Introduzione all'Informazione Quantistica

Gabriele Sicuro
Appunti dal corso tenuto dal prof. V.Giovannetti
Scuola Normale Superiore di Pisa
2012

Indice

Capitolo 1

Fondamenti

1.1. Introduzione

1.1.1. La teoria dell'Informazione

La teoria dell'Informazione quantistica si occupa dell'utilizzo di sistemi dalle caratteristiche quantistiche per l'elaborazione di dati. Essa può essere suddivisa in tre aree interconnesse fra loro, inerenti rispettivamente la *Quantum Computation*, la *Quantum Communication* e la *Quantum Cryptography*. Inoltre la *teoria dell'entanglement* e quella *dei sistemi quantici aperti* sono incluse nella teoria dell'informazione quali suoi aspetti fondamentali, mentre la *teoria del controllo* si occupa delle modalità di manipolazione di sistemi fisici quantistici.

La teoria dell'Informazione Quantistica (*Quantum Information Theory*) si occupa delle modalità di acquisizione, immagazzinamento e trasmissione dell'*informazione*, ovvero di una *possibilità di scelta* fra diverse configurazioni del sistema. Ovviamente un'informazione può essere trasportata, a livello pratico, in maniera molto diversa.

I primi lavori riguardanti gli aspetti matematici della teoria sono riconducibili a Bayes, ma occorre attendere gli anni Trenta del secolo precedente per una analisi più consapevole del problema, nata da questioni sia pratiche sia teoriche. I contributi più importanti in questo senso sono da associarsi a Turing, Church e von Neumann. Si pervenne in generale alla conclusione che si poteva distinguere fra problemi *computabili* e problemi *non computabili* (*tesi di Church–Turing*) ma anche i problemi computabili risultano suddivisibili in diverse *classi di complessità*, distinte procedendo con una discretizzazione del problema ed in seguito con l'individuazione del numero di processi fondamentali (ad esempio operazioni sui singoli bit) necessari per giungere ad una soluzione dello stesso. Questo tipo di valutazione è solitamente non banale e comunque tuttora non è ben definito il modo in cui è possibile distinguere tra problemi "semplici" e "difficili". La discretizzazione dell'informazione, inoltre, è necessaria perché la comunicazione sia stabile.

1.1.2. Qubit

Un *bit* è l'unità fondamentale della computazione ed è costituito da un dispositivo a due livelli in grado di immagazzinare un'informazione binaria, assumendo uno dei due valori ammessi (solitamente indicati con 0 e 1). In meccanica quantistica

possiamo considerare un sistema a due livelli $\mathcal{H} = \text{span}\{|0\rangle, |1\rangle\}$, dove $\langle 1|1\rangle = \langle 0|0\rangle = 1$, $\langle 0|1\rangle = 0$, ed utilizzarlo come bit, associando allo stato $|0\rangle$ il valore 0 e allo stato $|1\rangle$ il valore 1. A differenza del caso classico, tuttavia, il sistema ammette stati di sovrapposizione del tipo $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, $|\alpha|^2 + |\beta|^2 = 1$, $\alpha, \beta \in \mathbb{C}$. Un sistema quantico siffatto viene chiamato *qubit*. La sua introduzione negli anni Ottanta è nata dall'idea che problemi di difficile computazionabilità con tecniche classiche potessero invece risultare semplici con tecniche quantistiche.

1.1.3. Noise

La teoria della comunicazione classica è nata negli anni Quaranta con gli studi di Shannon sui problemi di trasmissione in presenza di rumore (*noise*). Tale rumore era introdotto, nella teoria originaria, da effetti estrinseci di malfunzionamento (inefficienze di trasmissione, disturbi nei cavi...). Oggi sono introdotti e studiati anche effetti *intrinseci*, dovuti alla natura quantistica del dispositivo in esame, in virtù dell'interesse verso le peculiarità del mondo dei quanti che permette effetti come il teletrasporto e l'entanglement (fondamentale in crittografia quantistica).

1.2. Stati puri e miscele

1.2.1. Matrice densità

Sia $|\psi\rangle$ un generico stato di un sistema associato ad uno spazio di Hilbert $\mathcal{H}$ di dimensione finita n, secondo l'interpretazione di Born la probabilità che una misura su $|\psi\rangle$ resituisca lo stato $|i\rangle$ è data da $P_i(\psi) \stackrel{\text{def}}{=} p\left(|\psi\rangle \to |i\rangle\right) = |\langle i|\psi\rangle|^2$.

Talvolta però non si hanno informazioni sufficienti per associare al sistema un'unica funzione d'onda (ad esempio a causa di possibili malfunzionamenti del dispositivo che prepara lo stato). Dunque è possibile che il sistema si trovi in uno di m possibili stati $|\psi_i\rangle$, ciascuno con probabilità p_i, $\sum_{i=1}^m p_i = 1$. L'insieme delle coppie $\{(|\psi_i\rangle, p_i)\}$ è detto *ensemble*. La relazione di Born dev'essere dunque riformulata come media sull'ensemble:

$$P_i\left(\{(|\psi_j\rangle, p_j)\}\right) = \sum_{k=1}^m p_k |\langle i|\psi_k\rangle|^2 = \langle i| \left(\sum_{k=1}^m p_k |\psi_k\rangle\langle\psi_k| \right) |i\rangle \stackrel{\text{def}}{=} \langle i|\rho|i\rangle. \tag{1.1}$$

L'operatore $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ è detto *matrice densità* e descrive sia le proprietà quantistiche sia quelle statistiche del sistema. L'operatore ρ è definito positivo, è autoaggiunto ed ha traccia unitaria. Inoltre si ha, per una generica osservabile,

$$\langle\Theta\rangle = \sum_k p_k \langle\psi_k|\Theta|\psi_k\rangle = \sum_{i=1}^n \sum_k p_k \langle i|\Theta|\psi_k\rangle\langle\psi_k|i\rangle = \text{tr}\left[\Theta\rho\right]. \tag{1.2}$$

Essendo ρ definito positivo, è sempre possibile diagonalizzarlo,

$$\rho = \sum_{l=1}^n \lambda_l |\chi_l\rangle\langle\chi_l|, \quad \sum_l \lambda_l = 1, \quad \langle\chi_l|\chi_k\rangle = \delta_{lk}, \quad \lambda_l \in [0,1].$$

Poiché è possibile variare strumenti di misura è chiaro (da un punto di vista fisico) che è possibile ottenere ρ utilizzando stati di riferimento diversi; si ottengono così

ensembles equivalenti ed indistinguibili. Tuttavia, solo in un ensemble è possibile avere ρ diagonale rispetto ad una base ortonormale. Se infatti $\mathcal{E}_1 = \{(|\psi_k\rangle, p_k)\}$, $m_1 = \dim \mathcal{E}_1$, ed $\mathcal{E}_2 = \{(|\phi_k\rangle, q_k)\}$, $m_2 = \dim \mathcal{E}_2$, sono due ensembles, possiamo supporre $m_1 = m_2 = m$ (se ad esempio $m_1 < m_2$ possiamo aggiungere stati in $\mathcal{E}_1$ corrispondenti a probabilità nulle). Se ρ_1 e ρ_2 sono le rispettive matrici densità, esse sono equivalenti se e solo se $\exists U = (u_{kl})_{kl}$ unitario tale che $\sqrt{p_k}|\psi_k\rangle = \sum_l u_{kl}\sqrt{q_l}|\phi_l\rangle$: sotto questa ipotesi si dimostra facilmente che $\rho_1 = U\rho_2 U^\dagger$. È possibile dunque ottenere ensembles da ensembles mediante isometrie.

Nel caso ideale, se l'apparato fornisce uno stato puro, allora $\rho = |\psi\rangle\langle\psi|$ (proiettore sullo stato $|\psi\rangle$). Osserviamo che in tal caso ρ è idempotente, $\rho^2 = \rho \Rightarrow \operatorname{tr}[\rho^2] = \operatorname{tr}[\rho] = 1$, mentre in generale $\operatorname{tr}[\rho^2] = \sum_{l=1}^n \lambda_l^2 \leq \operatorname{tr}[\rho] = 1$.

Per un operatore con le proprietà della matrice densità (ovvero semidefinito positivo a traccia unitaria) è sempre possibile associare un ensemble costruendo, ad esempio, la sua rappresentazione spettrale: dunque si può affermare che *un generico operatore è una matrice densità se e solo se è semidefinito positivo e a traccia unitaria*. L'insieme delle matrici densità su $\mathcal{H}$ si indica con $\sigma(\mathcal{H})$: tale insieme è *convesso*, essendo $p\rho_1 + (1-p)\rho_2 \in \sigma(\mathcal{H})$ per $p \in [0,1]$. I punti estremali di questo dominio sono stati puri (in quanto qualunque ρ è ottenibile come sovrapposizione di stati puri), con l'esclusione dei punti interni delle "facce piane".

Sfera di Bloch

Sia

$$\rho = \begin{pmatrix} p & \gamma \\ \gamma^* & 1-p \end{pmatrix} = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| + \gamma|0\rangle\langle 1| + \gamma^*|1\rangle\langle 0|,$$

con $p \in [0,1]$, una matrice densità per uno spazio bidimensionale $\mathcal{H} = \operatorname{span}\{|0\rangle, |1\rangle\}$ (tipo qubit). Nel caso di tale matrice gli elementi di diagonale sono detti *popolazioni*, mentre gli elementi fuori diagonale sono detti *di coerenza* e soddisfano la relazione $|\gamma| \leq \sqrt{p(1-p)}$ (dalla condizione $\det\rho \geq 0$). Introducendo le matrici di Pauli

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

possiamo scrivere $\rho = \frac{\sigma^0 + \mathbf{a}\cdot\boldsymbol{\sigma}}{2}$, dove $\mathbf{a} = (2\Re\gamma, -2\Im\gamma, 2p-1) \in \mathbb{R}^3$ con $|\mathbf{a}| \leq 1$ e $\boldsymbol{\sigma} = (\sigma^1, \sigma^2, \sigma^3)$. Il vettore $\mathbf{a}$, detto *vettore di Bloch*, individua lo stato del sistema: di conseguenza tutti gli stati del sistema sono associabili ai punti della palla chiusa di raggio unitario nello spazio $\mathbb{R}^3$, detta *sfera di Bloch*. Essa rappresenta graficamente lo spazio $\sigma(\mathcal{H})$; è chiusa e convessa, mentre per $|\mathbf{a}| = 1$ si ottengono stati puri, essendo in tal caso $\operatorname{tr}[\rho^2] = \operatorname{tr}[\rho] = 1$.

1.2.2. Decomposizione di Schmidt

In generale il nostro sistema quantistico è costituito da più sottosistemi: ad esempio, è possibile considerare una *coppia* di qubit, con spazio di singolo qubit $\mathcal{H}_A$ e $\mathcal{H}_B$ rispettivamente, il cui spazio di Hilbert sarà dato globalmente dal prodotto $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, nell'ipotesi di qubit non interagenti. Una base ortonormale completa in uno spazio dato dal prodotto tensore di $\mathcal{H}_A = \operatorname{span}\{|j\rangle_A\}_{j=1,\dots,d_A}$ e

$\mathcal{H}_B = \text{span}\{|l\rangle_B\}_{l=1,\dots,d_B}$ può essere ottenuta considerando i $d_A d_B$ vettori $\{|j\rangle_A \otimes |l\rangle_B\}_{\substack{j=1,\dots,d_A \\ l=1,\dots,d_B}}$ nell'ipotesi che $\{|j\rangle_A\}_{j=1,\dots,d_A}$ e $\{|l\rangle_B\}_{l=1,\dots,d_B}$ siano COS per i rispettivi spazi. Un generico vettore normalizzato $|\psi\rangle_{AB} \in \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ può dunque essere scritto

$$|\psi\rangle_{AB} = \sum_{j,l} \psi_{jl} |j\rangle_A \otimes |l\rangle_B.$$

In generale, dunque, uno stato siffatto non può essere messo nella forma $|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B$. Quando ciò è possibile si dice che $|\psi\rangle_{AB}$ è *separabile*; diversamente si dice che lo stato è *entanglato* (*entangled* – una più precisa definizione verrà data in seguito). Tuttavia, dato uno stato normalizzato generico $|\psi\rangle_{AB}$, è possibile decomporlo come

$$|\psi\rangle_{AB} = \sum_{k=1}^{r} \sqrt{\lambda_k} |\psi_k\rangle_A \otimes |\phi_k\rangle_B, \quad \sum_{k=1}^{r} \lambda_k = 1,$$

con $\{|\psi_k\rangle_A\}$ COS in $\mathcal{H}_A$ e $\{|\phi_k\rangle_B\}$ COS in $\mathcal{H}_B$. Nella precedente $r \leq \min\{d_A, d_B\}$ è detto *rango di Schmidt*. La decomposizione sopra è detta *decomposizione di Schmidt*. Per provare il precedente risultato, supponiamo di avere $d_A = d_B = d$ (ciò può essere ottenuto aggiungendo vettori ortonormali fittizi allo spazio di dimensione minore). Ricordiamo che, data una matrice quadrata $d \times d$ $\Psi = (\psi_{ij})_{ij}$ è sempre possibile decomporla come $\Psi = UDV$ (*singular value decomposition*), con $U = (u_{ij})_{ij}, V = (v_{ij})_{ij}$ unitarie e $D = (d_{ii}\delta_{ij})_{ij}$ diagonale semidefinita positiva con autovalori eguali a quelli di $\sqrt{\Psi^\dagger \Psi}$. Dunque $\psi_{jl} = \sum_k u_{jk} d_{kk} v_{kl}$. Adoperando questa relazione

$$\begin{aligned} |\psi\rangle_{AB} = \sum_{j,l} \psi_{jl} |j\rangle_A \otimes |l\rangle_B = \sum_k d_{kk} \left(\sum_j u_{jk} |j\rangle_A \right) \otimes \left(\sum_l v_{kl} |l\rangle_B \right) \\ = \sum_k d_{kk} |\psi_k\rangle_A \otimes |\phi_k\rangle_B. \quad (1.3) \end{aligned}$$

Nella precedente possiamo porre $\sqrt{\lambda_k} = d_{kk}$, osservando che $\sum_k \lambda_k = \text{tr}\left[D^2\right] = \text{tr}\left[\Psi^\dagger \Psi\right] = 1$ (essendo lo stato normalizzato). L'ortonormalità degli stati di base ottenuti si verifica facilmente come conseguenza dell'unitarietà di U e V.

