel.

Notably, BPPM's MI retains a resilient performance within this region, displaying minimal degradation. However, the tradeoff emerges that BPPM maintains its edge for effective communication within the specified error probability range, while this advantage may not extend to scenarios involving longer-distance transmissions. Other protocols might gain the upper hand in contexts involving extended distances.

4.5.5.2 MI for Constant Energy per Information Bit

This metric examines the equilibrium between energy consumption and the quality of information transmission, yielding information about the appropriate distribution of energy resources to ensure consistent data transmission quality in various scenarios. The relationship between the energy per information bit (EpB) and the number of photons per number of bit directly proportional.

$$\mathbf{EpB} = \frac{E}{B} \propto \frac{n}{log_2 K} \tag{4.2}$$

where E is the energy consumed in the system, B is the number of bits, n is the number of photons, and K is the number of codewords in a block.

The number of codewords for BPPM is:

$$K = n!, \tag{4.3}$$

where K is the number of codewords and n is the number of photons in BPPM.
The number of codewords for PPM is:

$$K = M, \tag{4.4}$$

where K is the nubmer of codewords and M is the number of time bins in PPM.
The number of codewords for General protocol is:

$$K = \frac{M!}{(M-n)!n!}, \tag{4.5}$$

where K is the numberof codewords, M is the number of time bins, and n is the number of photons in General protocol. A methodological approach is employed to

guarantee an impartial comparison of metrics across the three protocols. The energy per information bit is computed initially, taking into consideration the amount of photons and time bins specifically in the context of the BPPM protocol. Consequently, we observe the same number of photons in the General protocol as in BPPM protocol. Then, we determine the number of time bins necessary to accomplish this while keeping the photon number in the PPM protocol as one.

Table 4.6 illustrate the distribution of energy per information bit across three distinct modulation protocols. By studying the corresponding ratio across the three examined protocols, our study pinpointed three noteworthy values for investigating the mutual information quality. Specifically, we focused on 3 distinct ratio values, namely 0.43, and 0.39 and 0.33

Protocol	E/B = 0.43	E/B = 0.39	E/B = 0.33
BPPM	n = 11, M = 332	n = 14, M = 791	n = 19, M = 2359
PPM	n = 1, M = 5	n = 1, M = 6	n = 1, M = 8
General	n = 11, M = 30	n = 14, M = 44	n = 11, M = 74

Table 4.6: The table shows the corresponding number of photons and time bins for the values of energy per information bit constant among BPPM, PPM, and General protocol.

Figure 4.14 shows the comparative analysis of mutual information (MI) and the normalized mutual information (NMI) versus error probability for $0 \leq P_E \leq 0.1$ among the 3 protocols of BPPM, PPM, and General protocol based on the consistent ratio of EpB. The number of photons and the time bin in BPPM are adjusted proportionally to maintain uniform ratio across protocols. The analysis explicitly examines MI and NMI trends for the ratio $EpB = 0.43, 0.39, = 0.33$ indicated by graph A and B, C, D, and E and F, respectively.

Regarding mutual information, the general protocol and the BPPM first shows the comparable value for three examples (left-hand side of the figure). In contrast, BPPM maintained its information for a short distance within $P_E \leq 0.02$ for three cases, as indicated by the flatness of the curve, and then lost its information quantity relatively slower than General, which allowed BPPM to obtain the highest quantity of mutual information. However, General's information content subsequently deteriorated as the error probability increased over distance. The curve remains relatively flat for error probabilities between 0 and 0.02, indicating a stable performance due to error correction. In addition, PPM exhibits the lowest value of mutual information. In particular, when examining the impact of increasing error probability P_E on different protocols, we see a non-linear trend in both the BPPM and General protocols, wherein General mutual information undergo a significant increase while BPPM decline in a slower rate. The results show that BPPM consistently outperforms the other 2 protocols across all the tested values of EpB. It achieves the highest information transmission, surpassing PPM and the General protocol. Notably, BPPM exhibits a distinct characteristic in its mutual information curve.

