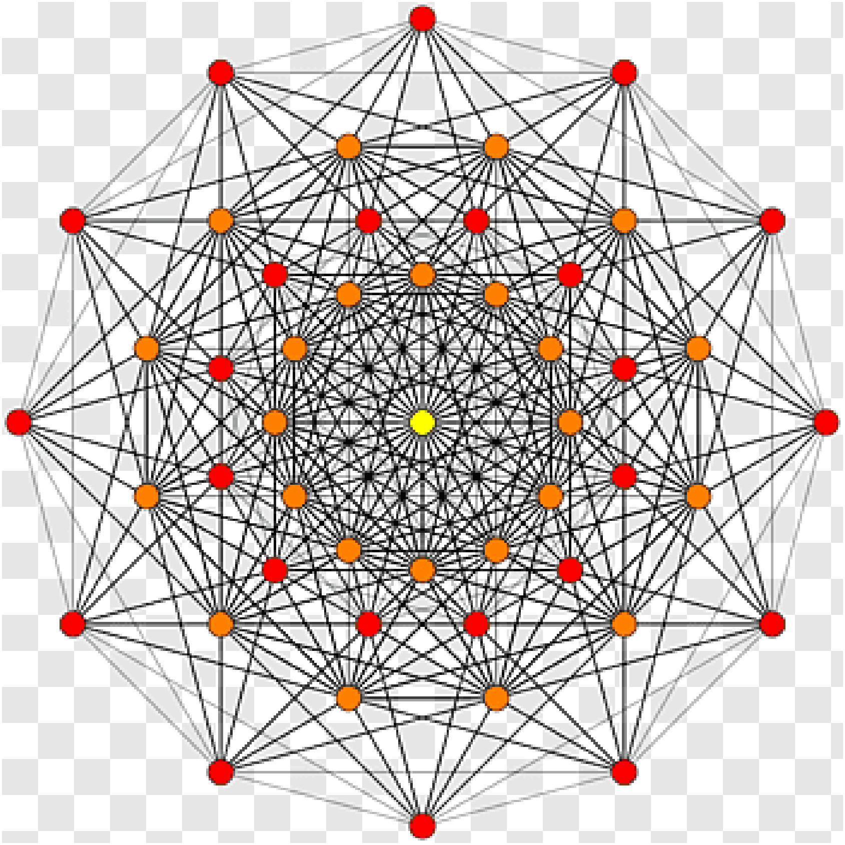
Entanglement-assisted zero-error communication

Stephen Adei

11 October 2023

Masterscriptie Wiskunde

Begeleiding: dr. Maris Ozols



Korteweg-de Vries Instituut voor Wiskunde
Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Universiteit van Amsterdam

Abstract

This thesis investigates the zero-error communication capacity of noisy classical channels in the presence of shared entanglement between the sender and receiver. It explores both entanglement-assisted and classical zero-error capacities for a family of channels constructed from symplectic hypergraphs. The main goal of this thesis is to understand the amplification of asymptotic zero-error capacity due to shared entanglement, as shown by Leung et al. [1], and to elaborate on their somewhat concise proof. Additionally, the thesis provides background material on linear algebra, quantum information, and finite fields, including the finite field trace. Finally we shortly elaborate on a connection between the symplectic graph $\mathrm{sp}(6, \mathbb{F}_2)$ and the E_7 root system, enabling the construction of orthogonal measurements and highlighting the interplay between symplectic graphs and the study of Lie algebras.

Titel: Entanglement-assisted zero-error communication
Auteur: Stephen Adei, stephen.adei@student.uva.nl, 11155000
Begeleiding: dr. Maris Ozols,
Einddatum: 11 October 2023

Korteweg-de Vries Instituut voor Wiskunde
Universiteit van Amsterdam
Science Park 904, 1098 XH Amsterdam
http://www.kdvi.uva.nl

Contents

1 Introduction

Entanglement-assisted communication is a subject of great interest as it offers the potential to overcome limitations in classical communication systems. By harnessing the power of entanglement, it is believed that information transmission can be significantly improved.

The capacity of a noisy channel can be viewed as the maximum achievable transmission rate for codes with vanishing error probability. In contrast, the zero-error capacity, denoted as C_0, is the maximum transmission rate achievable for codes with a perfect error-free performance. While C_0 appears to be a simpler characteristic than the general capacity, it poses computational challenges and raises unresolved questions. Our research focuses on this intriguing aspect and uncovers a remarkable phenomenon: the use of entangled states enables an increased number of error-free messages in certain classical channels. This finding highlights the potential of entanglement-assisted communication to enhance the transmission capabilities beyond what is achievable with unassisted communication.

In the exploration of this subject, one of our main findings is Theorem 6 which shows that the zero-error capacity C_0 of a channel constructed from the symplectic group over a binary field, $\mathrm{sp}(6, \mathbb{F}_2)$, is $\log 7$. However, the entanglement-assisted zero-error capacity C_0^E for the same channel stands at $\log 9$.

In Chapter 2, we provide the necessary background on linear algebra, quantum computing, and finite fields. We delve into key concepts such as hermitian and unitary matrices, the trace of a matrix, eigenvectors and eigenvalues, as well as the Kronecker product and partial trace operations used in quantum computing. Furthermore, we introduce the finite field trace, discussing its non-degeneracy, linearity, and surjectivity.

In Chapter 3, the concept of zero-error capacities in classical channels is discussed, focusing on confusability graphs and hypergraphs as ways of representing mutually confusable inputs. We present a protocol for entanglement-assisted communication and explore the conditions for achieving entanglement-assisted zero-error capacity. We also examine the connection between confusability graphs and zero-error capacity, highlighting the role of the independence number in determining the maximum number of error-free transmissions. The Lovász number is introduced as an upper bound for the Shannon capacity of hypergraphs, providing insights into the limits of the zero-error capacity.

In Chapter 4, a symplectic hypergraph channel is introduced using a binary symplectic vector space. Its confusability graph is based on the symplectic inner product that encodes orthogonality relationships between vectors in a binary symplectic space. Two simple examples, $\mathrm{sp}(2, \mathbb{F}_2)$ and $\mathrm{sp}(4, \mathbb{F}_2)$ are explored, after which we focus on the first non-trivial case, $\mathrm{sp}(6, \mathbb{F}_2)$.

Finally, in Chapter 5, we investigate the classical and entanglement-assisted zero-error capacities of the symplectic hypergraph channel. We establish that for the confusability graph $\mathrm{sp}(6, \mathbb{F}_2)$, the entanglement-assisted zero-error capacity exceeds the unassisted capacity. Specifically, we prove that the unassisted capacity is $\log 7$ using Haemers' bound, and in the entanglement-assisted case, the capacity is $\log 9$. Furthermore, we demonstrate a connection between $\mathrm{sp}(6, \mathbb{F}_2)$ and the E_7 root system, showing how the vertices of the symplectic graph correspond to elements of the root system. This connection allows us to construct orthogonal measurements used in the protocol for zero-error entanglement-assisted communication, highlighting the relationship between symplectic graphs and Lie algebras/Lie groups.

2 Background on linear algebra, quantum computing, and finite fields

2.1 Notions from linear algebra

We start by introducing some basic concepts from linear algebra and quantum computing based on [2, 3].

Let A be a matrix in $\mathbb{C}^{n\times n}$. We say that A is **hermitian** if $A^\dagger = A$, where $A^\dagger := \bar{A}^T = \overline{A^T}$. We say that A is **unitary** if $A^\dagger A = AA^\dagger = I$ where I denotes the **identity matrix**. The **trace** of A is defined as follows:

$$\mathrm{Tr}[A] = \sum_i \langle i| A |i\rangle ,$$

where $\langle i|$ and $|i\rangle$ denote the i'th standard basis row and column vectors of $\mathbb{C}^n$. It obeys the following **cyclic property**:

$$\mathrm{Tr}[AB] = \mathrm{Tr}[BA] \tag{2.1}$$

for any matrices A and B of compatible size.

If $A|\varphi\rangle = \lambda|\varphi\rangle$, we say that $|\varphi\rangle \in \mathbb{C}^n$ is an **eigenvector** of A with **eigenvalue** $\lambda \in \mathbb{C}$. The **eigendecomposition** of a matrix $A \in \mathbb{C}^{n\times n}$ is defined as

$$A = \sum_{i=1}^{n} \lambda_i |\varphi_i\rangle \langle\varphi_i| ,$$

where $|\varphi_i\rangle$ is the i'th eigenvector of A and λ_i is the corresponding eigenvalue. A matrix with such an eigendecomposition is called **diagonalisable**. Furthermore, a matrix A is diagonalisable exactly when $AA^\dagger = A^\dagger A$ (see Nielsen & Chuang [4]). Matrices with this property are called **normal**. Because of this we can say that a matrix is hermitian if it is diagonalisable and all its eigenvalues are real, i.e. $\lambda \in \mathbb{R}$. We say A is a **positive semi-definite** matrix if it is hermitian and all of its eigenvalues are non-negative, i.e., $\lambda_i \geq 0$. We denote the set of all positive semi-definite matrices on $\mathbb{C}^d$ by $\mathrm{PSD}(\mathbb{C}^d)$.

In quantum mechanics, the following operation that combines two matrices into a bigger one by multiplying them entry-wise plays a very important role.

Definition 1 (Kronecker product)**.** The **Kronecker product** or **tensor product** of a matrix $A \in \mathbb{C}^{p\times q}$ with a matrix $B \in \mathbb{C}^{r\times s}$ is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1q}B \\ \vdots & & \vdots \\ a_{p1}B & \cdots & a_{pq}B \end{bmatrix} ,$$

where $A \otimes B \in \mathbb{C}^{pr \times qs}$.

Conversely, we can make a big matrix on a tensor product space smaller by performing an operation that generalises the matrix trace introduced earlier.

Definition 2 (Partial trace)**.** For every linear operator $M_{AB} \in \mathbb{C}^{p\times p} \otimes \mathbb{C}^{r \times r}$, the **partial trace over** B is defined as follows:

$$\mathrm{Tr}_B[M_{AB}] := \sum_i (I_A \otimes \langle i|_B) M_{AB} (I_A \otimes |i\rangle_B),$$

where $|i\rangle_B$ is the standard basis of B and I_A is the identity operator on A.

Before we move on to quantum states we state another useful property from linear algebra. In quantum mechanics, the following lemma allows us to express states in different bases, depending on what is most convenient for the situation at hand. This can be useful when dealing with entangled systems. For example, we might be interested in changing the basis in which the state of a system is expressed to make certain quantities easier to compute. The proof of the lemma demonstrates the use of unitary matrices for changing bases, another common principal from quantum mechanics.

Lemma 1 (Orthogonal bases)**.** Let $|\varphi_1\rangle, \ldots, |\varphi_d\rangle$ be any orthonormal basis of $\mathbb{C}^d$. Then

$$\sum_{j=1}^{d} |j\rangle_A \otimes |j\rangle_B = \sum_{j=1}^{d} |\varphi_j\rangle_A \otimes \overline{|\varphi_j\rangle}_B. \tag{2.2}$$

Proof. Let $U := \sum_{i=1}^{d} |\varphi_i\rangle \langle i| \in \mathbb{C}^{d\times d}$ and note that U is unitary since

$$U^\dagger U = \left(\sum_{i=1}^{d} |i\rangle \langle\varphi_i|\right)\left(\sum_{j=1}^{d} |\varphi_j\rangle \langle j|\right) = \sum_{i,j=1}^{d} |i\rangle \langle\varphi_i|\varphi_j\rangle \langle j| = \sum_{i,j=1}^{d} \delta_{ij} |i\rangle\langle j| = \sum_{i=1}^{d} |i\rangle\langle i| = I \tag{2.3}$$

where we used the fact that $|\varphi_1\rangle, \ldots, |\varphi_d\rangle$ is an orthonormal basis. Note that

$$U|j\rangle = \sum_{i=1}^{d} |\varphi_i\rangle \langle i|j\rangle = \sum_{i=1}^{d} \delta_{ij} |\varphi_i\rangle = |\varphi_j\rangle \tag{2.4}$$

for every $j \in \{1, \ldots, d\}$. Using this,

$$\sum_{j=1}^{d} |\varphi_j\rangle_A \otimes \overline{|\varphi_j\rangle}_B = \sum_{j=1}^{d} U|j\rangle_A \otimes \overline{U}|j\rangle_B, \tag{2.5}$$

so our goal is to show that we can remove U and $\overline{U}$ from this expression.

