Alma Mater Studiorum · Università di Bologna

Dipartimento di Fisica e Astronomia "Augusto Righi"
Corso di Laurea in Fisica

Capacità di un canale di comunicazione: dalla definizione classica alle estensioni quantistiche

Relatore:
Prof. Lorenzo Piroli

Presentata da:
Giovanna Mariapia Monaldi

Anno Accademico 2024/2025

Sommario

Nella teoria dell'informazione classica una delle proprietà utilizzate per descrivere un canale di comunicazione è la capacità, ovvero la grandezza che quantifica l'informazione massima che può essere trasmessa efficacemente con un uso del canale. La teoria dell'informazione quantistica nasce dal tentativo di estendere a livello quantistico i risultati della teoria classica, questo avviene anche per il concetto di capacità. Usare un sistema quantistico per trasmettere informazione offre più alternative di quelle disponibili classicamente: si possono estrarre sia informazione classica che informazione quantistica, e la capacità è caratterizzata diversamente nei due casi. L'obiettivo di questa tesi è quello di studiare la natura dei canali quantistici per, nota la definizione operativa della capacità nella teoria dell'informazione classica, ottenere le espressioni della loro capacità classica e della loro capacità quantistica. Per questa trattazione, si parte dalla riformulazione della meccanica quantistica con il formalismo dell'operatore densità, formalismo che permette di modellizzare matematicamente i canali quantistici come mappe completamente positive che conservano la traccia. Successivamente si dimostra il teorema di Shannon, teorema che dà l'espressione operativa della capacità classica e punto di partenza per tutte le generalizzazioni quantistiche. Proseguendo quindi, per la capacità classica di un canale quantistico si affronta la distinzione dovuta alla possibilità di avere in ingresso nel canale stati separabili o stati entangled; per la capacità quantistica si ottiene la formula regolarizzata generale e si presenta una classe di canali per cui non serve la regolarizzazione.

Indice

Introduzione

Nel 1948 C. E. Shannon pubblicò un articolo [11] con cui diede inizio alla teoria dell'informazione. In tale articolo presentò una formulazione matematica, basata sulla teoria della probabilità, per parlare della codifica, della trasmissione e dell'estrazione dell'informazione. Shannon affermò che il problema principale della comunicazione è quello di riprodurre, correttamente o approssimativamente, in un punto un messaggio prodotto in un altro punto. A veicolare il messaggio è un sistema di comunicazione formato da una sorgente d'informazione, un codificatore, un canale e un decodificatore. Il canale di comunicazione assume particolare importanza, in quanto esso prende in ingresso l'informazione in un certo stato ma può modificare tale stato e restituirlo trasformato all'uscita. Comprendere le proprietà di un canale di comunicazione è quindi fondamentale per conoscere entro quali limiti è permessa la trasmissione d'informazione.

Il XX secolo fu anche il momento in cui si affermò la meccanica quantistica, framework teorico che permette di descrivere la natura a livello microscopico e che prevede effetti che non trovano alcun riscontro nella fisica classica. La meccanica quantistica ha una natura intrinsecamente probabilistica, in essa gioca un ruolo fondamentale l'incertezza, ammette che i sistemi si trovino in stati di sovrapposizione e prevede l'entanglement, un fenomeno per cui due o più sistemi possono presentare forti correlazioni tra loro.

A partire dal 1960 si iniziò a tener conto degli effetti quantistici nei sistemi di comunicazione e nel 1970 si svilupparono tecniche per controllare singolarmente gli atomi. Questo portò alla nascita della teoria dell'informazione quantistica, la teoria che studia come i sistemi quantistici possono essere usati per codificare ed elaborare l'informazione e che seguì uno sviluppo analogo alla teoria di Shannon classica. Un sistema di comunicazione quantistico presenta gli stessi elementi di un sistema di comunicazione classico ma singolarmente gli elementi che li compongono si comportano in maniera molto diversa. Il codificatore deve preparare un sistema quantistico in uno stato che rispetti la distribuzione di probabilità del messaggio che deve essere trasmesso. Il canale quantistico non può essere modellizzato solamente in termini probabilistici come quello classico ma deve poter rappresentare tutte le possibili trasformazioni che un sistema fisico quantistico può

subire e restituire quindi un sistema diverso. Ciò viene implementato matematicamente con mappe completamente positive che conservano la traccia; un esempio concreto di canale quantistico è rappresentato da un cavo di fibra ottica. Il decodificatore deve effettuare un'operazione di misura sul sistema quantistico, che può quindi venire perturbato. La teoria dell'informazione quantistica trova applicazione nella creazione di protocolli per la crittografia quantistica e nella computazione quantistica per la creazione degli stessi computer quantistici e per l'ideazione di algoritmi quantistici che possano essere implementati su essi. I concetti della teoria dell'informazione quantistica, inoltre, si applicano allo studio dei sistemi a molti corpi, per capire come l'informazione si propaga nei materiali quantistici, ma anche nella gravità quantistica, per capire cosa succede all'informazione che cade in un buco nero.

Il sistema quantistico più semplice è il *qubit*, un sistema quantistico a due livelli - di cui sono esempi lo spin di un elettrone, la polarizzazione di un fotone o atomi aventi uno stato fondamentale e uno stato eccitato. Il qubit è l'unità di misura dell'informazione quantistica ed è l'analogo del *bit* introdotto da Shannon, una grandezza binaria che può assumere come valori "0" o "1" e che rappresenta l' unità di misura classica.

Nella teoria quantistica dell'informazione, così come in quella classica, per comprendere come può avvenire la comunicazione, è di primaria importanza conoscere le proprietà dei canali quantistici. È questo l'argomento principale di questa tesi, in cui in particolare si vuole studiare la capacità dei canali quantistici: grandezza che misura la massima quantità d'informazione che può essere inviata attraverso un canale quantistico in modo tale che dallo stato in uscita si possa ricostruire in maniera attendibile il messaggio contenuto nello stato in entrata. Date le proprietà dei sistemi quantistici non previste classicamente, la teoria quantistica presenta molte novità rispetto a quella classica.
In dettaglio la tesi è così organizzata:

- Nel Capitolo 1 si presentano i risultati principali della meccanica quantistica in termini degli operatori densità. Questo formalismo permette di rappresentare gli stati come operatori, di costruire mappe per qualsiasi trasformazione che può avvenire sugli stati del sistema e di descrivere l'evoluzione di sistemi aperti. Inoltre si presentano l'entropia di Shannon e l'entropia di von Neumann, grandezze in termini delle quali sono formulati i risultati delle due teorie dell'informazione.

- Nel Capitolo 2 si dà la definizione classica di capacità di un canale e si dimostra il teorema con cui Shannon afferma che essa è pari alla massima informazione mutua tra la variabile in ingresso e quella in uscita nel canale.

- Nel Capitolo 3 si inizia l'analisi dei canali quantistici. Si parte dallo studio della loro possibilità di trasmettere informazione classica e si cerca un'espressione della capacità per questo utilizzo dei canali.
- Nel Capitolo 4 si analizza la trasmissione di informazione quantistica da parte dei canali quantistici, cercando un'espressione per quella che viene detta capacità quantistica.

Capitolo 1

Preliminari

Nella teoria dell'informazione quantistica si usa il formalismo dell'operatore densità perché permette di descrivere sistemi in stati puri, di descrivere sistemi che possono trovarsi in miscele statistiche di stati, di studiare l'evoluzione di sistemi aperti e, dato un sistema composto, di potersi ricondurre alla descrizione di una sola parte del sistema. Questo capitolo inizia presentando i risultati principali della meccanica quantistica con tale formalismo. Successivamente sono introdotte l'entropia di Shannon e l'entropia di von Neumann, concetti fondamentali rispettivamente della teoria dell'informazione classica e di quella quantistica, con le loro principali proprietà.

1.1 Richiami di meccanica quantistica

Nella meccanica quantistica, lo stato di un sistema è generalmente descritto da un vettore di stato $|\psi\rangle$ nello spazio di Hilbert $\mathcal{H}$. Una descrizione equivalente si può fare con l'*operatore densità*.

Dato uno stato puro $|\psi\rangle \in \mathcal{H}$ si introduce l'operatore densità:

$$\rho_\psi = |\psi\rangle \langle\psi| . \tag{1.1}$$

Esso ha le seguenti proprietà:

1. è limitato, con $\|\rho_\psi\| = 1$;
2. è hermitiano, $\rho = \rho^\dagger$;
3. è positivo, $\forall\, |\phi\rangle \in \mathcal{H}\ \ \langle\phi| \rho_\psi |\phi\rangle \geq 0$;
4. $\mathrm{tr}[\rho_\psi] = 1$;

5. è idempotente, $\rho_\psi^2 = \rho_\psi$.

Un sistema può però trovarsi in uno tra vari stati $|\psi_i\rangle$, in ciascuno rispettivamente con probabilità p_i tali che $\sum_i p_i = 1$, per tale sistema, l'operatore densità è definito come segue:

$$\rho = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| . \tag{1.2}$$

Si dice che tale espressione rappresenta uno *stato misto*. Esso ha tutte le proprietà elencate per il sistema in uno stato puro ad eccezione della (5).

Tutti i postulati della meccanica quantistica possono essere riformulati con gli operatori densità.

Postulato 1. *A ciascun sistema fisico quantistico è associato uno spazio di Hilbert $\mathcal{H}$. Lo stato del sistema è completamente descritto da un operatore densità ρ che agisce sullo spazio $\mathcal{H}$.*

Postulato 2. *Ogni grandezza osservabile A del sistema fisico si rappresenta con un operatore lineare hermitiano $A = A^\dagger$. Il valore di aspettazione dell'osservabile A su uno stato ρ si calcola come: $< A >= tr(\rho A)$.*

Postulato 3. *L'evoluzione temporale di un sistema quantistico chiuso è descritta da un operatore unitario U: $\rho(t) = U(t)\rho(0)U(t)^\dagger$.*

Postulato 4. *La misura quantistica è descritta da un insieme di operatori di misura $\{M_m\}$. Tali operatori agiscono sullo spazio che rappresenta il sistema su cui viene fatta la misura. L'indice m indica i possibili risultati di misura che si possono ottenere. Se il sistema è nello stato ρ subito prima della misura, la probabilità di ottenere il risultato m è data da:*

$$p(m) = tr(M_m^\dagger M_m \rho), \tag{1.3}$$

e lo stato del sistema dopo la misura diventa:

$$\frac{M_m \rho M_m^\dagger}{tr(M_m^\dagger M_m \rho)}. \tag{1.4}$$

Gli operatori di misura rispettano la relazione:

$$\sum_m M_m^\dagger M_m = \mathbb{I}. \tag{1.5}$$

Postulato 5. *Lo spazio di Hilbert per un sistema composto è il prodotto tensore degli spazi dei vari componenti,* $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. *Se ciascun sottosistema è preparato in uno stato* ρ_i, *lo stato globale del sistema sarà* $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$.

A proposito del Postulato 4 è necessario specificare che esistono diversi formalismi per la misura.
Quando gli operatori associati agli esiti di misura sono proiettori ortogonali, si parla di formalismo *PVM*[1].
Il formalismo più generale è il formalismo *POVM*[2]. In questo caso, a partire dagli operatori M_m si definiscono gli *elementi POVM* come:

$$E_m \equiv M_m^\dagger M_m, \tag{1.6}$$

i quali sono operatori positivi che soddisfano la relazione di completezza $\sum_m E_m = \mathbb{I}$. L'insieme $\{E_m\}$ costituisce un *POVM*. Il formalismo POVM include il formalismo PVM ed è quello che verrà usato in questo elaborato.

Gli operatori densità sono estremamente utili quando si ha un sistema composto e si vuole descrivere un solo sottosistema. Dati i sistemi A e B, il cui stato totale è descritto da ρ^{AB}, si definisce *operatore densità ridotto* per lo stato A:

$$\rho^A \equiv \mathrm{tr}_B(\rho^{AB}) = \sum_i \langle b_i| \rho^{AB} |b_i\rangle , \tag{1.7}$$

dove tr_B è la *traccia parziale* sul sistema B e $\{|b_i\rangle\}$ una base ortonormale dello spazio $\mathcal{H}_B$. Analogamente per la descrizione del solo sottosistema B si fa la traccia parziale di ρ^{AB} sul sistema A.
Dato un sistema composto AB si dice che lo stato che descrive il sistema totale è *separabile* se si può scrivere come prodotto tensoriale dei singoli stati dei sottosistemi $\rho^{AB} = \rho^A \otimes \rho^B$; se non è possibile scrivere lo stato totale come un prodotto tensoriale degli stati dei sottosistemi, si dice che lo stato totale *entangled*.