In order to acquire more information, the plot depicting normalised mutual information on the right side of the graph yields a comparable outcome as interpreted on the left side of the graph. The results indicate that the BPPM outperformed the established Quality Criterion (Qc), which is symbolised as $Q_c = 0.99$ and illustrated by the purple horizontal dashed line. The beneficial position of BPPM in effectively transmitting information across a channel with high levels of noise. As shown in the zoomed subplots of all three graph on the right-hand side, BPPM maintains the flatness of the curve but its robustness decreases as the number of photons increase, that the NMI of PPM coincide with BPPM's with lesser P_E with higher photon number. However, although PPM has higher NMI than other protocols do, the lost photon during transmission may lead to a time cost to send another photon carrying the same amount of mutual information once again, especially essential when the transmission is carred in space communication among planets, such as the Earth and Mars.
It is worth mentioning that BPPM's MI demonstrates a robust performance particular area, with low deterioration. However, it should be noted that while slower speed transmission is a tradeoff, BPPM still maintains its advantage in terms of successful communication within the given error probability range. However, it is important to acknowledge that this advantage may not be applicable in situations that include longer-distance transmissions. However, it can be observed that BPPM consistently outperforms General across the whole range of P_E.

4.5.6 Discussion and Conclusion of the Comparison

In mutual information for constant power, the analysis explicitly examines MI and NMI trends for error probabilities $P = 0.2$, $P = 0.05$, and $P = 0.01$, which allows us to evaluate how well each protocol maintains information transmission quality while keeping power levels consistent. The metrics reveal that BPPM achieves superior MI values compared to established quality benchmark $Q_c = 0.99$ at low error probabilities at $P = 0.01$. This phenomenon showcases BPPM's resilience and capability for effective communication over noisy channels, emphasizing its robust error correction mechanisms. However, while BPPM maintains an advantage in effective communication within specified error probability ranges, the advantage may not extend to longer-distance transmissions. Other protocols perform better in extended distances. In mutual information for constant energy per information bit, BPPM's consistent superiority over PPM and general protocol underscores its effectiveness in information transmission. Also, the distinct shape of BPPM's MI curve, initially flat and gradually decreasing, highlights its robustness to errors at low probabilities and its performance degradation with increasing error rates and photon counts. BPPM's flat mutual information curve suggests effective error correction capabilities, leading to stable performance within certain error probability ranges. Comparing BPPM with PPm and the general protocol reveals their linear mutual information decrease and lack of BPPM's flatness, emphasizing BPPM's unique characteristics. The plot of normalized mutual information provides a view of BPPM's information loss com-