Applicando la procedura sopra ad uno stato separabile si riottiene ovviamente la forma di partenza. Gli stati separabili hanno rango di Schmidt unitario, mentre $r > 1$ è condizione necessaria e sufficiente per la non separabilità. In effetti il rango di Schmidt è una "misura" della non separabilità di uno stato, o del suo entanglement. Gli stati puri sono solitamente una minoranza tra gli stati possibili.

1.2.3. Matrici densità ridotte

Se consideriamo un sistema composto da due sottosistemi A e B non interagenti, i cui spazi di Hilbert sono $\mathcal{H}_A$ ed $\mathcal{H}_B$ rispettivamente, ha senso chiedersi come estendere l'azione di un'osservabile Θ_A relativa al solo sistema A allo spazio $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Essendo osservabile, esiste una base in A in cui Θ_A è diagonale, ovvero

$\Theta_A = \sum_{i=1}^{d_A} \theta_i |\phi_i\rangle_A \langle\phi_i|$. La più ovvia estensione è

$$\Theta_A \otimes \mathbb{1}_B = \sum_{j=1}^{d_B} \sum_{i=1}^{d_A} \theta_i |\phi_i\rangle_A \langle\phi_i| \otimes |j\rangle_B \langle j| = \sum_{j=1}^{d_B} \sum_{i=1}^{d_A} \theta_i |\phi_i j\rangle_{AB} \langle\phi_i j|,$$

dove $\{|j\rangle_B\}_{j=1,\dots,d_B}$ è un COS in $\mathcal{H}_B$, mentre si è posto $|\phi_i j\rangle_{AB} \stackrel{\text{def}}{=} |\phi_i\rangle_A \otimes |j\rangle_B$.

La probabilità di ottenere lo stato l-esimo su A da un generico stato $|\psi\rangle_{AB}$ è dunque

$$P_l(|\psi\rangle_{AB}) = \sum_{j=1}^{d_B} |_{AB}\langle\psi|\phi_l j\rangle_{AB}|^2 = \sum_{j=1}^{d_B} \left({}_A\langle\phi_l| \otimes {}_B\langle j|\right) |\psi\rangle_{AB} \langle\psi| \left(|\phi_l\rangle_A \otimes |j\rangle_B\right)$$
$$\equiv {}_A\langle\phi_l|\rho_A|\phi_l\rangle_A, \quad (1.4)$$

ovvero la presenza del sistema B introduce una sorta di "degenerazione" su cui occorre sommare, mentre

$$\rho_A \stackrel{\text{def}}{=} \sum_{j=1}^{d_B} {}_B\langle j|\psi\rangle_{AB} \langle\psi|j\rangle_B \stackrel{\text{def}}{=} \operatorname{tr}_B [\rho_{AB}]$$

è detta *matrice densità ridotta* ed è ottenuta, come si vede, tramite una traccia parziale sugli elementi di base di B. Inoltre scrivendo $|\psi\rangle_{AB}$ in forma di Schmidt si può scrivere $\rho_A = \sum_{k=1}^{r} \lambda_k |\psi_k\rangle_A \langle\psi_k|$, ovvero le quantità λ_k che compaiono nella rappresentazione di Schmidt sono gli autovalori della matrice ρ_A.

Eseguendo analoghi ragionamenti su B si ottiene, relativamente allo stesso stato $|\psi\rangle_{AB}$, $\rho_B = \sum_{k=1}^{r} \lambda_k |\phi_k\rangle_B \langle\phi_k|$, ovvero le matrici ridotte hanno lo stesso spettro. Tale proprietà, tuttavia, vale solo per *stati puri*, mentre non vale per *stati misti*, la cui matrice densità è della forma $\rho = \sum_{j=1}^{m} p_j |\psi_j\rangle_{AB} \langle\psi_j|$, $\sum_{i=1}^{m} p_i = 1$, per i quali, eseguendo una traccia parziale, si ottiene $\rho_A = \sum_{i=1}^{m} p_i \rho_A^i$.

1.2.4. Purificazioni

In un sistema bipartito a stati puri come $|\psi\rangle_{AB}$ è possibile associare matrici di densità ridotte tipo $\rho_A = \operatorname{tr}_B [|\psi\rangle_{AB}\langle\psi|]$ e $\rho_B = \operatorname{tr}_A [|\psi\rangle_{AB}\langle\psi|]$. Mostriamo ora che, viceversa, $\forall \rho_A$ operatore con proprietà di matrice densità ridotta sul sottosistema A $\exists |\psi\rangle_{AB}$ (detto *purificazione*) tale che $\operatorname{tr}_B [|\psi\rangle_{AB}\langle\psi|] = \rho_A$. Ciò significa che è sempre possibile descrivere un processo in termini di stati puri, a prescindere dall'eventuale rumore presente o dalla procedura di misura. Considerando la matrice ρ_A nella base che la diagonalizza, $\rho_A = \sum_{j=1}^{d_A} \lambda_j |j\rangle_A \langle j|$, possiamo costruire un vettore in rappresentazione di Schmidt $|\psi\rangle_{AB} = \sum_{j=1}^{d_A} \sqrt{\lambda_j} |j\rangle_A \otimes |\phi_j\rangle_B$, dove $\{|\phi_j\rangle_B\}$ è un generico COS in B (se $\dim \mathcal{H}_B \leq \dim \mathcal{H}_A$ è sempre possibile aggiungere stati fittizi alla base di $\mathcal{H}_B$). È immediatamente evidente che $\operatorname{tr}_B [|\psi\rangle_{AB}\langle\psi|] = \rho_A$. Inoltre questo stato non è univocamente identificato in quanto tutti (e soli) gli stati $|\psi'\rangle_{AB} = (\mathbb{1}_A \otimes V_B) |\psi\rangle_{AB}$, con V_B operatore unitario su $\mathcal{H}_B$, hanno la stessa matrice densità ridotta. L'operatore V_B (che non fa altro che cambiare la base scelta su $\mathcal{H}_B$) è un esempio di *gate*.

1.3. Misure generalizzate e *Positive Operator Valued Measure*

È possibile riformulare la teoria degli osservabili in meccanica quantistica nell'ipotesi che il sistema, supposto ad esempio bipartito, sia composto dal sottosistema oggetto della misura e da un *sottosistema ancillare* non direttamente monitorato. Questo comporta naturalmente una riformulazione della regola di Born. Supponiamo che il sistema sia associato ad uno spazio di Hilbert $\mathcal{H}_{SA} = \mathcal{H}_S \otimes \mathcal{H}_A$, dove $\mathcal{H}_S$ è lo spazio di Hilbert relativo al nostro sottosistema di interesse, cui è associata la matrice densità ρ_S, mentre lo spazio $\mathcal{H}_A$ è associato al sistema ancillare, supposto nello stato puro $|0\rangle_A$. In tal caso $\rho_{SA} = \rho_S \otimes |0\rangle_A\langle 0|$. L'evoluzione avviene tramite un operatore unitario U_{SA}, mentre, data un'osservabile $\Theta_{SA} = \sum_j \theta_j |j\rangle_{SA}\langle j|$, la probabilità di ottenere l'autovalore θ_j risulta $P_j = {}_{SA}\langle j|U_{SA}\rho_{SA}U_{SA}^\dagger|j\rangle_{SA} \equiv {}_S\langle v_j|\rho_S|v_j\rangle_S = \mathrm{tr}\left[|v_j\rangle_S\langle v_j|\rho_S\right] \equiv \mathrm{tr}\left[E_j\rho_S\right]$, dove abbiamo definito $|v_j\rangle_S \overset{\text{def}}{=} {}_A\langle 0|U_{SA}^\dagger|j\rangle_{SA}$ e in particolare abbiamo introdotto gli *operatori hermitiani semidefiniti positivi*

$$E_j \overset{\text{def}}{=} |v_j\rangle_S\langle v_j| = {}_A\langle 0|U_{SA}^\dagger|j\rangle_{SA}\langle j|U_{SA}|0\rangle_A, \qquad \sum_j E_j = \mathbb{1}_S,$$

detti *positive operator valued measure* (POVM). La regola

$$P_j = \mathrm{tr}\left[E_j\rho_S\right]$$

generalizza la regola di Born.

Dato un set di operatori semidefiniti positivi tali che $\sum_j E_j = \mathbb{1}_S$, è possibile associarvi una POVM. Essendo infatti gli operatori semidefiniti positivi, l'operatore $\sqrt{E_j}$ è ben definito e, supponendo il sistema inizializzato nello stato $|\psi\rangle_S$ e l'ancella nello stato $|0\rangle_A$, possiamo introdurre un operatore unitario U_{SA} descrivendone l'azione sui vettori $|\psi\rangle_S \otimes |0\rangle_A$, ovvero imponendo $U_{SA}\left(|\psi\rangle_S \otimes |0\rangle_A\right) = \sum_j \sqrt{E_j}|\psi\rangle_S \otimes |j\rangle_A$ (si prova banalmente che il prodotto scalare è preservato). La probabilità P_m è data da ${}_S\langle\psi|\sqrt{E_m}^\dagger\sqrt{E_m}|\psi\rangle_S = {}_S\langle\psi|E_m|\psi\rangle_S = \mathrm{tr}\left[E_m\rho_S\right]$. Lo stato (normalizzato) dopo la misura è dunque $\frac{\sqrt{E_m}|\psi\rangle_S}{\sqrt{{}_S\langle\psi|E_m|\psi\rangle_S}} \Rightarrow \rho' = \frac{\sqrt{E_m}\rho\sqrt{E_m}^\dagger}{\mathrm{tr}\left[E_m\rho\right]}$: si noti che la relazione è non lineare.

Le misure proiettive si ottengono considerando $E_j = |j\rangle_S\langle j|$ (caso in cui l'ancella è assente).

1.4. Distanza fra stati

1.4.1. Premessa

La distanza indotta dal prodotto scalare nello spazio di Hilbert $\mathcal{H}$ non è utilizzabile per distinguere gli stati fisici. In spazi di Hilbert di dimensione finita si introduce una distanza sulla falsariga della teoria delle distribuzioni di probabilità discrete, dove, date due distribuzioni $\{p_x\}_x$ e $\{q_x\}_x$, dove l'indice x corre su un dominio discreto X, si definisce la *distanza di Kolmogorov* $\mathcal{D}(p_x, q_x) = \frac{1}{2}\sum_x |p_x - q_x| \in [0, 1]$. Si definisce anche la *fedeltà* tra due distribuzioni come $\mathcal{F}(p_x, q_x) = \left(\sum_x \sqrt{p_x q_x}\right)^2 \in [0, 1]$.

1.4.2. Trace distance

Siano $\rho, \tau \in \sigma(\mathcal{H})$. Definiamo la *distanza traccia (trace distance)* tra i due stati come

$$D(\rho,\tau) \stackrel{\text{def}}{=} \frac{\text{tr}\,[|\rho-\tau|]}{2}, \quad |\rho| \stackrel{\text{def}}{=} \sqrt{\rho^\dagger \rho}. \tag{1.5}$$

La precedente verifica banalmente tutte le proprietà di una distanza, eccettuata la disuguaglianza triangolare che proveremo esplicitamente. Occorre osservare preliminarmente che $D(\rho,\tau) \stackrel{\text{def}}{=} \frac{\text{tr}[|\rho-\tau|]}{2} = \max_P \text{tr}\,[P(\rho-\tau)]$, dove $P = P^\dagger$ è un generico proiettore. Essendo $\rho - \tau$ hermitiano e a traccia nulla, infatti, possiamo scegliere una base in cui $\rho - \tau = A - B$, con $A = \alpha_i \delta_{ij}$, $B = \beta_i \delta_{ij}$ tali che[1] $\alpha_i = \lambda_i \theta(\lambda_i) \geq 0, \beta_i = -\lambda_i \theta(-\lambda_i) \geq 0$, λ_i i-esimo autovalore di $\rho - \tau$. Si ha $A = A^\dagger$, $B = B^\dagger$, $A, B \geq 0$, e $AB = 0$; dunque $(A-B)^\dagger(A-B) = (A-B)^2 = A^2 + B^2 = (A+B)^2 \Rightarrow |\rho - \tau| = A + B$, da cui $D(\rho,\tau) = \text{tr}\,[A]$. D'altra parte, $\text{tr}\,[A] \geq \text{tr}\,[PA]$ $\forall P$ proiettore. Dunque $D(\rho,\tau) = \text{tr}\,[A] \geq \text{tr}\,[PA] \geq \text{tr}\,[P(A-B)] \equiv \text{tr}\,[P(\rho-\tau)]$. Se consideriamo il proiettore P_A sul supporto di A, è possibile provare l'implicazione inversa. Per cui, detto $\bar{P}$ il proiettore che satura la diseguaglianza, $D(\rho,\tau) = \text{tr}\left[\bar{P}(\rho-\tau)\right] = \text{tr}\left[\bar{P}(\rho-\upsilon)\right] + \text{tr}\left[\bar{P}(\upsilon-\tau)\right] \leq D(\rho,\upsilon) + D(\upsilon,\tau)$, che prova la validità della diseguaglianza triangolare.