Note that $U|j\rangle = \sum_{i=1}^{d} U_{ij}|i\rangle$ where $U_{ij} := \langle i| U |j\rangle$ is the ij-th entry of U. Inserting this in eq. (2.5) gives us

$$\begin{aligned}
\sum_{j=1}^{d}\left(\sum_{i=1}^{d} U_{ij}|i\rangle\right)_A \otimes \left(\sum_{k=1}^{d} \overline{U_{kj}}|k\rangle\right)_B &= \sum_{i,j,k=1}^{d} U_{ij}\overline{U_{kj}}\,|i\rangle_A \otimes |k\rangle_B && (2.6)\\
&= \sum_{i,k=1}^{d} \delta_{ik}\,|i\rangle_A \otimes |k\rangle_B && (2.7)\\
&= \sum_{i=1}^{d} |i\rangle_A \otimes |i\rangle_B\,, && (2.8)
\end{aligned}$$

where we used the fact that $\sum_{j=1}^{d} U_{ij}\overline{U_{kj}} = \sum_{j=1}^{d} U_{ij}(U^\dagger)_{jk} = \langle i| UU^\dagger |k\rangle = \langle i| I |k\rangle = \delta_{ik}$ since U is unitary. □

2.2 Quantum states and measurements

The following section covers quantum states and how to measure them to extract classical information. We will need these concepts when describing entanglement-assisted communication protocols later on. We'll start by defining some key terms and then we'll get into how these concepts interact with each other.

Definition 3 (Density operator)**.** A **density operator** is a positive semi-definite linear operator (PSD) with trace one. We denote the set of density operators on a Hilbert space $\mathcal{H}$ by

$$D(\mathcal{H}) = \{\rho \in \mathrm{PSD}(\mathcal{H}) \mid \mathrm{Tr}[\rho] = 1\}.$$

In the world of quantum mechanics, a density operator is like a rule book that tells us about the **state** of a quantum system which is described by a density operator acting on the underlying Hilbert space $\mathcal{H}$ of the quantum system. A state is identified with its corresponding density operator. In quantum mechanics, we distinguish between pure states and mixed states. The former, **pure states**, can be represented in the form $\rho = |\varphi\rangle\langle\varphi|$, which effectively correspond to Hilbert space vectors $|\varphi\rangle \in \mathcal{H}$ of unit length. Pure states are those where we have complete knowledge about the quantum system, and they can be represented by a state vector in a Hilbert space, a complex vector space with an inner product. The state vector uniquely determines the pure state, and all observables of the state can be derived from the state vector.

The latter, **mixed states**, are not expressible in this particular form. Mixed states are used when we only have partial knowledge about the quantum system. These states cannot be described by a single state vector, as we will see later on. Instead, they are represented by a density matrix, which is a positive semi-definite, hermitian operator acting on a Hilbert space, with trace equal to one. The density matrix for a mixed state can be thought of as a probabilistic mixture of pure states.

If $|\varphi\rangle$ is a unit vector in Hilbert space $\mathcal{H}$ then

$$P = |\varphi\rangle\langle\varphi|$$

is the rank-1 **orthogonal projection** or the **orthogonal projector** onto the one-dimensional space $\mathbb{C}|\varphi\rangle$. More generally, a matrix $P \in \mathrm{PSD}(\mathcal{H})$ is an orthogonal projection if

$$P^2 = P.$$

To clarify, orthogonal projections, such as $P = |\varphi\rangle\langle\varphi|$, are indeed examples of density operators which uniquely describe pure quantum states. They also represent the simplest observations, projecting onto one-dimensional subspaces of the underlying Hilbert space $\mathcal{H}$. However, these observations can be more general and may also involve orthogonal projections onto higher-dimensional subspaces or even arbitrary positive semi-definite matrices, thus enabling a more comprehensive understanding of the quantum system. The only way to access information in a quantum system is by making a measurement on the system. This is where **Positive Operator-Valued Measures (POVMs)** come into play. A POVM is a collection of measurement operators that act on a quantum system. Each operator corresponds to a possible outcome of the measurement.

Definition 4 (Positive Operator-Valued Measure (POVM) or measurement)**.** A **POVM** or **measurement** on a Hilbert space $\mathcal{H}$ with outcomes in some finite set Ω is a function

$$\mu : \Omega \to \mathrm{PSD}(\mathcal{H}) \text{ such that } \sum_{x\in\Omega} \mu(x) = I.$$

If all $\mu(x)$ are orthogonal projectors, then we say that μ is **projective**. We will denote by $\mathscr{M}(\mathcal{H}, \Omega)$ the set of all measurements on $\mathcal{H}$ with outcomes in Ω.

We can interpret these measurements using Born's rule.

Definition 5 (Born's rule)**.** If we measure a quantum system in state $\rho \in D(\mathcal{H})$ using a measurement μ, then the probability of outcome $x \in \Omega$ is given by **Born's rule**:

$$p(x) := \mathrm{P}(\text{outcome } x) = \mathrm{Tr}[\mu(x)\rho].$$

When $\rho = |\varphi\rangle\langle\varphi|$ is pure, we can equivalently write

$$p(x) = \langle\varphi|\,\mu(x)\,|\varphi\rangle$$

by using the cyclic property of trace, see eq. (2.1).

Born's rule is like a quantum dice roll that produces random measurement outcomes. It's the rule that helps us to determine the outcomes of quantum measurements and learn some information about the system. Now, consider that we're not dealing with just a single quantum system, but multiple ones. Each of these systems has its own Hilbert space. When we're interested in the behavior of these systems as a whole, we need to combine these individual state spaces.

Definition 6 (Composing systems)**.** For a quantum system composed of n subsystems with Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, the overall Hilbert space is given by the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$.

When dealing with a single quantum system, you describe its state using a density operator acting on a Hilbert space specific to that system. This is a fundamental aspect of quantum mechanics. Once you start working with multiple quantum systems, you have to consider a density operator on the combined Hilbert space of all the systems. When the state of each subsystem is independent from each other, the overall state is constructed as the tensor product of the individual states.

Definition 7 (Product state)**.** A state $\rho \in D(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$ is called a **product state** if

$$\rho = \rho_1 \otimes \cdots \otimes \rho_n,$$

for $\rho_i \in D(\mathcal{H}_i), i = 1, \ldots, n$. A state that is not a product state is called **correlated**.

The states of these combined systems can take on different forms. We distinguish between them as follows.

Definition 8 (Separable and entangled states)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two Hilbert spaces. A quantum state $\rho_{AB} \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ is called **separable** or **unentangled** (between systems A and B) if it is a convex combination of product states, i.e., if there is a probability distribution $(p_i)_{i \in I}$ and states $\rho_{A,i} \in D(\mathcal{H}_A), \rho_{B,i} \in D(\mathcal{H}_B)$ such that

$$\rho_{AB} = \sum_{i \in I} p_i \; \rho_{A,i} \otimes \rho_{B,i}.$$

A state that is not separable is called **entangled**.

In the world of quantum systems, measurements often play a crucial role, and this is even more true when we consider the joint system and apply a **partial measurement**.

Definition 9 (Born's rule for a partial measurement)**.** If a joint system is in state ρ_{AB} and we apply a measurement $\mu_A : \Omega_A \to \mathrm{PSD}(\mathcal{H}_A)$ on A, the probability of outcome $x \in \Omega_A$ is calculated as follows:

$$p(x) = \mathrm{Tr}\left[(\mu_A(x) \otimes I_B)\rho_{AB}\right].$$

In particular, if $\rho_{AB} = |\phi\rangle \langle\phi|_{AB}$ is a pure state then

$$p(x) = \langle\phi|_{AB} \left(\mu_A(x) \otimes I_B\right) |\phi\rangle_{AB}.$$

Furthermore, each subsystem of a joint system can be assigned its own state, known as the **reduced state**.

Definition 10 (Reduced state)**.** Given a state ρ_{AB} on AB, we define its **reduced state** on subsystem A by $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$, where Tr_B denotes the partial trace from Definition 2.

Let's consider a situation where we apply a measurement to a subsystem of a joint system and obtain an outcome. The state of the remaining subsystem after this measurement can be calculated as follows:

Definition 11 (Post-measurement state on the remaining subsystem)**.** Suppose a quantum system is in state ρ_{AB}, and we apply measurement $\mu_A : \Omega_A \to \mathrm{PSD}(\mathcal{H}_A)$ on A and obtain outcome $x \in \Omega_A$. Then the **post-measurement state** of the remaining system B after the measurement is given by

$$\rho_B(x) = \frac{\mathrm{Tr}_A[(\mu_A(x) \otimes I_B)\rho_{AB}]}{\mathrm{Tr}[(\mu_A(x) \otimes I_B)\rho_{AB}]} = \frac{\mathrm{Tr}_A[(\mu_A(x) \otimes I_B)\rho_{AB}]}{p(x)}.$$

2.3 Finite fields

This section provides an introduction to finite fields and the concept of finite field trace, along with an explanation of their fundamental properties. These concepts will be applied in Chapter 4 to create a certain classical communication channel based on a vector space over a finite field. The main advantage of using finite fields to describe a channel is that one can then deploy tools from linear algebra to analyze its transmission capabilities.

A **finite field** is a structure that consists of a finite set of elements and two operations, called addition and multiplication. The elements of the field must satisfy certain properties, such as existence of an additive and multiplicative identity, the ability to add and multiply any two elements together, and the existence of additive and multiplicative inverses. The two operations must be compatible, in the sense that multiplication distributes over addition.

The notation $\mathbb{F}_p$ or $\mathrm{GF}(p)$ is used to indicate a finite field with p elements, where the "GF" stands for "Galois field," named after Evariste Galois, who first studied finite fields. A finite field with p elements, where p is a prime number, is isomorphic to the ring of integers modulo p:

$$\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z},$$

giving us an infinite family of familiar examples. However, there are additional finite fields that are not of this form.

An **extension field** $\mathbb{F}_q$ is a field that contains another field, called the base field $\mathbb{F}_{q'}$, as a subfield. In general the sizes q and q' of an extension and base field are related follows: $q := (q')^m$, for some **extension degree** $m \in \mathbb{N}$. We will only be interested in the case when $q' = p$, a prime. Moreover, in Chapter 4 we will mostly consider the special case when $p = 2$. From now on we will use the convention that

$$q = p^m, \tag{2.9}$$

where p is any prime, and treat $\mathbb{F}_q = \mathbb{F}_{p^m}$ as an extension field over the base field $\mathbb{F}_p$. Note that in general, a finite field of size q exists if only if q is a prime power.

There are a number of interpretations of an extension field and its elements. One way to see them is as polynomials with degree of at most $m-1$ with coefficients in the base field:

$$\mathbb{F}_q = \{\sum_{i=0}^{m-1} a_i X^i \mid a_i \in \mathbb{F}_p\}.$$

The **degree** $\deg(f(X))$ of a polynomial $f(X) = \sum_{i=0}^{m-1} a_i X^i$ is the largest n such that $a_n \neq 0$. We say a_n is the **leading coefficient** of $f(X)$. If a polynomial has leading coefficient 1, we call the polynomial **monic**. It is easy to see that when $a_i = 0$ for $i = 1, \ldots, m-1$ the polynomial reduces to $f(X) = a_0 \in \mathbb{F}_p$ and we get exactly the subfield $\mathbb{F}_p \subseteq \mathbb{F}_q$, the set of all constant polynomials f. Note that for such a polynomial $\deg(f) = 0$.

Let us briefly discuss how the two field operations are defined when we interpret the extension field elements as polynomials. The operation of addition in the field $\mathbb{F}_q$ is defined as follows:

$$\left(\sum_{i=0}^{m-1} a_i \cdot X^i\right) + \left(\sum_{i=0}^{m-1} b_i \cdot X^i\right) = \left(\sum_{i=0}^{m-1} (a_i + b_i) \cdot X^i\right).$$

Before we move on to defining multiplication, we need to introduce the concept of **irreducible** and **minimal** polynomials.

Definition 12 (Irreducible polynomial)**.** An **irreducible polynomial** is a polynomial $f(X) = \sum_{i=0}^{m} a_i X^i \in \mathbb{F}_q$ which cannot be factored into the product of two polynomials $g, h \in \mathbb{F}_q$ such that $f(X) = g(X)h(X)$ where $\deg(g), \deg(h) < \deg(f)$.

A closely related concept is that of minimal polynomials.

Definition 13. We say a polynomial f is the **minimal polynomial of** α **over** $\mathbb{F}_p$, often denoted as $f_{\mathbb{F}_p}^{\alpha}$ or f^{α}, if f is monic and of minimal degree m such that $f(\alpha) = 0$. If this is the case, then we must have for any other polynomial g that if $g(\alpha) = 0$, then $m \leq \deg(g)$.

The following lemma shows that these two concepts are closely related.

Lemma 2. Consider an irreducible polynomial f. The polynomial f serves as the minimal polynomial for any root α that satisfies the equation $f(x) = 0$. Conversely, a minimal polynomial of is irreducible.

Proof. Let α be the root of an irreducible polynomial f. Let's assume that f is not the minimal polynomial of α. Therefore a minimal polynomial $f^{\alpha}(x)$ of α with $\deg(f^{\alpha}) < \deg(f)$ such that $f^{\alpha}(\alpha) = 0$. We can then write

$$f = f^{\alpha} \cdot h + r,$$

where Applying polynomial Euclidean algorithm we might decompose f as follows :

$$\deg(r) < \deg(f^{\alpha}). \tag{2.10}$$

We can distinguish two cases:

Case 1: $r \equiv 0$**.** In this case we get

$$f = f^{\alpha} \cdot h,$$

giving us a factorization of f, which contradicts the assumption that f is irreducible. **Case 2:** $r \not\equiv 0$**.** Since $f(\alpha) = f^{\alpha}(\alpha) = 0$, we get $r(\alpha) = 0$. But since $\deg(r) < \deg(f^{\alpha})$, we find α is also a root of r, which is of lower degree than f^{α}, thus contradicting the assumption that f^{α} is a minimal polynomial. As both assumptions lead to a contradiction, f^{α} is irreducible.