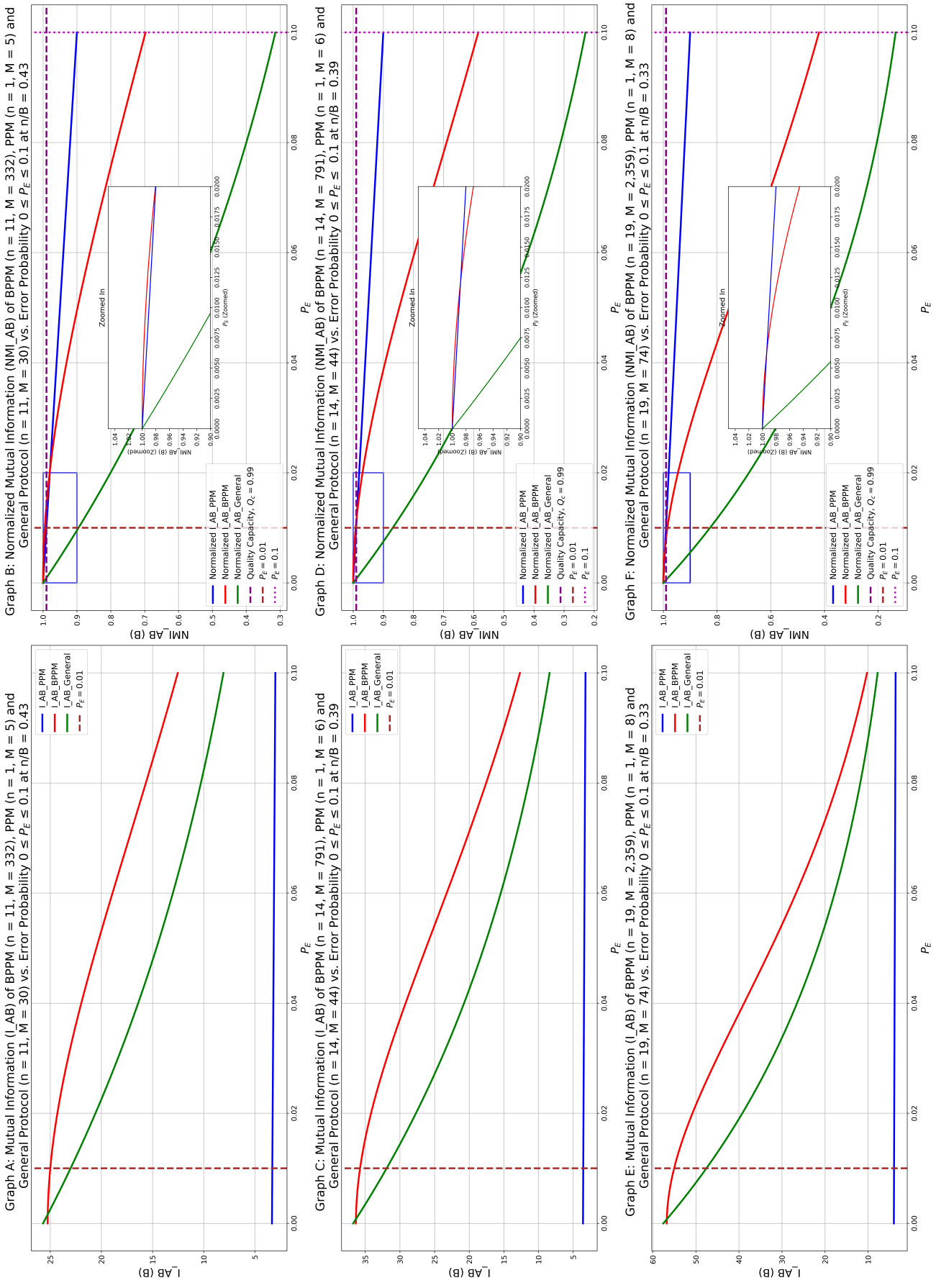

Figure 4.14: The diagram depicts the mutual information and normalized mutual information versus error probability for $0 \leq P_E \leq 0.1$ among the 3 protocols of BPPM, PPM, and General protocol based on the same energy values per information bit. The value of $\frac{E}{B} = 0.43, 0.39, 0.39$ were selected.

pared to other protocols under varying error probabilities and photon counts.
Overall, the 2 metrics collectively underline BPPM's unique characteristics and its suitability in transmission. While it excels in low error probabilities and demonstrates robust error correction capabilities, the advantage diminishes at longer distances. BPPM's consistent superiority and distinct curve shape in consistent superiority and distinct curve shape in the constant energy per information bit analysis further emphasize its effectiveness.
The comparison of normalized mutual information across the two analyses provides comprehensive information into BBPM's performance under various conditions, aiding in the design and optimization of energy-efficient communication systems.

4.5.7 More on Metrics Comparison

Results of 2 more metrics were generated for comparison: mutual information per photon, mutual information per time bin, and to further analyze the performance of different communication protocols. The two metrics provide information into the efficiency and tradeoffs of each protocol in utilizing photons and time bins for information transmission under different conditions. MI per photon quantifies the information transmitted per individual photon used in communication. The metric can gauge how effectively each protocol utilizes photons to convey information. MI per time bin measures the information transmitted per unit of time, allowing us to understand which protocol optimally allocates time bins to achieve the highest information transmission efficiency.

Protocol	M = 332
BPPM	$1 \leq n \leq 12$, $1 \leq M \leq 332$
PPM for 1 photon	$1 \leq M \leq 332$,
OOK	$1 \leq n \leq 166$, $1 \leq M \leq 332$

Table 4.7: The table shows the corresponding number of photons and time bins for the same number of time bins among BPPM, PPM, and General.

Figure 4.15 illustrates the two metrics evaluating the performance of BPPM, PPM, and OOK protocol in single-photon transmission.
The figure depicts 2 metrics comparison of BPPM ($3 \leq n \leq 7$), PPM ($4 \leq n \leq 76$) and OOK ($3 \leq n \leq 38$) at Probability P = 0 (right-handed side) and P = 0.1 (right-handed side). Graph A and Graph B show mutual information per photon ($\frac{I_{AB}}{n}$) Vs. Number of time bins (M). Graph C and Graph D show mutual information per number of time bins ($\frac{I_{AB}}{M}$) Vs. the number of time bins (M). The photon number's specific range on three protocols is chosen since their corresponding number of time bins are comparable. Logarithmic scales are used in plots to show linear relationships between variables, and both large and small values can be shown clearly. In mutual information per photon Vs. the number of time bins, the value in PPM increase more significantly than that in BPPM as the number of time bins increases, while the value in OOK is constant under both situations of error probability of