Proprietà della trace distance

Assegnati due stati puri $\rho = |\psi\rangle\langle\psi|$ e $\tau = |\phi\rangle\langle\phi|$, si ha $D(\rho,\tau) = \max_P \text{tr}\,[P(\rho-\tau)] = 1 - |\langle\psi|\phi\rangle|^2$, dove si è considerato che il proiettore $P = |\zeta\rangle\langle\zeta|$ che massimizza la quantità $|\langle\zeta|\psi\rangle|^2 - |\langle\zeta|\phi\rangle|^2$ è ottenuto per $|\zeta\rangle = |\psi\rangle$.

Si prova inoltre che

$$p_x \stackrel{\text{def}}{=} \text{tr}\,[E_x \rho], \quad q_x \stackrel{\text{def}}{=} \text{tr}\,[E_x \tau] \Longrightarrow D(\rho,\tau) = \max_{\text{POVM}} \mathcal{D}(p_x, q_x).$$

Ciò equivale a dire che la *trace distance* esprime la massima distanza tra le statistiche ottenibili mediante misure qualsivoglia.

Dato U operatore unitario, si ha

$$D(U\rho U^\dagger, U\tau U^\dagger) = D(\rho,\tau).$$

Sia data una trasformazione lineare Φ sullo spazio degli operatori (*superoperatore*), tale che goda delle seguenti proprietà (*trace-preserving positive quantum operator*):

- Φ preserva la traccia, $\text{tr}\,[\rho] = \text{tr}\,[\Phi[\rho]]$;
- se $\Theta \geq 0$ allora $\Phi[\Theta] \geq 0$;
- Φ è lineare sullo spazio degli operatori.

Allora si prova che tale operatore è *contrattivo*

$$D(\Phi[\rho], \Phi[\tau]) \leq D(\rho,\tau);$$

[1] Abbiamo indicato con $\theta(x) = \begin{cases} 1 & \text{se } x > 0, \\ 0 & \text{altrimenti} \end{cases}$ la funzione di Heaviside.

in particolare, dato un sistema composito $\mathcal{H}_A \otimes \mathcal{H}_B$ una mappa $\Phi[\cdot] \stackrel{\text{def}}{=} \text{tr}_B[\cdot]$ è del tipo sopra e dunque $D(\rho_A, \tau_A) \leq D(\rho_{AB}, \tau_{AB})$. Il segno di uguaglianza vale solo se $\rho_{AB} = \rho_A \otimes \rho_B$ e $\tau_{AB} = \tau_A \otimes \rho_B$.

Se sono date inoltre ρ e $\tau = \sum_j p_j \tau_j$, allora $D(\rho, \tau) \leq \sum_j p_j D(\rho, \tau_j)$.

Trace distance e qubit

Consideriamo il caso di due qubit con matrici densità $\rho = \frac{1+\mathbf{a}\cdot\boldsymbol{\sigma}}{2}$ e $\tau = \frac{1+\mathbf{b}\cdot\boldsymbol{\sigma}}{2}$ rispettivamente. Tramite calcolo diretto si vede che $D(\rho, \sigma) = \frac{\|\mathbf{a}-\mathbf{b}\|}{2}$, dunque la distanza è massima per punti antipodali, per i quali vale 1.

1.4.3. Fidelity

Il concetto di *fidelity* può essere introdotto nel contesto della meccanica quantistica utilizzando la definizione

$$F(\rho, \tau) = \left(\text{tr} \left[\sqrt{\sqrt{\rho}\tau\sqrt{\rho}} \right] \right)^2 \in [0, 1].$$

Se $\tau = |\psi\rangle\langle\psi|$, allora $F(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\rho|\psi\rangle$. La complessa espressione precedente è una funzione simmetrica nei suoi argomenti (benché ciò non risulti evidente). Inoltre per una trasformazione Φ *trace-preserving* come quelle introdotte sopra si ha $F(\Phi[\rho], \Phi[\tau]) \geq F(\rho, \tau)$; infine $F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|^2$ e $F(\rho, \tau) = 1 \Leftrightarrow \rho = \tau$ (criterio necessario e sufficiente per l'identità tra stati).

Valgono i seguenti teoremi.

Teorema (Uhlmann)**.** *Dati due stati quantistici ρ, τ, vale la seguente identità:*

$$F(\rho, \tau) = \max_{|\psi\rangle_\rho,\, |\phi\rangle_\tau} |_\rho\langle\psi|\phi\rangle_\tau|^2,$$

dove $|\psi\rangle_\rho$ e $|\phi\rangle_\tau$ sono purificazioni di ρ e τ rispettivamente.

Corollario. *Dati due stati quantistici ρ, τ e due loro purificazioni $|\psi\rangle_\rho$ e $|\phi\rangle_\tau$ rispettivamente,*

$$F(\rho, \tau) \geq |_\rho\langle\psi|\phi\rangle_\tau|^2.$$

Il precedente corollario mostra che, dato un sottosistema a del sistema A ed un sottosistema b del sistema B, a e b sono solitamente più fedeli di A e B; inoltre il teorema di Uhlmann esplicita la simmetria del funzionale introdotto ad inizio paragrafo.

La *fidelity* si ottiene operativamente utilizzando il risultato

$$p_x \stackrel{\text{def}}{=} \text{tr}[E_x\rho], \quad q_x \stackrel{\text{def}}{=} \text{tr}[E_x\tau] \Longrightarrow F(\rho, \tau) = \min_{\text{POVM}} \mathcal{F}(p_x, q_x).$$

La *fidelity* è adoperata per valutare l'efficienza di un *quantum computer* tramite confronto del risultato del calcolo reale e di quello ideale. È possibile provare che

$$1 - \sqrt{F(\rho, \tau)} \leq D(\rho, \tau) \leq \sqrt{1 - F(\rho, \tau)}$$

(per la disuguaglianza di destra, basta osservare che per due purificazioni $|\psi\rangle_\rho$ e $|\phi\rangle_\tau$ si ha $D(\rho, \tau) \leq D(|\psi\rangle_\rho\langle\psi|, |\phi\rangle_\tau\langle\phi|) \leq \sqrt{1 - |_\rho\langle\psi|\phi\rangle_\tau|^2} \leq \sqrt{1 - F(\rho, \tau)}$).

Incidentalmente osserviamo che si prova che $A(\rho, \tau) = \arccos F(\rho, \tau) \geq 0$ ha le proprietà di una distanza.

Capitolo 2

Dinamica dei sistemi aperti

2.1. Rappresentazione di Kraus e rappresentazione di Stinespring

È noto che in meccanica quantistica l'evoluzione di uno stato avviene mediante operatori unitari generati dall'hamiltoniana del sistema. Nel caso dei sistemi aperti, l'evoluzione è ancora determinata da trasformazioni unitarie e le matrici densità evolvono come $\rho \mapsto \tilde{\rho} = U(t)\rho U^\dagger(t)$. La trasformazione U *preserva la purezza di uno stato*, essendo $\mathrm{tr}\left[\tilde{\rho}^2\right] = \mathrm{tr}\left[\rho^2\right]$. Sia in $t = 0$ data la matrice densità ρ_S del sistema S e sia $|0\rangle_E\langle 0|$ lo stato dell'environment E. L'hamiltoniana avrà la forma $H_{SE} = H_S + H_E + H_{\mathrm{int}}$, dove H_{int} contiene i termini di accoppiamento tra sistema ed environment. Dunque $\rho_S \otimes |0\rangle_E\langle 0| \mapsto U_{SE}(t)\left(\rho_S \otimes |0\rangle_E\langle 0|\right) U_{SE}^\dagger(t)$; se $H_{\mathrm{int}} \equiv 0$, allora $U_{SE} = U_S U_E$. Poiché ci interessa solo la dinamica di S, occorre tracciare via l'environment, ottenendo la cosiddetta *rappresentazione di Stinespring*; ovvero, per un tempo t generico,

$$\begin{aligned}\rho_S(t) &= \mathrm{tr}_E\left[U_{SE}(t)\left(\rho_S \otimes |0\rangle_E\langle 0|\right) U_{SE}^\dagger(t)\right] \quad \text{(rappresentazione di Stinespring)} \\ &= \sum_{l=1}^{d_E} \left[{}_E\langle l|U_{SE}(t)|0\rangle_E\right] \rho_S \left[{}_E\langle 0|U_{SE}^\dagger(t)|l\rangle_E\right] \equiv \sum_{l=1}^{d_E} M_{l,S}(t)\rho_S M_{l,S}^\dagger(t), \\ &\qquad M_{l,S}(t) \stackrel{\mathrm{def}}{=} {}_E\langle l|U_{SE}(t)|0\rangle_E, \quad \sum_{l=1}^{d_E} M_{l,S}^\dagger(t) M_{l,S}(t) = \mathbb{1}_S.\end{aligned}$$

Gli operatori $M_{l,S}(t)$ sono detti *operatori di Kraus*. Condizione necessaria e sufficiente perché $\{M_{l,S}(t)\}_l$ sia un set di operatori di Kraus è che venga soddisfatta la condizione $\sum_l M_{l,S}^\dagger(t) M_{l,S}(t) = \mathbb{1}_S$; $\rho_S(t) = \sum_l M_{l,S}(t)\rho_S M_{l,S}^\dagger(t)$ è detta *rappresentazione di Kraus*.

Teorema. *Ogni superoperatore* Φ *rappresentabile in forma di Kraus come* $\Phi_S[\Theta](t) = \sum_{l=1}^{d_E} M_{l,S}(t)\Theta_S M_{l,S}^\dagger(t)$ *è esprimibile in rappresentazione di Stinespring.*

Dimostrazione. Sia $\Phi_S[\Theta](t) = \sum_{l=1}^{d_E} M_{l,S}(t)\Theta_S M_{l,S}^\dagger(t)$, con $M_{l,S}(t)$ operatori di Kraus. Consideriamo un environment fittizio di dimensione d_E e *definiamo* gli operatori U_{SE} mediante la loro azione su uno stato $|\psi\rangle_S \otimes |0\rangle_E$: $U_{SE}(t)\left(|\psi\rangle_S \otimes |0\rangle_E\right) =$

$\sum_{l=1}^{d_E} (M_{l,S}(t)|\psi\rangle_S) \otimes |l\rangle_E$. È immediato verificare che

$$\Phi_S\left[|\psi\rangle_S\langle\psi|\right] = \mathrm{tr}_E\left[U_{SE}\left(|\psi\rangle_S\langle\psi| \otimes |0\rangle_E\langle 0|\right) U_{SE}^\dagger\right]$$

Inoltre $U_{SE}^\dagger U_{SE} = \sum_l M_{l,S}^\dagger M_{l,S} = \mathbb{1}_S$. □

Osserviamo che esistono infinite rappresentazioni di Stinespring, potendo sostituire $|0\rangle_E \to V_E|0\rangle_E$, con V_E operatore unitario sull'environment.

La rappresentazione di Stinespring fornisce un criterio per scrivere l'evoluzione di osservabili Θ_S relative al sistema in esame in presenza dell'environment:

$$\Theta_S(t) = \mathrm{tr}_E\left[U_{SE}(t)\left(\Theta_S \otimes \mathbb{1}_E\right) U_{SE}^\dagger(t)\right].$$

In questo modo operatori positivi sono mappati in operatori positivi con medesima traccia, ovvero la mappa è *positiva* e *trace-preserving* e dunque *contrattiva*.

2.2. Mappe CPT

2.2.1. CPT e teorema di Kraus

La mappa $\rho_S(t) = \Phi[\rho_S(t)]$ introdotta nel paragrafo precedente, oltre ad essere positiva, è *completamente positiva*. Supponiamo infatti di avere un terzo sistema T *non interagente con l'environment* E, tale che S e T siano inizializzati nello stato ρ_{ST}. L'evoluzione del sistema $S+T$, sotto queste ipotesi, è $\rho_{ST}(t) = \mathrm{tr}_E\left[U_{SE}(t)\left(\rho_{ST} \otimes |0\rangle_E\langle 0|\right) U_{SE}^\dagger(t)\right]$: la nuova matrice $\rho_{ST}(t)$ deve godere delle proprietà di matrice densità $\forall t$; dunque la mappa deve conservare le proprietà di positività anche estendendone l'azione al sistema $S+T$, qualunque sia T non interagente con l'environment. In generale è possibile associare ad un sistema un environment quando si dispone di una mappa che sia CPT (*complete positive, trace preserving*), ovvero

- lineare;
- positiva;
- *trace preserving*;
- completamente positiva.

Osserviamo che esistono mappe positive non completamente positive. Se consideriamo ad esempio un qubit S, l'operatore $\Phi_S\colon \rho \mapsto \rho^T$ è positivo; tuttavia, accoppiando al qubit di partenza un secondo qubit T e considerando lo stato globale $|\Phi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, sia $\Phi_{ST} = \Phi_S \otimes \mathbb{1}_T$, si ha

$$|\Phi^+\rangle\langle\Phi^+| = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\Phi_{ST}} \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

che non è una matrice densità accettabile non essendo definita positiva.