We prove the converse by contradiction. Consider a minimal polynomial f^{α} of α. Suppose $f^{\alpha} = g \cdot h$ is reducible. Then $g(\alpha) = 0$ or $h(\alpha) = 0$, where of course $\deg(g), \deg(h) < \deg(f^{\alpha})$. This contradicts a condition. □

Note that if $f(X)$ is a monic polynomial of $\deg(f) = m$ then $f(X) = 0$ is equivalent to the following

$$X^m = -a_{m-1}X^{m-1} - \cdots - a_1 X - a_0, \tag{2.11}$$

This can be used to define a multiplicative structure in $\mathbb{F}_q$ Indeed multiplication of polynomials is determined by the rule

$$(a_i X^i) \cdot (a_j X^j) = (a_i \cdot b_i) X^{i+j},$$

giving us

$$\left(\sum_{i=0}^{m-1} a_i \cdot X^i\right) \cdot \left(\sum_{i=0}^{m-1} b_i \cdot X^i\right) = \sum_{k=0}^{2m-2} \left(\sum_{i+j=k} a_i b_j\right) \cdot X^k.$$

While this polynomial likely has degree larger than $m-1$, it can always be reduced below m by repeatedly applying eq. (2.11).

Here is an example of such reduction when multiplying the elements of $\mathbb{F}_q$ with $q = 2^3$. Take irreducible polynomial $f(X) = X^3 + X + 1$. Since

$$f(X) = 0 \Longleftrightarrow X^3 = X + 1,$$

we can reduce the following product of two degree-2 polynomials from a degree-4 polynomial to a degree-2 one as follows:

$$\begin{aligned}(1 + X^2)(1 + X + X^2) &= 1 + X + X^2 + X^2 + X^3 + X^4 \\ &= 1 + X + (X + 1) + X(X + 1) \\ &= X + X^2,\end{aligned}$$

which again lies in $\mathbb{F}_q$.

Another interpretation of an extension field is the following. Let $\alpha \in \mathbb{F}_q$ be a root of an irreducible polynomial of degree m. Then the following set consists of complex numbers:

$$\mathbb{F}_p(\alpha) = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} \mid n \in \mathbb{Z}_{\geq 0}, a_0, \ldots, a_n \in \mathbb{F}_p\}$$

equipped with usual multiplication and and addition is isomorphic to $\mathbb{F}_q$ is a subfield of $\mathbb{F}_q$. We define the following evaluation homomorphism for $a \in \mathbb{F}_q$

$$\Phi : \mathbb{F}_q \to \mathbb{F}_p(\alpha) : f(X) \mapsto f(a). \tag{2.12}$$

Lemma 3. Let α be the root of f^α, a monic irreducible polynomial over $\mathbb{F}_p$. Then

$$\mathbb{F}_p(\alpha) \cong \mathbb{F}_q =: \mathbb{F}_p[X]/(f^\alpha).$$

Proof. First, note that Φ in eq. (2.12) is surjective. Furthermore, by Lemma 2 we can conclude that f^α generates $\mathrm{Ker}(\Phi)$, since any other polynomial g with $g(\alpha) = 0$ can be divided by f^α. By the first homomorphism theorem we then find the desired result. □

It is easy to see that Φ preserves the multiplication and addition structure.

This lemma gives us the freedom to choose our interpretation of the finite field based on the property we are proving. To read more on rings and minimal polynomials we refer to [5].

Using Theorem 12.5 from [5] we see that the following property holds for all $x \in \mathbb{F}_q$:

$$x^q = x, \tag{2.13}$$

which will be useful for us later.

2.4 Finite field trace

One of the key concepts we will encounter in the Chapter 4 is the finite field trace. It will play significant role in our analysis of symplectic hypergraphs. The finite field trace, in the context of our investigation, serves as a tool for understanding the non-degeneracy of the *symplectic form* which is introduced in Section 4.1. The following lemmas walk us through proving non-degeneracy of the finite field trace, which will be a key ingredient later.

Lemma 4. Let $\mathbb{F}_q$ be a finite field with $q = p^m$ elements, where p is a prime number and m is a positive integer. Then, for all $a, b \in \mathbb{F}_q$ and integers $n \geq 0$, the following holds:

$$(a \pm b)^{p^n} = a^{p^n} \pm b^{p^n}.$$

Proof. We first prove the identity $(a+b)^{p^n} = a^{p^n} + b^{p^n}$ by induction on n. The base case $n = 1$ follows from the binomial theorem, since $(a+b)^p = a^p + \binom{p}{1} a^{p-1} b + \cdots + b^p = a^p + b^p$ using the fact that p is a prime and divides all binomial coefficients $\binom{p}{k}$ for $0 < k < p$.

Now, suppose that the identity holds for $n = k$, i.e., $(a + b)^{p^k} = a^{p^k} + b^{p^k}$ for all $a, b \in \mathbb{F}_q$. We need to show that it also holds for $n = k + 1$. We have:

$$\begin{aligned}(a + b)^{p^{k+1}} &= (a + b)^{p^k \cdot p} \\ &= ((a + b)^{p^k})^p \\ &= (a^{p^k} + b^{p^k})^p \\ &= a^{p^{k+1}} + b^{p^{k+1}}\end{aligned}$$

where we used the inductive assumption and the base case. This completes the induction step, and so the identity holds for all n. Note that by a similar argument it also follows that

$$(a+b+c)^{p^n} = (a+(b+c))^{p^n} = a^{p^n} + (b+c)^{p^n} = a^{p^n} + b^{p^n} + c^{p^n},$$

where $a, b, c \in \mathbb{F}_q$.

The identity $(a-b)^{p^n} = a^{p^n} - b^{p^n}$ can be proved in a similar way, using the binomial theorem and the fact that p divides all binomial coefficients $\binom{p}{k}$ for $0 < k < p$. We omit the details. □

Additionally, we require a property that addresses the number of solutions for a polynomial. By applying a universal theorem of algebra, we can establish the ensuing lemma.

Lemma 5 (Number of solutions of a polynomial [6])**.** If $f(x)$ is a polynomial of degree m with coefficients in a field $\mathbb{F}_p$, then the equation $f(x) = 0$ can have at most m distinct solutions in $\mathbb{F}_q \cong \mathbb{F}_p(\alpha)$.

Recognizing the non-trivial nature of the finite field trace is key as it reflects the function's distinctiveness and linear properties, instrumental in advanced computations within the finite field context. We define the finite field trace as follows.

Definition 14 (Finite field trace)**.** The **finite field trace** on $\mathbb{F}_{p^m}$, as an extension over $\mathbb{F}_p$, is defined as

$$\mathrm{Tr}(x) := x + x^p + x^{p^2} + \cdots + x^{p^{m-1}} = \sum_{i=0}^{m-1} x^{p^i}$$

for any $x \in \mathbb{F}_{p^m}$.

Lemma 6 (Trace is not trivial)**.** There exists $x \in \mathbb{F}_q$ such that $\mathrm{Tr}[x] \neq 0$.

Proof. Note by Lemma 5 that the polynomial $\mathrm{Tr}[x] = x + x^p + x^{p^2} + \cdots + x^{p^{m-1}} = 0$ has at most p^{m-1} solutions. Since $\mathbb{F}_p(\alpha)$ contains only p^m elements, we conclude that the there exists $x \in \mathbb{F}_p(\alpha)$ such that $\mathrm{Tr}[x] \neq 0$. □

To ensure an element lies in a finite field, one can check if the requirements of the following statement are met.

Lemma 7 (Identifying base field elements)**.** Let $x \in \mathbb{F}_q$. Then $x^p = x$ if and only if $x \in \mathbb{F}_p$.

Proof. First we prove the statement from right to left ($\Leftarrow$). Fermat's Little Theorem states that if p is a prime and a is an integer not divisible by p, then $a^{p-1} \equiv 1 \pmod{p}$. Applying this to the finite field $\mathbb{F}_p \cong \mathbb{Z}_p$, we have that if $x \in \mathbb{F}_p$, then $x^{p-1} \equiv 1 \pmod{p}$. Multiplying both sides of the equation by x gives us the desired result.

Now we prove the other implication from left to right ($\Rightarrow$). For this part, we shall work with the second definition of finite fields as described in 2.3. Suppose that $x \in \mathbb{F}_q$

is such that $x^p = x$. We need to show that $x \in \mathbb{F}_p$. Let us rewrite the equation as a polynomial:

$$f(x) := x^p - x = 0.$$

According to Lemma 5, this polynomial has at most p solutions. As we showed, we have proven that all elements $x \in \mathbb{F}_p$ are solutions to this equation. Consequently, it is not possible for any other field element to satisfy it. □

It is obvious from this definition that $\mathrm{Tr}(x) \in \mathbb{F}_{p^m}$. However, it turns out that in fact the stronger statement $\mathrm{Tr}(x) \in \mathbb{F}_p \subseteq \mathbb{F}_{p^m}$ is also true.

Lemma 8 (The trace takes values in the base field)**.** $\mathrm{Tr}(x) \in \mathbb{F}_p$ for all $x \in \mathbb{F}_q$.

Proof. Using Lemma 7 and eq. (2.13), it is enough to show that $(\mathrm{Tr}(x))^p = \mathrm{Tr}(x)$. Choose arbitrary $x \in \mathbb{F}_q$, the following holds:

$$\begin{aligned}
\mathrm{Tr}[x]^p &= (x + x^p + x^{p^2} + \cdots + x^{p^{m-2}} + x^{p^{m-1}})^p \\
&= x^p + x^{p^2} + x^{p^3} + \cdots + x^{p^{m-1}} + x^{p^m} \\
&= x^p + x^{p^2} + x^{p^3} + \cdots + x^{p^{m-1}} + x \\
&= x + x^p + x^{p^2} + x^{p^3} + \cdots + x^{p^{m-1}} \\
&= \mathrm{Tr}[x],
\end{aligned}$$

where we made use of Lemma 4 in the second equality. This finishes the proof □

Furthermore, we have that the trace is linear over the base field.

Lemma 9 (The trace is linear over the base field)**.** $\mathrm{Tr}(ax + by) = a\mathrm{Tr}(x) + b\mathrm{Tr}(y)$ for all $x, y \in \mathbb{F}_q$ and $a, b \in \mathbb{F}_p$.

Proof. Notice that for arbitrary $c \in \mathbb{F}_p$, we have $c^p = c$, which implies that $c^{p^i} = c$ for every $i \in \mathbb{N}$. We find from Definition 14 that

$$\begin{aligned}
\mathrm{Tr}[ax + by] &= (ax + by) + (ax + by)^p + (ax + by)^{p^2} + \cdots + (ax + by)^{p^{m-1}} \\
&= ax + by + (ax)^p + (by)^p + (ax)^{p^2} + (by)^{p^2} + \cdots + (ax)^{p^{m-1}} + (by)^{p^{m-1}} \\
&= ax + (ax)^p + (ax)^{p^2} + \cdots + (ax)^{p^{m-1}} + by + (by)^p + (by)^{p^2} + \cdots + (by)^{p^{m-1}} \\
&= ax + a^p x^p + a^{p^2} x^{p^2} + \cdots + a^{p^{m-1}} x^{p^{m-1}} + by + b^p y^p + b^{p^2} y^{p^2} + \cdots + b^{p^{m-1}} y^{p^{m-1}} \\
&= a\mathrm{Tr}(x) + b\mathrm{Tr}(y).
\end{aligned}$$

□

The linearity of the finite field trace plays a pivotal role in establishing its surjectivity, ensuring the trace covers the entire base field.

Lemma 10. The finite field trace is onto.

Proof. By Lemma 6 there exists $x \in \mathbb{F}_q$ such that $\mathrm{Tr}[x] \neq 0$. By linearity (Lemma 9), $\mathrm{Tr}[ax] = a\mathrm{Tr}[x]$ for any $a \in \mathbb{F}_p$, implying that the range of the trace is the entire base field $\mathbb{F}_p$. □

The following lemma then utilises this property to establish that distinct elements in $\mathbb{F}_q$ have distinct trace maps.

Lemma 11. For any element $b \in \mathbb{F}_q$, we define the following map

$$L_b(x) := \mathrm{Tr}[bx].$$

Then $b \neq c \in \mathbb{F}_q \implies L_b \neq L_c$.