Un sistema quantistico può essere soggetto a diversi processi che ne determinano l'evoluzione. Tali processi possono essere descritti mediante il formalismo dei *canali quantistici*, ovvero mappe CTCP[3] $\mathcal{E} : \rho \to \rho'$ che soddisfano le seguenti proprietà:

1. Sono lineari, $\mathcal{E}(\alpha\rho_1 + \beta\rho_2) = \alpha\mathcal{E}(\rho_1) + \beta\mathcal{E}(\rho_2)$.

[1] Dall'inglese *Projection-Valued Measurement*.
[2] Dall'inglese *Positive Operator-Valued Measurement*.
[3] Dall'inglese *Completely Positive Trace-Preserving*

2. L'output è ancora hermitiano, $\rho = \rho^\dagger \Rightarrow \mathcal{E}(\rho) = \mathcal{E}(\rho)^\dagger$.

3. Conservano la positività, $\rho \geq 0 \Rightarrow \mathcal{E}(\rho) \geq 0$.

4. Conservano la traccia, $\mathrm{tr}(\mathcal{E}(\rho)) = \mathrm{tr}(\rho)$.

Per definizione, si richiede che queste mappe siano completamente positive, cioè che ogni loro estensione sia ancora positiva.

I canali quantistici permettono di descrivere l'evoluzione di sistemi aperti. Per vedere come, consideriamo un sistema principale nello stato ρ^S che interagisce con un ambiente nello stato ρ^A. Il sistema complessivamente è chiuso, per cui lo stato $\rho^S \otimes \rho^A$ evolve con un operatore unitario U. Se si vuole conoscere lo stato del sistema principale dopo l'evoluzione, è sufficiente fare la traccia parziale rispetto all'ambiente:

$$\mathcal{E}(\rho^S) = \mathrm{tr}_A[U(\rho^S \otimes \rho^A)U^\dagger]. \tag{1.8}$$

Analogamente se il sistema totale è in uno stato entangled.
I canali quantistici, inoltre, possono essere rappresentati in una forma detta *operator-sum representation*, che permette di scrivere la mappa $\mathcal{E}$ come:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger, \tag{1.9}$$

con gli operatori $\{E_k\}$ che sono detti *operatori di Kraus* e che rispettano la relazione $\sum_k E_k^\dagger E_k = \mathbb{I}$.

1.2 Entropia di Shannon

Sia X una variabile casuale. Siano i valori x che tale variabile può assumere appartenenti all'alfabeto $\mathcal{X}$ e distribuiti secondo la distribuzione di probabilità $p(x)$. Si definisce l'*entropia* per misurare l'incertezza di una variabile casuale.

Definizione 1.1. (Entropia di Shannon) L'*entropia di Shannon* di una variabile casuale discreta X con distribuzione di probabilità $p(x)$ è definita come:

$$H(X) \equiv -\sum_x p(x) \log p(x), \tag{1.10}$$

con 'log' che si riferisce al logaritmo in base 2, questa convenzione verrà usata nel resto dell'elaborato.

L'entropia si misura in bits. Per convenzione si considera $0 \log 0 \equiv 0$.
L'entropia di una variabile X che ammette solo due esiti, uno con probabilità p e l'altro con probabilità $1 - p$, prende il nome di *entropia binaria*:

$$H_{bin}(p) \equiv H(p) \equiv -p \log p - (1 - p) \log(1 - p). \tag{1.11}$$

L'entropia di Shannon rispetta le seguenti proprietà.

Teorema 1.1. (Proprietà dell'entropia di Shannon)

(1) È non negativa: $H(X) \geq 0$, per ogni variabile casuale discreta X.

(2) È concava rispetto alla distribuzione di probabilità $p(x)$.

(3) Si annulla se e solo se X è una variabile deterministica.

(4) Ammette un valore massimo:

$$H(X) \leq \log |\mathcal{X}|, \tag{1.12}$$

con $|\mathcal{X}|$ numero di elementi dell'alfabeto $\mathcal{X}$. L'uguaglianza si ha se e solo se X ha una distribuzione di probabilità uniforme.

Quando si ha una coppia di variabili casuali X e Y si possono definire altre grandezze.

Definizione 1.2. (Entropia congiunta) L'*entropia congiunta* di una coppia di variabili casuali discrete (X, Y) con distribuzione di probabilità congiunta $p(x, y)$ è definita come:

$$H(X, Y) \equiv -\sum_{x,y} p(x, y) \log p(x, y). \tag{1.13}$$

Essa misura l'incertezza totale sulla coppia (X, Y).
Quando si è a conoscenza del valore di Y, l'incertezza sulla coppia (X, Y) diminuisce di una quantità $H(Y)$, per l'incertezza rimanente su X allora si definisce:

Definizione 1.3. (Entropia condizionata) L'*entropia di X condizionata* dalla conoscenza di Y è:

$$H(X|Y) \equiv H(X, Y) - H(Y). \tag{1.14}$$

Si definisce anche una grandezza che permette di conoscere la quantità di informazione che due variabili condividono:

Definizione 1.4. (Mutua informazione) Siano X e Y due variabili casuali discrete con distribuzione di probabilità congiunta $p(x, y)$. La *mutua informazione* si definisce come:

$$I(X : Y) \equiv H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y). \tag{1.15}$$

Enunciamo, senza dimostrare, alcune proprietà delle grandezze appena definite.

Teorema 1.2. (Proprietà fondamentali)

(1) $H(X, Y) = H(Y, X)$; $\quad I(X : Y) = I(Y : X)$.

(2) $H(Y|X) \geq 0$ *e quindi* $I(X : Y) \leq H(Y)$, *l'uguaglianza vale se e solo se* Y *è una funzione di* X.

(3) $H(X) \leq H(X, Y)$, *l'uguaglianza vale se e solo se* Y *è una funzione di* X.

(4) ***Subadditività****:* $H(X, Y) \leq H(X) + H(Y)$ *l'uguaglianza vale se e solo se* X *e* Y *sono variabili casuali indipendenti.*

(5) $H(Y|X) \leq H(Y)$ *e quindi* $I(X : Y) \geq 0$, *l'uguaglianza vale se e solo se* X *e* Y *sono variabili casuali indipendenti.*

(6) ***Il condizionamento riduce l'entropia****:* $H(X|Y, Z) \leq H(X|Y)$.

Le grandezze sopra definite verificano delle disuguaglianze che vengono poi estese in teoria dell'informazione quantistica, le riportiamo in seguito.
Una sequenza di variabili casuali $X_1 \to X_2 \to \ldots$, si dice *sequenza di Markov* se la variabile X_{n+1}, data la variabile X_n, è indipendente da $X_1 \ldots X_{n-1}$, ovvero:

$$p(X_{n+1} = x_{n+1}|X_n = x_n, \ldots, X_1 = x_1) = p(X_{n+1} = x_{n+1}|X_n = x_n). \tag{1.16}$$

Teorema 1.3. (Disuguaglianza di processazione dei dati) *Sia* $X \to Y \to Z$ *una catena di Markov. Allora:*

$$H(X) \geq I(X : Y) \geq I(X : Z). \tag{1.17}$$

Vale l'uguaglianza nella prima parte se, dato Y *si può ricostruire* X.

1.3 Entropia di von Neumann

In teoria dell'informazione quantistica, l'analogo dell'entropia di Shannon è l'*entropia di von Neumann*. Essa si definisce sugli operatori densità, che prendono il ruolo delle distribuzioni di probabilità della descrizione classica.

Definizione 1.5. (Entropia di von Neumann) Sia un sistema quantistico preparato in uno stato ρ. Allora la sua *entropia di von Neumann* è definita come:

$$S(\rho) \equiv -\mathrm{tr}(\rho \log \rho). \tag{1.18}$$

Se ρ ha un set di autovalori λ_x allora l'entropia diventa $S(\rho) = -\sum_x \lambda_x \log \lambda_x$, per cui si considera $0 \log 0 \equiv 0$.

L'entropia di von Neumann ha molte proprietà utili.

Teorema 1.4. (Proprietà dell'entropia di von Neumann)

(1) È non negativa: $S(\rho) \geq 0$. L'uguaglianza vale se e solo se lo stato è puro.

(2) In uno spazio di Hilbert d-dimensionale, ammette un valore valore massimo:

$$S(\rho) \leq \log d, \tag{1.19}$$

l'uguaglianza vale solo se il sistema è nello stato completamente misto $\mathbb{I}/d$.

(3) È concava: $S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$.

(4) Considerato il sistema composto AB in uno stato puro, allora vale $S(A) = S(B)$.

(5) Siano p_i probabilità e siano i supporti degli stati ρ_i ortogonali tra loro. Allora:

$$S\left(\sum_i p_i \rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i). \tag{1.20}$$

(6) Siano p_i probabilità, siano $|i\rangle$ stati ortogonali per un sistema A, siano ρ_i operatori densità per un sistema B. Allora:

$$S\left(\sum_i p_i \left|i\right\rangle \left\langle i\right| \otimes \rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i). \tag{1.21}$$

Si definiscono:

Definizione 1.6. (Entropia di von Neumann congiunta) L'*entropia congiunta* per un sistema composto AB è definita come:

$$S(A, B) \equiv -\mathrm{tr}(\rho^{AB} \log(\rho^{AB})), \tag{1.22}$$

con ρ^{AB} matrice densità del sistema AB.

L'entropia congiunta di von Neumann verifica le seguenti disuguaglianze:

- **Subadditività**: $S(A,B) \leq S(A) + S(B)$. L'ugualianza vale se lo stato totale del sistema AB è separabile, $\rho^{AB} = \rho^A \otimes \rho^B$.
- **Disuguaglianza di Araki-Lieb**: $S(A,B) \geq |S(A) - S(B)|$.

Le due disuguaglianze precedenti si possono estendere a più di due sistemi, così si ottiene la disuguaglianza di *subadditività forte*:

$$S(A,B,C) + S(B) \leq S(A,B) + S(B,C). \tag{1.23}$$

L'analogo quantistico dell'entropia condizionata è:

Definizione 1.7. (Entropia di von Neumann condizionata) L'entropia sul sistema A condizionata dalla conoscenza del sistema B è data da:

$$S(A|B) \equiv S(A,B) - S(B). \tag{1.24}$$

Ancora, analogamente al caso classico si definisce:

Definizione 1.8. (Mutua informazione quantistica) Per due sistemi quantistici A e B, la mutua informazione si definisce come:

$$I(A:B) \equiv S(A) + S(B) - S(A,B) \tag{1.25}$$

$$= S(A) - S(A|B) = S(B) - S(B|A). \tag{1.26}$$

Per molte applicazioni sono utili ulteriori proprietà dell'entropia quantistica racchiuse nel seguente teorema:

Teorema 1.5. *(1)* (Il condizionamento riduce l'entropia.) *Sia ABC un sistema quantistico composto. Allora $S(A|B,C) \leq S(A|B)$.*

(2) (Eliminare un sottosistema quantistico non aumenta mai la mutua informazione.) *Sia ABC un sistema quantistico composto. Allora $I(A:B) \leq I(A:B,C)$.*

(3) (I canali quantistici non aumentano la mutua informazione.) *Sia AB un sistema quantistico composto e $\mathcal{E}$ una quantum operation sul sistema B che conserva la traccia. Sia $I(A:B)$ la mutua informazione prima che $\mathcal{E}$ sia applicata al sistema B e $I(A':B')$ la mutua informazione dopo l'applicazione di $\mathcal{E}$. Allora $I(A':B') \leq I(A:B)$.*

Capitolo 2

Capacità classica di un canale

Nell'articolo del 1948 [11], Shannon affrontò due problemi principali, tra questi dimostrò qual è il massimo *rate* con cui si può trasmettere informazione in maniera affidabile attraverso un canale di comunicazione. Tale risultato è presentato nel teorema di codifica di canale rumoroso, in cui si dà una procedura effettiva per determinare la capacità di un canale rumoroso. L'obiettivo di questo capitolo è arrivare a dimostrare questo teorema fondamentale della teoria dell'informazione classica, sul quale si basano poi le varie estensioni quantistiche.

2.1 Modello di canale classico

La trasmissione di informazione tra un mittente ed un destinatario - spesso chiamati Alice e Bob in teoria dell'informazione - avviene efficacemente se il secondo riesce a risalire esattamente al messaggio mandato dal primo. L'informazione passa attraverso un canale che presenta rumore, ovvero che può disturbare la trasmissione, per questo motivo è necessario utilizzare un codice per la comunicazione che permetta a Bob di riottenere il messaggio mandato da Alice anche se modificato durante la trasmissione.

Definizione 2.1. (Canale di comunicazione) Un *canale di comunicazione* $\mathcal{N}$ è un sistema che, dato un alfabeto di input $\mathcal{X}$ e un alfabeto di output $\mathcal{Y}$, per ogni segnale di ingresso $x \in \mathcal{X}$ rilascia un segnale $y \in \mathcal{Y}$ secondo una distribuzione condizionata $p(y|x)$.