Teorema (Rappresentazione di Kraus)**.** *Sia data la mappa* CPT $\Phi_S\colon \mathcal{L}(\mathcal{H}_S) \to \mathcal{L}(\mathcal{H}_S)$, $\mathcal{L}(\mathcal{H}_S)$ *spazio degli operatori lineari su* $\mathcal{H}_S$*. Allora esiste un set di operatori* $\{M_{l,S}\}_l$ *tale che* $\sum_l M_{l,S}^\dagger M_{l,S} = \mathbb{1}_S$ *e* $\sum_l M_{l,S}\rho M_{l,S}^\dagger = \Phi_S[\rho]$*; viceversa, se* $\Phi[\rho] = \sum_k M_k\rho M_k^\dagger$*, con* $\sum_k M_k^\dagger M_k = \mathbb{1}$*, allora* Φ *è una mappa* CPT*.*

Dimostrazione. Dato un sistema S ed un sistema ancillare A, detta $d = \dim\mathcal{H}_S$, sia $\dim\mathcal{H}_A = d$; essendo $\{|l\rangle_S\}_l$ e $\{|l\rangle_A\}_l$ due COS di $\mathcal{H}_S$ ed $\mathcal{H}_A$ rispettivamente, definiamo lo stato massimamente entanglato, detto *stato di Choi–Jamiołkowski*, $\frac{|\mathbb{1}\rangle_{SA}}{\sqrt{d}} = \frac{1}{\sqrt{d}}\sum_l |l\rangle_S \otimes |l\rangle_A$. Consideriamo ora $\Phi_S \otimes \mathbb{1}_A$ e due stati $|\psi\rangle_S = \sum_l \alpha_l|l\rangle_S \in \mathcal{H}_S$ e $|\psi^*\rangle_A = \sum_l \alpha_l^*|l\rangle_A \in \mathcal{H}_A$. Allora

$$ {}_A\langle\psi^*|\left(\Phi_S\otimes\mathbb{1}_A\right)\left(\frac{|\mathbb{1}\rangle_{SA}\langle\mathbb{1}|}{d}\right)|\psi^*\rangle_A = \sum_{ll'}\alpha_l\alpha_{l'}^*\frac{\Phi_S\left[|l\rangle_S\langle l'|\right]}{d} = \frac{1}{d}\Phi_S\left[|\psi\rangle_S\langle\psi|\right]. $$

Abbiamo così espresso l'azione della mappa in funzione della contrazione dell'azione della stessa su uno stato massimamente entanglato. In virtù della convessità del dominio e della linearità della mappa si generalizza la precedente agli stati non puri.

Essendo la mappa completamente positiva, l'operatore $\Phi_S\otimes\mathbb{1}_A$ è positivo. Inoltre $(\Phi_S\otimes\mathbb{1}_A)\left[\frac{|\mathbb{1}\rangle_{SA}\langle\mathbb{1}|}{d}\right] = \sum_{k=1}^{d^2}\lambda_k|k\rangle_{SA}\langle k|$, dove $\{|k\rangle_{SA}\}_k$ è un COS nello spazio $\mathcal{H}_S\otimes\mathcal{H}_A$. Da quanto detto sopra, $\Phi_S\left[|\psi\rangle_S\langle\psi|\right] = d\sum_{k=1}^{d^2}\lambda_{kA}\langle\psi^*|k\rangle_{SA}\langle k|\psi^*\rangle_A$. Se ora definiamo l'operatore lineare

$$ M_{k,S}\colon |\psi\rangle_S \mapsto \sqrt{d\lambda_k}\,{}_A\langle\psi^*|k\rangle_{SA} \Rightarrow \Phi_S\left[|\psi\rangle_S\langle\psi|\right] = \sum_{k=1}^{d^2} M_{k,S}|\psi\rangle_S\langle\psi|M_{k,S}^\dagger. $$

Per concludere la dimostrazione, osserviamo che la proprietà di *trace preserving* garantisce che $\sum_k M_{k,S}^\dagger M_{k,S} = \mathbb{1}_S$.

Per dimostrare l'implicazione inversa, sia $\Phi[\rho] = \sum_k M_k\rho M_k^\dagger$, con $\sum_k M_k^\dagger M_k = \mathbb{1}$. La mappa preserva la traccia ($\mathrm{tr}\left[\Phi[\rho]\right] = \mathrm{tr}\left[\sum_k M_k^\dagger M_k\rho\right] = \mathrm{tr}\left[\rho\right]$) ed è positiva ($\sum_k\langle\psi|M_k\rho M_k^\dagger|\psi\rangle \geq 0$). Inoltre è *completamente positiva*, essendo la sua azione su uno spazio ampliato

$$ {}_{ES}\langle\psi|(\mathbb{1}_E\otimes\Phi_S)[\rho_{ES}]|\psi\rangle_{ES} = {}_{ES}\langle\psi|(\mathbb{1}_E\otimes M_k)\rho_{ES}(\mathbb{1}_E\otimes M_k^\dagger)|\psi\rangle_{ES} \equiv \langle\phi|\rho_{ES}|\phi\rangle_{ES}, $$

dove $|\phi\rangle_{ES} \stackrel{\text{def}}{=} (\mathbb{1}_E\otimes M_k^\dagger)|\psi\rangle_{ES}$. □

2.2.2. Regole di composizione

Siano Φ_1 e Φ_2 due trasformazioni CPT; ogni combinazione convessa $\alpha\Phi_1 + (1-\alpha)\Phi_2$, $\alpha\in[0,1]$, è anch'essa CPT: infatti, utilizzando le rappresentazioni di Kraus per $\Phi_1 \to \{M_k\}$ e $\Phi_2 \to \{N_l\}$ per ciascuna mappa, si ha $\left[\alpha\Phi_1 + (1-\alpha)\Phi_2\right][\rho] = \alpha\sum_k M_k\rho M_k^\dagger + (1-\alpha)\sum_l N_l\rho N_l^\dagger$, da cui può essere ottenuta una nuova rappresentazione di Kraus in termini di $\{\sqrt{\alpha}M_k, \sqrt{1-\alpha}N_l\}$.

Le misture di trasformazioni CPT sono ancora trasformazioni CPT: una CPT nella forma

$$ \Phi[\rho] = \sum_i p_i U_i\rho U_i^\dagger, \tag{2.1} $$

con U_i unitario, può essere ottenuta adoperando, ad esempio, un dispositivo di misura imperfetto.

Se si ha una trasformazione CPT Φ, che gode della proprietà $\Phi[\mathbb{1}] = \mathbb{1}$, essa è detta *unitale*: tutte le mappe (2.1) sono unitali, ma non vale il viceversa (eccettuato il caso dei qubit, dove tutte le mappe unitali hanno la forma (2.1)).

Possiamo concatenare due trasformazioni CPT *in serie* scrivendo $(\Phi_1 \circ \Phi_2)[\rho] = \Phi_1[\Phi_2[\rho]]$. Da un punto di vista fisico, le trasformazioni composte così ottenute non possono non essere CPT, essendo sperimentalmente realizzabili. Occorre ricordare che in generale $\Phi_1 \circ \Phi_2 \neq \Phi_2 \circ \Phi_1$. Le trasformazioni CPT sono dunque un semigruppo rispetto all'operazione di composizione[1].

Le trasformazioni CPT sono dette anche *mappe* o *canali*. Date due mappe CPT Φ e Φ', esse sono dette *unitariamente equivalenti* se esistono due operatori unitari U, V, tali che $\Phi' = U \circ \Phi \circ V$.

2.2.3. Mappe separabili ed EBC

Le mappe CPT possono essere anche composte *in parallelo*: dati due sistemi A, su cui opera la trasformazione Φ, e B, su cui opera la trasformazione Ψ, è possibile definire $\Phi \otimes \Psi \colon \sigma(\mathcal{H}_A \otimes \mathcal{H}_B) \to \sigma(\mathcal{H}_A \otimes \mathcal{H}_B)$. Una trasformazione siffatta opera *localmente* su ciascuna delle due parti del sistema globale senza introdurre interazione. Una combinazione $\sum_i p_i \Phi_i \otimes \Psi_i$, con $\{p_i\}$ probabilità, è detta *separabile*: una trasformazione siffatta non può creare *entanglement*. Per meglio precisare tale affermazione, chiamiamo *separabile* uno stato ρ_{AB} se e solo se è possibile scrivere $\rho_{AB} = \sum_j p_j |\psi_j\rangle_A\langle\psi_j| \otimes |\phi_j\rangle_B\langle\phi_j|$; uno stato non separabile è detto *entanglato*. Osserviamo che lo spazio degli stati separabili è convesso ed è possibile ottenere stati separabili da stati separabili operando localmente. Tra le trasformazioni separabili esiste uno spazio convesso di operazioni, dette LOCC (*local operation and classical communication*, ovvero trasformazioni locali con comunicazione classica) che conservano la separabilità ma sono contenute propriamente nell'insieme delle trasformazioni separabili. I canali separabili ammettono una decomposizione di Kraus del tipo $\Phi[\rho_{AB}] = \sum_i M_{i,A} \otimes N_{i,B}\ \rho_{AB}\ {M_{i,A}}^\dagger \otimes {N_{i,B}}^\dagger$. Le trasformazioni precedenti, tuttavia, possono essere scritte come combinazioni convesse di trasformazioni locali (ovvero come LOCC) solo in alcuni casi.

Supponiamo ora che due sistemi A ed B siano entanglati e vediamo quali trasformazioni Φ_A su A permettono di rompere l'entanglement e quali lo preservano operando con $\Phi_A \otimes \mathbb{1}_B$. Si dice che Φ_A è un canale *entanglement-breaking* (EBC) se per un qualche B (che possiamo assumere isodimensionale ad A) $\Phi_A \otimes \mathbb{1}_B$ mappa uno stato entanglato in uno stato separabile. Si prova che tutti e soli gli EBC possono essere scritti nella *forma di Holevo*:

$$\Phi[\rho] = \sum_j \rho_j \operatorname{tr}[E_j \rho],$$

[1] In generale, l'inversa di una CPT non ha senso fisico. Si prova in particolare che solo le trasformazioni unitarie ammettono inverse che sono trasformazioni fisiche. Le misture non sono invertibili: ciò è plausibile se si ricorda che le trasformazioni CPT sono in generale contraenti, a meno che non siano unitarie.

dove $\{E_j\}_j$ è un set di POVM e ρ_j sono matrici densità (il processo di misura tramite EBC è detto *criptoclassico*). In generale un EBC non è invertibile; inoltre una combinazione seriale di EBC è ancora un EBC.

2.2.4. Evoluzione degli osservabili

In maniera analoga a quanto avviene con i sistemi isolati, è possibile far evolvere gli osservabili con una mappa $\Phi_H \colon \Theta \mapsto \Phi_H[\Theta]$, legata alla mappa di evoluzione per ρ nella maniera seguente:

$$\langle\Theta\rangle = \operatorname{tr}\left[\Theta\Phi[\rho]\right] = \operatorname{tr}\left[\Theta\sum_j M_j\rho M_j^\dagger\right] = \operatorname{tr}\left[\sum_j M_j^\dagger\Theta M_j\rho\right] \equiv \operatorname{tr}\left[\Phi_H[\Theta]\rho\right],$$
$$\Phi_H[\Theta] \stackrel{\text{def}}{=} \sum_j M_j^\dagger\Theta M_j.$$

In generale Φ_H è positiva ma non *trace preserving*, $\sum_j M_j M_j^\dagger \neq \mathbb{1}$.

2.3. Canali quantistici

Consideriamo in questo paragrafo mappe operanti su qubit, ovvero su sistemi con spazio di Hilbert bidimensionale la cui base ortonormale verrà indicata con $\{|0\rangle, |1\rangle\}$.

2.3.1. Amplitude damping channel

Sia $\rho = \left(\begin{smallmatrix} p & \gamma \\ \gamma^* & 1-p \end{smallmatrix}\right)$ lo stato di un qubit. Consideriamo l'azione di una mappa Φ_η tale che

$$\begin{pmatrix} p & \gamma \\ \gamma^* & 1-p \end{pmatrix} \xrightarrow{\Phi} \begin{pmatrix} p+(1-\eta)(1-p) & \sqrt{\eta}\gamma \\ \sqrt{\eta}\gamma^* & \eta(1-p) \end{pmatrix}, \quad \eta \in [0,1].$$

Come si vede, l'operatore altera le popolazioni e diminuisce le coerenze; in generale avremmo potuto alterare le coerenze con un fattore diverso $\sqrt{\eta'}$, purché $\eta' \leq \eta$. Essa è tale che $\Phi_{\eta=1}[\rho] = \rho$ e $\Phi_{\eta=0}[\rho] = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Una possibile rappresentazione di Kraus per l'operatore dato è

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\eta} \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & \sqrt{1-\eta} \\ 0 & 0 \end{pmatrix}.$$

È possibile costruire una rappresentazione di Stinespring pensando il sistema accoppiato con un environment E costituito da un altro qubit mediante una hamiltoniana $H_{SE} \equiv H_{\text{int}} = \sigma_S^+ \otimes \sigma_E^- + \sigma_S^- \otimes \sigma_E^+$, da cui $U = \mathrm{e}^{i\tau H_{SE}}$ per τ fissato[2]. Osserviamo infine che $\Phi_\eta \circ \Phi_{\eta'} = \Phi_{\eta\eta'}$.

[2] **Esercizio:** trovare la relazione tra τ ed η.

2.3.2. Dephasing channel, phase flip channel

Il *canale defasatore* è della forma

$$\rho = \begin{pmatrix} p & \gamma \\ \gamma^* & 1-p \end{pmatrix} \xrightarrow{\Phi} \begin{pmatrix} p & \sqrt{\eta}\gamma \\ \sqrt{\eta}\gamma^* & 1-p \end{pmatrix}, \quad \eta \in [0,1].$$

Come si vede la mappa, che è CPT, altera le coerenze come il precedente canale (*decoerenza*) ma lascia invariate le popolazioni. Per $\eta \to 0$ la sfera di Bloch si riduce progressivamente al segmento $[-1,1]$ dell'asse z. Poiché le popolazioni non cambiano, ciò può avvenire anche senza scambi di energia con l'esterno. La rappresentazione di Kraus di questo canale è

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\eta} \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{1-\eta} \end{pmatrix};$$

oppure è possibile scrivere la rappresentazione alternativa (*phase flip channel*) $\rho \to \alpha\rho + (1-\alpha)\sigma^3\rho\sigma^3$, con $\alpha = \frac{1+\sqrt{\eta}}{2} \in [0,1]$, ovvero $M_1 = \sqrt{\alpha}\mathbb{1}$ e $M_2 = \sqrt{1-\alpha}\sigma^3$ (il sistema rimane in se stesso con probabilità α, mentre è trasformato con probabilità $1-\alpha$). In rappresentazione di Stinespring $U = \mathrm{e}^{i\tau(|1\rangle_S\langle 1|\otimes\sigma^1_E)}$, dove l'environment E è costituito, come solito, da un qubit $|0\rangle_E$.