Proof. The trace is a linear map (Lemma 9), thus, the map L_b is linear as well. Assume $b \neq c$, then $b - c \neq 0$. Now by Lemma 6 we can choose an $a \in \mathbb{F}_q$ such that $\mathrm{Tr}[a] \neq 0$ and define $a' := (b-c)^{-1}a$. Now we find

$$\mathrm{Tr}[ba'] - \mathrm{Tr}[ca'] = \mathrm{Tr}[(b-c)a'] = \mathrm{Tr}[a] \neq 0.$$

Which leads us to

$$\mathrm{Tr}[ba'] - \mathrm{Tr}[ca'] = L_b(a') - L_c(a') \neq 0 \implies L_b \neq L_c$$

as claimed. □

Using the established lemmas, we can now prove the non-degeneracy of the finite field trace.

Lemma 12 (The trace is non-degenerate)**.** Let $x \in \mathbb{F}_q$. If $\mathrm{Tr}(xy) = 0$ for all $y \in \mathbb{F}_q$ then $x = 0$.

Proof. By Definition 14 we have

$$\begin{aligned} \mathrm{Tr}(xy) &= xy + (xy)^p + (xy)^{p^2} + \cdots + (xy)^{p^{m-1}} \\ &= xy + x^p y^p + x^{p^2} y^{p^2} + \cdots + x^{p^{m-1}} y^{p^{m-1}}. \end{aligned}$$

Evidently $\mathrm{Tr}(xy) = 0$ when $x = 0$.

On the other hand, by Lemma 11 we know that $L_0 \neq L_x$ when $x \neq 0$. This means that for any $0 \neq x \in \mathbb{F}_q$ there exists a $y \in \mathbb{F}_q$ such that $L_0(y) \neq L_x(y)$. In other words,

$$0 = \mathrm{Tr}[0y] \neq \mathrm{Tr}[xy].$$

We conclude that if $x \neq 0$ then $\mathrm{Tr}[xy] \neq 0$ for some $y \in \mathbb{F}_q$. □

3 Zero-error capacities

This chapter introduces classical channels and their memoryless property. In it we define confusability graphs, and later hypergraphs, to represent relationships between confusable inputs of a channel. We need these concepts to understand what entanglement-assisted zero-error capacity entails. The chapter also presents a protocol for entanglement-assisted communication. First we establish the conditions under which density operators and positive semi-definite operators can achieve the entanglement-assisted zero-error capacity. Next, we show that certain properties of a hypergraph allow for a one-shot zero-error communication protocol with a capacity equal to the number of hyperedges in the representation. These properties offer valuable insights into the entanglement-assisted zero-error capacity of classical channels and the connection between confusability graphs and zero-error capacity.

3.1 Classical zero-error capacity

In this section, we embark on a journey into the realm of zero-error capacities, a concept in information theory. We delve into the maximum number of different messages that can be transmitted through a channel without any decoding errors, pushing the boundaries of reliable communication. We start by examining classical channels and their confusability graphs, visual representations that reveal the intricate relationships between input symbols. Additionally, we briefly explore the Lovász number as an upper bound for the Shannon capacity of hypergraphs, to get a better idea of limits of error-free transmission.

We say a channel is **memoryless** when consecutive uses of it are independent. As such, a memoryless classical channel $\mathcal{N}$ is entirely described by the **conditional probability distribution** $\mathcal{N}(y|x)$ of output y for each input x. That is, under input x the output of the channel is y with probability $\mathcal{N}(y|x)$ [7]. Given two channels $\mathcal{N}_1$ and $\mathcal{N}_2$, the channel attained by a single use of each channel is written as their tensor product, $\mathcal{N}_1 \otimes \mathcal{N}_2$, indicating that its conditional probability matrix is the tensor product of the respective conditional probability distributions. Analogously, n uses of a channel $\mathcal{N}$ is denoted by $\mathcal{N}^{\otimes n}$.

Throughout the thesis we will use X and Y to denote the set of input and output symbols of a channel $\mathcal{N}$ of interest. We say that two input symbols $x_1, x_2 \in X$ are **confusable** if there exists an output $y \in Y$ such that $\mathcal{N}(y|x_1) > 0$ and $\mathcal{N}(y|x_2) > 0$. The **confusability graph** of $\mathcal{N}$ is a graph $G(\mathcal{N}) = (V, E)$ where the set of vertices $V = X$ correspond to distinct input symbols of the channel $\mathcal{N}$, and two vertices are joined by an edge if said vertices are confusable. The confusability graph of $\mathcal{N}_1 \otimes \mathcal{N}_2$

$$\begin{array}{ccccc} & & \text{input} & & \\ & 1 & 2 & 3 & 4 \\ \text{a} & \frac{7}{10} & 0 & \frac{5}{8} & 0 \\ \text{b} & \frac{1}{10} & \frac{1}{2} & 0 & 0 \\ \text{c} & \frac{1}{5} & \frac{1}{2} & \frac{1}{4} & 0 \\ \text{d} & 0 & 0 & \frac{1}{8} & 1 \end{array}$$

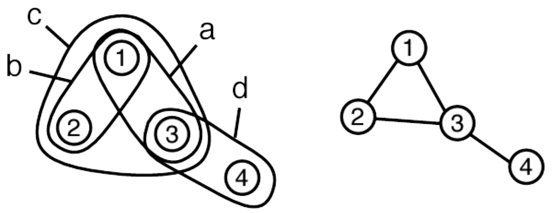


Figure 3.1: Left to right: Conditional probability matrix of channel $\mathcal{N}$; its hypergraph $H(\mathcal{N})$ with labelled hyperedges; confusability graph $G(\mathcal{N})$. Inputs 1 and 4 form a maximum non-confusable set. This example is taken from [8].

is based on those of $\mathcal{N}_1$ and $\mathcal{N}_2$ in the following way: $G(\mathcal{N}_1 \otimes \mathcal{N}_2) = G(\mathcal{N}_1) \boxtimes G(\mathcal{N}_2)$, where $\boxtimes$ denotes the **strong graph product** [1].

Definition 15 (Strong graph product)**.** The **strong product** of graphs $G_1, \ldots, G_n$ is a graph $G_1 \boxtimes \cdots \boxtimes G_n$ whose vertices are the n-tuples $V(G_1) \times \cdots \times V(G_n)$ and distinct vertices $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are joined by an edge if they are entry-wise confusable, i.e., for each $j \in \{1, \ldots, n\}$ either $a_j = b_j$ or $a_j b_j \in E(G_j)$. Likewise, we define the **strong power** of graph G by $G^{\boxtimes 1} := G$ and $G^{\boxtimes n} := G \boxtimes G^{\boxtimes(n-1)}$.

A **clique** in a graph is a subset of vertices where every vertex is joined by an edge to every other vertex in the subset. In a confusability graph, a connection between two vertices means that the corresponding inputs are confusable, i.e., they can produce the same output. Therefore, a clique in a confusability graph represents a set of inputs that are mutually confusable, as each input is connected to every other input in the set.

For this thesis, I have decided to use **confusability hypergraphs** instead of confusability graphs. The main reason for this is that the focus of my research is on the confusable inputs of a channel, which are better represented by hypergraphs. A set of inputs that are mutually confusable in a confusability graph is represented by a clique, i.e., an edge between each pair of vertices in the subset. In contrast, a confusability hypergraph represents this set using just a single hyperedge. This means that although a confusability hypergraph may appear more complex at first, it actually has significantly fewer edges compared to a confusability graph.

Definition 16 (Confusability hypergraph)**.** The **confusability hypergraph** of a channel $\mathcal{N}$ is a hypergraph $H(\mathcal{N}) = (V, E)$ where the set of vertices V correspond to the distinct input symbols of the channel $\mathcal{N}$. Each output symbol $y \in Y$, is represented by a hyperedge $y := e_x \subseteq V$ such that $e_x = \{x \in V : \mathcal{N}(y|x) > 0\}$. In other words, the hyperedge contains all input symbols that can produce the same output symbol y.

Suppose that in a channel $\mathcal{N}$, the input symbols 1, 2, and 3 are confusable and produce the same output $y \in Y$. Then, in the confusability hypergraph of $\mathcal{N}$, we add a hyperedge that contains all three vertices $\{1, 2, 3\}$, representing the fact that these three input symbols are pair-wise confusable. We do not add hyperedges for the subsets

$\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, because these subsets are already included in the largest subset $\{1, 2, 3\}$. More precisely, E consists of maximal subsets of pair-wise confusable inputs, where each subset represents the confusable inputs for every output $y \in Y$. This means that without loss of generality we can take the output set Y of the channel $\mathcal{N}$ to coincide with the set of hyperedges E of its confusability hypergraph $H(\mathcal{N})$. Similarly, the set X of input symbols for $\mathcal{N}$ coincides with the set of vertices V of $H(\mathcal{N})$:

$$X = V(H(\mathcal{N})), \qquad\qquad Y = E(H(\mathcal{N})). \tag{3.1}$$

Remark. The claim made regarding the equivalence between the output set Y of the channel $\mathcal{N}$ and the set of hyperedges E of its confusability hypergraph $H(\mathcal{N})$ raises a concern. To illustrate this, consider the scenario where we have $X = 1, 2$ and $Y = a, b$ with channel probabilities $\mathcal{N}(a|1) = \mathcal{N}(a|2) = \mathcal{N}(b|1) = \mathcal{N}(b|2) = 1/2$. In this case, it becomes necessary to include the hyperedge $\{1, 2\}$ twice in the hypergraph representation. This duplication accounts for the fact that inputs 1 and 2 can be confused in two distinct ways: either of them can yield the outputs a and b. Consequently, our set of hyperedges is treated as a "multiset," which, fortunately, does not introduce any additional complications throughout the study.

The hypergraph of $\mathcal{N}_1 \otimes \mathcal{N}_2$ is based on those of $\mathcal{N}_1$ and $\mathcal{N}_2$ in the following way: $H(\mathcal{N}_1 \otimes \mathcal{N}_2) = H(\mathcal{N}_1) \boxtimes H(\mathcal{N}_2)$, where $\boxtimes$ denotes the **strong hypergraph product**.

Definition 17 (Strong hypergraph product)**.** The **strong product** of hypergraphs $H_1, \ldots, H_n$ is a hypergraph $H_1 \boxtimes \cdots \boxtimes H_n$ whose vertices are the n-tuples $V(H_1) \times \cdots \times V(H_n)$ and distinct vertices $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are joined by a hyperedge if they are entry-wise confusable, i.e, for each $j \in \{1, \ldots, n\}$ either $a_j = b_j$ or $a_j, b_j \in e$ for some $e \in E(H_j)$. More precisely, for every pair of hyperedges $e_1 \in E(H_1)$ and $e_2 \in E(H_2)$, we define a new hyperedge $e_1 \times e_2$ in $H_1 \boxtimes H_2$. Likewise, we define the **strong power** of hypergraph H by $H^{\boxtimes 1} := H$ and $H^{\boxtimes n} := H \boxtimes H^{\boxtimes (n-1)}$.

The notion of independence number is introduced as a key measure in determining the zero-error capacity, capturing the maximum number of messages that can be transmitted without any confusability.

Definition 18 (Independence number)**.** An **independent set** of a hypergraph is a subset of its vertices where no two vertices are contained in the same hyperedge. The **independence number** $\alpha(H)$ of hypergraph H is the maximum size of an independent set of H.

We can interpret the independence number as the largest number of inputs that lead to distinct outputs with certainty. This brings finally brings to the notion of **zero-error capacity**.

Definition 19 (Zero-error capacity)**.** Let $M_0(\mathcal{N})$ denote the maximum number of different messages which can be sent with a single use of $\mathcal{N}$ with zero probability of decoding error. If $H(\mathcal{N})$ denotes the confusability hypergraph of $\mathcal{N}$ then

$$M_0(\mathcal{N}) = \alpha(H(\mathcal{N})),$$

since $\alpha(H(\mathcal{N}))$ is the maximum number of messages which are mutually non-confusable. The **zero-error capacity** of $\mathcal{N}$ is then

$$C_0(\mathcal{N}) := \lim_{n\to\infty} \frac{1}{n} \log M_0(\mathcal{N}^{\otimes n})$$

where "log" denotes the base-2 logarithm.

Shannon observed [9] that M_0 and C_0 depend solely on the confusability graph of the channel since $M_0(\mathcal{N}) = \alpha(G(\mathcal{N}))$ and $C_0(\mathcal{N}) = \log \Theta(G(\mathcal{N}))$ where

$$\Theta(G) := \lim_{n\to\infty} \sqrt[n]{\alpha(G^{\boxtimes n})},$$

is known as the **Shannon capacity** of G. This analogy extends to hypergraphs of the channel as well. Evidently, $\Theta(H) \geq \alpha(H)$, because $H^{\boxtimes n}$ has an independent set of size $\alpha(H)^n$. In particular, this implies that

$$C_0(\mathcal{N}) \geq \log M_0(\mathcal{N}). \tag{3.2}$$

In general however, $\Theta(H)$ can be larger than $\alpha(H)$. The simplest example is the 5-cycle (hyper)graph C_5 for which $\alpha(C_5) = 2$ but $\Theta(C_5) = \sqrt{5}$ [10].