Un canale può essere *discreto* e *senza memoria* quando sia l'alfabeto di input che quello di output sono finiti e ogni utilizzo del canale risulta indipendente dagli altri.

Un codice per il canale $\mathcal{N}$ si costruisce partendo da un insieme di messaggi $\{1, \dots, M\}$ che, tramite una funzione di codifica $C^n : \{1, \dots, M\} \to \mathcal{X}^n$, viene associato ad un

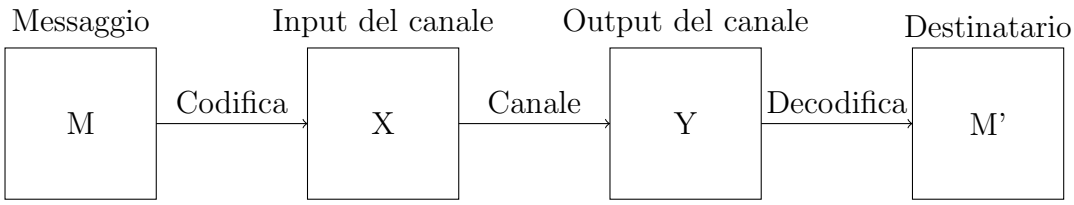

Figura 2.1: Schema di un sistema di comunicazione che permette di creare un codice. [Immagine ripresa da [8].]

insieme di sequenze di simboli di ingresso del canale $x^n = (x_1 x_2 \dots x_n)$. Ognuna di queste sequenze generate è detta parola di codice (*codeword*). Ciascuna sequenza attraversa il canale, il quale introduce rumore, e produce in uscita una sequenza $y^n \in \mathcal{Y}^n$. Ricevuta y^n, il destinatario applica su essa una funzione di decodifica $D^n : \mathcal{Y}^n \to \{1, \dots, M\}$.
Per la coppia C^n, D^n si definisce la *probabilità di errore*:

$$p(C^n, D^n) \equiv \max_{m \in \{1,\dots,M\}} p(D^n(Y^n) \neq m | X^n = C^n(m)). \tag{2.1}$$

Inoltre, per un codice di lunghezza n che permette di inviare M messaggi si definisce il *rate* come:

$$R = \frac{\log M}{n}, \tag{2.2}$$

esso misura la quantità media di informazione veicolata da ogni uso del canale. Si dice che un rate R è *affidabile* se esiste una famiglia di codici $(\lceil 2^{nR} \rceil, n)$ per cui la probabilità di errore tenda a zero per $n \to \infty$. Si definisce *capacità* $C(\mathcal{N})$ di un canale rumoroso $\mathcal{N}$ come l'estremo superiore di tutti i rate affidabili per quel canale.

2.2 Sequenze tipiche

Un modo per vedere se la decodifica di un segnale possa essere accettata è quello di valutare se le sequenza in ingresso e in uscita dal canale risultano *congiuntamente tipiche*. Di seguito riportiamo alcune definizioni e alcuni teoremi, senza dimostrarli, per capire questo concetto.
Una sorgente di informazione è modellizzata in termini probabilistici, mediante una variabile aleatoria X che può assumere valori scelti da un alfabeto $\mathcal{X}$, secondo una distribuzione di probabilità $p(x)$.

Definizione 2.2. (Sequenza ε-tipica) Considerata una sorgente che emette le variabili $X_1, X_2, \dots$ indipendenti e identicamente distribuite (*i.i.d.*), dato $\varepsilon > 0$, una sequenza

$x^n = (x_1 x_2 ... x_n)$ si definisce *ε-tipica* se:

$$2^{-n(H(X)+\varepsilon)} \leq p(x^n) \leq 2^{-n(H(X)-\varepsilon)} \tag{2.3}$$

O, alternativamente, se:

$$\left| \frac{1}{n} \log \frac{1}{p(x^n)} - H(X) \right| \leq \varepsilon. \tag{2.4}$$

L'insieme delle sequenze ε-tipiche di lunghezza n si indica con $T(n, \varepsilon)$.
Utilizzando la legge dei grandi numeri si dimostrano una serie di proprietà per le sequenze ε-tipiche, racchiuse nel seguente:

Teorema 2.1. (delle sequenze tipiche)

(1) Sia fissato $\varepsilon > 0$. Allora per ogni $\delta > 0$, per n sufficientemente grande, la probabilità che una sequenza sia ε-tipica è almeno $1 - \delta$.

(2) Per ogni coppia di $\varepsilon > 0$ e $\delta > 0$ fissati, per n sufficientemente grande, il numero $|T(n, \varepsilon)|$ di sequenze ε-tipiche soddisfa:

$$(1 - \delta) 2^{n(H(X)-\varepsilon)} \leq |T(n, \varepsilon)| \leq 2^{n(H(X)+\varepsilon)}. \tag{2.5}$$

(3) Sia $S(n)$ una collezione contenente al massimo 2^{nR} sequenze di lunghezza n, con $R < H(X)$ fissato. Allora per ogni $\delta > 0$ e per n abbastanza grande,

$$\sum_{x^n \in S(n)} p(x^n) \leq \delta. \tag{2.6}$$

Si rimanda a [8] per la dimostrazione.

Definizione 2.3. (Sequenze congiuntamente ε-tipiche) La coppia di sequenze (x^n, y^n), con $x^n = (x_1 x_2 \dots x_n)$ e $y^n = (y_1 y_2 \dots y_n)$, si dice *congiuntamente ε-tipica* se valgono contemporaneamente:

$$\begin{aligned} 2^{-n(H(X)+\varepsilon)} &\leq p(x^n) \leq 2^{-n(H(X)-\varepsilon)}, \\ 2^{-n(H(Y)+\varepsilon)} &\leq p(y^n) \leq 2^{-n(H(Y)-\varepsilon)}, \\ 2^{-n(H(X,Y)+\varepsilon)} &\leq p(x^n, y^n) \leq 2^{-n(H(X,Y)-\varepsilon)}, \end{aligned} \tag{2.7}$$

con $H(X, Y)$ entropia congiunta.

Ci sono circa $2^{nH(X)}$ sequenze ε-tipiche tra le sequenze fatte a partire dalla variabile X e circa $2^{nH(Y)}$ sequenze ε-tipiche tra quelle fatte a partire dalla variabile Y. Il numero di sequenze congiuntamente ε-tipiche è però pari a $2^{nH(X,Y)}$, quindi non tutte le coppie di sequenze (x^n, y^n) tipiche singolarmente sono anche congiuntamente tipiche.

2.3 Teorema di codifica di un canale

Dopo aver visto come si costruisce un modello di canale classico con rumore e il concetto di sequenze tipiche, si può procedere con la dimostrazione del teorema di Shannon sulla codifica di canale rumoroso. La formula finale che si ottiene per la capacità del canale, permette di vedere il suo calcolo come un problema di ottimizzazione.

Teorema 2.2. (di Shannon sulla codifica di un canale rumoroso) *La capacità di un canale rumoroso $\mathcal{N}$ è data da:*

$$C(\mathcal{N}) = \max_{p(x)} I(X:Y) \tag{2.8}$$

dove il massimo è preso tra tutte le possibili distribuzioni di probabilità per la variabile casuale X all'ingresso del canale, mentre Y è la corrispondente variabile casuale in uscita.

Dimostrazione. Seguiamo l'argomentazione presente in [9] per delineare una dimostrazione di questo risultato.

Consideriamo il canale $\mathcal{N}$ caratterizzato dalla distribuzione di probabilità condizionata $p(y|x)$ e assumiamo che ciascun singolo simbolo in ingresso nel canale sia descritto dalla variabile casuale $X = \{x, p(x)\}$.
Si costruisce un codice casuale fatto di parole di codice di lunghezza n e rate R, scegliendo 2^{nR} parole di codice dall'insieme X^n. Questo codice è conosciuto sia da Alice che da Bob.
Ciascuna parola di codice è trasmessa con n usi indipendenti del canale e la probabilità condizionata $p(y|x)$ agisce ogni volta. Per ogni uso del canale si crea quindi una distribuzione congiunta (X, Y); per estensione, per n usi si ha la distribuzione congiunta (X^n, Y^n).
Quando Bob riceve una sequenza y^n, tenta di risalire alla sequenza di ingresso corrispondente verificando se esiste una sequenza x^n che sia congiuntamente tipica con y^n. Se tale sequenza esiste ed è unica, la usa per decodificare; altrimenti decodifica arbitrariamente.
Per ciascuna coppia $\delta > 0$ e $\varepsilon > 0$ una coppia (x'^n, y^n) è congiuntamente ε-tipica con probabilità almeno $1 - \delta$. Bisogna capire cosa fare quando si ha più di una sequenza congiuntamente tipica con y^n.
Sia prodotta in input la sequenza x^n e successivamente x'^n. Sia y^n la sequenza di output corrispondente alla sequenza trasmessa x^n. Dall'indipendenza di x^n e x'^n si ha: $p(x^n, x'^n) = p(x^n)p(x'^n)$ e $p(x'^n, y^n) = p(x'^n)p(y^n)$.

Detto $N_{s.t.}$ il numero di tutte le possibili sequenze congiuntamente ε-tipiche e usando la terza delle equazioni (2.7), si ha:

$$N_{s.t.}2^{-n(H(X,Y)+\varepsilon)} \leq \sum_{(x^n,y^n)s.t.} p(x^n,y^n) \leq 1 \quad \Longrightarrow \quad N_{s.t.} \leq 2^{n(H(X,Y)+\varepsilon)} \tag{2.9}$$

Una sequenza x'^n è ε-tipica con $p(x'^n) \leq 2^{-n(H(X)-\varepsilon)}$. Analogamente, una sequenza y^n è ε-tipica con $p(y^n) \leq 2^{-n(H(Y)-\varepsilon)}$. Ne consegue:

$$\sum_{(x'^n,y^n)s.t.} p(x'^n,y^n) = \sum_{(x'^n,y^n)s.t.} p(x'^n)p(y^n) \leq N_{s.t.}2^{-n(H(X)-\varepsilon)}2^{-n(H(Y)-\varepsilon)} \tag{2.10}$$

$$\leq 2^{n(H(X,Y)+\varepsilon)}2^{-n(H(X)-\varepsilon)}2^{n(H(Y)-\varepsilon)} \tag{2.11}$$

$$= 2^{-n(H(X)+H(Y)-H(X,Y)-3\varepsilon)} \tag{2.12}$$

$$= 2^{-n(I(X:Y)-3\varepsilon)}. \tag{2.13}$$

L'espressione (2.13) fornisce un limite superiore alla probabilità di avere una coppia indipendente (x'^n, y^n) congiuntamente ε-tipica, moltiplicandola per il numero totale di parole di codice, si ottiene la probabilità che qualsiasi altra sequenza di ingresso oltre x^n sia congiuntamente tipica con y^n:

$$2^{nR}2^{-n(I(X:Y)-3\varepsilon)} = 2^{n(R-I(X:Y)+3\varepsilon)}. \tag{2.14}$$

L'obiettivo è cercare di minimizzare questa probabilità per avere una sola sequenza di ingresso che sia congiuntamente tipica con y^n e quindi ottimizzare la decodifica. Scegliendo $R = I(X:Y) - c$, con $c > 0$, e nel limite $n \to \infty$, la probabilità di errore durante la decodifica va a zero.

La probabilità d'errore tendente a zero è stata ottenuta rispetto a una media sulle parole di codice - poiché tutte le parole di codice x^n sono equiprobabili per costruzione - e sui codici - scegliendone uno casuale. Quindi, deve esistere una sequenza di codici che, per $n \to \infty$, hanno probabilità d'errore piccola. In tali codici poi si può ottenere che la probabilità d'errore vada asintoticamente a 0 per ogni parola di codice, riducendo opportunamente il numero di parole stesse del codice. Se ne conclude che il rate:

$$R = I(X:Y) - o(1) \tag{2.15}$$

è raggiungibile asintoticamente con probabilità d'errore trascurabile, quindi è affidabile.

La mutua informazione non dipende solamente dalla probabilità condizionata $p(y|x)$ che caratterizza il canale, ma anche dalla distribuzione della variabile di ingresso $p(x)$.