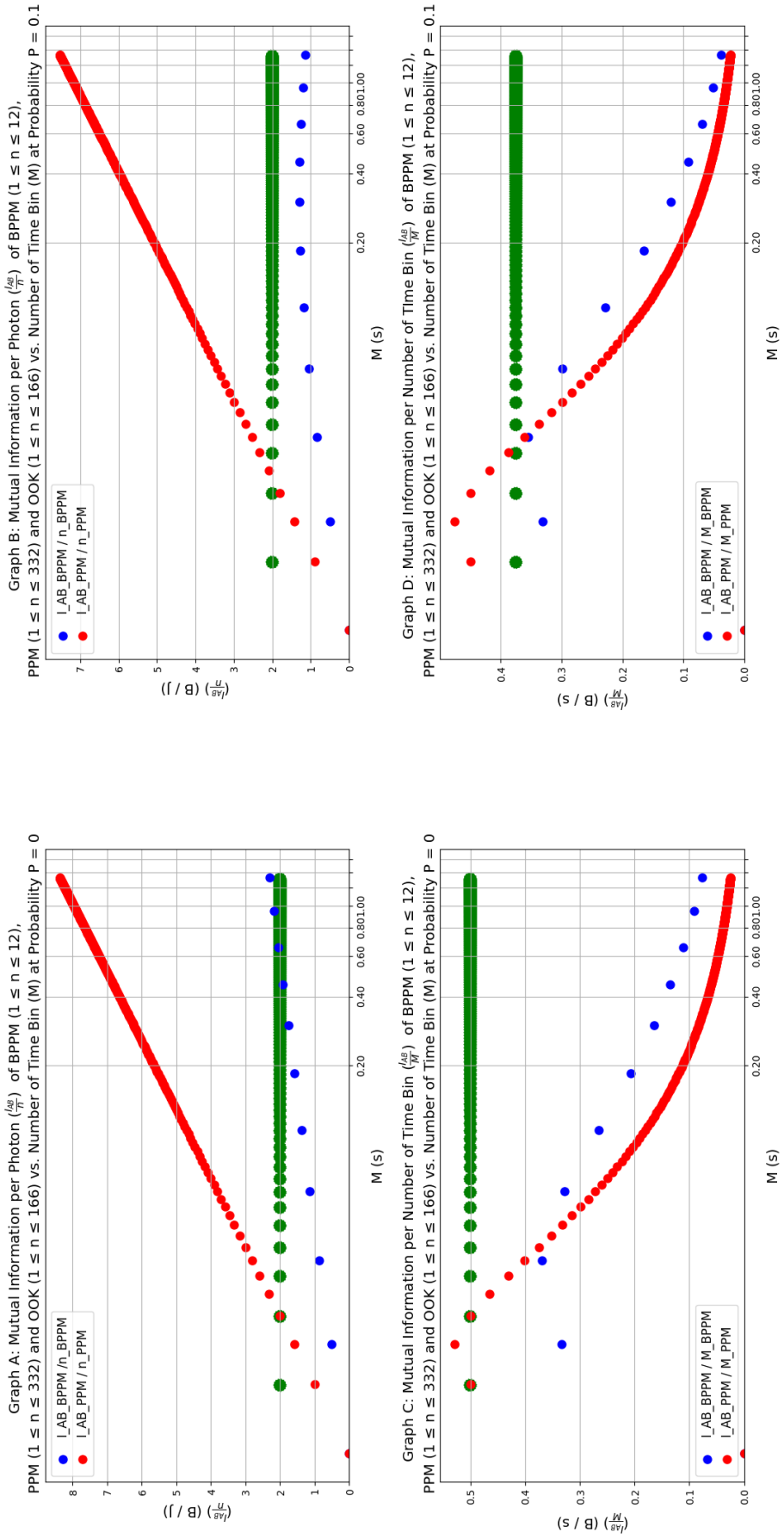

Figure 4.15: The figure depicts 2 metrics comparison of BPPM ($3 \leq n \leq 7$), PPM ($4 \leq n \leq 76$) and OOK ($3 \leq n \leq 38$) at Probability P = 0 (right-handed side) and P = 0.1 (right-handed side): Graph A and Graph B show mutual information per photon ($\frac{I_{AB}}{n}$) Vs. Number of time bins (M). Graph C and Graph D show mutual information per number of time bins ($\frac{I_{AB}}{M}$) vs. the number of time bins (M).