Per esemplificare il canale precedente, supponiamo che l'environment inizializzato nello stato $|E_0\rangle$ sia associato al nostro qubit e che si abbia la mappa

$$\begin{cases} |0\rangle \otimes |E_0\rangle \to |0\rangle \otimes |E_0\rangle \\ |1\rangle \otimes |E_0\rangle \to |1\rangle \otimes |E_1\rangle \end{cases},$$

dove $|E_1\rangle$ è un diverso stato dell'environment (ovvero, la lettura del qubit modifica l'environment). Dunque, se lo stato iniziale del qubit è $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, allora $|\psi\rangle\otimes|E_0\rangle \to \alpha|0\rangle\otimes|E_0\rangle + \beta|1\rangle\otimes|E_1\rangle$, $\rho = |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1| + \alpha\beta^*\langle E_0|E_1\rangle|0\rangle\langle 1| + \alpha^*\beta\langle E_1|E_0\rangle|1\rangle\langle 0|$, con $|\langle E_0|E_1\rangle| \leq 1$, per cui l'alterazione dell'environment induce sfasamento.

2.3.3. Bit-flip channel e phase-bit flip channel

Consideriamo un canale $\rho \to \alpha\rho + (1-\alpha)\sigma^1\rho\sigma^1$, $\alpha \in [0,1]$. Tale canale, detto *bit-flip channel*, è *unitariamente equivalente* al *phase-flip channel*. Se introduciamo infatti la base $|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$, σ^1 agisce su tale base come σ^3 sulla base $\{|0\rangle, |1\rangle\}$; dunque un cambio di base permette di ottenere una mappa dall'altra. In letteratura è studiata anche il *phase-bit flip channel*, $\rho \to \alpha\rho + (1-\alpha)\sigma^2\rho\sigma^2$, $\alpha \in [0,1]$: anche questa mappa è unitariamente equivalente alle precedenti.

2.3.4. Considerazioni finali sulle mappe di qubit

Nel caso di un qubit uno stato trasforma secondo la relazione generale

$$\rho = \frac{\mathbb{1} + \mathbf{a}\cdot\boldsymbol{\sigma}}{2} \mapsto \rho' = \frac{\mathbb{1} + \mathbf{a}'\cdot\boldsymbol{\sigma}}{2}, \quad \mathbf{a}' = T\mathbf{a} + \mathbf{b};$$

se la mappa è unitale, $\rho' = U\rho U^\dagger$, $U \in SU(2)$, T è una rotazione e $\mathbf{b} = \mathbf{0}$, ovvero la mappa è una rotazione nello spazio di Bloch. Se due canali sono unitariamente

equivalenti, allora essi hanno le stesse proprietà ed è utile studiare le classi di equivalenza. In particolare due mappe $(T, \mathbf{b})$ e $(T', \mathbf{b}')$ sono equivalenti se esistono due matrici $O_1, O_2 \in SO(3)$ tali che $(T', \mathbf{b}') = (O_2 T O_1, O_2 \mathbf{b})$. In virtù della singular value decomposition, esisterà una coppia di matrici ortogonali tali che $T \sim D$, con D matrice diagonale. I suoi autovalori specificano le proprietà della mappa se è unitale e si trovano nel tetraedro di vertici $(1,1,1), (1,1,-1), (1,-1,1), (-1,1,1)$.

2.4. Master equation

Cerchiamo un'equazione che descriva l'evoluzione di ρ in maniera continua nel tempo. Questa equazione sarà in generale non lineare ed avrà la forma di un'equazione di evoluzione $\dot{\rho} = F[\rho]$, dove F è un certo funzionale. Vogliamo in particolare determinare sotto quali condizioni F è un funzionale lineare in grado di generalizzare l'equazione $i\hbar\dot{\rho} = [H, \rho]$. Poniamo quindi

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] + \mathfrak{L}[\rho],$$

dove il funzionale $\mathfrak{L}[\cdot]$ è detto *forma di Lindblad*; esso è dato dall'espressione

$$\mathfrak{L}[\rho] = \sum_l \left(2L_l \rho L_l^\dagger - L_l^\dagger L_l \rho - \rho L_l^\dagger L_l \right),$$

dove gli operatori L_l sono detti *operatori di Lindblad* che descrivono processi dissipativi in cui l'environment ha dinamiche sufficientemente rapide da ridistribuire velocemente le informazioni senza che il sistema se ne riappropri; inoltre l'equazione può essere derivata a livello microscopico (specificando i tempi caratteristici del sistema). Proviamo che la precedente è l'unica forma ammessa perché la mappa sia CPT in una maniera piuttosto formale. Sia Φ_t una famiglia di mappe tali che $\Phi_t[\rho(0)] = \rho(t)$ ed inoltre

- $\Phi_{t=0} = \mathbb{1}$;
- $\mathrm{tr}\,[\Theta \Phi_t[\rho(0)]]$ è continua in t per ogni osservabile Θ;
- Φ_t forma un semigruppo;
- Φ_t sono CPT.

Si prova che la seconda condizione garantisce che esista il limite

$$\dot{\rho}(t) = \lim_{\epsilon \to 0} \frac{\rho(t+\epsilon) - \rho(t)}{\epsilon} = \lim_{\epsilon \to 0} \left(\frac{\Phi_\epsilon - \mathbb{1}}{\epsilon} \right) [\rho(t)] \equiv F[\rho(t)],$$

dove $F = \lim_{\epsilon \to 0} \frac{\Phi_\epsilon - \mathbb{1}}{\epsilon}$ è un operatore lineare.

Scriviamo ora Φ_t in rappresentazione di Kraus, $\Phi_t[\Theta] = \sum_k M_k(t) \Theta M_k^\dagger(t)$. Se la dimensione dello spazio di Hilbert $\mathcal{H}$ considerato è d, allora possiamo introdurre una base ortogonale nello spazio degli operatori lineari $\mathcal{L}(\mathcal{H})$ $\{E_j\}_{j=0,\dots d^2-1}$, con la condizione $E_0 = \mathbb{1}$ e tale che $\mathrm{tr}\,[E_i E_j] = 0$ se $i \neq j$. Allora $M_k(t) = \sum_{j=0}^{d^2-1} c_{kj}(t) E_j$, da cui

$$\Phi_t[\Theta] = \sum_{j,j'=0}^{d^2-1} \chi_{jj'}(t) E_j \Theta E_{j'}^\dagger, \quad \chi_{jj'}(t) \stackrel{\text{def}}{=} \sum_k c_{kj}(t) c_{kj'}^*(t).$$

La matrice $\chi(t)$ è hermitiana definita positiva $\forall t$. Utilizzando la precedente forma

$$F[\Theta] = \dot{\chi}_{00}(0)\Theta + \left(\sum_{j=1}^{d^2-1} \dot{\chi}_{j0}(0)E_j\right)\Theta + \Theta\sum_{j'=1}^{d^2-1} \dot{\chi}_{0j'}(0)E_{j'}^\dagger + \sum_{j,j'=1}^{d^2-1} \dot{\chi}_{jj'}(0)E_j\Theta E_{j'}^\dagger$$
$$= A\Theta + \Theta A^\dagger + \sum_{j,j'=1}^{d^2-1} \dot{\chi}_{jj'}(0)E_j\Theta E_{j'}^\dagger, \quad A \stackrel{\text{def}}{=} \frac{\dot{\chi}_{00}(0)}{2} + \sum_{j=1}^{d^2-1} \dot{\chi}_{j0}(0)E_j,$$

dove si è sfruttato il fatto che $\dot{\chi}_{0j}^*(0) = \dot{\chi}_{j0}(0)$. Dovendo Φ_t preservare la traccia, allora $\forall t \ \operatorname{tr}[\dot{\rho}_t] = \frac{\mathrm{d}\operatorname{tr}[\rho(t)]}{\mathrm{d}t} = 0$. Poiché le matrici densità sono una base lo spazio dei funzionali lineari, si ha che $\forall\Theta \in \mathcal{L}(\mathcal{H})$

$$0 = \operatorname{tr}\left[\left(A + A^\dagger + \sum_{j,j'=1}^{d^2-1} \dot{\chi}_{jj'}(0)E_{j'}^\dagger E_j\right)\Theta\right] \Rightarrow A + A^\dagger = -\sum_{j,j'=1}^{d^2-1} \dot{\chi}_{jj'}(0)E_{j'}^\dagger E_j.$$

Scrivendo $A = \frac{A+A^\dagger}{2} + \frac{A-A^\dagger}{2} = \frac{A+A^\dagger}{2} + \frac{iH}{\hbar}$, con $H = -i\hbar\frac{A-A^\dagger}{2}$ hermitiano, è possibile scrivere

$$F[\Theta] = -\frac{i}{\hbar}[H,\Theta] + \sum_{j,j'=1}^{d^2-1} \frac{\dot{\chi}_{jj'}(0)}{2}\left(2E_j\Theta E_{j'}^\dagger - E_{j'}^\dagger E_j\Theta - \Theta E_{j'}^\dagger E_j\right).$$

La matrice $\dot{\chi}_{jj'}(0)$ è hermitiana definita positiva (ricordiamo infatti che $\chi_{jj'}(0) = \delta_{j0}\delta_{j'0}$, per cui per $j,j' = 1,\dots,d^2-1$ si ottiene $\dot{\chi}_{jj'}(0) = \lim_{\epsilon\to 0}\frac{\chi_{jj'}(\epsilon)}{\epsilon} \geq 0$). Esiste perciò una matrice unitaria U che la diagonalizza, $\dot{\chi}(0) = UDU^\dagger$, D diagonale. Inserendo questa decomposizione si ottiene la forma di Lindblad. Si noti che $\operatorname{tr}[[H,\rho]] = \operatorname{tr}[\mathfrak{L}[\rho]] = 0$ e che il numero di operatori di Lindblad è al più d^2-1.

2.4.1. Esempio

Supponiamo di avere un sistema a due livelli $|0\rangle, |1\rangle$, dove $|0\rangle$ è il livello fondamentale, e sia

$$\dot{\rho}(t) = -i[H,\rho] + \gamma\left(2\sigma^-\rho\sigma^+ - \sigma^+\sigma^-\rho - \rho\sigma^+\sigma^-\right), \quad \hbar = 1,$$

dove $\gamma > 0$, $\sigma^+ = |1\rangle\langle 0|$ e $\sigma^- = |0\rangle\langle 1|$ ed $H = \frac{\omega}{2}(|1\rangle\langle 1| - |0\rangle\langle 0|)$ (equazioni ottiche di Bloch). Passando in rappresentazione di interazione $\tilde{\rho}(t) = U^\dagger\rho U$, $U = \mathrm{e}^{itH}$, si ottiene

$$\dot{\tilde{\rho}}(t) = \gamma\left(2\sigma^-\tilde{\rho}\sigma^+ - \sigma^+\sigma^-\tilde{\rho} - \tilde{\rho}\sigma^+\sigma^-\right)$$

ovvero, espandendo nella base canonica $\tilde{\rho} = p|0\rangle\langle 0| + c|0\rangle\langle 1| + c^*|1\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ e sostituendo nella master equation si ottiene un sistema di equazioni differenziali per coerenze e popolazioni:

$$\begin{cases}\dot{p} = -2\gamma p \\ \dot{c} = -\gamma c\end{cases} \Rightarrow \dot{\tilde{\rho}}(t) = \begin{pmatrix} p(0)\,\mathrm{e}^{-2\gamma t} & c(0)\,\mathrm{e}^{-\gamma t} \\ c^*(0)\,\mathrm{e}^{-\gamma t} & 1 - p(0)\,\mathrm{e}^{-2\gamma t}\end{pmatrix}$$

che è un *amplitude damping channel*.

Capitolo 3

Entanglement

3.1. Osservazioni generali

Abbiamo visto come l'entanglement sia caratterizzante dei sistemi quantistici. In generale gli stati (puri) di un sistema composto non sono infatti separabili. L'entanglement è inoltre una risorsa importante per la computazione quantistica e può essere utilizzato in diversi modi.

Sia dato un sistema bipartito $A+B$; abbiamo già chiarito che $|\psi\rangle_{AB}$ è separabile se esistono vettori $|\psi'\rangle_A$ e $|\psi''\rangle_B$ tali che $|\psi\rangle_{AB} = |\psi'\rangle_A \otimes |\psi''\rangle_B$; uno stato ρ_{AB} è separabile se è sovrapposizione convessa di stati puri separabili, da cui segue che lo spazio degli stati separabili è convesso. In generale, ricorrendo alla decomposizione di Schmidt, uno stato è separabile se ha rango di Schmidt unitario. Definiamo *stato massimamente entanglato* uno stato $|\psi\rangle_{AB}$ i cui coefficienti di Schmidt sono dati da $\lambda_i = \frac{1}{d}$, con $d = \max\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$ (supporremo in seguito che $\dim \mathcal{H}_A = \dim \mathcal{H}_B$). Esso ha di conseguenza una matrice densità $\rho_A = \operatorname{tr}_B [|\psi\rangle_{AB}\langle\psi|] = \frac{\mathbb{1}_A}{d}$. Poiché le trasformazioni locali (e le LOCC in particolare) cambiano localmente la base ma non alterano i coefficienti di Schmidt, esse non alterano l'entanglement.