Quanto ottenuto finora vale per qualsiasi $p(x)$, quindi si può stimare il rate massimo con cui si ha trasmissione affidabile nel canale, che coincide con la capacità:

$$C \equiv \max_{p(x)} I(X:Y). \tag{2.16}$$

Per concludere la dimostrazione dobbiamo dimostrare che per $R > C$ non si può avere trasmissione di informazione con probabilità d'errore tendente a zero.
Per farlo consideriamo un codice arbitrario che abbia 2^{nR} parole di codice e supponiamo che esse siano uniformemente distribuite, ovvero ciascuna ha una probabilità pari a 2^{-nR}. Associamo a questo codice la variabile aleatoria $\tilde{X}^n = \{x^n, p(x^n) = 2^{-nR}\}$. L'entropia di Shannon per questa variabile sarà:

$$H(\tilde{X}^n) = nR. \tag{2.17}$$

Ciascuna parola di codice richiede n usi del canale per essere trasmessa, per i quali si crea la distribuzione congiunta $(\tilde{X}^n, \tilde{Y}^n)$. Poiché i vari usi del canale sono indipendenti vale:

$$p(y_1 y_2 \dots y_n | x_1 x_2 \dots x_n) = p(y_1|x_1)p(y_2|x_2)\dots p(y_n|x_n). \tag{2.18}$$

Per cui, l'entropia condizionata si può scrivere come:

$$H(\tilde{Y}^n|\tilde{X}^n) = \langle -\log p(y^n|x^n)\rangle = \sum_i \langle -\log p(y_i|x_i)\rangle = \sum_i H(\tilde{Y}_i|\tilde{X}_i), \tag{2.19}$$

dove $\langle\cdot\rangle$ indica la media rispetto all'opportuna distribuzione. Per ciascuna posizione i, $\tilde{X}_i$ e $\tilde{Y}_i$ sono le variabili aleatorie corrispondenti alla i-esima lettera della parola di codice, rispettivamente in ingresso e in uscita dal canale, definite tramite distribuzioni marginali rispetto alla distribuzione delle parole di codice.
Per la subadditività dell'entropia di Shannon si ha:

$$H(\tilde{Y}^n) \leq \sum_i H(\tilde{Y}_i). \tag{2.20}$$

Allora:

$$\begin{aligned} I(\tilde{Y}^n : \tilde{X}^n) &= H(\tilde{Y}^n) - H(\tilde{Y}^n|\tilde{X}^n) && (2.21)\\ &\leq \sum_i \Big(H(\tilde{Y}_i) - H(\tilde{Y}_i|\tilde{X}_i)\Big) && (2.22)\\ &= \sum_i I(\tilde{Y}_i : \tilde{X}_i) \leq nC. && (2.23) \end{aligned}$$

Usando la simmetria della mutua informazione si ha anche:

$$I(\tilde{X}^n : \tilde{Y}^n) = H(\tilde{X}^n) - H(\tilde{X}^n|\tilde{Y}^n) \tag{2.24}$$
$$= nR - H(\tilde{X}^n|\tilde{Y}^n) \leq nC. \tag{2.25}$$

Avere decodifica affidabile significa che, per $n \to \infty$, noto il messaggio finale non si ha più incertezza sul messaggio iniziale, quindi deve essere $\frac{1}{n}H(\tilde{X}^n|\tilde{Y}^n) \to 0$. Se si ha trasmissione con probabilità d'errore asintoticamente nulla allora la (2.25) diventa:

$$R \leq C + o(1). \tag{2.26}$$

Non si può quindi avere trasmissione affidabile per $R > C$.

Abbiamo dimostrato che:

$$C \equiv \max_{p(x)} I(X : Y) \tag{2.27}$$

è il più alto dei rate affidabili, quindi la capacità del canale. □

La mutua informazione è una funzione concava, ha un massimo locale unico. È questo che rende possibile, per molti canali, il calcolo della capacità.

Capitolo 3

Informazione classica in canali quantistici

Nel capitolo precedente è stata proposta l'espressione per la capacità di un canale classico. Anche un sistema quantistico può essere usato come canale di comunicazione e, in generale, questo potrebbe portare ad avere rate di trasmissione maggiori di quelli ottenibili con un canale classico [14]. Per ottenere risultati analoghi al teorema di Shannon, si trattano separatamente il caso in cui un canale quantistico trasmette informazione classica e il caso in cui trasmette informazione quantistica. In questo capitolo si affronta il primo caso, cercando una formula per la capacità classica di un canale quantistico. Una prima generalizzazione quantistica del teorema di Shannon è il teorema di Holevo-Schumacher-Westmoreland, valido se gli stati in ingresso nel canale sono separabili. Successivamente, si discute brevemente la possibilità di usare stati entangled tra di loro all'ingresso del canale [4, 2], quindi di capacità classica entanglement-assistita.

3.1 Codifica dell'informazione in stati quantistici

Prima di procedere con la trattazione della capacità classica di un canale quantistico, vediamo come tale informazione possa essere codificata in un sistema quantistico.
Data una variabile aleatoria $X = \{x, p(x)\}$ associata ad una sorgente classica, si considera che Alice possa associare ad ogni valore di x un certo stato ρ_x di un sistema quantistico. Alice potrà così mandare un certo messaggio a Bob scegliendo lo stato in cui preparare il suo sistema dall'insieme $\Sigma = \{\rho_x, p(x)\}$. Bob può estrarre informazione classica sotto forma della variabile Y effettuando una misura con elementi POVM $\{E_y\}$ sugli stati che riceve. La quantità che misura il guadagno di informazione su X che si

ottiene a partire da Y è la mutua informazione, ma poiché Bob può scegliere quale set di operatori utilizzare per massimizzare il guadagno di informazione ottenuto, si definisce l'*informazione accessibile* dell'insieme Σ:

$$I_{acc}(\Sigma) \equiv \max_{\{E_y\}} I(X:Y). \tag{3.1}$$

Questa è limitata superiormente dal *teorema di Holevo.*

Teorema 3.1. (Limite di Holevo) *Sia lo stato preparato scelto dall'insieme* $\{\rho_x, p_x\}$*, con* $(0,\ldots,n)$ *set di indici per* x*. Sia* $\{E_y\} = \{E_0,\ldots,E_m\}$ *il set di operatori POVM con cui viene fatta la misura sullo stato, sia* Y *il risultato della misura; allora il limite di Holevo afferma che per qualsiasi misura fatta sullo stato:*

$$I(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x) \tag{3.2}$$

con $\rho = \sum_x p_x \rho_x$*.*

La quantità di destra della disuguaglianza 3.2 è detta *quantità* χ *di Holevo.*

Dimostrazione. Siano Q, P, M tre sistemi:

- Q è il sistema quantistico i cui possibili stati sono gli stati ρ_x.
- P è il sistema di preparazione. Si considera avere una base ortonormale $\{|x\rangle\}$, tali x possono prendere i valori del set di indici $0,\ldots,n$.
- M può essere visto come il sistema corrispondente allo strumento di misura. Ha una base $\{|y\rangle\}$, con y che può assumere i valori del set di indici $0,\ldots,n$ in base alla misura.

Per semplificare la notazione, di seguito indicheremo l'entropia di von Neumann di uno stato di un certo sistema, specificando solamente il sistema considerato.

Lo stato iniziale del sistema totale PQM è:

$$\rho^{PQM} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|. \tag{3.3}$$

Sia $\mathcal{E}$ un canale quantistico che agisce solamente sul sistema QM:

$$\mathcal{E}(\sigma \otimes |0\rangle\langle 0|) \equiv \sum_y \sqrt{E_y}\sigma\sqrt{E_y} \otimes |y\rangle\langle y| \tag{3.4}$$

con σ un generico stato di Q e $|0\rangle$ lo stato iniziale di M.
Inizialmente non c'è correlazione tra M e il resto del sistema, quindi: $I(P : Q) = I(P : Q, M)$. Richiamando il Teorema 1.5 si ha: $I(P : Q, M) \geq I(P' : Q', M')$ e $I(P' : Q', M') \geq I(P' : M')$. Dove le lettere P, Q, M indicano i tre sistemi prima dell'applicazione di $\mathcal{E}$ e P', Q', M' sono i sistemi successivamente. Mettendo insieme questi risultati si ottiene:

$$I(P' : M') \leq I(P : Q). \tag{3.5}$$

Si consideri il membro di destra della (3.5). Uno stato del sistema PQ è descritto da:

$$\rho^{PQ} = \sum_x p_x |x\rangle \langle x| \otimes \rho_x, \tag{3.6}$$

per cui si ha: $S(P) = H(p_x)$; $S(Q) = S(\rho)$ con $\rho = \sum_x p_x \rho_x$ e - dal Teorema 1.4 - $S(P, Q) = H(p_x) + \sum_x p_x S(\rho_x)$. Da cui:

$$I(P : Q) = S(P) + S(Q) - S(P, Q) = S(\rho) - \sum_x p_x S(\rho_x). \tag{3.7}$$

Si consideri il membro di sinistra della 3.5. Dopo l'azione di $\mathcal{E}$, il sistema totale è nello stato:

$$\rho^{P'Q'M'} = \sum_{xy} p_x |x\rangle \langle x| \otimes \sqrt{E_y} \rho_x \sqrt{E_y} \otimes |y\rangle \langle y| . \tag{3.8}$$

Facendo la traccia parziale sul sistema Q' e notando che vale $p(x, y) = p_x p(y|x) = p_x \mathrm{tr}(\rho_x E_y) = p_x \mathrm{tr}(\sqrt{E_y} \rho_x \sqrt{E_y})$, rimane:

$$\rho^{P'M'} = \sum_{xy} p(x, y) |x\rangle \langle x| \otimes |y\rangle \langle y| . \tag{3.9}$$

L'ultima espressione permette di vedere che $I(P' : M') = I(X : Y)$. Sostituendo le espressioni trovate per i due membri della 3.5, si vede che è dimostrata la tesi (3.2). □

Dal teorema del limite di Holevo, segue un importante corollario:

Corollario 1. *Dato un insieme $\{\rho_x, p(x)\}$, in cui gli stati ρ_x sono stati di un sistema rappresentato da uno spazio di Hilbert $\mathcal{H}$ d-dimensionale, allora vale:*

$$I_{acc}(\{\rho_x, p(x)\}) \leq \log dim(\mathcal{H}). \tag{3.10}$$

Il limite di Holevo mostra quindi che da n qubit non è possibile ottenere più di n bit di informazione classica. Tuttavia ciò non rappresenta un ostacolo per il vantaggio quantistico, in quanto è stato dimostrato, sia teoricamente che sperimentalmente, che l'uso di stati quantistici - grazie alla sovrapposizione quantistica - consente di utilizzare l'informazione in modi non permessi classicamente. [7]

3.2 Sottospazi tipici

Se per parlare di capacità classica si usa il concetto di sequenze tipiche, quando si usano stati di sistemi quantistici si parla di *sottospazi tipici*. Di seguito ne vengono riportate definizione e proprietà.
Sia ρ un operatore densità che ammette decomposizione ortogonale: $\rho = \sum_x p(x) |x\rangle \langle x|$, quindi $|x\rangle$ sono autovettori ortonormali e $p(x)$ autovalori di ρ. Poiché sono non negativi e $\mathrm{tr}(\rho) = 1$, gli autovalori $p(x)$ corrispondono ad una distribuzione di probabilità. Da ciò: $H(p(x)) = S(\rho)$. Per gli indici x si può parlare di sequenze $x_1 \dots x_n$ ϵ-tipiche:

$$\left| \frac{1}{n} \log \left(\frac{1}{p(x_1) \dots p(x_n)} \right) - S(\rho) \right| \leq \epsilon. \tag{3.11}$$

A una sequenza ϵ-tipica si associa lo stato ϵ-tipico $|x_1\rangle |x_2\rangle \dots |x_n\rangle$.

Definizione 3.1. (Sottospazio ϵ-tipico) Si definisce *sottospazio ϵ-tipico*, il sottospazio generato da tutti gli stati ϵ-tipici $|x_1\rangle |x_2\rangle \dots |x_n\rangle$. Viene indicato con $T(n, \epsilon)$.

Il proiettore sul sottospazio ϵ-tipico $T(n, \epsilon)$ ha la seguente espressione:

$$P(n, \epsilon) = \sum_{x^n \epsilon - tipiche} |x_1\rangle \langle x_1| \otimes |x_2\rangle \langle x_2| \otimes \dots |x_n\rangle \langle x_n| . \tag{3.12}$$

Enunciamo l'analogo quantistico del Teorema 2.1.

Teorema 3.2. (dei sottospazi tipici)

(1) Sia fissato $\epsilon > 0$. Allora per ogni $\delta > 0$, per n sufficientemente grande, vale:

$$tr(P(n, \epsilon)\rho^{\otimes n}) \geq 1 - \delta. \tag{3.13}$$

(2) Per ogni coppia di $\epsilon > 0$ e $\delta > 0$ fissati, per n sufficientemente grande, la dimensione del sottospazio ϵ-tipico $|T(n, \epsilon)| = tr(P(n, \epsilon))$ di $T(n, \epsilon)$ soddisfa:

$$(1 - \delta) 2^{n(S(\rho) - \epsilon)} \leq |T(n, \epsilon)| \leq 2^{n(S(\rho) + \epsilon)}. \tag{3.14}$$

(3) Sia $S(n)$ un proiettore su un qualsiasi sottospazio di $H^{\otimes n}$ di dimensione al massimo 2^{nR}, con $R < S(\rho)$ fissato. Allora per ogni $\delta > 0$ e per n sufficientemente grande:

$$tr(S(n)\rho^{\otimes n}) \leq \delta. \tag{3.15}$$

3.3 Capacità per stati prodotto

Le grandezze introdotte nei paragrafi precedenti, tra cui l'informazione accessibile e la quantità di Holevo, sono il punto di partenza per dimostrare il teorema di Holevo-Schumacher-Westmoreland. Tale teorema costituisce una prima estensione quantistica del teorema di codifica di un canale rumoroso di Shannon, ma permette di ottenere una formula per la capacità classica di un canale quantistico solo se gli stati in ingresso nel canale sono stati separabili $\rho_1 \otimes \rho_2 \ldots$, si parla quindi di *capacità per stati prodotto*, indicata con $C^{(1)}(\mathcal{E})$ - dove $\mathcal{E}$ rappresenta il canale. Per generalizzare l'espressione della capacità classica serve una regolarizzazione su più usi del canale.