Although computing the independence number of a hypergraph is conceptually simple, it is NP-hard. As such, there is no known algorithm to compute the Shannon capacity of a hypergraph. There is however a popular upper bound thanks to Lovász [10]. The **Lovász number** $\vartheta(H)$ of H. An efficiently computable quantity which satisfies $\vartheta(H) \geq \alpha(H)$ and

$$\vartheta(H_1 \boxtimes H_2) = \vartheta(H_1)\vartheta(H_2). \tag{3.3}$$

Because of this, we find $\vartheta(H) \geq \Theta(H)$ [10].

3.2 Entanglement-assisted zero-error capacity

The entanglement-assisted one-shot zero-error capacity $M_0^E(\mathcal{N})$ of a classical channel $\mathcal{N}$ represents the maximum number of messages that can be transmitted without error using entanglement assistance and a single use of the channel. Here is a formal definition.

Definition 20 (Entanglement-assisted one-shot zero-error capacity)**.** Let $\mathcal{N}$ be a classical channel with input set X and output set $Y = E(H(\mathcal{N})) \subseteq 2^X$. The **entanglement-assisted one-shot zero-error capacity** of $\mathcal{N}$, denoted $M_0^E(\mathcal{N})$, is the largest integer m such that there exist dimensions $d_A, d_B \geq 1$, a quantum state $|\phi\rangle_{AB} \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, measurements $\mu_i \in \mathscr{M}(\mathbb{C}^{d_A}, X)$ for all $i \in [m]$ and $\nu_e \in \mathscr{M}(\mathbb{C}^{d_B}, [m])$ for all $e \in E(H(\mathcal{N}))$ such that

$$\langle\phi|_{AB} \left(\mu_i(x)_A \otimes \nu_e(j)_B\right) |\phi\rangle_{AB} = \delta_{ij} \tag{3.4}$$

for all $i, j \in [m]$, $x \in X$ and $e \in E(H(\mathcal{N}))$ such that $x \in e$.

Let's break this definition down. Alice and Bob share an arbitrary finite-dimensional entangled state $|\phi\rangle_{AB}$ with local dimensions d_A and d_B, on which each can perform arbitrary local measurements which they have agreed on before the start of the protocol. Alice has a message $i \in [m]$ she wants to send. Alice chooses a set of measurements $\mu_i \in \mathscr{M}(\mathbb{C}^{d_A}, X)$ for $i \in [m]$ and arbitrary d_A, with outcomes found in set X. This measurement consists of operators $\mu_i(x)$ where $x \in X$, i.e., $\mu_i = \{\mu_i(x) : x \in X\}$. If Alice's message is i, she measures her half of the shared state $|\phi\rangle_{AB}$ with measurement μ_i. She obtains outcome $x \in X$ with probability $p_i(x) = \langle\phi|_{AB} (\mu_i(x)_A \otimes I_B) |\phi\rangle_{AB}$. After the measurement she transmits the outcome x through the channel $\mathcal{N}$.

On the receiving end, the output of the channel is a hyperedge, say e_x, containing x and possibly several other $x' \in X$. For a given hyperedge e_x, Bob performs the measurement $\nu_{e_x} \in \mathscr{M}(\mathbb{C}^{d_B}, [m])$ on his system B and gets outcome $j \in [m]$. Since this is a zero-error protocol, we must have that $i = j$ with probability 1. Therefore, we need to have

$$\langle\phi|_{AB} \left(\mu_i(x)_A \otimes \nu_{e_x}(j)_B\right) |\phi\rangle_{AB} = \delta_{ij} \tag{3.5}$$

for all $i, j \in [m]$, where δ_{ij} denotes the Kronecker delta function.

The following theorem provides an alternative characterisation of the one-shot entanglement-assisted zero-error capacity $M_0^E(\mathcal{N})$ of any classical channel $\mathcal{N}$.

Theorem 1 (Theorem 9 in [11]). For any classical channel $\mathcal{N}$ with inputs X and outputs Y, the entanglement-assisted zero-error capacity $M_0^E(\mathcal{N})$ is the largest $m \in \mathbb{N}$ such that there exists a density operator ρ_B and positive semi-definite operators $\beta_i(x)$ for all $i \in [m]$ and $x \in X$, on some Hilbert space, such that

$$\forall i \in [m] : \sum_{x \in X} \beta_i(x) = \rho_B.$$

Furthermore, for any pair of distinct messages $i \neq i' \in [m]$ and confusable inputs $x, x' \in X$ (i.e., inputs for which there is a hyperedge $e \in E(H(\mathcal{N}))$ such that $x, x' \in e$),

$$\mathrm{Tr}[\beta_i(x)\beta_{i'}(x')] = 0.$$

Consequently, $M_0^E(\mathcal{N})$ only depends on the confusability hypergraph $H(\mathcal{N})$.

Proof. We call the shared entangled state ρ_{AB}. Without loss of generality, to send message $i \in [m]$, Alice performs a measurement $\mu_i = \{\mu_i(x) : x \in X\}$ on her system, so with probability $p_i(x) = \mathrm{Tr}\left[(\mu_i(x)_A \otimes I_B)\rho_{AB}\right]$ she obtains outcome $x \in X$. Conditional on the knowledge of i and x, the residual state of Bob's system is

$$\rho_{B,i}(x) = \frac{\mathrm{Tr}_A[(\mu_i(x) \otimes I_B)\rho_{AB}]}{p_i(x)}.$$

Letting $\beta_i(x) := p_i(x)\rho_{B,i}(x) = \mathrm{Tr}_A[(\mu_i(x) \otimes I_B)\rho_{AB}]$, for all messages i we have

$$\sum_{x \in X} \beta_i(x) = \sum_{x \in X} \mathrm{Tr}_A[(\mu_i(x) \otimes I_B)\rho_{AB}] = \mathrm{Tr}_A\left[\sum_{x \in X} (\mu_i(x) \otimes I_B)\rho_{AB}\right] = \mathrm{Tr}_A[\rho_{AB}] =: \rho_B,$$

since $\sum_{x\in X}\mu_i(x) = I_A$. This means that Bob cannot determine which message Alice sent without receiving the output of the classical channel, which shows that causality is respected. On the other hand, any set of positive operators $\beta_i(x)$, with $i \in [m]$, that meets this requirement for a particular ρ_B can be accomplished by selecting an appropriate ρ_{AB} and generalized measurements.

To prove the second half of the theorem we need to use the assumption that Bob is able to perfectly recover Alice's message i.

The achievement of zero-error capacity in entanglement-assisted communication relies on the measurements used by Alice and Bob. Assuming the protocol indeed achieves zero-error communication, through joint measurements, Alice's measurements affect the state of Bob's subsystem, while Bob's measurements depend on the outcomes obtained by Alice. The entanglement shared between them enables the establishment of correlations and the utilisation of quantum properties for reliable information transmission without any error. Let's break down the protocol.

Alice puts the measurement outcome x into the channel $\mathcal{N}$. Bob obtains the channel output

$$e_x := y \in Y,$$

in addition to a quantum state left in his half of the entangled system. This information is captured by a bipartite classical-quantum state on Bob's side given by

$$\sigma_{B,i} := \sum_{x\in X, y\in Y} \mathcal{N}(y|x)\, |y\rangle\, \langle y| \otimes \beta_i(x).$$

The encoding works if and only if Bob can distinguish perfectly between all $\sigma_{B,i}$, i.e. for all distinct $i, i' \in [c]$,

$$\begin{aligned}
0 &= \mathrm{Tr}[\sigma_{B,i}\sigma_{B,i'}] \\
&= \sum_{x,x'\in X, y,y'\in Y} \mathcal{N}(y|x)\mathcal{N}(y'|x')\, \langle y|y'\rangle\, \mathrm{Tr}[\beta_i(x)\beta_{i'}(x')] \\
&= \sum_{x,x'\in X, y,y'\in Y} \mathcal{N}(y|x)\mathcal{N}(y'|x')\delta_{yy'}\mathrm{Tr}[\beta_i(x)\beta_{i'}(x')] \\
&= \sum_{x,x'\in X} \left(\sum_{y\in Y} \mathcal{N}(y|x)\mathcal{N}(y|x')\right) \mathrm{Tr}[\beta_i(x)\beta_{i'}(x')] \\
&= \sum_{x,x':\{x,x'\}\in E(G)} \left(\sum_{y\in Y} \mathcal{N}(y|x)\mathcal{N}(y|x')\right) \mathrm{Tr}[\beta_i(x)\beta_{i'}(x')].
\end{aligned}$$

Note that the in the final equation the sum is taken only over confusable inputs x, x', which satisfy

$$\mathcal{N}(y|x)\mathcal{N}(y|x') > 0.$$

So for these confusable inputs we must have

$$\mathrm{Tr}[\beta_i(x)\beta_{i'}(x')] = 0,$$

when $i \neq i'$, to be able to make the sought after distinction. □

Now that we have a clear understanding of what the entanglement-assisted one-shot zero-error capacity $M_0^E(\mathcal{N})$ of a classical channel $\mathcal{N}$ is, we can define the **(asymptotic) entanglement-assisted zero-error capacity.**

Definition 21 (Entanglement-assisted zero-error capacity)**.** The **entanglement-assisted zero-error capacity** of a classical channel $\mathcal{N}$ is

$$C_0^E(\mathcal{N}) := \lim_{n\to\infty} \frac{1}{n} \log M_0^E(\mathcal{N}^{\otimes n}).$$

Similar to eq. (3.2), it is the case that

$$C_0^E(\mathcal{N}) \geq \log M_0^E(\mathcal{N}). \tag{3.6}$$

Much like in the unassisted scenario, C_0^E only depend on the confusability (hyper)graph of the corresponding channel $\mathcal{N}$ [10]. Because of this, we will use the entanglement-assisted zero-error capacity of (hyper)graphs and of channels interchangeably:

$$M_0^E(H) = M_0^E(\mathcal{N}), \qquad\qquad C_0^E(H) = C_0^E(\mathcal{N}). \tag{3.7}$$

In [10] we see that hypergraphs H exist with $M_0^E(H) > M_0(H)$. Subsequently, we see in [7, 12] that the Lovász bound is also applicable to entanglement-assisted quantities, since $M_0^E(H) \leq \vartheta(H)$ and $C_0^E(H) \leq \log \vartheta(H)$.

3.3 Protocol for entanglement-assisted communication

This section delves into a protocol for entanglement-assisted communication and its implications. We present a theorem by Cubitt et al. [11] that provides insights into the zero-error capacity of classical channels and the relationship between graph properties and capacity.

The use of orthonormal representations connects the study of entanglement-assisted communication with concepts and techniques from graph theory. By considering orthonormal representations, we can gain insights into the distinctness and separability of vertices within the graph structure. We will use them to design efficient coding schemes for an entanglement-assisted communication protocols. The proof describes an example of said protocol.

Definition 22. (Orthonormal representation) A d-dimensional **orthonormal representation** of a hypergraph $H = (V, E)$ is a function $\gamma : V(H) \to \mathbb{C}^d$ that assigns unit vectors to the vertices of H such that for each hyperedge $e \in E$ and pair of vertices $u, v \in e$, the vectors $|\gamma(u)\rangle$ and $|\gamma(v)\rangle$ are orthogonal, i.e., $\langle\gamma(u)|\gamma(v)\rangle = 0$.

Theorem 2 (Theorem 11 in [11], adapted for hypergraphs)**.** Suppose that $H = (V, E)$ is a hypergraph with an orthonormal representation $\gamma : V(H) \to \mathbb{C}^d$, and that the vertices of H can be partitioned into m hyperedges $\{E_1, \ldots, E_m\} \subseteq E$, each of size d. Then

$$M_0^E(H) = \vartheta(H) = m$$

where $\vartheta(H)$ is the Lovász theta, see Section 3.1. In particular, the entanglement-assisted one-shot zero-error capacity $M_0^E(H)$ can be achieved by using a rank-d maximally entangled state.

Proof. First, we establish $M_0^E(H) \geq m$ by describing an entanglement-assisted protocol in which Alice and Bob share the rank-d maximally entangled state $|\phi\rangle_{AB} := \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |j\rangle_A \otimes |j\rangle_B$, with $|j\rangle$ the computational basis vectors for each party. The m hyperedges of size d, which partition the vertices of the graph, correspond to m complete orthonormal bases for $\mathbb{C}^d$ given by $\mathcal{B}_i = \{|\gamma(x)\rangle : \forall x \in E_i\}$ for $i \in [m]$. Alice and Bob agree on this partitioning before any measurements are made.