$P = 0$ and $P = 0.1$. In mutual information per number of time bin Vs the number of time bins, PPM, and BPPM reach a certain peak at M = 2, reaching 0.53 and 0.37 respectively at $P = 0$, and 0.48 and 0.35 respectively at $P = 0.1$. However, mutual information per the number of time bins in both protocols decreases as the number of time bins increases till they approach zero, with BPPM decreasing relatively slower than PPM. Similarly, the corresponding value in OOK stays constant as the number of time bins increases. PPM exhibits a more significant increase in mutual information per photon with increasing time bins than BPPM because PPM adopts single-photon telecommunication. At the same time, BPPM could have multiple photon transmission such that its information per photon decreases as the number of photons increase. In Both BPPM and PPM, the data rate can be increased by adding more pulses within a fixed time window. Initially, adding more pulses increases the data rate without causing significant overlap or confusion between adjacent pulses. This increment leads to a peak. After that, as the pulses start to be more closely packed to each other with the increase in pulse density, it becomes more challenging to distinguish and detect individual pulses accurately. This effect causes errors in pulse detection and a decrease in $\frac{I_{AB}}{M}$.

The trends in communication protocols have significant implications for the design and optimization of communication systems. Protocols should aim to transmit data reliably while minimizing energy consumption. This trend may involve optimizing transmission power based on the expected error probabilities.

Figure 4.16: The diagram described the telecommunication between Earth and Mars via Satellite in deep space.

Let the distance between Mars and Earth be d

$$5.6 \times 10^7 \quad \text{km} \leq d \leq 6.45 \times 10^8 \quad \text{km} \tag{4.6}$$

Considering the speed of light in optimal situation, the time for message transmis-

sion, t

$$4 \text{ minutes } 20 \text{ seconds } (260\text{s}) \leq t \leq 24 \text{ minute } 40 \text{ seconds } (1{,}480 \text{ seconds}) \quad (4.7)$$

When contemplating the transmission of a message from Earth to Mars using a satellite, it is necessary to assess the energy efficiency and error correction capabilities associated with Pulse Position Modulation (PPM) and Binary Pulse Position Modulation (BPPM).
Undoubtedly, the use of Pulse Position Modulation (PPM) enables the transmission of normalised information content with greater reliability compared to BPPM. Nonetheless, it is important to acknowledge the limitation of PPM, which lies in its ability to transmit only a single photon that carries all the information bits during the transmission process. The presence of lost or added errors in the transmission of photons can be highly problematic, particularly given the vast distance between Earth and Mars, spanning millions of kilometres. In perfect conditions, it takes from 4 to 24 minutes to transmit a single message between the two planets. If engaging in a process of reciprocal message exchange, all transmissions will be duplicated. If the photon becomes lost, it necessitates the repetition of all operations.
Moreover, it is worth noting that BPPM has the potential to possess a greater amount of information compared to Pulse Position Modulation (PPM), particularly when considering its error correcting capabilities within a specific error probability threshold. When considering the power ratio and energy per bit ratio, it can be shown that BPPM demonstrates greater energy efficiency compared to PPM. Hence, from a practical standpoint, it may be argued that BPPM demonstrates more reliability in the transmission of messages across extended distances.

5

Conclusion

We have investigated the capacity of a single photon to carry information through polarization and time ordering based on the patented protocol Beyond Pulse Position Modulation (BPPM), and we have compared BPPM information quality on various metrics against PPM, OOK, and General protocol using the Numpy library for numerical computation and matplotlib library for visualization. The result shows that BPPM preserves the highest information quality among other protocols over longer distances with limited power within error probability $P_E \leq 0.01$ over the quality criterion $Q_c = 0.99$, yet a relatively slower transmission speed in single-photon telecommunication.

Since we spend most of our time examining diverse metrics for comparative analysis and due to the limited scope of the master's thesis project, we would further work on developing an error correction technique to improve the reliability of BBPM. All the future works are put into Appendix for reference.

6

Appendix

6.1 Source Coding

Coding is a major application area of information theory, which can be used to remove unsystematic redundancy from message signals and induce systematic redundancy to correct errors caused by non-perfect practical channels. Ideal communication systems provide a meaningful basis for comparing the performance of realizable systems, and an understanding of coding and its impact on the design and performance of communication systems requires an understanding of basic concepts in information theory as mentioned in the previous chapter.