3.3.1 Teorema di Holevo-Schumacher–Westmoreland

Teorema 3.3. (di Holevo-Schumacher–Westmoreland (HSW)) *Sia $\mathcal{E}$ un canale quantistico che conserva la traccia. Si definisce:*

$$\chi(\mathcal{E}) \equiv \max_{\{p_j,\rho_j\}} [S(\mathcal{E}(\sum_j p_j\rho_j)) - \sum_j p_j S(\mathcal{E}(\rho_j))], \tag{3.16}$$

dove il massimo è preso tra tutti gli insiemi $\{p_j, \rho_j\}$ di possibili stati di input ρ_j del canale. Allora $\chi(\mathcal{E})$ è la capacità di uno stato prodotto per il canale $\mathcal{E}$, quindi $\chi(\mathcal{E}) = C^{(1)}(\mathcal{E})$.

Dimostrazione. Si segue la discussione data in [8] per illustrare la dimostrazione.
Siano ρ_j gli input del canale $\mathcal{E}$ e $\sigma_j \equiv \mathcal{E}(\rho_j)$ i corrispondenti output. Si costruisce un codice le cui parole di codice sono il prodotto (tensore) di n stati ρ_j, dove l'insieme di indici $\{j\}$ ha distribuzione di probabilità p_j. Alice può mandare a Bob messaggi M scelti dall'insieme $\{1, \ldots, 2^{nR}\}$; ad ogni messaggio è associata una parola di codice: $\rho_{M_1} \otimes \rho_{M_2} \otimes \cdots \otimes \rho_{M_n}$, con $M_1, \ldots, M_n \in \{j\}$. Per comodità scriviamo $\rho_M \equiv \rho_{M_1} \otimes \rho_{M_2} \otimes \cdots \otimes \rho_{M_n}$ e $\sigma_M = \mathcal{E}^{\otimes n}(\rho_M)$.
Quando Bob riceve uno stato σ_M, per risalire al messaggio mandatogli da Alice, effettua su esso una misura usando un set di operatori POVM. Si suppone che tale set sia costruito con un elemento E_M per ogni messaggio M, più altri possibili elementi che vengono racchiusi in un unico E_0 tale che: $E_0 = \mathbb{I} - \sum_{M\neq 0} E_M$. Il messaggio M viene decodificato efficacemente con probabilità $\mathrm{tr}(\sigma_M E_M)$, quindi la probabilità di errore nel riconoscere M sarà: $p_M^e \equiv 1 - \mathrm{tr}(\sigma_M E_M)$.

Per dimostrare l'esistenza di codici ad alto rate con $p_M^e \to 0 \ \forall M$ per $n \to \infty$, si dimostra prima che ci sono codici ad alto rate per cui la probabilità *media* di errore p_{av} tende a zero per n sufficientemente grande.

La probabilità media di errore, supponendo che Alice produca messaggi scegliendo uniformemente dall'insieme $\{1, \ldots, 2^{nR}\}$, è definita come:

$$p_{av} \equiv \frac{\sum_M p_M^e}{2^{nR}} = \frac{\sum_M (1 - \mathrm{tr}(\sigma_M E_M))}{2^{nR}}. \tag{3.17}$$

Prima di tutto bisogna costruire un set di operatori POVM $\{E_M\}$ che agevoli la misura.

Iniziamo considerando $\epsilon > 0$ e definendo $\bar{\sigma} \equiv \sum_j p_j \rho_j$ e P come il proiettore sul sottospazio ϵ-tipico di $\sigma^{\bar{\otimes} n}$. Dal Teorema 3.2 segue che:

$$\mathrm{tr}(\bar{\sigma}^{\otimes n}(\mathbb{I} - P)) \leq \delta. \tag{3.18}$$

Si definisce $\bar{S} \equiv \sum_j p_j S(\sigma_j)$. Inoltre, considerando la decomposizione spettrale di $\sigma_j = \sum_k \lambda_k^j \left|e_k^j\right\rangle \left\langle e_k^j\right|$, si ha:

$$\sigma_M = \sum_K \lambda_K^M \left|E_K^M\right\rangle \left\langle E_K^M\right|, \tag{3.19}$$

con $K = (K_1, \ldots, K_n)$; $\lambda_K^M \equiv \lambda_{K_1}^{M_1} \lambda_{K_2}^{M_2} \ldots \lambda_{K_n}^{M_n}$ e $\left|E_K^M\right\rangle \equiv \left|e_{K_1}^{M_1}\right\rangle \left|e_{K_2}^{M_2}\right\rangle \ldots \left|e_{K_n}^{M_n}\right\rangle$.

Si definisce P_M come il proiettore sullo spazio generato dagli stati $\left|E_K^M\right\rangle$ i cui λ_K^M corrispondenti soddisfano:

$$\left| \frac{1}{n} \log \frac{1}{\lambda_K^M} - \bar{S} \right| \leq \epsilon. \tag{3.20}$$

Per la legge dei grandi numeri, $\forall \delta > 0$ per n sufficientemente grande, vale $\mathbf{E}(\mathrm{tr}(\sigma_M P_M)) \geq 1 - \delta$, con $\mathbf{E}(\cdot)$ che indica il valore di aspettazione fatto tenendo conto della distribuzione delle parole di codice ρ_M rispetto ai diversi codici casuali possibili. Quindi, $\forall M$ vale:

$$\mathbf{E}[\mathrm{tr}(\sigma_M (I - P_M))] \leq \delta. \tag{3.21}$$

Per la definizione (3.20) si ha che la dimensione del sottospazio su cui proietta P_M è al massimo $2^{n(\bar{S}+\epsilon)}$, ovvero:

$$\mathbf{E}(\mathrm{tr}(P_M)) \leq 2^{n(\bar{S}+\epsilon)}. \tag{3.22}$$

A questo punto si definisce l'operatore POVM corrispondente al messaggio M:

$$E_M \equiv \left(\sum_{M'} P P_{M'} P \right)^{-1/2} P P_M P \left(\sum_{M'} P P_{M'} P \right)^{-1/2}. \tag{3.23}$$

Per tutti gli operatori vale $\sum_M E_M \leq \mathbb{I}$, perciò si costruisce $E_0 = \mathbb{I} - \sum_M E_M$ per completare il set.

Per dare un'idea, E_M è uguale a P_M a meno di piccole correzioni, e la misura di Bob consiste nel vedere se l'output del canale rientra effettivamente nello spazio in cui P_M proietta.

Tornando a p_{av}, essa ammette un limite superiore (non diamo qui la dimostrazione di come si ottiene):

$$p_{av} \leq \frac{1}{2^{nR}} \sum_M \left[3\mathrm{tr}(\sigma_M(\mathbb{I} - P)) + \sum_{M' \neq M} \mathrm{tr}(P\sigma_M P P_{M'}) + \mathrm{tr}(\sigma_M(\mathbb{I} - P_M)) \right]. \tag{3.24}$$

Serve calcolare il valore di aspettazione di p_{av} su tutti i codici casuali possibili. Per costruzione si ha: $\mathbf{E}(\sigma_M) = \bar{\sigma}^{\otimes n}$ e che σ_M e $P_{M'}$ sono indipendenti per $M' \neq M$. Perciò si ottiene:

$$\mathbf{E}(p_{av}) \leq 3\mathrm{tr}(\bar{\sigma}^{\otimes n}(\mathbb{I} - P)) + (2^{nR} - 1)\mathrm{tr}(P\bar{\sigma}^{\otimes n} P \mathbf{E}(P_1)) + \mathbf{E}(\mathrm{tr}(\sigma_1(\mathbb{I} - P_1))). \tag{3.25}$$

Sostituendo la (3.18) e la (3.21) si ha:

$$\mathbf{E}(p_{av}) \leq 4\delta + (2^{nR} - 1)\mathrm{tr}(P\bar{\sigma}^{\otimes n} P \mathbf{E}(P_1)). \tag{3.26}$$

Ma $P\bar{\sigma}^{\otimes n} P \leq 2^{-n(S(\bar{\sigma})-\epsilon)}\mathbb{I}$ e da (3.22) si ha:

$$\mathbf{E}(p_{av}) \leq 4\delta + (2^{nR} - 1)2^{-n(S(\bar{\sigma})-\bar{S}-2\epsilon)}. \tag{3.27}$$

Con $R < S(\bar{\sigma}) - \bar{S}$, si ha $\mathbf{E}(p_{av}) \to 0$ per $n \to \infty$. Ma $S(\bar{\sigma}) - \bar{S}$ è la χ di Holevo, quindi, per come è definita in (3.16), è analogo scrivere $R < \chi(\mathcal{E})$.
Deve esistere una sequenza di codici di rate R tali che $p_{av} \to 0$ all'aumentare di n, allora per ogni $\eta > 0$ fissato e n sufficientemente grande:

$$p_{av} = \frac{\sum_M p_M^e}{2^{nR}} \leq \eta. \tag{3.28}$$

Quest'ultima espressione, per un codice, è vera se almeno metà dei messaggi del codice rispetta $p_M^e < 2\eta$. Eliminando da un codice di questo tipo le parole di codice con $p_M^e < 2\eta$, si ottiene un nuovo codice con rate R, con $2^{n(R-1/n)}$ parole di codice e con $p_M^e < 2\eta$ per ogni messaggio M.
Mettendo insieme i risultati ottenuti, abbiamo dimostrato che esistono codici con rate $R < \chi(\mathcal{E})$ che usano stati prodotto come input e permettono trasmissione efficace attraverso il canale $\mathcal{E}$.

Per completare del tutto la dimostrazione, bisogna far vedere che per $R > \chi(\mathcal{E})$ non è possibile avere trasmissione efficace nel canale. Usando la disuguaglianza di Fano, il limite di Holevo e la subadditività dell'entropia, si dimostra che, quando $R > \chi(\mathcal{E})$, la probabilità media di errore non tende a zero per $n \to \infty$. Non si può quindi fare la stessa procedura di sopra per avere un codice in cui ogni messaggio ha una bassa probabilità d'errore, cioè un codice con trasmissione affidabile.

Si conclude allora che $\chi(\mathcal{E})$ è il massimo rate che permette trasmissione affidabile, rappresentando quindi la capacità del canale quando esso prende in input stati prodotto. □

Per alcuni canali quantistici, la quantità $\chi(\mathcal{E})$ è additiva, cioè:

$$\chi(\mathcal{E}^{\otimes n}) = n\chi(\mathcal{E}) \quad \forall n, \tag{3.29}$$

e la capacità classica del canale coincide con la capacità per stati prodotto: $C(\mathcal{E}) = C^{(1)}(\mathcal{E})$. Ad esempio, ciò avviene per i canali entanglement-breaking, canali che distruggono l'entanglement con altri sistemi, per i quali l'additività è garantita [12], oppure per il depolarizing channel. Tuttavia, l'additività di $\chi(\mathcal{E})$ non è valida in generale, come dimostrato da Hastings [4], e questo porta ad avere una formula regolarizzata per la capacità classica di un canale quantistico:

$$C(\mathcal{E}) \equiv \lim_{n\to\infty} \frac{1}{n}\chi(\mathcal{E}^{\otimes n}), \tag{3.30}$$

[15].

3.3.2 Esempio per il depolarizing channel

Un esempio in cui la capacità per stati prodotto può essere calcolata esplicitamente è fornito dal *depolarizing channel* con parametro p, un canale quantistico che, per i qubit, agisce nel seguente modo:

$$\mathcal{E}(\rho) = p\rho + (1-p)\frac{\mathbb{I}}{2}. \tag{3.31}$$

Considerato l'insieme di stati quantistici $\{p_j, |\psi_j\rangle\}$, con $|\psi_j\rangle \in \mathcal{H}$ stati puri e $\dim(\mathcal{H}) = 2$, si ha:

$$\mathcal{E}(|\psi_j\rangle\langle\psi_j|) = p\,|\psi_j\rangle\langle\psi_j| + (1-p)\frac{\mathbb{I}}{2}. \tag{3.32}$$

Gli autovalori di questo stato sono $(1+p)/2$ e $(1-p)/2$ indipendentemente dallo stato puro in entrata. Da ciò si ottiene che:

$$S(\mathcal{E}(|\psi_j\rangle\langle\psi_j|)) = H\left(\frac{1+p}{2}\right), \tag{3.33}$$

con $H(\cdot)$ entropia binaria, non dipende da $|\psi_j\rangle$.