To encode the message i, Alice measures her half of the shared state along the basis $\mathcal{B}_i^c$. This is done by projecting her half of the shared state onto the basis states of $\mathcal{B}_i^c$. Specifically, for each basis state $\overline{|\gamma(x)\rangle} \in \mathcal{B}_i^c$, Alice performs the projection

$$\mu_i(x) := \overline{|\gamma(x)\rangle \langle\gamma(x)|}.$$

The probability that Alice obtains outcome $x \in X$ is then given by

$$p_i(x) = \mathrm{Tr}[(\mu_i(x) \otimes I_B)\rho_{AB}],$$

where ρ_{AB} is the density operator corresponding to the shared state. If the outcome corresponds to $x \in E_i (\subseteq V)$, Bob's subsystem is left in the state

$$\rho_{B,i}(x) = |\gamma(x)\rangle \langle\gamma(x)| = \frac{\overline{\langle\gamma(x)|}\, |\phi\rangle \langle\phi|\, \overline{|\gamma(x)\rangle}}{p_i(x)} = \frac{\mathrm{Tr}_A[(\mu_i(x) \otimes I_B)\rho_{AB}]}{\mathrm{Tr}[(\mu_i(x) \otimes I_B)\rho_{AB}]}.$$

Thus, if the outcome corresponds to x, Bob's subsystem is left in the state $|\gamma(x)\rangle$. This follows from combining Lemma 1 and Definition 11.

The output $e_x \in E$ Bob receives from $\mathcal{N}$, is a hyperedge that contains x, which is not necessarily one of the hyperedges from the original partitioning that Alice and Bob agreed on. Thus, Bob's subsystem must be in one of the corresponding sets of *mutually* orthogonal states $\{|\gamma(x')\rangle : x' \in e_x\}$. Therefore, he can perform a projective measurement on his subsystem to determine exactly which state he has, from which he can deduce x and, as a direct consequence, the symbol $i \in [m]$ which Alice chose, with certainty. This establishes that $M_0^E(H) \geq m$.

Second, to obtain $M_0^E(H) = \vartheta(H) \leq m$, note that $\vartheta(H)$ can only increase if edges are removed from H. Since H contains m hyperdedges of size d that cover all vertices, we can remove all other hyperedges from H to obtain the hypergraph $\bar{K}_m \boxtimes K_d$, where K_d and $\bar{K}_m$ are the complete and empty graphs on d and m vertices, respectively. Since $\vartheta(\bar{K}_m) = m$ and $\vartheta(K_d) = 1$, and the Lovász theta is multiplicative under strong graph product, where we make use of eq. (3.3).

$$\vartheta(H) \leq \vartheta(\bar{K}_m \boxtimes K_d) = \vartheta(\bar{K}_m)\vartheta(K_d) = m.$$

Using the result from [7, 12] that $M_0^E(H) \leq \vartheta(H)$, and putting both parts together, we get the desired claim $M_0^E(H) = \vartheta(H) = m$. □

The following theorem extends Theorem 2 from one-shot capacity M_0^E to the asymptotic capacity C_0^E. Its proof follows directly from Theorem 2 and the properties of $\vartheta(H)$.

Theorem 3. If a hypergraph $H = (V, E)$ has a d-dimensional orthonormal representation and there are m disjoint hyperedges $e_i \in E$ such that $|e_i| = d$ and $\bigcup_{i=1}^m e_i = V$, then $C_0^E(H) = \log m$.

Proof. From eq. (3.6) and Theorem 2 we get that

$$C_0^E(H) \geq \log M_0^E(H) = \log m. \tag{3.8}$$

Conversely,

$$C_0^E(H) \leq \log \vartheta(H) = \log m. \tag{3.9}$$

where the first inequality was established in [7, 12] and the equality in Theorem 2. □

4 Symplectic hypergraphs

In this chapter we'll take a closer look at symplectic hypergraphs that encode orthogonality relations among vectors in a symplectic space. The vertices of a symplectic hypergraph are points in symplectic space and they represent the possible inputs of a channel, while the hyperedges represent the possible outputs. Symplectic hypergraphs are important in the context of entanglement-assisted communication because the first example of a classical channel whose entanglement-assisted capacity exceeds the non-assisted capacity is based on a symplectic hypergraph known as $\mathrm{sp}(6, \mathbb{F}_2)$. By studying the structure and properties of this graph, we aim to gain insights into the capacity and performance of the corresponding entanglement-assisted communication channel.

4.1 Symplectic spaces

To begin, we define the notion of a non-degenerate symplectic form. A symplectic form is a bi-linear map that satisfies certain properties, such as skew-symmetry and non-degeneracy. These properties capture the symmetries and information-carrying properties of a communication channel. The canonical symplectic form is a specific example that we will focus on in our analysis.

Definition 23. Let V be a vector space over a field K. A **non-degenerate symplectic form** is a bi-linear map $S : V \times V \to K$ which is

- **skew-symmetric**: $S(u, v) = -S(v, u)$ for all $u, v \in V$,
- **non-degenerate**: if $S(u, v) = 0$ for all $v \in V$, then $u = 0$.

If K has characteristic 2, an extra requirement is that $S(u, u) = 0$ for all $u \in V$.

Next we define a symplectic space as a vector space equipped with a non-degenerate symplectic form.

Definition 24. A (non-degenerate) **symplectic space** (V, S) is a vector space V equipped with a (non-degenerate) symplectic form.

The following theorem tells us that the dimension of any symplectic space must always be even, and the symplectic form S can be taken to be the **canonical symplectic form**

$$\sigma(u, v) := u^T \begin{pmatrix} 0 & \mathbb{1}_m \\ -\mathbb{1}_m & 0 \end{pmatrix} v \tag{4.1}$$

where $\mathbb{1}_m$ is the $m \times m$ identity matrix.

Theorem 4 (Isometry of canonical symplectic space)**.** Any non-degenerate symplectic space with finite dimensional vector space V is isomorphic to the canonical symplectic space (V, σ). This implies that the dimension of any symplectic space is even.

From now on we will only consider binary symplectic spaces $V = \mathbb{F}_2^{2m}$ with canonical symplectic form σ. We will call two vectors $u, v \in V$ **orthogonal** if $\sigma(u, v) = 0$. Similarly, we say that a set of distinct vectors $v_1, \ldots, v_k \in V$ are **mutually orthogonal** if $\sigma(v_i, v_j) = 0$ for all $i, j \in \{1, \ldots, k\}$ such that $i \neq j$.

4.2 Symplectic hypergraph channel

Next, we introduce the concept of a symplectic hypergraph that will later be used to construct a classical communications channel. A symplectic hypergraph is a structure that represents the orthogonality relationships between vectors in a symplectic space. Based on Definition 16, we will then construct a channel whose confusability graph is a symplectic hypergraph.

Definition 25 (Symplectic hypergraph)**.** For any integer $m \geq 1$, the **symplectic hypergraph** $\mathrm{sp}(2m, \mathbb{F}_2)$ is the hypergraph (V, E) with vertices $V = \mathbb{F}_2^{2m} \setminus \{0\}$, i.e., all non-zero vectors in $\mathbb{F}_2^{2m}$, and hyperedges E that correspond to maximal sets of mutually orthogonal vectors in V.

Let's look at what these symplectic graphs look like for $m = 1$ and $m = 2$.

Example 1 ($m = 1$)**.** The vertices of $\mathrm{sp}(2, \mathbb{F}_2)$ are exactly the non-zero vectors of $\mathbb{F}_2^2$, giving us

$$a = (0, 1), \qquad b = (1, 0), \qquad c = (1, 1). \tag{4.2}$$

In $\mathrm{sp}(2, \mathbb{F}_2)$, there are precisely three hyperedges, and each hyperedge contains exactly one of the previously mentioned points:

$$E(\mathrm{sp}(2, \mathbb{F}_2)) := \{\{(0, 1)\}, \{(1, 0)\}, \{(1, 1)\}\},$$

since each point is only orthogonal to itself, using the canonical symplectic form (see Table 4.1). This gives us 3 orthogonal sets of 1 vertex each.

	a	b	c
a	0	1	1
b	1	0	1
c	1	1	0

Table 4.1: Symplectic inner products between the vectors in eq. (4.2).

	1	2	3	a	4	5	6	b	7	8	9	c	10	11	12
1	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
2	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
3	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
a	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
4	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0
5	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0
6	1	1	0	0	1	1	0	1	0	0	1	1	0	0	1
b	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
7	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0
8	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0
9	1	1	0	1	0	0	1	0	1	1	0	1	0	0	1
c	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
10	0	1	1	1	1	0	0	1	1	0	0	0	0	1	1
11	1	0	1	1	0	1	0	1	0	1	0	0	1	0	1
12	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0

Table 4.2: Symplectic inner products between vectors from eq. (4.3). Taken from [13]

Example 2 ($m = 2$)**.** The graph for $\mathrm{sp}(4, \mathbb{F}_2)$ is more intriguing. We can label its vertices by non-zero bitstrings as follows:

$$\begin{aligned} 1 &= (0,1,0,0), & 2 &= (0,1,0,1), & 3 &= (0,0,0,1), \\ a &= (1,0,0,0), & b &= (1,0,1,0), & c &= (0,0,1,0), \\ 4 &= (1,1,0,0), & 5 &= (1,1,0,1), & 6 &= (1,0,0,1), \\ 7 &= (1,1,1,0), & 8 &= (1,1,1,1), & 9 &= (1,0,1,1), \\ 10 &= (0,1,1,0), & 11 &= (0,1,1,1), & 12 &= (0,0,1,1). \end{aligned} \tag{4.3}$$

The symplectic inner products between these vectors are listed in Table 4.2. In this case, the set of hyperedges consists entirely of three-vertex sets:

$$\begin{aligned} E(\mathrm{sp}(4, \mathbb{F}_2)) = \{&\{1,a,4\}, \{2,a,5\}, \{3,a,6\}, \{1,b,7\}, \{2,b,8\}, \\ &\{3,b,9\}, \{1,c,10\}, \{2,c,11\}, \{3,c,12\}, \{4,8,12\}, \\ &\{5,7,12\}, \{6,7,11\}, \{4,9,11\}, \{5,9,10\}, \{6,8,10\}\}. \end{aligned}$$

The corresponding hypergraph is depicted in Fig. 4.1.

Theorem 5. The $2^{2m}-1$ vertices of $\mathrm{sp}(2m, \mathbb{F}_2)$ can be partitioned into 2^m+1 hyperedges of size $2^m - 1$.

Proof. Such a partition of the symplectic hypergraph is known as a **symplectic spread**, and it's existence is well established [14]. We give an easy to follow construction from [15]. Another proof is given in terms of commuting sets of Pauli matrices in [16, 17].

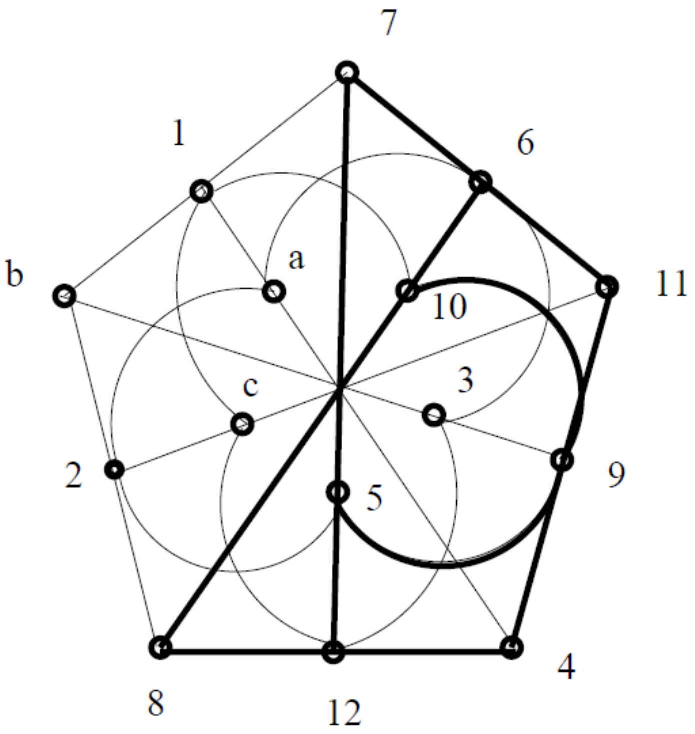


Figure 4.1: The vectors from equation eq. (4.3) exhibit an underlying geometric structure, where the points correspond to these vectors. This structure is further represented by a collection of continuous lines or arcs, each of which signifies a maximal orthogonal set of vectors [13]

Let $N = 2^m$. We will identify the vertices of $\text{sp}(2m, \mathbb{F}_2)$ with the non-zero vectors in $\mathbb{F}_N^2$ using the following procedure. First, note that the elements $(w, x) \in \mathbb{F}_N^2$ consist of polynomials $w, x \in \mathbb{F}_N$ as by Section 2.3. As such, we can write them as

$$w = a_0 + a_1 X + \cdots + a_{m-1} X^{m-1}$$

and

$$x = b_0 + b_1 X + \cdots + b_{m-1} X^{m-1},$$

for some coefficients $a_i, b_i \in \mathbb{F}_2$.