Se i due sistemi sono costituiti da qubit, gli stati massimamente entanglati sono i cosiddetti *stati di Bell* $|\Psi^{\pm}\rangle = \frac{|0\rangle|1\rangle \pm |1\rangle|0\rangle}{\sqrt{2}}$ e $|\Phi^{\pm}\rangle = \frac{|0\rangle|0\rangle \pm |1\rangle|1\rangle}{\sqrt{2}}$ che formano peraltro una base ortonormale nello spazio $\mathcal{H}_A \otimes \mathcal{H}_B$; inoltre

$$|\Phi^-\rangle = (\mathbb{1}_A \otimes \sigma^1)|\Psi^-\rangle, \quad |\Psi^+\rangle = (\mathbb{1}_A \otimes \sigma^3)|\Psi^-\rangle, \quad |\Phi^+\rangle = (\mathbb{1}_A \otimes \sigma^2)|\Psi^-\rangle.$$

Su uno spazio siffatto introduciamo l'operazione di CNOT: tale operazione consiste nell'invertire lo stato del qubit B se e soltanto se il qubit A è nello stato $|1\rangle$. Una mappa siffatta può creare (e distruggere) entanglement: si può provare in generale che ogni trasformazione non scrivibile come prodotto di trasformazioni locali può creare e distruggere entanglement

3.2. Quantum Bell Telephone e paradosso EPR

Supponiamo che due osservatori, Alice e Bob, dispongano ciascuno di un qubit e che il sistema complessivo sia entanglato, $|\Psi^-\rangle_{AB} = \frac{|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B}{\sqrt{2}}$. Se Alice applica σ^3 sul suo qubit essa ottiene con probabilità $\frac{1}{2}$ ciascuno dei due possibili

risultati, inferendo automaticamente cosa otterrà Bob all'atto della misura. Questo tipo di correlazione non ha nessuna particolarità rispetto ad un comportamento classico. Tuttavia osserviamo che, detti $|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$, è possibile scrivere lo stato $|\Psi^-\rangle_{AB} = \frac{|+-\rangle_{AB} - |-+\rangle_{AB}}{\sqrt{2}}$. Inoltre gli stati $|\pm\rangle$ sono autostati di σ^1 con autovalori ± 1, per cui se Alice esegue una misura σ^1 ottiene $|+\rangle_A$ o $|-\rangle_A$ con probabilità $\frac{1}{2}$ e ciò gli permette come prima di inferire lo stato del qubit di Bob. Qui emerge la peculiarità dell'entanglement: lo scenario non viene alterato cambiando la base in cui si esegue la misura[1].

3.2.1. Quantum Bell Telephone

Supponiamo che Alice e Bob condividano inizialmente n qubit nello stato $|\Psi^-\rangle$; lo stato iniziale è dunque $(|\Psi^-\rangle)^{\otimes n}$. Si utilizza il seguente protocollo di misura: se Alice vuole comunicare l'informazione a allora misurerà mediante σ^3 tutti i qubit; se vorrà invece comunicare b misurerà con σ^1. Risulta che Bob, pur disponendo di n qubit, non potrà capire che tipo di misure sta eseguendo Alice, dato che lo stato da lui ottenuto fornirà (sia che Bob misuri con σ^1 sia che misuri con σ^3) la stessa statistica e dunque egli non potrà discriminare se la sua stringa è del tipo $|+\rangle|-\rangle|-\rangle|+\rangle\cdots$ o $|1\rangle|0\rangle|1\rangle|1\rangle\ldots$. Questo esempio mostra che non è sufficiente eseguire una serie di trasformazioni di natura esclusivamente locale, pur in presenza di entanglement.

3.2.2. Paradosso EPR (nella formulazione di Bohm)

Definiamo, seguendo Einstein, un *elemento di realtà* come una qualunque informazione ottenibile da un sistema a prescindere dall'esperimento che si esegue. Una teoria si dice *completa* se contiene tutti gli elementi di realtà. Infine richiediamo che non sia possibile avere influenze causali tra esperimenti con distanze di tipo spazio. Se Alice e Bob condividono uno stato entanglato $|\Psi^-\rangle_{AB}$, Alice può determinare lo stato del qubit di Bob (che sia $|\pm\rangle$ o $|1\rangle/|0\rangle$) eseguendo una misura σ^1 o σ^3 sul suo qubit: tale fatto fa sì che il risultato dell'operatore σ^1 o σ^3 sul qubit di Bob (a seconda della scelta di Alice) sia un elemento di realtà per Bob già prima della misura; tuttavia *non esiste* un autostato comune a σ^1 e σ^3. Dunque la scelta di Alice determina gli elementi di realtà presso Bob, a prescindere dalla distanza. Ciò sembrava suggerire che la Meccanica quantistica non fosse una teoria completa ed esistessero *variabili nascoste*.

3.2.3. Disuguaglianza di Bell CHSH

Siano ora λ (con una certa distribuzione $p(\lambda)$) le variabili nascoste in un sistema bipartito $A + B$ in cui $A(\lambda), A'(\lambda)$ sono osservabili del sottosistema A, $B(\lambda), B'(\lambda)$ sono osservabili del sottosistema B, $A(\lambda), A'(\lambda), B(\lambda), B'(\lambda) \in [-1, 1]$. Supponiamo che l'introduzione delle variabili nascoste permetta di descrivere completamente la

[1] Se Alice usasse il canale completamente defasante otterrebbe invece stati separabili distruggendo le correlazioni.

dinamica del sistema e dunque[2]

$$\begin{aligned}
&\langle AB\rangle \stackrel{\text{def}}{=} \int p(\lambda)A(\lambda)B(\lambda)\,\mathrm{d}\,\lambda \\
\Rightarrow &\left|\langle AB\rangle - \langle AB'\rangle\right| = \left|\langle AB(1 \pm A'B')\rangle - \langle AB'(1 \pm A'B)\rangle\right| \leq 2 \pm \left(\langle A'B\rangle + \langle A'B'\rangle\right) \\
&\Leftrightarrow \left|\langle AB\rangle - \langle AB'\rangle\right| + \left|\langle A'B\rangle + \langle A'B'\rangle\right| \leq 2 \\
&\Leftrightarrow \boxed{|\langle\mathcal{C}\rangle| \stackrel{\text{def}}{=} \left|\langle AB\rangle - \langle AB'\rangle + \langle A'B\rangle + \langle A'B'\rangle\right| \leq 2} \quad (3.1)
\end{aligned}$$

dove si è usato il fatto che $1 \pm \langle A'B\rangle$ e $1 \pm \langle A'B'\rangle$ sono positivi. La precedente *disuguaglianza di Bell*, detta *disuguaglianza CHSH*, può essere testata considerando due qubit e come operatori

$$A \mapsto \mathbf{a}\cdot\boldsymbol{\sigma}_A, \quad A' \mapsto \mathbf{a}'\cdot\boldsymbol{\sigma}_A, \quad B \mapsto \mathbf{b}\cdot\boldsymbol{\sigma}_B, \quad B' \mapsto \mathbf{b}'\cdot\boldsymbol{\sigma}_B, \quad |\mathbf{a}| = |\mathbf{a}'| = |\mathbf{b}| = |\mathbf{b}'| = 1,$$

da valutarsi sullo stato $|\Psi^-\rangle_{AB}$ definito sopra. Si ottiene in particolare che

$$_{AB}\langle\Psi^-|(\mathbf{a}\cdot\boldsymbol{\sigma}_A)\otimes(\mathbf{b}\cdot\boldsymbol{\sigma}_B)|\Psi^-\rangle_{AB} = -\mathbf{a}\cdot\mathbf{b}.$$

Se ad esempio si considera $\mathbf{a} \perp \mathbf{a}'$, $\mathbf{b} \perp \mathbf{b}'$ e $\arccos \mathbf{a}\cdot\mathbf{b} = \frac{\pi}{4}$ si ottiene una violazione della disuguaglianza (la quantità a destra della (3.1) è pari a $2\sqrt{2} > 2$). L'esperimento (1982) ha confermato la violazione.

È interessante osservare che gli stati separabili non violano le disuguaglianze, così come loro combinazioni convesse del tipo $\rho_{AB} = \sum_k p_k |\psi_k\rangle_A\langle\psi_k| \otimes |\phi_k\rangle_B\langle\phi_k|$: in effetti una forma del genere equivale all'assunzione di "variabili nascoste" con distribuzione $\{p_k\}$.

Segnaliamo infine una seconda diseguaglianza, detta *disuguaglianza di Tsirelson* in cui si prova che in meccanica quantistica vale sempre la relazione $\|\mathcal{C}\| \leq 2\sqrt{2}$ nell'ipotesi che $A^2 = A'^2 = B^2 = B'^2 = \mathbb{1}$ e $[A, B] = [A, B'] = [A', B] = [A', B'] = 0$. Così[3]

$$\mathcal{C}^2 = 4\mathbb{1} + [A, A'] \otimes [B, B'] \Rightarrow \left\|\mathcal{C}^2\right\| \leq 4 + \left\|[A, A'] \otimes [B, B']\right\| \leq 4 + 4\|A\|\|A'\|\|B\|\|B'\| = 8.$$

Essendo $\|\mathcal{C}^2\| = \|\mathcal{C}\|^2$, si ha la disuguaglianza cercata.

Esistono teorie che permettono di violare anche la disuguaglianza di Tsirelson e non permettono lo scambio di segnali a velocità superluminare.

3.2.4. GHZ

L'argomento GHZ è un'osservazione che permette di discriminare tra la meccanica quantistica ed una generica teoria che preveda l'esistenza di variabili nascoste. Lo stato GHZ consiste di tre qubit condivisi da Alice, Bob e Charlie nello stato $|\mathrm{GHZ}\rangle \stackrel{\text{def}}{=} \frac{|000\rangle + |111\rangle}{\sqrt{2}}$. Tracciando via il sottosistema C si ottiene $\mathrm{tr}_C\left[|\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|\right] = \frac{|00\rangle\langle 00|}{2} +$

[2] La dimostrazione che segue è valida anche considerando $\langle\cdot\rangle = \int \cdot\, p(\lambda, A, A', B, B')\,\mathrm{d}\,\lambda\,\mathrm{d}\,A\,\mathrm{d}\,A'\,\mathrm{d}\,B\,\mathrm{d}\,B'$.

[3] Ricordiamo che $\|\mathcal{C}\| = \sup_{\||\psi\rangle\|=1} \|\mathcal{C}|\psi\rangle\|$.

$\frac{|11\rangle\langle 11|}{2}$ che è uno stato separabile (ciò evidenzia quanto l'entanglement sia fragile). Inoltre

$$\sigma^1\otimes\sigma^2\otimes\sigma^2|\mathrm{GHZ}\rangle = -|\mathrm{GHZ}\rangle, \quad \sigma^2\otimes\sigma^1\otimes\sigma^2|\mathrm{GHZ}\rangle = -|\mathrm{GHZ}\rangle,$$
$$\sigma^2\otimes\sigma^2\otimes\sigma^1|\mathrm{GHZ}\rangle = -|\mathrm{GHZ}\rangle, \quad \sigma^1\otimes\sigma^1\otimes\sigma^1|\mathrm{GHZ}\rangle = |\mathrm{GHZ}\rangle.$$

Ciò significa che valgono le seguenti relazioni per le componenti (m_1, m_2, m_3) di ciascuno stato

$$m_1^A m_2^B m_2^C = -1, \quad m_2^A m_1^B m_2^C = -1, \quad m_2^A m_2^B m_1^C = -1, \quad m_1^A m_1^B m_1^C = 1,$$

che è evidentemente un set di equazioni incompatibili.

3.3. Criteri di separabilità

3.3.1. Criterio della trasposta parziale

Abbiamo già visto che l'operazione di trasposizione $T_B\colon \rho_B \mapsto \rho_B^T$ (rispetto ad una base fissata) è un'operazione positiva ma non completamente positiva, $\mathbb{1}_A\otimes T_B \not\geq 0$. Tuttavia, se ρ_{AB} è separabile, allora lo è anche lo stato trasformato. Dunque se per un certo $\tilde\rho_{AB}$ si ha $(\mathbb{1}_A\otimes T_B)[\tilde\rho_{AB}] \not\geq 0$ allora $\tilde\rho_{AB}$ è necessariamente entanglato (il criterio è dunque sufficiente ma non necessario).

Si dice che uno stato ρ_{AB} è PPT (*partial positive transpose*) se $(\mathbb{1}_A\otimes T_B)[\rho_{AB}] \geq 0$. Esistono stati entanglati PPT; se A e B sono qubit, gli stati separabili sono i soli stati PPT. Analogamente v'è coincidenza tra l'insieme degli stati PPT e l'insieme degli stati separabili se un sistema è un qubit e l'altro è un *qutrix*, ovvero un sistema a tre livelli. In tali casi il criterio della trasposta parziale è necessario e sufficiente. Se ρ_{AB} è entanglato ma PPT si parla di *bound entanglement*.

Stati di Werner

Sia $|\Phi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ ed introduciamo gli stati di Werner $\rho_W = p|\Phi^+\rangle\langle\Phi^+| + \frac{1-p}{4}\mathbb{1}\otimes\mathbb{1}$. Al variare di p è possibile passare da stati massimamente entanglati ($p=1$) a stati separabili ($p=0$). Adoperando il criterio sopra si vede che per $\frac{1}{3} < p < 1$ si ha entanglement. La fedeltà $F(\rho_W, |\Phi^+\rangle) = \langle\Phi^+|\rho_W|\Phi^+\rangle = \frac{1+3p}{4}$, ovvero $F > \frac{1}{2}$ per $p > \frac{1}{3}$, $F < \frac{1}{2}$ per $p < \frac{1}{3}$.

3.3.2. Criterio di riduzione

Introduciamo il superoperatore $\Lambda\colon \Theta \mapsto \mathrm{tr}\,[\Theta]\,\mathbb{1} - \Theta$. Tale mappa è positiva ma non completamente positiva. Sia $(\mathbb{1}_A\otimes\Lambda_B)[\rho_{AB}] = \rho_A\otimes\mathbb{1}_B - \rho_{AB}$: si prova che se $\rho_A\otimes\mathbb{1}_B - \rho_{AB} \not\geq 0$ lo stato ρ_{AB} è entanglato. Questo criterio (sufficiente ma non necessario) è invece necessario e sufficiente se si hanno due qubit o un qubit e un qutrix. Inoltre vale il seguente fondamentale teorema.