Per massimizzare la χ di Holevo è dunque sufficiente massimizzare $S(\mathcal{E}(\sum_j p_j\rho_j)) = S(\sum_j p_j\mathcal{E}(|\psi_j\rangle\langle\psi_j|))$. Tale risultato si ottiene scegliendo $|\psi_j\rangle$ che formino una base ortonormale di $\mathcal{H}$ e che siano equiprobabili, così la capacità per stati prodotto di tale canale sarà:

$$C^1(\mathcal{E}) = 1 - H\left(\frac{1+p}{2}\right). \tag{3.34}$$

3.4 Capacità classica entanglement-assistita

La trasmissione di informazione classica attraverso un canale quantistico può avvenire anche se il mittente e il destinatario condividono a priori stati entangled. In questo contesto si definisce la *capacità classica entanglement-assistita* del canale, indicata con $C_E(\mathcal{E})$, per indicare la massima informazione che può essere scambiata con un uso del canale.
Sia $\mathcal{E}$ un canale quantistisco che agisce su stati del sistema Q. È stato dimostrato in [2] che la capacità classica del canale, quando Alice e Bob hanno accesso a un numero illimitato di coppie di stati entangled, è data dalla massimizzazione della mutua informazione quantistica tra input e output del canale:

$$C_E(\mathcal{E}) = \max_{\rho \in Q} S(\rho) + S(\mathcal{E}(\rho)) - S((\mathcal{E} \otimes \mathbb{I})\rho^{QR}), \tag{3.35}$$

dove si introduce un sistema ausiliario R tale che lo stato congiunto ρ^{QR} sia puro e abbia come ridotto su Q lo stato di input ρ.
A differenza della (3.30), questa espressione non richiede che sia fatto un limite su infiniti usi del canale e ricorda molto l'espressione classica (2.8).

Capitolo 4

Informazione quantistica in canali quantistici

Un canale quantistico può essere usato per trasmettere stati quantistici, in tal caso, il rate più alto con cui il canale può effettuare attendibilmente tale trasmissione è detto *capacità quantistica*. Ad oggi, esiste un'espressione regolarizzata per tale quantità, per ottenerla, sono stati introdotti strumenti teorici estremamente importanti. In questo capitolo verranno inizialmente introdotti tali strumenti. In particolare, vengono introdotte l'informazione coerente e le disuguaglianze da essa verificate, analoghe a quelle verificate classicamente dalla mutua informazione - grandezza in termini della quale si introduce la capacità classica. A seguire, presentando la massima informazione coerente come limite superiore e limite inferiore della capacità quantistica , si giunge alla sua espressione regolarizzata dovuta a Lloyd, Shor e Devetak [6, 13, 3]. Infine, si analizza una classe di canali per cui la massima informazione coerente è additiva, e per cui quindi la capacità quantistica presenta un'espressione esatta, i canali degradabili.

4.1 Informazione coerente

Per proseguire, consideriamo il seguente sistema. Sia Q il sistema quantistico soggetto all'azione del canale $\mathcal{E}$. Si considera che esso faccia parte di un sistema più grande RQ, con R sistema di riferimento, tale che lo stato iniziale $|\Psi^{RQ}\rangle$ sia puro. Inoltre si considera un sistema ambiente E, inizialmente in uno stato puro, tale che esista un operatore unitario U^{QE}, estensione del canale $\mathcal{E}$:

$$\mathcal{E}(\rho^Q) = \mathrm{tr}_E \left[U^{QE}(\rho^Q \otimes |0^E\rangle \langle 0^E|) U^{QE\dagger} \right] . \tag{4.1}$$

Inizialmente il sistema totale RQE è nello stato: $|\Psi^{RQE}\rangle = |\Psi^{RQ}\rangle \otimes |0^E\rangle$. Poiché l'evoluzione complessiva è unitaria, anche lo stato finale sarà puro: $|\Psi^{RQ'E'}\rangle = (\mathbb{I}^R \otimes U^{QE} |\Psi^{RQE}\rangle)$.

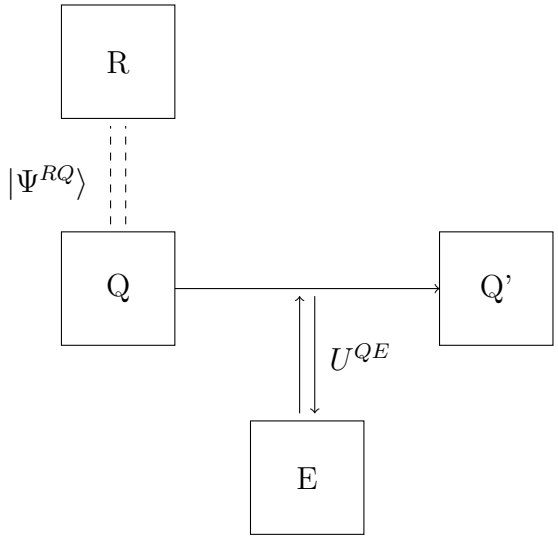

Figura 4.1: Sistemi considerati per la descrizione della trasmissione di uno stato del sistema Q attraverso il canale $\mathcal{E}$. [Immagine ripresa da [1]].

Per quantificare il rumore apportato sullo stato ρ^Q dal canale, si definisce l'*entropia di scambio*:

$$S_e(\rho^Q, \mathcal{E}) \equiv S(\rho^{RQ'}) = S(\rho^{E'}), \tag{4.2}$$

è l'entropia del sistema RQ dopo l'evoluzione di Q, la seconda uguaglianza è dovuta al fatto che il sistema RQE è ancora in uno stato puro dopo l'evoluzione.

In generale, dato uno stato puro $|\psi\rangle$ di un sistema quantistico, si definisce *fidelity* F di uno stato arbitrario ρ del sistema come:

$$F = \langle\psi| \rho |\psi\rangle . \tag{4.3}$$

Per misurare quanto il canale conserva l'entanglement tra Q e R, si definisce l'*entanglement fidelity*:

$$F_e \equiv F(\rho^Q, \mathcal{E}) = \langle\Psi^{RQ}| \rho^{RQ'} |\Psi^{RQ}\rangle , \tag{4.4}$$

essa è tale che $0 \leq F_e \leq 1$.

Ora si può definire l'*informazione coerente*:

$$I_e \equiv I(\rho^Q, \mathcal{E}) = S(\rho^{Q'}) - S(\rho^{RQ'}) \tag{4.5}$$

$$= S(\rho^{Q'}) - S_e. \tag{4.6}$$

Essa dipende solamente dallo stato iniziale ρ^Q e dal canale $\mathcal{E}$ che agisce su esso: $\mathcal{E}(\rho^Q) = \rho^{Q'}$.
L'entropia di scambio e l'entanglement fidelity sono legati da una disuguaglianza di cui non riportiamo la dimostrazione, la *disuguaglianza quantistica di Fano*:

Teorema 4.1. (Disuguaglianza quantistica di Fano) *Sia ρ lo stato di un sistema quantistico e sia $\mathcal{E}$ un canale quantistico che conserva la traccia. Allora:*

$$S_e(\rho, \mathcal{E}) \leq H(F_e(\rho, \mathcal{E})) + (1 - F_e(\rho, \mathcal{E})) \log(d^2 - 1), \tag{4.7}$$

dove $H(\cdot)$ denota l'entropia binaria di Shannon e d è la dimensione del sistema quantistico in cui vive ρ.

4.1.1 Disuguaglianza quantistica di processazione dei dati

L'informazione coerente gioca un ruolo fondamentale nella teoria dell'informazione quantistica: essa infatti soddisfa disuguaglianze che costituiscono l'analogo quantistico di quelle valide per la mutua informazione classica, la quale, come visto, è alla base della definizione della capacità classica di un canale. Tra queste disuguaglianze vi è quella di *processazione quantistica dei dati*, che si applica a un processo quantistico che avviene in due fasi: $\rho \xrightarrow{\mathcal{E}_1} \rho' \xrightarrow{\mathcal{E}_2} \rho''$.

Teorema 4.2. (Disuguaglianza di processazione quantistica dei dati) *Sia ρ lo stato di un sistema quantistico e siano $\mathcal{E}_1$ e $\mathcal{E}_2$ due canali quantistici che conservano la traccia. Allora:*

$$S(\rho) \geq I(\rho, \mathcal{E}_1) \geq I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1). \tag{4.8}$$

Dimostrazione. Per alleggerire la notazione, di seguito scriviamo l'entropia di von Neumann di uno stato del generico sistema A come: $S(\rho^A) = S(A)$.

Facciamo uso di quattro sistemi per la dimostrazione:

- Q sistema quantistico che entra nei due canali;
- R sistema di riferimento;
- E_1 che permette di estendere l'azione di $\mathcal{E}_1$;
- E_2 che permette di estendere l'azione di $\mathcal{E}_2$.

Si parte dalla prima disuguaglianza:

$$I(\rho, \mathcal{E}_1) = S(Q') - S(R, Q') = \tag{4.9}$$
$$= S(Q') - S(E_1') = \tag{4.10}$$
$$= S(R, E_1') - S(E_1') \leq \tag{4.11}$$
$$\leq S(R) + S(E_1') - S(E_1') = \tag{4.12}$$
$$= S(R) = S(Q) = S(\rho), \tag{4.13}$$

con ρ stato del sistema Q. Dove sono state usata la proprietà per sistemi composti in stati puri e la subadditività dell'entropia di von Neumann. Per la seconda disuguaglianza, si sfrutta la subadditività forte:

$$S(R, E_1'', E_2'') + S(E_1'') \leq S(R, E_1'') + S(E_1'', E_2''), \tag{4.14}$$

i sistemi doppiamente primati rappresentano i sistemi dopo entrambe le evoluzioni. Lo stato totale del sistema $RQ''E_1''E_2''$ è ancora puro, quindi:

$$S(R, E_1'', E_2'') = S(Q''). \tag{4.15}$$

Dalla definizione di scambio di entropia:

$$S(E_1'') = S(E_1') = S(\rho, \mathcal{E}_1), \tag{4.16}$$

e analogamente:

$$S(E_1'', E_2'') = S(\rho, \mathcal{E}_2 \circ \mathcal{E}_1). \tag{4.17}$$

Inoltre, poiché $\mathcal{E}_2$ non agisce su RE_1', si ha:

$$S(R, E_1') = S(Q') = S(R, E_1''). \tag{4.18}$$

Mettendo insieme (4.15), (4.16), (4.18), (4.17), si ottiene:

$$S(Q'') + S(\rho, \mathcal{E}_1) \leq S(Q') + S(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) \tag{4.19}$$
$$S(Q'') - S(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) \leq S(Q') - S(\rho, \mathcal{E}_1). \tag{4.20}$$

Si nota che entrambi i lati corrispondo alla definizione di informazione coerente: $I(\rho, \mathcal{E}_1) = S(Q') - S(\rho, \mathcal{E}_1)$, mentre $I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) = S(Q'') - S(\rho, \mathcal{E}_2 \circ \mathcal{E}_1)$. Quindi: $I(\rho, \mathcal{E}_1) \geq I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1)$. □

Questa disuguaglianza mostra che l'informazione coerente non aumenta nonostante essa venga processata [10].

4.2 Limite superiore della capacità quantistica

Dopo aver definito l'informazione coerente e le sue proprietà, si può procedere con l'analisi della capacità quantistica. In questo paragrafo, viene presentata una prova del limite superiore della capacità quantistica di un canale, nella forma proposta in [1]. In particolare, riportiamo la dimostrazione del limite superiore nel caso in cui l'operazione di codifica sugli stati quantistici venga fatta con un operatore unitario, ma è valida in generale.

Nel Capitolo 2 è stato illustrato lo schema di una comunicazione classica, adesso serve presentare lo schema di una comunicazione quantistica. Una sorgente quantistica Σ emette uno stato quantistico ρ_s che agisce sullo spazio H_s e che viene codificato in uno stato ρ_c con un'*operazione di codifica* $\mathcal{C}$. Lo stato codificato attraversa il canale $\mathcal{E}$ e subisce una trasformazione. All'uscita del canale, lo stato ottenuto ρ_o viene decodificato con un'*operazione di decodifica* $\mathcal{D}$. La trasmissione avviene *attendibilmente* se esiste una serie di codici codifica-decodifica $(\mathcal{C}^n, \mathcal{D}^n)$ che permette di ottenere

$$\lim_{n\to\infty} F_e(\rho_s^n, \mathcal{D}^n \circ \mathcal{E}^{\otimes n} \circ \mathcal{C}^n) = 1, \tag{4.21}$$

con ρ_s^n si intende uno stato, emesso dalla sorgente, che agisce sullo spazio $H_s^{\otimes n}$.