Let us first establish a correspondence between $\mathbb{F}_N$ and $\mathbb{F}_2^m$ as binary vector spaces. Note that

$$\text{Tr}[wx] = \text{Tr}\left[\sum_{i=0}^{m-1} a_i X^i \sum_{j=0}^{m-1} b_j X^j\right] = \sum_{i,j=0}^{m-1} a_i b_j \text{Tr}[X^i X^j] = a^T M b \tag{4.4}$$

where $a^T := (a_0, \ldots, a_{m-1})$, $b := (b_0, \ldots, b_{m-1})^T$, and $M_{ij} := \text{Tr}[X^{i+j}]$ is a symmetric $m \times m$ matrix over $\mathbb{F}_2$. This is a non-degenerate inner product in $\mathbb{F}_2^m$. Indeed, the map $(w, x) \mapsto \mathrm{Tr}[wx]$ is symmetric, bilinear (see Lemma 9), and non-degenerate (see Lemma 12). Hence, up to some invertible transformation, it is equivalent to the standard inner product in $\mathbb{F}_2^m$. Namely, there exists an invertible $m \times m$ matrix P over $\mathbb{F}_2$ such that $M = P^T P$ and

$$\mathrm{Tr}[wx] = a^T P^T P b. \tag{4.5}$$

The following is an explicit bijection between $\mathbb{F}_N$ and $\mathbb{F}_2^m$ as linear spaces over $\mathbb{F}_2$:

$$\Gamma : \mathbb{F}_N \to \mathbb{F}_2^m : \sum_{i=0}^{m-1} a_i X^i \mapsto Pa \tag{4.6}$$

where $a = (a_0, \ldots, a_{m-1})^T$. According to eq. (4.5), it preserves inner products in the sense that $\mathrm{Tr}[wx] = \Gamma(w)^T \Gamma(x)$. Moreover, $\Gamma(0) = 0$.

We can extend the above map Γ from $\mathbb{F}_N \to \mathbb{F}_2^m$ to $\mathbb{F}_N^2 \to \mathbb{F}_2^{2m}$ in the following way:

$$\Gamma : \mathbb{F}_N^2 \to \mathbb{F}_2^{2m} : (w, x) \mapsto (Pa, Pb). \tag{4.7}$$

Since P is invertible, $\ker(\Gamma) = (0, 0)$ so Γ is surjective and hence an isomorphism. Furthermore, for $(w, x), (y, z) \in \mathbb{F}_N^2$ we find

$$\begin{aligned}
\Gamma((w, x) + (y, z)) &= \Gamma((w + y, x + z)) \\
&= (P(a + c), P(b + d)) \\
&= (P(a), P(b)) + (P(c), P(d)) \\
&= \Gamma(w, x) + \Gamma(y, z),
\end{aligned}$$

where

$$\begin{aligned}
y &= c_0 + c_1 X + \cdots + c_{m-1} X^{m-1}, \\
z &= d_0 + d_1 X + \cdots + d_{m-1} X^{m-1}.
\end{aligned}$$

This completes our identification of the vertices of $\mathrm{sp}(2m, \mathbb{F}_2)$ and the elements of $\mathbb{F}_N^2$.

Next, consider the following map $\mathbb{F}_N^2 \to \mathbb{F}_2$:

$$\sigma_N((w, x), (y, z)) := \mathrm{Tr}[wz + xy],$$

where $\mathrm{Tr} : \mathbb{F}_N \to \mathbb{F}_2$ is the finite field trace introduced in Definition 14, and show that this is a symplectic form. We show below that σ_N is bilinear, skew-symmetric and non-degenerate:

- *σ_N is bilinear:* For any $(w_1, x_1), (w_2, x_2), (y, z) \in \mathbb{F}_N^2$ and $\alpha, \beta \in \mathbb{F}_N$, we have

  $$\begin{aligned}
  \sigma_N(\alpha(w_1, x_1) + \beta(w_2, x_2), (y, z)) &= \mathrm{Tr}[(\alpha w_1 + \beta w_2)z + (\alpha x_1 + \beta x_2)y] \\
  &= \alpha \mathrm{Tr}[w_1 z + x_1 y] + \beta \mathrm{Tr}[w_2 z + x_2 y] \\
  &= \alpha \sigma_N((w_1, x_1), (y, z)) + \beta \sigma_N((w_2, x_2), (y, z)),
  \end{aligned}$$

  which follows by Lemma 9. Thus σ_N is linear in its first argument. Similarly, we can show that σ_N is linear in its second argument.

- σ_N *is skew-symmetric:* For any $(w,x),(y,z) \in \mathbb{F}_N^2$, we have

$$\begin{aligned}\sigma_N((w,x),(y,z)) + \sigma_N((y,z),(w,x)) &= \mathrm{Tr}[wz+xy] + \mathrm{Tr}[yx+zw] \\ &= 2\mathrm{Tr}[wz+xy] = 0.\end{aligned}$$

  Therefore, $\sigma_N((w,x),(y,z)) = -\sigma_N((y,z),(w,x))$ for all $(w,x),(y,z) \in \mathbb{F}_N^2$.

- σ_N *is non-degenerate:* This follows directly from the finite field trace being non-degenerate as shown in Lemma 12.

Thus σ_N is a symplectic form. Now, let us show that the symplectic spaces $(\mathbb{F}_2^{2m},\sigma)$ and $(\mathbb{F}_N^2,\sigma_N)$ are isomorphic. Recall that Γ defined in eq. (4.7) induces an isomorphism between the two spaces. To show that the symplectic form σ_N is preserved by Γ, note the following:

$$\begin{aligned}\sigma_N((w,x),(y,z)) &= \mathrm{Tr}[wz+xy] \\ &= \mathrm{Tr}[wz] + \mathrm{Tr}[xy] \\ &= a^T P^T P d + b^T P^T P c \\ &= \begin{pmatrix} Pa \\ Pb \end{pmatrix}^T \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} Pc \\ Pd \end{pmatrix} \\ &= \sigma((Pa,Pb),(Pc,Pd)) \\ &= \sigma(\Gamma(w,x),\Gamma(y,z)),\end{aligned}$$

where σ is the canonical symplectic form as described in eq. (4.1). This allows us to describe the partition for the non-zero elements of $(\mathbb{F}_2^{2m},\sigma)$ in terms of one for those of $(\mathbb{F}_N^2,\sigma_N)$.

Denoting the multiplicative group (of order $N-1$) in $\mathbb{F}_N$ by $\mathbb{F}_N^\times := \mathbb{F}_N \setminus \{0\}$, the cells of a partition of the non-zero elements of $\mathbb{F}_N^2$ are:

$$\pi_a = \{(x,ax) : x \in \mathbb{F}_N^\times\} \quad (a \in \mathbb{F}_N), \tag{4.8}$$

$$\pi_{N+1} = \{(0,x) : x \in \mathbb{F}_N^\times\}. \tag{4.9}$$

It is easy to check that these $N+1$ cells of size $N-1$ partition $\mathbb{F}_N^2$. Moreover, if (x,ax) and (y,ay) are in the same cell, then

$$\sigma_N((x,ax),(y,ay)) = \mathrm{Tr}[xay + axy] = \mathrm{Tr}[2xay] = \mathrm{Tr}[0] = 0.$$

Therefore, each cell is a clique with respect to the symplectic inner product σ_N. By applying the bijection Γ^{-1} we can translate this partition of $(\mathbb{F}_N^2,\sigma_N)$ back to the original symplectic space $(\mathbb{F}_2^{2m},\sigma)$. □

The **symplectic hypergraph channel** $\mathcal{N}$ is described by the confusability hypergraph $H(\mathcal{N}) = \mathrm{sp}(2m,\mathbb{F}_2)$ where $m \geq 1$ is any integer. Recall from eq. (3.1) that the input and output sets of $\mathcal{N}$ are $X := V(H(\mathcal{N}))$ and $Y := E(H(\mathcal{N}))$, respectively.

According to Theorem 5, the number of inputs to $\mathcal{N}$ is $|X| = 2^{2m} - 1$ and the number of outputs is $|Y| = 2^m + 1$. Under input $x \in X$ the output of the channel is any $y \in Y$ such that $x \in y$, and it occurs with some probability $\mathcal{N}(y|x) > 0$ such that $\sum_{y:x \in y} \mathcal{N}(y|x) = 1$ [7].

From now on we will focus on the specific case of $m = 3$ or $\mathrm{sp}(6, \mathbb{F}_2)$.

5 The zero-error capacities of the symplectic hypergraph channel

The next two sections prove that for channels with confusability graph $\mathrm{sp}(6, \mathbb{F}_2)$, the entanglement-assisted zero-error capacity is larger than the unassisted one. More specifically,

Theorem 6. *$C_0(\mathrm{sp}(6, \mathbb{F}_2)) = \log 7$, however $C_0^E(\mathrm{sp}(6, \mathbb{F}_2)) = \log 9$.*

The first part of this theorem, $C_0(\mathrm{sp}(6, \mathbb{F}_2)) = \log 7$ is a special case of a result from [18], where $m = 3$. To prove this we will need **Haemers' bound**.

Theorem 7 (Haemers' bound)**.** *Let H be a hypergraph and let M_{uv} with $u, v \in V(H)$ be a matrix with entries in any field K. We say that matrix M* **fits** *H if $M_{uu} \neq 0$ and $M_{uv} = 0$ when there is no $e \in E(H)$ such that $u, v \in e$, meaning there is no hyperedge which contains both u and v. Then $\Theta(H) \leq R(G) := \min\{\mathrm{rank}(M) : M \text{ fits } G\}$, where $\Theta(H)$ is the Shannon capacity of H. In particular, we have $C_0(H) \leq \log R(G)$.*

The proof for this theorem is outlined as follows.

5.1 Classical case

Theorem 8. *$C_0(\mathrm{sp}(2m, \mathbb{F}_2)) = \log(2m + 1)$.*

Proof. To establish an upper bound, we can construct a matrix over the field $K = \mathbb{F}_2$ that satisfies the requirements of $\mathrm{sp}(2m, \mathbb{F}_2)$ and has a rank of $2m + 1$, after which we will use Haemers' bound to deduce the sought after inequality.

We first define a subspace U_m as follows:

$$U_m := \{v \in \mathbb{F}_2^{2m+1} : \langle v, v\rangle\} = 0.$$

This subspace, residing within $\mathbb{F}_2^{2m+1}$, possesses a dimension of $2m$. Importantly, it is characterised by the property that each vector within U_m consists of an even number of entries equal to one. When we consider the standard inner product $\langle ., .\rangle$ on the vector space $\mathbb{F}_2^{2m+1}$ and apply it to the subspace U_m, something interesting happens. The standard inner product on U_m exhibits properties similar to those of a non-degenerate symplectic form. In other words, there exists an isomorphism

$$T : (\mathbb{F}_2^{2m}, \sigma) \to (U_m, \langle ., .\rangle)$$

that maps the vector space $(\mathbb{F}_2^{2m}, \sigma)$ with the canonical symplectic form σ to the subspace $(U_m, \langle ., .\rangle)$ equipped with the standard inner product $\langle ., .\rangle$, such that

$$\forall u, v \in \mathbb{F}_2^{2m} : \sigma(u, v) = \langle T(u), T(v) \rangle.$$

Consider the vector $\mathbf{1} \in \mathbb{F}_2^{2m+1}$ which has 1 in every entry. Note that for all $v \in U_m$ we then have $\langle \mathbf{1}, v \rangle = 0$. For all $u, v \in \mathbb{F}_2^{2m}$ we define the fitting matrix M as follows:

$$\begin{aligned} M_{uv} &:= \langle \mathbf{1} + T(u), \mathbf{1} + T(v) \rangle \\ &= \langle \mathbf{1}, \mathbf{1} \rangle + \langle \mathbf{1}, T(v) \rangle + \langle T(u), \mathbf{1} \rangle + \langle T(u), T(v) \rangle \\ &= 1 + \sigma(u, v). \end{aligned}$$

Since $\sigma(u, v) = 1$ if and only if u and v are not both contained in the same set, the matrix M fits $\mathrm{sp}(2m, \mathbb{F}_2)$. Since it is the Gram matrix of a set of $(2m+1)$-dimensional vectors (where the entry at position i, j corresponds to the inner product of the i-th and j-th vectors, according to an arbitrary ordering of the set), its rank is at most $2m + 1$. Consequently, by Haemers' bound, we have

$$C_0(\mathrm{sp}(2m, \mathbb{F}_2)) \leq \log(2m + 1).$$

For the matching lower bound, consider the standard basis e_i for $i \in \{1, \ldots, 2m + 1\}$ in $\mathbb{F}_2^{2m+1}$. Let $f_i := e_i + \mathbf{1}$. It can be observed that $\langle f_i, f_j \rangle = 1 - \delta_{ij}$, where δ_{ij} is the Kronecker delta. Therefore, f_i belongs to U_m and $\langle T^{-1}(f_i), T^{-1}(f_j) \rangle = 1 - \delta_{ij}$. As a result, the set

$$\{T^{-1}(f_i) : i \in \{1, \ldots, 2m + 1\}\}$$

forms an independent set of size $2m+1$ in $\mathrm{sp}(2m, \mathbb{F}_2)$, implying $\alpha(\mathrm{sp}(2m, \mathbb{F}_2)) \geq 2m+1$. Consequently, we have $C_0(\mathrm{sp}(2m, \mathbb{F}_2)) = \log(2m + 1)$, and the upper bound on the zero-error capacity is achieved by a code of block length one. □

In coding theory, the term "block length" refers to the length of the codewords used in the code. A code with a "block length one" means that each codeword in the code consists of a single symbol or character.