Teorema. *Lo stato ρ_{AB} è separabile se e solo se $\forall\Lambda_B > 0$ si ha $(\mathbb{1}_A\otimes\Lambda_B)[\rho_{AB}] \geq 0$.*

Corollario. *Se lo stato ρ_{AB} è entanglato, allora esiste un superoperatore Λ_B positivo tale che $(\mathbb{1}_A\otimes\Lambda_B)[\rho_{AB}] \not\geq 0$.*

3.3.3. Criterio di maggiorazione

Un altro criterio sufficiente ma non necessario è il cosiddetto *criterio di maggiorazione*. Consideriamo le distribuzioni classiche $\{p_j\}_j$ e $\{q_j\}_j$ ordinate in modo che per $i \leq j$ $p_i \geq p_j$ e $q_i \geq q_j$. Scriveremo che la distribuzione $\{q_j\}_j$ è *più ordinata* della distribuzione $\{p_j\}_j$, $\{p_j\}_j \prec \{q_j\}_j$, se $\forall k$ $\sum_{j=1}^{k} p_j \leq \sum_{j=1}^{k} q_j$. Ora, se $\rho_{AB} = \sum_j p_j |\psi_j\rangle_{AB}\langle\psi_j|$ e $\rho_A = \mathrm{tr}_B[\rho_{AB}] = \sum_j q_j |\phi_j\rangle_A\langle\phi_j|$ (è possibile aggiungere zeri all'insieme $\{q_j\}_j$ in modo che essa abbia lo stesso numero di elementi di $\{p_j\}_j$), si ha che se ρ_{AB} è separabile, allora $\{p_j\}_j \prec \{q_j\}_j$, per cui $\{p_j\}_j \not\prec \{q_j\}_j$ implica che ρ_{AB} è entanglato. Gli stati classici, dunque, sono più ordinati localmente di quanto non lo siano globalmente, mentre gli stati entanglati possono essere molto ordinati nel complesso ma disordinati localmente. Uno stato $|\Phi^+\rangle_{AB}\langle\Phi^+|$ viola il criterio di cui sopra.

3.4. Misure di entanglement

In generale, una buona misura di entanglement deve soddisfare quattro requisiti:

1. $\mathcal{E}(\rho) \geq 0$, $\mathcal{E}(\rho) = 0 \Leftrightarrow$ ρ separabile e $\mathcal{E}(|\psi_{\text{max. ent.}}\rangle) = \mathcal{E}_{\max}$.
2. $\mathcal{E}(\Phi_{\text{LOCC}}[\rho]) \leq \mathcal{E}(\rho)$;
3. $\mathcal{E}(\rho \otimes \sigma) \leq \mathcal{E}(\rho) + \mathcal{E}(\sigma)$;
4. $\mathcal{E}(\alpha\rho + (1-\alpha)\tau) \leq \alpha\mathcal{E}(\rho) + (1-\alpha)\mathcal{E}(\tau)$, $\alpha \in [0,1]$.

3.4.1. *Entanglement witness*

Sulla base della convessità dello spazio degli stati separabili, è possibile "tagliare" lo spazio degli stati in maniera da separare alcuni (e soli) stati non separabili. È possibile definire un'osservabile (per ciascun taglio) che assuma valori positivi da un lato e negativi dall'altro, con soli e tutti gli zeri sul taglio. Inoltre tale operatore Θ è detto *entanglement witness* se $\forall\rho_{AB}$ separabile si ha $\mathrm{tr}[\Theta\rho_{AB}] \geq 0$ ed esiste ρ_{AB} non separabile tale che $\mathrm{tr}[\Theta\rho_{AB}] < 0$ (le disuguaglianze di Bell sono un esempio di questo tipo di approccio).

3.4.2. *Entanglement distillation*

Supponiamo di avere due qubit, uno gestito da Alice e l'altro da Bob, entanglati tra loro, ma non necessariamente massimamente entanglati. Ci chiediamo se è possibile ottenere uno stato massimamente entanglato. Ciò è fattibile eseguendo una *distillazione di entanglement* purché si disponga di due copie del sistema, $\rho_{AB}^{\otimes 2}$, utilizzando l'entanglement di una copia per "aumentare" l'entanglement dell'altra. Se ρ_{AB} è uno stato di Werner $\rho_{AB} = p|\Phi^+\rangle\langle\Phi^+| + \frac{1-p}{4}\mathbb{1}$, $\frac{1}{3} < p < 1$, indichiamo con $A_1 + B_1$ e $A_2 + B_2$ le due coppie di qubit entanglate condivise da Alice e Bob. Alice e Bob possono eseguire un CNOT sul sistema $A_1 + A_2$ e $B_1 + B_2$ rispettivamente, utilizzando A_1, B_1 come bit di controllo. Fatto questo, occorre eseguire una misura proiettiva su A_1, B_1: se tale misura restituisce lo stesso risultato per i due bit, allora $F(\rho_{A_2B_2}, |\Phi^+\rangle)$ risulta aumentata; viceversa, se le due misure proiettive

restituiscono risultato opposto il protocollo è fallito. Iterando ancora si ha che, asintoticamente, si ottiene uno stato massimamente entanglato con probabilità 1. Partendo da N coppie, detto $m(N)$ il numero di qubit con *fidelity* maggiore di una certa $\bar{F}$ prefissata alla fine del processo, si ha che $\lim_N \frac{m(N)}{N} \stackrel{\text{def}}{=} \mathcal{E}_D(\rho_{AB})$, proprietà del solo stato di partenza, detto *entanglement di distillazione*. Se lo stato iniziale è separabile la definizione precedente restituisce correttamente zero, mentre si ha 1 se lo stato di partenza è già massimamente entanglato. Esistono tuttavia stati *non distillabili*, ovvero non separabili ma con entanglement di distillazione nullo. È stato provato che tali stati hanno come sottoinsieme gli stati non separabili PPT, ma non è noto se vale il viceversa.

3.4.3. *Entanglement cost*

Se Alice e Bob hanno N stati massimamente entanglati, ci si chiede se è possibile costruire uno stato definito ρ_{AB}. Se $m(N)$ è il numero massimo di copie di ρ_{AB} ottenibile da N stati massimamente entanglati, allora si definisce $\mathcal{E}_C(\rho_{AB}) = \sup_{\text{LOCC}} \lim_N \frac{N}{m(N)}$. Si prova che $\mathcal{E}_C(\rho_{AB}) \geq \mathcal{E}_D(\rho_{AB})$ (gli stati di tipo *bound entanglement* soddisfano la diseguaglianza stretta). Infine $\mathcal{E}_C(\rho_{AB}) \geq 0$ se e solo se ρ_{AB} è entanglato; se ρ_{AB} è uno stato puro, allora $\mathcal{E}_C(\rho_{AB}) = \mathcal{E}_D(\rho_{AB})$.

Definendo $\mathcal{E}_C$ ed $\mathcal{E}_D$ rispetto non alle trasformazioni LOCC ma alle trasformazioni *separabili* (ovvero che preservano la separabilità) si ha che $\bar{\mathcal{E}_C} = \bar{\mathcal{E}_D}$, dove abbiamo indicato con la barra il fatto che le definizioni sono intese rispetto alle trasformazioni separabili (Brandau, Plenio, 2008).

3.4.4. *Entanglement formation*

Supponiamo di avere uno stato puro $|\psi\rangle_{AB} = \sum_j \sqrt{\lambda_j}|\psi_j\rangle_A \otimes |\phi_j\rangle_B$. I valori λ_j hanno caratteristiche di distribuzione di probabilità e ciò può essere utilizzato per estrarre una qualche misura di entanglement. In particolare, più la distribuzione è piccata più lo stato è separabile. Si introduce così l'*entropia di Shannon* $H(\lambda_i) = -\sum_i \lambda_i \log_2 \lambda_i \in [0, \log_2 d]$ (la massima entropia si ha per $\lambda_i = \frac{1}{d}$). In questo modo uno stato separabile ha entropia nulla, mentre uno stato massimamente entanglato ha entropia massima. Se due matrici densità hanno lo stesso spettro, esse hanno dunque stessa entropia di Shannon. Se $\rho_A = \text{tr}_B\left[|\psi\rangle_{AB}\langle\psi|\right]$, si definisce $S(\rho_A) \stackrel{\text{def}}{=} H(\lambda_i)$ *entropia di von Neumann*. Si definisce *entanglement di formazione* $\mathcal{E}_F(\rho_{AB}) \stackrel{\text{def}}{=} S(\rho_A) = S(\rho_B)$.

Per uno stato puro l'entanglement di formazione è uguale al costo di entanglement. Per stati non puri, $\rho_{AB} = \sum_l q_l |\psi_l\rangle_{AB}\langle\psi_l|$, $\sum_l q_l = 1$, $q_l \in [0,1]$, essendo la decomposizione di ρ_{AB} non univoca, si definisce $\mathcal{E}_F = \inf_{q_l, |\psi_l\rangle_{AB}} \sum_l q_l \mathcal{E}_F(|\psi_l\rangle_{AB}\langle\psi_l|)$.

Osserviamo che in generale $\mathcal{E}_F(\rho_{AB}) \neq \mathcal{E}_C(\rho_{AB})$: è stato provato (Hasting, 2008) che[4] $\mathcal{E}_C(\rho_{AB}) = \lim_n \frac{\mathcal{E}_F(\rho_{AB}^{\otimes n})}{n}$.

[4] In generale $\mathcal{E}_F(\rho_{AB}^{\otimes 2}) \geq 2\mathcal{E}_F(\rho_{AB})$.

Esempio

Per una coppia di qubit si calcola l'entropia di formazione $\mathcal{E}_F(\rho_{AB})$ come

$$\mathcal{E}_F(\rho_{AB}) = \mathfrak{E}(C(\rho_{AB})), \quad \mathfrak{E}(x) \stackrel{\text{def}}{=} H_2\left(\frac{1+\sqrt{1-x^2}}{2}\right),$$
$$H_2(y) \stackrel{\text{def}}{=} -y\log_2 y - (1-y)\log_2(1-y), \quad C(\rho) \stackrel{\text{def}}{=} \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\},$$

dove la funzione C è detta *concurrence* e $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ sono gli autovalori dell'operatore

$$X(\rho) = \sqrt{\sqrt{\rho}\sigma^2 \otimes \sigma^2\rho^*\sigma^2 \otimes \sigma^2\sqrt{\rho}};$$

per uno stato puro $C(\rho_{AB}) = \sqrt{2(1 - \text{tr}\,[\rho_{AB}^2])}$.

Capitolo 4

Macchine quantistiche

4.1. Macchine impossibili

In meccanica quantistica esistono macchine impossibili da costruire ed inoltre esiste una gerarchia tra queste macchine, ovvero l'impossibilità di costruire alcune macchine determina l'impossibilità di costruire macchine più complesse e dalle diverse funzioni.

4.1.1. *Quantum cloner*

Un quantum cloner è una macchina capace di restituire come output due copie di uno stato ρ in ingresso, qualunque sia lo stato in ingresso. Una macchina del genere non è ammissibile in quanto l'operazione $\rho \mapsto \rho \otimes \rho$ *non è lineare*. Considerando ad esempio uno stato puro $|\psi\rangle_S$, vogliamo che l'operatore U_{SAE} esegua l'operazione (utilizzando la rappresentazione di Stinespring) $|\psi\rangle_S \otimes |0\rangle_A \otimes |E\rangle_E \mapsto |\psi\rangle_S \otimes |\psi\rangle_A \otimes |E_\psi\rangle_E$, dove $|E\rangle_E$ è l'environment e $|0\rangle_A$ è il sistema ancillare su cui si scrive; dovendo però U_{SAE} preservare i prodotti scalari, si ha che

$$\begin{cases} |\psi\rangle_S \otimes |0\rangle_A \otimes |E\rangle_E \mapsto |\psi\rangle_S \otimes |\psi\rangle_A \otimes |E_\psi\rangle_E \\ |\phi\rangle_S \otimes |0\rangle_A \otimes |E\rangle_E \mapsto |\phi\rangle_S \otimes |\phi\rangle_A \otimes |E_\phi\rangle_E \end{cases} \quad \to \langle\phi|\psi\rangle = |\langle\phi|\psi\rangle|^2 \langle E_\phi|E_\psi\rangle$$

che è un assurdo dovendo valere per *ogni* coppia di stati in ingresso. È però legittimo chiedersi quale livello di fedeltà $F(\rho, \rho') \geq 1 - \epsilon$ fra i due output ρ e ρ' può essere ottenuto: si prova che si può rendere ϵ tanto più piccolo quanto più grande è la dimensionalità del sistema in ingresso. Le macchine possibili che massimizzano la fedeltà sono dette *optimal quantum cloners*. Non potendo esistere un quantum cloner non è possibile (per le stesse ragioni) costruire un dispositivo che restituisca N copie da un singolo stato in ingresso. Se però si hanno $n \to +\infty$ stati in ingresso, si possono avere m stati in uscita (ricostruzione tomografica).

Osserviamo incidentalmente che l'esistenza di un quantum cloner permetterebbe di violare il principio di complementarità, potendo eseguire su ciascuna copia una misura arbitrariamente precisa di due variabili complementari (ad esempio posizione e impulso).

4.1.2. Teletrasporto classico

Un processo di teletrasporto classico consiste nella lettura in qualche modo di uno stato quantistico ρ che viene descritto da un set di dati inviati per mezzo di un canale classico ad una seconda macchina in grado di ricostruire lo stato. Questa macchina è impossibile in quanto l'informazione classica è sempre copiabile e dunque sarebbe possibile costruire, utilizzando più volte l'informazione classica, un quantum cloner.