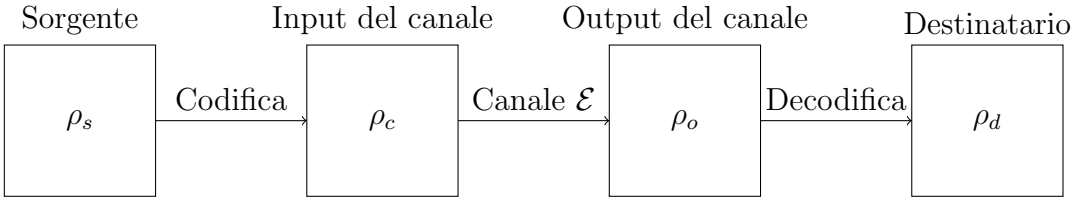

Figura 4.2: Schema di un sistema di comunicazione quantistica. [Immagine ripresa da [1].]

Definiamo il *tasso di entropia* di una sorgente Σ come:

$$S(\Sigma) \equiv \limsup_{n\to\infty} \frac{S(\rho_s^n)}{n}. \tag{4.22}$$

Consideriamo che un canale può trasmettere in maniera attendibile se ha un rate pari a $R = S(\Sigma)$.

Per la dimostrazione del limite superiore della capacità, serve il seguente lemma:

Lemma 1. *Sia $\mathcal{E}$ un canale quantistico e ρ uno stato quantistico. Allora, per qualsiasi altro canale quantistico $\mathcal{D}$ vale:*

$$S(\rho) \leq I(\rho, \mathcal{E}) + 2 + 4(1 - F_e(\rho, \mathcal{D} \circ \mathcal{E})) \log d. \tag{4.23}$$

Prima di procedere con la dimostrazione, ritorniamo a considerare i tre sistemi Q soggetto all'evoluzione di $\mathcal{E}$, R e E; indicati con le lettere primate per riferirsi al post evoluzione. Dalle relazioni sull'entropia di stati puri e per la subadditività vale:

$$S(\rho^Q) = S(\rho^R) = S(\rho^{Q'E'}) \leq S(\rho^{Q'}) + S(\rho^{E'}). \tag{4.24}$$

Riscrivibile come:

$$S(\rho^Q) - S(\mathcal{E}(\rho^Q)) \leq S_e(\rho^Q, \mathcal{E}). \tag{4.25}$$

Dimostrazione. Dalla disuguaglianza di processazione dei dati: $I(\rho, \mathcal{E}) \geq I(\rho, \mathcal{D} \circ \mathcal{E})$, quindi:

$$\begin{aligned} S(\rho) - I(\rho, \mathcal{E}) &\leq S(\rho) - I(\rho, \mathcal{D} \circ \mathcal{E}) && (4.26)\\ &= S(\rho) - S((\mathcal{D} \circ \mathcal{E})\rho) + S_e(\rho, \mathcal{D} \circ \mathcal{E}) && (4.27)\\ &\leq S_e(\rho, \mathcal{D} \circ \mathcal{E}) + S_e(\rho, \mathcal{D} \circ \mathcal{E}) = 2S_e(\rho, \mathcal{D} \circ \mathcal{E}). && (4.28) \end{aligned}$$

L'ultimo passaggio è stato ottenuto utilizzando la (4.25) con $\mathcal{D} \circ \mathcal{E}$.
Ora, applicando la disuguaglianza quantistica di Fano per $S_e(\rho, \mathcal{D} \circ \mathcal{E})$, si ottiene:

$$\begin{aligned} S(\rho) - I(\rho, \mathcal{E}) &\leq 2\left[H(F_e(\rho, \mathcal{D} \circ \mathcal{E})) + (1 - F_e(\rho, \mathcal{D} \circ \mathcal{E}))\log(d^2 - 1)\right] && (4.29)\\ &\leq 2 + 4\,(1 - F_e(\rho, \mathcal{D} \circ \mathcal{E}))\log d. && (4.30) \end{aligned}$$

In cui alla fine sono state usate $H(F_e(\rho, \mathcal{D} \circ \mathcal{E})) \leq 1$ perché qui $H(\cdot)$ è l'entropia binaria di Shannon e $\log(d^2 - 1) \leq 2\log d$.
Abbiamo dunque concluso la dimostrazione. □

Il limite superiore della capacità quantistica si può ricavare in contesti diversi: quando la codifica è unitaria, quando la codifica non è unitaria o per classi di codifica più ristrette. Qui riportiamo come si ottiene il limite nel caso in cui la codifica è unitaria; anche le mappe non unitarie, aggiungendo un sistema ambiente, si possono estendere a mappe unitarie.

4.2.1 Codifiche Unitarie

Per proseguire, si definisce la massima informazione coerente:

$$\mathcal{Q}^{(1)}(\mathcal{E}) \equiv \max_{\rho} I(\rho, \mathcal{E}). \tag{4.31}$$

Dimostreremo che il limite superiore della capacità è dato da:

$$\mathcal{Q}(\mathcal{E}) \equiv \lim_{n\to\infty} \frac{\mathcal{Q}^{(1)}(\mathcal{E}^{\otimes n})}{n}. \tag{4.32}$$

Si consideri un sistema di comunicazione quantistica in cui l'operazione di codifica è un'operazione unitaria.

Teorema 4.3. *Sia una sorgente quantistica $\Sigma = (H_s, \rho_s)$, sia $\mathcal{U}^n$ una serie di operazioni di codifica unitarie per la sorgente e sia $\mathcal{E}$ il canale quantistico. Supponiamo che esista una serie di operazioni di decodifica $\mathcal{D}^n$ tali che:*

$$\lim_{n\to\infty} F_e(\rho_s^n, \mathcal{D}^n \circ \mathcal{E}^{\otimes n} \circ \mathcal{U}^n) = 1. \tag{4.33}$$

Allora:

$$S(\Sigma) \leq \mathcal{Q}(\mathcal{E}). \tag{4.34}$$

Dimostrazione. Data l'unitarietà di $\mathcal{U}^n$, si ha:

$$I(\rho_s^n, \mathcal{E}^{\otimes n} \circ \mathcal{U}^n) = I(\mathcal{U}^n(\rho_s^n), \mathcal{E}^{\otimes n}) \leq \mathcal{Q}^{(1)}(\mathcal{E}^{\otimes n}), \tag{4.35}$$

dove la disuguaglianza segue dalla definizione (4.31). Applicando prima il Lemma 1 e poi la (4.35), si ottiene:

$$\begin{aligned} S(\rho_s^n) &\leq I(\rho_s^n, \mathcal{E}^{\otimes n} \circ \mathcal{U}^n) + 2 + 4(1 - F_e(\rho_s^n, \mathcal{D}^n \circ \mathcal{E}^{\otimes n} \circ \mathcal{U}^n)) \log d^n & (4.36)\\ &\leq \mathcal{Q}^{(1)}(\mathcal{E}) + 2 + 4n(1 - F_e(\rho_s^n, \mathcal{D}^n \circ \mathcal{E}^{\otimes n} \circ \mathcal{U}^n)) \log d. & (4.37) \end{aligned}$$

Dividendo tutto per n:

$$\frac{S(\rho_s^n)}{n} \leq \frac{\mathcal{Q}^{(1)}(\mathcal{E}^{\otimes n})}{n} + \frac{2}{n} + 4(1 - F_e(\rho_s^n, \mathcal{D}^n \circ \mathcal{E}^{\otimes n} \circ \mathcal{U}^n)) \log d. \tag{4.38}$$

Facendo il limite superiore per $n \to \infty$ per entrambi i lati, si ha:

$$\limsup_{n\to\infty} \frac{S(\rho_s^n)}{n} \leq \limsup_{n\to\infty} \frac{\mathcal{Q}^{(1)}(\mathcal{E}^{\otimes n})}{n} = \mathcal{Q}(\mathcal{E}), \tag{4.39}$$

che conlude la prova facendo vedere che il rate $S(\Sigma)$ è limitato superiormente da $\mathcal{Q}(\mathcal{E})$. □

Dalla (4.34) si ha che anche il più alto dei rate affidabili per il canale, quindi la capacità quantistica per definizione, è limitato superiormente.

4.3 Limite inferiore della capacità quantistica

Un canale quantistico può essere usato per svolgere diversi compiti: per generare entanglement, per simulare un canale senza rumore, per trasmettere entanglement ... Ciascuna modalità di utilizzo prevede un rate massimo,ma si dimostra che ci sono modi di utilizzo che condividono lo stesso rate massimo, a cui ci si riferisce in generale con il nome di capacità quantistica del canale.

Dato un canale quantistico $\mathcal{N}$, siano $\mathcal{Q}_s(\mathcal{N})$ la capacità del canale di trasmettere sottospazi, $\mathcal{Q}_e(\mathcal{N})$ la capacità del canale di trasmettere entanglement e $E(\mathcal{N})$ la capacità del canale di generare entanglement. In [3] si dimostra che queste tre si equivalgono per cui, basta trovare che una delle tre è inferiormente limitata per avere tale limitazione anche sulle altre.
Per la generazione di entanglement si dimostra:

Teorema 4.4. *Sia $\mathcal{N}$ un canale quantistico. Ogni rate R tale che $0 \leq R \leq I(\rho, \mathcal{N})$ è raggiungibile per la generazione di entanglement con $\mathcal{N}$.*

Per brevità, si rimanda a [5] per la dimostrazione, che coinvolge l'uso di vari lemmi e delle nozioni di distanza tra matrici; ma vediamo che data la non additività dell'informazione coerente, dal Teorema 4.4 segue che per n usi del canale un rate $R < \frac{1}{n}\max_\rho I(\rho, \mathcal{N}^{\otimes n})$ è ancora raggiungibile. Per definizione di capacità bisogna prendere il limite superiore dei rate raggiungibili, quindi:

$$\limsup_{n\to\infty} \frac{1}{n} \max_\rho I(\rho, \mathcal{N}^{\otimes n}) = \limsup_{n\to\infty} \frac{\mathcal{Q}^{(1)}(\mathcal{N}^{\otimes n})}{n} \leq E(\mathcal{N}). \tag{4.40}$$

Data l'uguaglianza di $E(\mathcal{N})$ con le capacità sopra definite, questo limite inferiore è vero in generale per la capacità dei canali quantistici.

4.4 Formula di Lloyd-Shor-Devetak

Dalle dimostrazioni delle sezioni precedenti, si vede che si può esprimere la capacità quantistica di un canale in termini della massima informazione coerente. La formula richiede la regolarizzazione su n usi del canale e viene detta *formula di Lloyd-Shor-Devetak*, dagli autori dei risultati che hanno portato ad ottenerla:

$$\mathcal{Q}(\mathcal{E}) = \lim_{n\to\infty} \frac{1}{n} \mathcal{Q}^{(1)}(\mathcal{E}^{\otimes n}), \tag{4.41}$$

con $\mathcal{Q}^{(1)}\mathcal{E} = \max_\rho I(\rho, \mathcal{E})$. Così come la mutua informazione ha un ruolo chiave nella definizione della massima quantità di informazione che può essere trasmessa in un canale

classico, l'informazione coerente - che misura quanto un canale quantistico è in grado di conservare uno stato quantistico - è fondamentale per i canali quantistici.

4.5 Esempio per canali degradabili

Esiste una classe di canali quantistici, detti *canali degradabili*, per cui la capacità quantistica è calcolabile esattamente. Per tali canali, si dimostra che la massima informazione coerente $\mathcal{Q}^{(1)}(\mathcal{E})$ è additiva e quindi la capacità coinciderà con essa. Rientrano nella classe dei canali degradabili: l'erasure channel, l'amplitude damping channel e alcuni canali gaussiani.

Sia un canale quantistico $\mathcal{E} : \rho^Q \to \rho^{Q'}$ e sia la sua estensione unitaria U^{QE} tale che $\mathcal{E}(\rho^Q) = \mathrm{tr}_E \left[U^{QE}(\rho^{QE})U^{QE\dagger}\right]$. Si definisce il suo canale complementare $\mathcal{E}_c : \rho^Q \to \rho^{E'}$ tale che $\mathcal{E}_c = \mathrm{tr}_{Q'} \left[U^{QE}(\rho^{QE})U^{QE\dagger}\right]$.