5.2 Entanglement-assisted case

In this section we explain the second part of Theorem 6, namely that $C_0^E(\mathrm{sp}(6, \mathbb{F}_2)) = \log 9$. Afterwards we give a more explicit description of the protocol for $m = 3$.

Lemma 13. There is an orthonormal representation of $\mathrm{sp}(6, \mathbb{F}_2)$ of dimension 7.

Proof. A complete orthonormal representation of $\mathrm{sp}(6, \mathbb{F}_2)$, which consists of 9 orthonormal bases in $\mathbb{R}^7$, is presented in Appendix A of [1]. It remains to verify that this representation possesses the desired properties (see Definition 22). Notably, it is interesting to observe that the representation comprises vectors from the root system E_7. In Appendix B of [1], a more insightful description and proof of the representation in relation to E_7 is provided.

Given that the 63 vertices of $\mathrm{sp}(6, \mathbb{F}_2)$ can be partitioned into 9 cliques, each containing 7 vertices (as stated in Theorem 5), it follows from Theorem 3 that $C_0^E(\mathrm{sp}(6, \mathbb{F}_2)) = \log 9$. Moreover, since we have previously shown that $C_0(\mathrm{sp}(6, \mathbb{F}_2)) = \log(2 \cdot 3 + 1) = \log 7$, we have successfully concluded the proof of Theorem 6. □

Next, let's take a closer look at what our channel and the corresponding entanglement-assisted protocol described in Section 3.3 looks like for $m = 3$. By Theorem 5 we know that the $2^{2\cdot 3} - 1 = 63$ vertices of $\mathrm{sp}(6, \mathbb{F}_2)$ can be partitioned into $2^3 + 1 = 9$ cliques, each consisting of $2^3 - 1 = 7$ vertices. By Lemma 13 in [1] we know that $\mathrm{sp}(6, \mathbb{F}_2)$ has an orthonormal representation γ in $\mathbb{R}^7$. By using this orthonormal representation, we can construct a set of nine orthonormal measurements $\{\mu^i(x) \in \mathbb{R}^{7\times 7} : x \in X(i)\}$ for Alice. If Alice wants to transmit message $i = 1, \ldots, 9$, she performs the measurement μ^i with operators

$$\mu^i(x) := \overline{|\gamma(x)\rangle \langle \gamma(x)|}, \qquad x \in X(i).$$

The measurement outcome x lies in the i-th clique $X(i) \subset V(\mathrm{sp}(6, \mathbb{F}_2)) = \mathbb{F}_2^6 \setminus \{0\}$ and is used by Alice as the input to the classical channel. The output of the channel is an arbitrary hyperedge e_x that contains Alice's input x. For each hyperedge e_x, the corresponding Bob's measurement $\{\nu^{e_x}(j) \in \mathbb{R}^{7\times 7} : j = 1, \ldots, 7\}$ is constructed from the same orthonormal representation γ:

$$\nu^{e_x}(j) := |\gamma(e_x(j))\rangle \langle \gamma(e_x(j))| \,, \qquad j = 1, \ldots, 7,$$

where $e_x(j) \in V(\mathrm{sp}(6, \mathbb{F}_2))$ is the j-th vertex in the hyperedge e_x output by the channel.

At the beginning of the protocol Alice and Bob share the maximally entangled state $|\psi\rangle_{AB} = \frac{1}{\sqrt{7}} \sum_{i=1}^{7} |i\rangle_A \otimes |i\rangle_B$, on which they perform the following measurements:

$$\begin{aligned} \text{Alice:} &\quad \mu^i(x) \otimes \mathbb{I}, \\ \text{Bob:} &\quad \mathbb{I} \otimes \nu^{e_x}(j), \end{aligned}$$

where $\mathbb{I}$ is the identity operator. They obtain the following outcome probabilities:

$$p(x, j|i, e_x) = \langle \psi|_{AB} \left(\mu^i(x)_A \otimes \nu^{e_x}(j)_B \right) |\psi\rangle_{AB} \,.$$

This protocol satisfies the conditions of Section 3.3 and can transmit any message $i = 1, \ldots, 9$ with zero error, so it achieves zero-error entanglement-assisted capacity $\log 9$ for the classical channel whose confusability hypergraph is $\mathrm{sp}(6, \mathbb{F}_2)$.

5.3 Connection to E_7

The term “root system” refers to a set of vectors with special properties. It was introduced by mathematician Wilhelm Killing while studying Lie algebras and Lie groups. The name “root” was chosen to highlight the analogy with the roots of polynomial equations. Just like polynomial roots are crucial in understanding the equation, the roots in a root system play a fundamental role in characterising the structure and symmetries of Lie algebras and Lie groups. Root systems are a valuable tool to investigate the symmetries and algebraic properties of these structures.

The Lie group E_7 has 126 continuous connected components, which form the vertices of the corresponding root system. The symplectic graph $\mathrm{sp}(6, \mathbb{F}_2)$ is closely related to the E_7 root system, which is a finite, orthogonal representation of the Lie algebra E_7 in $\mathbb{R}^7$. This is because the vertices of $\mathrm{sp}(6, \mathbb{F}_2)$ can be identified with the components of the E_7 root system. The dimension of the E_7 root system is 7, which is the same as the dimension of the orthonormal representation of $\mathrm{sp}(6, \mathbb{F}_2)$ that we have been using. This connection to the E_7 root system is what makes it possible to construct orthogonal measurements that satisfy the conditions of Section 3.3 and achieve zero-error entanglement-assisted communication for the symplectic graph $\mathrm{sp}(6, \mathbb{F}_2)$. The connection of the symplectic graph $\mathrm{sp}(6, \mathbb{F}_2)$ to the E_7 root system not only allows us to use the properties of the E_7 root system to analyze the graph, but also highlights the connections between the study of symplectic graphs and the study of Lie algebras and Lie groups.

Popular Summary

Communication channels are the pathways we use to transmit information from one person to another. These channels can be spoken language, written text, radio waves, or computer networks. The goal of communication is to guarantee that the information sent by the sender is accurately received and understood by the receiver. However, communication channels are often imperfect – they can be affected by various factors that introduce errors or distortions into the transmitted information. This is where the concept of noisy channels comes into play.

A noisy channel refers to a communication channel where interference, noise, or errors can corrupt the message during its transmission. This can occur due to factors like signal attenuation, electrical interference, or external sources of disturbance. In the field of information theory, researchers study the capacity of a noisy channel, specifically focusing on the zero-error capacity. The zero-error capacity of a channel refers to the maximum rate of reliable communication, where the receiver can decode the message with absolute accuracy and without any errors, even in the presence of noise. Understanding the zero-error capacity helps us determine the limits of reliable communication over noisy channels and develop coding and decoding techniques that can maximize the accuracy and robustness of the transmitted information.

Quantum physics offers a surprising way to boost channel capacity by sharing entanglement between the sender and the receiver. Entanglement is when particles become intertwined and their properties become highly correlated, even if they are far apart. By sharing entangled particles, the sender and receiver in a communication channel can increase channel capacity if they can manipulate them properly. Furthermore, by repeatedly utilising a communication channel, researchers have noted the possibility of heightened communication efficiency as opposed to singular usage.

In this thesis, as well as the study of communication channels more broadly, certain abstract structures called finite fields play a significant role. Finite fields provide a generalised framework for arithmetic operations, similar to numbers, but with specific properties and limitations. We use them to construct and analyse a channel whose zero-error capacity is increased when the sender and receiver share entanglement.

Bibliography

[1] D. Leung, L. Mancinska, W. Matthews, M. Ozols, and A. Roy, “Entanglement can Increase Asymptotic Rates of Zero-Error Classical Communication over Classical Channels,” *Communications in Mathematical Physics*, vol. 311, no. 1, pp. 97–111, 2012.

[2] S. Khatri and M. M. Wilde, “Principles of Quantum Communication Theory: A Modern Approach,” 11 2020. [Online]. Available: http://arxiv.org/abs/2011.04672

[3] M. Walter and M. Ozols, “Lectures Notes on Quantum Information Theory,” University of Amsterdam, Tech. Rep., 2022.

[4] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge University Press, 2010.

[5] H. W. Lenstra and F. Oort, “Ringen en lichamen,” Universiteit van Amsterdam, Tech. Rep., 2014.

[6] R. J. McEliece, *Finite Fields for Computer Scientists and Engineers*, ser. The Kluwer International Series in Engineering and Computer Science. Boston, MA: Springer US, 1987, vol. 23. [Online]. Available: http://link.springer.com/10.1007/978-1-4613-1983-2

[7] S. Beigi, “Entanglement-assisted zero-error capacity is upper bounded by the Lovasz theta function,” *Physical Review A*, 2 2010. [Online]. Available: http://arxiv.org/abs/1002.2488http://dx.doi.org/10.1103/PhysRevA.82.010303

[8] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, “Improving zero-error classical communication with entanglement,” *Physical Review Letters*, 11 2009. [Online]. Available: http://arxiv.org/abs/0911.5300http://dx.doi.org/10.1103/PhysRevLett.104.230503

[9] C. E. Shannon, “The zero error capacity of a noisy channel,” *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.

[10] L. Lovász, “On the Shannon Capacity of a Graph,” *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.

[11] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, “Zero-error channel capacity and simulation assisted by non-local correlations,” *IEEE Transactions on Information Theory*, 3 2010. [Online]. Available: http://arxiv.org/abs/1003.3195http://dx.doi.org/10.1109/TIT.2011.2159047

[12] R. Duan, S. Severini, and A. Winter, “Zero-error communication via quantum channels, non-commutative graphs and a quantum Lovasz theta function,” *IEEE Transactions on Information Theory*, 2 2010. [Online]. Available: http://arxiv.org/abs/1002.2514http://dx.doi.org/10.1109/TIT.2012.2221677

[13] M. Planat and M. Saniga, “On the Pauli graphs of N-qudits,” Institut FEMTO-ST, CNRS, DÂ´epartement LPMO, Tech. Rep., 2007.

[14] R. H. Dye, “Partitions and Their Stabilizers for Line Complexes and Quadrics (*),” *Annali di Matematica*, 1977.

[15] S. Ball and J. Bamberg, “Symplectic Spreads,” *Designs Codes and Cryptography*, vol. 32, pp. 9–14, 2004.

[16] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, “A new proof for the existence of mutually unbiased bases *,” Electrical Engineering Department, UCLA, Tech. Rep., 2001.

[17] J. Lawrence, Ë. Caslav Brukner, and A. Zeilinger, “Mutually unbiased binary observable sets on N qubits,” *Physical Review A*, 2018.

[18] R. Peeters, “ORTHOGONAL REPRESENTATIONS OVER FINITE FIELDS AND THE CHROMATIC NUMBER OF GRAPHS,” *Combinatorica*, vol. 16, no. 3, pp. 417–431, 1996.

[19] R. Yuan, “A Brief Introduction to POVM Measurement in Quantum Communications,” Beijing University of Posts and Telecommunications, Tech. Rep., 2022.

[20] P. Hayden, “Capacities Enhanced by Entanglement,” *Encyclopedia of Mathematical Physics: Five-Volume Set*, pp. 418–423, 1 2006.

[21] J. Briët, H. Buhrman, M. Laurent, T. Piovesan, and G. Scarpa, “Entanglement-assisted zero-error source-channel coding,” *IEEE Transactions on Information Theory*, 8 2013. [Online]. Available: http://arxiv.org/abs/1308.4283http://dx.doi.org/10.1109/TIT.2014.2385080

[22] G. Mullen and C. Mummert, *Finite Fields and Applications*, ser. The Student Mathematical Library. Providence, Rhode Island: American Mathematical Society, 11 2007, vol. 41. [Online]. Available: http://www.ams.org/stml/041

[23] M. M. Wilde, “From Classical to Quantum Shannon Theory,” Hearne Institute for Theoretical Physics, Tech. Rep., 6 2011. [Online]. Available: http://arxiv.org/abs/1106.1445http://dx.doi.org/10.1017/9781316809976.001

[24] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, “Improving zero-error classical communication with entanglement,” *Physical Review Letters*, vol. 104, no. 23, 6 2010.

[25] R. W. Fitzgerald, “Invariants of trace forms over finite fields of characteristic 2,” *Finite Fields and their Applications*, vol. 15, no. 2, pp. 261–275, 4 2009.

[26] K. Schäcke, “On the Kronecker Product,” Ph.D. dissertation, University of Waterloo, 2013.

[27] Z. Tang and Z. X. Wan, “Symplectic graphs and their automorphisms,” *European Journal of Combinatorics*, vol. 27, no. 1, pp. 38–50, 1 2006.

[28] J. Watrous, “The Theory of Quantum Information,” University of Waterloo, Tech. Rep., 2018.

[29] C. F. Van Loan, “The ubiquitous Kronecker product,” *Journal of Computational and Applied Mathematics*, vol. 123, pp. 85–100, 2000. [Online]. Available: www.elsevier.nl/locate/cam