Information from a source that produced different symbols according to some probability scheme could be described by the entropy $H(X)$. Since entropy has units of bits per symbol, the symbol rate must be known in order to specify the source information rate in bits per second. In other words, the source information rate R, is given by

$$R_s = rH(x)\mathbf{bit/s} \tag{6.1}$$

where $H(X)$ is the source entropy in bits per symbol and r is the symbol rate in symbols per second.

6.2 Hamming Code

We can understand how codes can detect and correct errors from a geometric point of view. A binary codeword with a sequence of 1's and 0's has n symbols in length. The weights, namely Hamming weight $w(s_j)$ of codeword s_j is defined as the number if 1's in that codeword. We also define the number of positions in which s_i and s_j as Hamming distance $d(s_i, s_j)$ or d_{ij}.

Hamming distance can be written in terms of Hamming weight as

$$d_{ij} = w(s_i \bigoplus s_j)$$

where the symbol $\oplus$ denotes modulo-2-addition, which is binary addition without a carry.

In a broader sense, Hamming distance represents the sum of corresponding elements that differ between two vectors.

$$d(s_i, s_j) = \frac{1}{n} \sum_{n=1}^{N} s_i - s_j$$

A minimum decode can always correct as many as e errors, where e is the largest integer not to exceed

$$\frac{1}{2}(d_m - 1)$$

where d_m is the minimum distance between codewords. If d_m is odd, all received words can be assigned to a codeword. However, if d_m is even, a received word can lie halfway between two codewords. in this case, even though errors can be detected, they cannot be corrected.

6.3 Reed-Solomon Codes

The Reed-Solomon (RS) codes are the non-binary codes, they are important for the use in communication systems where errors appear in bursts rather than independent random errors.

RS codes were discovered by Reed and Solomon in 1960. The encoding process assumes a code of RS(N,K) which results in N code words of length N symbols each storing K symbols of data, being generated, that are then sent over an erasure channel. The non-binary BCH block codes have $2^m(\{0, 1, 2, \ldots, 2^m - 1\})$ symbols with block length $n = 2^m - 1$, which can be extended to $n = 2^m$ or $m = 2^m + 1$. RS codes can correct up to e_0 errors within a block of n symbols by using $n - k = n - 2e_0 = 2^m - 1 - 2e_0$ parity symbols. Or it can locate and correct up to $\frac{t}{2}$ erroneous symbols at unknown locations. As an erasure code, it can correct up to t erasures at location that are known and provided to the algorithm, or it can detect and correct combinations of errors and erasures.

RS code can achieve the maximum number of error correction by finding the largest possible $d_{min} = 2e_0 + 1$
Any combination of K code words received at the other end is enough to reconstruct all of the N code words. The code rate is generally set to $\frac{1}{2}$ unless the channel's erasure likelihood can be adequately modelled and is seen to be less. In conclusion, N is usually 2K , meaning that at least half of all the code words sent must be received in order to reconstruct all of the code words sent.

We construct the encoding matrix E with dimension $k \times n$ where k is the number of check packets and n is the number of data packets.

$$E = \begin{bmatrix} 1 & 1 & 6 \\ 4 & 3 & 2 \\ 5 & 2 & 2 \\ 5 & 3 & 4 \\ 4 & 2 & 4 \end{bmatrix}$$

When the $n \times 1$ vector of data words is multiplied by E, an $k \times 1$ vector of check values is produced.

$$\begin{bmatrix} 1 & 1 & 6 \\ 4 & 3 & 2 \\ 5 & 2 & 2 \\ 5 & 3 & 4 \\ 4 & 2 & 4 \end{bmatrix} \times \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 4 \\ 3 \\ 2 \end{bmatrix}$$

The values contained in the 8 'packets' that are sent: (0,4), (1,5), (2,6),(3,3),(4,5),(5,4),(6,3),(7,2). We observe that each of the 8 packets has an identifier that allows the recipient to determine exactly which packets of a Forward Error Correction (FEC) group have been received.

$$V = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 4 \\ 1 & 3 & 5 \\ 1 & 4 & 6 \\ 1 & 5 & 7 \\ 1 & 6 & 2 \\ 1 & 7 & 3 \end{bmatrix}$$

We have a transformed matrix D from Vandermonde matrix.