Definizione 4.1. (Canale degradabile) Un canale quantistico $\mathcal{E}$ si dice degradabile se esiste una mappa $\mathcal{T} : \rho^{Q'} \to \rho^{E'}$ tale che:

$$\mathcal{E}_c = \mathcal{T} \circ \mathcal{E}. \tag{4.42}$$

Teorema 4.5. (Additività per un canale degradabile) *Siano $\mathcal{N}$ e $\mathcal{M}$ due canali quantistici degradabili. Allora vale:*

$$Q^{(1)}(\mathcal{N} \otimes \mathcal{M}) = Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{M}). \tag{4.43}$$

Dimostrazione. La disuguaglianza $Q^{(1)}(\mathcal{N} \otimes \mathcal{M}) \geq Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{M})$ è valida in generale. Sia ρ^{Q_1} l'operatore densità che massimizza $I(\rho, \mathcal{N})$ e sia ρ^{Q_2} l'operatore densità che massimizza $I(\rho, \mathcal{M})$. Scegliamo lo stato prodotto $\rho = \rho^{Q_1} \otimes \rho^{Q_2}$ come stato del sistema composto su cui agisce il canale $\mathcal{N} \otimes \mathcal{M}$ e valutiamo l'informazione coerente per esso:

$$\begin{aligned}
I(\rho^{Q_1} \otimes \rho^{Q_2}, \mathcal{N} \otimes \mathcal{M}) &= S((\mathcal{N} \otimes \mathcal{M})(\rho^{Q_1} \otimes \rho^{Q_2})) - S((\mathcal{N} \otimes \mathcal{M})_c(\rho^{Q_1} \otimes \rho^{Q_2})) = && (4.44)\\
&= S((\mathcal{N} \otimes \mathcal{M})(\rho^{Q_1} \otimes \rho^{Q_2})) - S((\mathcal{N}_c \otimes \mathcal{M}_c)(\rho^{Q_1} \otimes \rho^{Q_2})) = && (4.45)\\
&= S(\mathcal{N}(\rho^{Q_1}) \otimes \mathcal{M}(\rho^{Q_2})) - S(\mathcal{N}_c(\rho^{Q_1}) \otimes \mathcal{M}_c(\rho^{Q_2})) = && (4.46)\\
&= S(\mathcal{N}(\rho^{Q_1})) + S(\mathcal{M}(\rho^{Q_2})) - S(\mathcal{N}_c(\rho^{Q_1})) - S(\mathcal{M}_c(\rho^{Q_2})) = && (4.47)\\
&= I(\rho^{Q_1}, \mathcal{N}) + I(\rho^{Q_2}, \mathcal{M}) = && (4.48)\\
&= Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{M}), && (4.49)
\end{aligned}$$

dove l'ultima uguaglianza vale per come sono stati scelti gli stati ρ^{Q_1} e ρ^{Q_2}. A questo punto:

$$Q^{(1)}(\mathcal{N}\otimes\mathcal{M}) = \max_{\rho^{Q_1Q_2}} I(\rho^{Q_1Q_2}, \mathcal{N}\otimes\mathcal{M}) \tag{4.50}$$

$$\geq I(\rho^{Q_1}\otimes\rho^{Q_2}, \mathcal{N}\otimes\mathcal{M}) = Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{M}). \tag{4.51}$$

La disuguaglianza $Q^{(1)}(\mathcal{N}\otimes\mathcal{M}) \leq Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{M})$ è valida solo per i canali degradabili.
Consideriamo lo stato $\rho^{Q_1Q_2}$ ingresso del canale $\mathcal{N}\otimes\mathcal{M}$ come stato che massimizza l'informazione coerente di tale canale. Vale:

$$\mathcal{Q}^{(1)}(\mathcal{N}\otimes\mathcal{M}) = I(\rho^{Q_1Q_2}, \mathcal{N}\otimes\mathcal{M}) = \tag{4.52}$$

$$= S(Q_1', Q_2') - S(R, Q_1', Q_2') = \tag{4.53}$$

$$= S(Q_1', Q_2') - S(E_1', E_2') = \tag{4.54}$$

$$= S(Q_1') + S(Q_2') - I(Q_1' : Q_2') - S(E_1') - S(E_2') + I(E_1' : E_2') = \tag{4.55}$$

$$= S(Q_1') - S(E_1') + S(Q_2') - S(E_2') - [I(Q_1' : Q_2') - I(E_1' : E_2')] \tag{4.56}$$

$$\leq S(Q_1') - S(E_1') + S(Q_2') - S(E_2') = \tag{4.57}$$

$$= I(\rho^{Q_1}, \mathcal{N}) + I(\rho^{Q_2}, \mathcal{M}). \tag{4.58}$$

L'uguaglianza in (4.52) segue dalla scelta dello stato $\rho^{Q_1Q_2}$, la (4.53) e la (4.54) seguono dalla definizione di informazione coerente (per esprimere l'entropia di von Neumann si utilizza la stessa convenzione usata nel Teorema 4.2). Nel passare all'espressione (4.55) si applica la definizione di entropia di von Neumann congiunta. Nell'applicare la disuguaglianza tra la (4.56) e la (4.57) si usa il concetto di canale degradabile insieme alla proprietà (3) del Teorema 1.5. Infatti date le operazioni $\mathcal{T}_1 : \rho^{Q_1'} \to \rho^{E_1'}$ e $\mathcal{T}_2 : \rho^{Q_2'} \to \rho^{E_2'}$, gli stati dei sistemi Q_1' e Q_2' condividono più informazione prima dell'applicazione delle mappe $\mathcal{T}_1$ e $\mathcal{T}_2$, cioè vale $I(Q_1' : Q_2') \geq I(E_1' : E_2')$, quindi la quantità $[I(Q_1' : Q_2') - I(E_1' : E_2')]$ è positiva. Si può proseguire vedendo:

$$\mathcal{Q}^{(1)}(\mathcal{E}) \leq I(\rho^{Q_1}, \mathcal{N}) + I(\rho^{Q_2}, \mathcal{M}) \tag{4.59}$$

$$\leq \max_{\rho^{Q_1'}} I(\rho^{Q_1}, \mathcal{N}) + \max_{\rho_{Q_2'}} I(\rho^{Q_2}, \mathcal{M}) \tag{4.60}$$

$$= Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{M}). \tag{4.61}$$

È conclusa così la dimostrazione. □

L'additività della massima informazione coerente per canali degradabili fa quindi sì che per n usi del canale: $\mathcal{Q}^{(1)}(\mathcal{E}^{\otimes n}) = n\mathcal{Q}^{(1)}(\mathcal{E})$. Se ne conclude che questa classe di canali ha una capacità quantistica ben definita: $\mathcal{Q}(\mathcal{N}) = \mathcal{Q}^{(1)}(\mathcal{N})$.

Conclusioni

In questa tesi sono stati ripercorsi i risultati principali del lavoro iniziato nell'ultimo decennio del XX secolo, nell'ambito della teoria dell'informazione quantistica, per comprendere concettualmente e qualitativamente la capacità dei canali quantistici. Affrontare tale problema ha portato ad ottenere risposte anche molto diverse rispetto a ciò che si conosceva in teoria dell'informazione classica ed ancora oggi molte questioni rimangono aperte, tra cui la mancata equivalenza tra la capacità per un uso del canale e la capacità asintotica.

La prima differenza tra il contesto classico e quello quantistico è che, mentre classicamente la capacità di un canale è unica e coincide con la massima mutua informazione condivisa tra ingresso e uscita del canale, per un canale quantistico ci sono diverse capacità da considerare, dipendenti dall'uso che viene fatto del canale e dalle risorse che si sceglie di utilizzare per la trasmissione dell'informazione. Una seconda differenza, nonché la principale difficoltà che si incontra nel cercare generalizzazioni quantistiche del concetto di capacità, è dovuta alla non additività delle grandezze che determinano le varie capacità. Ovvero, la somma dei valori che tali grandezze assumono indipendentemente su due canali non coincide con il valore valutato sull'uso congiunto di due canali. Tuttavia, prendere come caso studio specifiche classi di canali quantistici permette di ottenere risultati più precisi.

Nei primi due capitoli della presente tesi sono stati presentati il formalismo quantistico necessario per proseguire e i ben consolidati risultati classici. Nel terzo e nel quarto capitolo sono state analizzate la capacità classica per un canale quantistico, la capacità entanglement assistita e la capacità quantistica.

La capacità classica di un canale quantistico si ha quando il mittente manda informazione classica al destinatario attraverso un canale quantistico rumoroso. Il ruolo della mutua informazione è in questo contesto ricoperto dalla quantità χ di Holevo, quantità che limita superiormente la massima informazione mutua presente tra la variabile in ingresso e quella in uscita del canale. Il teorema di Holevo-Schumacher-Westmoreland dimostra che quando gli stati in ingresso nel canale sono separabili, la massima quantità

di Holevo coincide con la capacità per stati prodotto. La quantità di Holevo non è però uguale alla capacità classica per canali quantistici in generale, a causa della sua non additività; si ricorre quindi alla regolarizzazione della massima quantità di Holevo su molti usi del canale. I canali di entanglement-breaking e i depolarizing channels, citati nel terzo capitolo di questo lavoro, ma anche gli *unital channels* - canali che lasciano l'identità uguale a se stessa - sono però un esempio per cui la massima quantità di Holevo è additiva e quindi la capacità classica è ben definita.

La capacità classica entanglement-assistita si considera per la trasmissione di informazione classica tra mittente e destinatario che condividono stati entangled prima della trasmissione. Si calcola massimizzando la mutua informazione quantistica tra gli stati in ingresso e in uscita del canale ed è l'unica, tra le capacità considerate, che non presenta una formula regolarizzata, presentandosi così come la più semplice da computare e la più simile alla capacità classica.

La capacità quantistica è un limite per la quantità di informazione quantistica, codificata in stati quantistici, trasmissibile attraverso un canale. Per valutarla, si utilizza l'informazione coerente, grandezza che quantifica la perdita di informazione di uno stato durante la sua evoluzione in un canale. Poiché in generale la massima informazione coerente è fortemente non additiva, l'espressione della capacità quantistica, detta formula di Lloyd-Shor-Devetak si può ottenere soltanto con la regolarizzazione su molti usi del canale. È stato addirittura dimostrato teoricamente che due canali con capacità nulla singolarmente, possono avere una capacità non nulla se utilizzati insieme, un fenomeno detto superattivazione. Per la classe di canali degradabili, si dimostra l'additività della massima informazione coerente che quindi rappresenta proprio la loro capacità quantistica.

Continuare a comprendere la capacità dei canali quantistici, migliorandone le espressioni e trovando risposta alle domande aperte è fondamentale non solo a livello teorico, ma anche per applicazioni pratiche, per sfruttare i sistemi quantistici nello sviluppo di nuove tecnologie per la comunicazione e la crittografia quantistica.

Bibliografia

[1] H. Barnum, M. A. Nielsen, and B. Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, 57:4153–4175, Jun 1998. URL: https://link.aps.org/doi/10.1103/PhysRevA.57.4153, doi:10.1103/PhysRevA.57.4153.

[2] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, 2002. doi:10.1109/TIT.2002.802612.

[3] I. Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, 2005. doi:10.1109/TIT.2004.839515.

[4] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255–257, 2009. doi:10.1038/nphys1224.

[5] P. Hayden, M. Horodecki, J. Yard, and A. Winter. A decoupling approach to the quantum capacity. *Open Systems & Information Dynamics*, 15, 03 2007. doi:10.1142/S1230161208000043.

[6] S. Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613–1622, Mar 1997. URL: https://link.aps.org/doi/10.1103/PhysRevA.55.1613, doi:10.1103/PhysRevA.55.1613.

[7] D. Maslov, J.-S. Kim, S. Bravyi, T. J. Yoder, and S. Sheldon. Quantum advantage for computations with limited space. *Nature Physics*, 17:894–897, 2021. doi:10.1038/s41567-021-01271-7.

[8] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[9] J. Preskill. Quantum Information, Chapter 10. Quantum Shannon Theory. Lecture notes, Institute for Quantum Information and Matter, California Institute of Technology, 1998. https://preskill.caltech.edu/ph219/.

[10] B. Schumacher and M. A. Nielsen. Quantum data processing and error correction. *Phys. Rev. A*, 54:2629–2635, Oct 1996. URL: https://link.aps.org/doi/10.1103/PhysRevA.54.2629, doi:10.1103/PhysRevA.54.2629.

[11] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi:10.1002/j.1538-7305.1948.tb01338.x.

[12] P. W. Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics*, 43(9):4334–4340, 2002. arXiv:https://pubs.aip.org/aip/jmp/article-pdf/43/9/4334/19183079/4334_1_online.pdf, doi:10.1063/1.1498000.

[13] P. W. Shor. The quantum channel capacity and coherent information. Lecture notes, MSRI Workshop on Quantum Computation, 2002.

[14] M. M. Wilde. From Classical to Quantum Shannon Theory, 6 2011. arXiv:1106.1445, doi:10.1017/9781316809976.001.

[15] M. M. Wilde, A. Winter, and D. Yang. Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, 2014. doi:10.1007/s00220-014-2122-x.