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 6 \\ 4 & 3 & 2 \\ 5 & 2 & 2 \\ 5 & 3 & 4 \\ 4 & 2 & 4 \end{bmatrix}$$

For a collection of n packets including both data and check packets have been received, we extract n rows of the D matrix corresponding the n received packets. We call this $n \times n$ matrix $D^{'}$.

$$D^{'} = \begin{bmatrix} 1 & 1 & 6 \\ 4 & 3 & 2 \\ 5 & 2 & 2 \end{bmatrix}$$

Inverting $D^{'}$ yields $D^{'-1}$.

$$D'^{-1} = \begin{bmatrix} 5 & 5 & 2 \\ 5 & 7 & 3 \\ 3 & 3 & 3 \end{bmatrix}$$

The receiver knows the algorithms by which the check packets were constructed as follow:

$$\begin{bmatrix} 1 & 1 & 6 \\ 4 & 3 & 2 \\ 5 & 2 & 2 \end{bmatrix} \times \begin{bmatrix} d_0 \\ d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix}$$

Multiplying the received check value by D'^{-1} recovers the original values.

$$\begin{bmatrix} 5 & 5 & 2 \\ 5 & 7 & 3 \\ 3 & 3 & 3 \end{bmatrix} \times \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Bibliography

[1] Pulse modulation - definition, types, block diagrams, pulse modulation width.

[2] Romain Alléaume, François Treussart, Gaëtan Messin, Yannick Dumeige, Jean-François Roch, Alexios Beveratos, Rosa Brouri-Tualle, Jean-Philippe Poizat, and Philippe Grangier. Experimental open-air quantum key distribution with a single-photon source. *New Journal of physics*, 6(1):92, 2004.

[3] DARRELL L Ash. A comparison between ook/ask and fsk modulation techniques for radio links. *RF Monothilics Inc, Dallas, TX, USA, Tech. Rep*, 1992.

[4] David Barnett, David Groth, and Jim McBee. *Cabling: the complete guide to network wiring*. John Wiley & Sons, 2006.

[5] Richard E Blahut. *Algebraic codes for data transmission*. Cambridge university press, 2003.

[6] Ian F Blake, Gérard Cohen, and Mikhail Deza. Coding with permutations. *Information and Control*, 43(1):1–19, 1979.

[7] Iliya Bouyukliev and Valentin Bakoev. A method for efficiently computing the number of codewords of fixed weights in linear codes. *Discrete Applied Mathematics*, 156(15):2986–3004, 2008.

[8] S Butman, Joseph Katz, and J Lesh. Bandwidth limitations on noiseless optical channel capacity. *IEEE Transactions on Communications*, 30(5):1262–1264, 1982.

[9] LUCID Vision Labs Modern Machine Vision Cameras. Polarization - a property of light, Jan 2018.

[10] Yen-Kuang Chen. Challenges and opportunities of internet of things. In *17th Asia and South Pacific design automation conference*, pages 383–388. IEEE, 2012.

[11] Sheri Edwards. Elements of information theory, thomas m. cover, joy a. thomas, john wiley & sons, inc.(2006), 2008.

[12] Marco Giordani and Michele Zorzi. Non-terrestrial networks in the 6g era: Challenges and opportunities. *IEEE Network*, 35(2):244–251, 2020.

[13] Ali Grami. Chapter 9 - information theory. In Ali Grami, editor, *Introduction to Digital Communications*, pages 377–408. Academic Press, Boston, 2016.

[14] Vehbi C Gungor and Frank C Lambert. A survey on communication networks for electric system automation. *Computer Networks*, 50(7):877–897, 2006.

[15] Dan Hoey. The on-line encyclopedia of integer sequences, May 2016.

[16] Abu Jahid, Mohammed H Alsharif, and Trevor J Hall. A contemporary survey on free space optical communication: Potentials, technical challenges, recent advances and research direction. *Journal of Network and Computer Applications*, 200:103311, 2022.

[17] Farzan Jazaeri, Arnout Beckers, Armin Tajalli, and Jean-Michel Sallese. A review on quantum computing: From qubits to front-end electronics and cryogenic mosfet physics. In *2019 MIXDES-26th International Conference" Mixed Design of Integrated Circuits and Systems"*, pages 15–25. IEEE, 2019.

[18] Christoforos Kachris, Konstantinos Kanonakis, and Ioannis Tomkos. Optical interconnection networks in data centers: recent trends and future challenges. *IEEE Communications Magazine*, 51(9):39–45, 2013.

[19] Malaram Kumhar and Jitendra Bhatia. Emerging communication technologies for 5g-enabled internet of things applications. *Blockchain for 5G-Enabled IoT: The new wave for Industrial Automation*, pages 133–158, 2021.

[20] J Lesh. Capacity limit of the noiseless, energy-efficient optical ppm channel. *IEEE Transactions on Communications*, 31(4):546–548, 1983.

[21] R McEliece. Practical codes for photon communication. *IEEE Transactions on Information Theory*, 27(4):393–398, 1981.

[22] Kostas Pentikousis. In search of energy-efficient mobile networking. *IEEE Communications Magazine*, 48(1):95–103, 2010.

[23] John Pierce. Optical channels: Practical limits with photon counting. *IEEE Transactions on Communications*, 26(12):1819–1821, 1978.

[24] Mirko Pittaluga, Mariella Minder, Marco Lucamarini, Mirko Sanzaro, Robert I Woodward, Ming-Jun Li, Zhiliang Yuan, and Andrew J Shields. 600-km repeater-like quantum communications with dual-band stabilization. *Nature Photonics*, 15(7):530–535, 2021.

[25] Joschka Roffe. Quantum error correction: an introductory guide. *Contemporary Physics*, 60(3):226–245, 2019.

[26] Nicolas Sangouard and Hugo Zbinden. What are single photons good for? *Journal of Modern Optics*, 59(17):1458–1464, 2012.

[27] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[28] Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[29] Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1):3–55, 2001.

[30] Peter W Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Physical review letters*, 85(2):441, 2000.

[31] Tristan BH Tentrup, Thomas Hummel, Tom AW Wolterink, Ravitej Uppu, Allard P Mosk, and Pepijn WH Pinkse. Transmitting more than 10 bit with a single photon. *Optics express*, 25(3):2826–2833, 2017.

[32] The Open University. Exploring communications technology.

[33] Bingrui Wang, Pingping Chen, Yi Fang, and Francis CM Lau. The design of vertical rs-crc and ldpc code for ship-based satellite communications on-the-move. *IEEE access*, 7:44977–44986, 2019.

[34] William K Wootters and Wojciech H Zurek. The no-cloning theorem. *Physics Today*, 62(2):76–77, 2009.

[35] Raymond W Yeung. *A first course in information theory*. Springer Science & Business Media, 2002.

[36] Yequn Zhang. Advanced coding techniques for fiber-optic communications and quantum key distribution. 2015.

[37] Rodger Ziemer and William H Tranter. *Principles of communications: system modulation and noise*. John Wiley & Sons, 2006.

Glossary

Conditional Entropy The quantification of information required to characterise the result of a random variable Y, given the knowledge of another random variable X, which is denoted as $H(Y|X)$ The function $H(Y|X)$ is the average uncertainty of the received symbol given that X is transmitted.$H(X|Y)$ is sometimes called equivocation, which measures the average uncertainty of the transmitted symbol after a symbol has been received.. 38

Entropy The measure of self-information associated with a random variable. It is the average uncertainty in information associated with a discrete random variable X with a probability mass function $p(x_j)$ is defined as the entropy $H(X)$. $H(X)$ is the average uncertainty of the source, whereas H(Y) is the average uncertainty of the received symbol. . 35

General Protocol A protocol where a fixed number of photons is arbitrarily placed among M time bins.. 20

Information Bit (IB) A measure of the data carried by a particular entity, such as a symbol, photon, or time bin, during transmission inside a communication system. IB is measured when error probability does not exist in the communication system.. 52

Mutual Information (MI) also referred to as information gain, measures the mutual dependence of two variables in a probabilistic system. It quantifies the amount of knowledge garnered about one random variable by observing another. MI is concerned with the statistical relationship between two random variables in bits and is particularly relevant when error probability is present in a communication system.. 33

Number of Photons A short hand to number of photons per symbol.. 13

Number of Time Bins A short hand to number of time bins per symbol.. 2

Quality Criterion (Q_c) A benchmark delineates a threshold for effective communication across the protocols.. 84

Raw Information Rate A communication protocol's maximum potential data transmission rate under ideal conditions, i.e., without error correction, overhead, or encoding.. 81

Time Bin A specific time interval or window within which multiple events or signals can be detected or measured.. 2, 11