4.1.3. Telefono di Bell

È possibile provare che vale la seguente catena di implicazioni:

telet. classico $\Rightarrow$ quantum cloning $\Rightarrow$ misura di oss. complementari $\Rightarrow$ telefono di Bell.

Per giustificare il precedente risultato, ricordiamo che nel caso del telefono di Bell era impossibile per Bob distinguere due stati $|\psi\rangle$ e $|\phi\rangle$ con sicurezza (ottenendo per i due casi statistiche identiche). Disponendo di un quantum cloner, tuttavia, Bob potrebbe copiare il sistema n volte. Se $|\psi\rangle$ e $|\phi\rangle$ sono due stati diversi, allora $(\langle\psi|)^{\otimes n}(|\phi\rangle)^{\otimes n} = (\langle\psi|\phi\rangle)^n \to 0$ per $n \to \infty$, permettendo asintoticamente di distinguere i due stati.

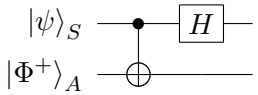
4.2. Macchine possibili e superdense coding

4.2.1. Teletrasporto quantistico

Il teletrasporto quantistico permette di trasferire un sistema quantistico mediante un canale quantistico (entanglement) ed un canale classico. Supponiamo che Alice e Bob dispongano di uno stato entanglato, tipo $|\Phi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$; sia lo stato che Alice vuol trasmettere del tipo $|\psi\rangle_S = \alpha|0\rangle_S + \beta|1\rangle_S$. Alice accoppia anzitutto il suo stato con il suo qubit A mediante la mappa U_{SA}:

$$U_{SA}|00\rangle_{SA} \mapsto \frac{|0\rangle_S+|1\rangle_S}{\sqrt{2}} \otimes |0\rangle_A, \quad U_{SA}|10\rangle_{SA} = \frac{|0\rangle_S-|1\rangle_S}{\sqrt{2}} \otimes |1\rangle_A,$$
$$U_{SA}|01\rangle_{SA} = \frac{|0\rangle_S+|1\rangle_S}{\sqrt{2}} \otimes |1\rangle_A, \quad U_{SA}|10\rangle_{SA} = \frac{|0\rangle_S-|1\rangle_S}{\sqrt{2}} \otimes |0\rangle_A,$$

che può essere scritta come una mappa CNOT seguita da una rotazione locale detta *trasformazione di Hadamard* H: $H|0\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, $H|1\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$:



Si ottiene dunque

$$|\psi\rangle_S \otimes |\Phi^+\rangle_{AB} = \alpha|0\rangle_S \frac{|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B}{\sqrt{2}} + \beta|1\rangle_S \frac{|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B}{\sqrt{2}} \xrightarrow{\text{CNOT}}$$
$$\alpha|0\rangle_S \frac{|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B}{\sqrt{2}} + \beta|1\rangle_S \frac{|1\rangle_A|0\rangle_B + |0\rangle_A|1\rangle_B}{\sqrt{2}} \xrightarrow{H}$$
$$\alpha \frac{|0\rangle_S + |1\rangle_S}{\sqrt{2}} \frac{|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B}{\sqrt{2}} + \beta \frac{|0\rangle_S - |1\rangle_S}{\sqrt{2}} \frac{|1\rangle_A|0\rangle_B + |0\rangle_A|1\rangle_B}{\sqrt{2}}.$$

A questo punto Alice esegue due misure sulla base canonica (mediante σ^3) su S ed A; possono verificarsi quattro possibili casi, indicati con $(a, b) =$ (misura su S, misura su A):

- $(0,0) \to |\phi\rangle_B = \alpha|0\rangle_B + \beta|1\rangle_B = |\psi\rangle_B$;
- $(0,1) \to |\phi\rangle_B = \alpha|1\rangle_B + \beta|0\rangle_B = \sigma^1|\psi\rangle_B$;
- $(1,0) \to |\phi\rangle_B = \alpha|0\rangle_B - \beta|1\rangle_B = \sigma^3|\psi\rangle_B$;
- $(1,1) \to |\phi\rangle_B = \alpha|1\rangle_B - \beta|0\rangle_B = -i\sigma^2|\psi\rangle_B$.

In ogni caso, dunque, Bob può recuperare lo stato $|\psi\rangle$ dopo che Alice gli avrà comunicato l'esito delle misure, in modo da comprendere esattamente che operatore applicare al suo stato. Notiamo che il teletrasporto dello stato da Alice a Bob comporta la distruzione dello stesso da parte di Alice, coerentemente col teorema *no-cloning*; la linea di comunicazione classica, inoltre, svolge un ruolo essenziale. Infine, la trasmissione è totalmente efficace se e solo se lo stato adoperato come ausiliario è massimamente entanglato, mentre negli altri casi la fidelity tra stato di partenza e stato trasmesso è minore di 1. Dunque due bit di informazione ed un E-bit (entanglement) permettono la trasmissione di un qubit.

4.2.2. *Superdense coding*

Il protocollo di *superdense coding* permette di inviare due bit partendo da un E-bit. Abbiamo visto che esistono quattro stati di Bell ottenibili l'uno dall'altro mediante rotazioni locali. Gli stati di Bell sono inoltre ortonormali e dunque perfettamente discriminabili da Bob. Alice può quindi scegliere di lasciare inalterato il suo qubit entanglato in uno stato di Bell con quello di Bob o eseguire una delle tre operazioni possibili per ottenere un diverso stato di Bell. Se Alice invia in seguito il suo qubit a Bob, egli potrà discriminare lo stato di Bell a sua disposizione mediante semplici misure proiettive. Dunque un totale di quattro possibili informazioni distinte (equivalenti a due bit) sono inviabili sfruttando l'entanglement.

Capitolo 5

Computazione quantistica

5.1. Macchina di Turing

5.1.1. Macchina di Turing

La *macchina di Turing*, ideata da Alan Turing nel 1930, consiste di un artefatto matematico in cui una stringa illimitata di simboli viene letta da un puntatore, collegato ad una testa (*head*) dotata di memoria interna finita (ovvero a numero di stati possibili finito, tra cui uno *stato di partenza* S ed uno *di arresto* H), capace di leggere, memorizzare, cancellare e scrivere. L'alfabeto di simboli da scrivere nelle caselle è finito ed uno dei simboli è sempre da associare alla *casella vuota*; un carattere speciale viene collocato nella prima casella del nastro ed associato allo stato S. Per far funzionare la macchina occorre un *programma* predefinito, ovvero una lista finita di istruzioni, ciascuna delle quali dipende dallo stato in cui si trova la macchina e dal valore letto sul nastro e restituisce un nuovo stato per la macchina ed un nuovo valore da scrivere sul nastro e il numero di passi $(0, -1, +1)$ di cui si deve muovere la testina sul nastro per l'avanzamento dell'operazione. Possiamo per semplicità limitarci a considerare un alfabeto binario.

Data ora una funzione $f\colon \{0,1\}^n \mapsto \{0,1\}^m$, occorre scrivere un *programma specifico* per f: se tale programma non è scrivibile o se non è possibile scrivere un programma che calcoli f in un numero finito di passi il problema è *indecidibile*. Un tipico problema indecidibile è il *problema della fermata* (*halting problem*): *non esiste una macchina di Turing capace, data una qualunque altra macchina di Turing ed un suo input, di stabilire se l'elaborazione terminerà o no.*

Congettura di Church–Turing. *La classe delle funzioni computabili tramite una macchina di Turing è uguale alla classe delle funzioni computabili tramite una qualunque procedura algoritmica.*

Una *macchina di Turing universale* (UTM) è una macchina di Turing programmabile. Una *probabilistic Turing machine* (PTM) è una macchina di Turing capace, all'occorrenza, di generare numeri random.

Congettura forte di Church–Turing. *Ogni modello computazionale può essere simulato da una PTM usando risorse extra con piccolo costo (ovvero con un costo di risorse che scala polinomialmente nella dimensione dell'input)*

NOT		AND		OR	
Input	Output	Input	Output	Input	Output
1	0	0 0	0	0 0	0
0	1	0 1	0	0 1	1
		1 0	0	1 0	1
		1 1	1	1 1	1

La teoria dell'informazione ha messo in crisi questa seconda congettura pur non contraddicendo la prima.

Un problema risolubile in un numero di passi polinomiale nella taglia dell'input è detto di classe **P**. Addizione e moltiplicazione sono esempi di problemi **P**. Si definisce problema **NP** un problema decidibile e tale che, dato l'input x e la soluzione $f(x)$, è possibile stabilire in un tempo polinomiale se la soluzione è corretta. Sicuramente $\mathbf{P} \subseteq \mathbf{NP}$ ma non è ancora noto se $\mathbf{P} = \mathbf{NP}$ o $\mathbf{P} \neq \mathbf{NP}$. Esiste una sottoclasse di problemi **NP** detti ***NP-completi***: se tali problemi sono risolvibili in tempi t, allora tutti i problemi **NP** sono risolubili in tempi polinomiali in t. In particolare, se i problemi **NP**-completi sono risolubili in un tempo polinomiale, $\mathbf{NP} = \mathbf{P}$.

Un problema risolubile (pur in tempi crescenti) con un numero di celle polinomiale è detto **PSPACE**. Si ha $\mathbf{P} \subseteq \mathbf{NP} \subseteq \mathbf{PSPACE}$.

Esiste anche la classe di complessità **BPP** (*bounded-error probabilistic polynomial*) relativa alle PTM e la sua equivalente quantistica **BQP** (*bounded-error quantistic polynomial*), $\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP}$.

5.1.2. *Universal gate set*

È possibile stabilire se un problema è di classe **P** o no valutando il numero di gate necessari per risolverlo. Si può dimostrare che l'analisi di un gate array è equivalente all'analisi di una macchina di Turing che risolve lo stesso problema. Esistono quattro gate fondamentali che formano il *set universale*: AND, OR, NOT, COPY. Il gate COPY duplica l'informazione di un bit, mentre gli altri operano come in tabella. Per costruire la generica funzione $F\colon \{0,1\}^m \to \{0,1\}^n$ vediamo come implementare una sua componente $f\colon \{0,1\}^m \to \{0,1\}$. Sia $S_0 = \{x \in \{0,1\}^m\colon f(x) = 0\}$ e $S_1 = \{x \in \{0,1\}^m\colon f(x) = 1\}$. Introduciamo la funzione $g_x(y) = \delta_{xy}\ \forall x \in \{0,1\}^m$. Allora, detto $S_1 = \{\tilde{x}_1, \ldots, \tilde{x}_d\}$ si ha $f(x) = g_{\tilde{x}_1}(x)$ OR $\cdots$ OR $g_{\tilde{x}_d}(x)$. I proiettori $g_x(y)$ possono essere costruiti secondo la regola

$$x = (1,1,\ldots,1) \Rightarrow g_x(y) = y_1 \text{ AND } y_2 \cdots \text{ AND } y_n$$
$$x = (0,1,\ldots,1) \Rightarrow g_x(y) = \text{NOT}(y_1) \text{ AND } y_2 \cdots \text{ AND } y_n \quad \ldots$$

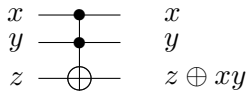
Ogni funzione può essere dunque riprodotta utilizzando AND, OR, NOT, COPY che è perciò un set universale. Inoltre è possibile ridurre il set ai soli operatori AND, NOT, COPY, essendo x OR y = NOT(NOT(x) AND NOT(y)). Introducendo il gate $\cdot$ NAND $\cdot \stackrel{\text{def}}{=}$ NOT($\cdot$ AND $\cdot$) è possibile utilizzare solo NAND, COPY per ottenere ogni altro gate. Infine, se è possibile inizializzare a piacimento lo stato y, si può utilizzare un unico gate, NAND&NOT, che ha come input x ed y e come output NOT(y) e x NAND y.

5.1.3. Principio di Landauer

Esiste un legame tra l'irreversibilità di alcuni dei gate precedenti e il secondo principio della termodinamica, stabilito dal seguente

Principio di Landauer. *Sia dato un computer classico in equilibrio termodinamico a temperatura T. Allora ogni trasformazione logicamente irreversibile produce un aumento dell'entropia $\Delta S \geq k_B \ln 2$ (equivalentemente $\Delta Q \geq k_B T \ln 2$).*

Si può provare che il precedente equivale al secondo principio della termodinamica. Si prova inoltre che l'unico processo logicamente irreversibile di natura fondamentale è la cancellazione di un bit. Inoltre non è possibile fare computazione reversibile utilizzando solo gate ad uno o due bit. Esiste un gate a tre bit *reversibile* detto *Toffoli gate*



in cui, indicando con $\oplus$ la somma binaria, si ha $(x, y, z) \mapsto (x, y, z \oplus xy)$. Ammettendo di poter inizializzare a piacere il gate di Toffoli, esso fornisce una porta universale (COPY si ottiene, ad esempio, ponendo $y = 1$ e $z = 0$). Un altro esempio di porta universale è la *porta di Fredkin*, in cui $(0, y, z) \mapsto (0, y, z)$ e $(1, y, z) \mapsto (1, z, y)$

5.2. Gate quantici

In computazione quantistica si sfrutta la natura quantistica di sistemi a due livelli per costruire memorie ed eseguire operazioni. Pertanto l'informazione è depositata in uno spazio di Hilbert $\mathcal{H} = \{|0\rangle, |1\rangle\}^{\otimes n}$ per n qubit. Le trasformazioni sui qubit si intendono in ogni caso unitarie, ovvero del tipo $U = \mathrm{e}^{-\frac{i}{\hbar}\int_0^t H(\tau)\,\mathrm{d}\,\tau}$.

5.2.1. Operazioni a un qubit

Le operazioni ad un qubit sono rappresentate da matrici di $SU(2)$ nella forma $U = \mathrm{e}^{i\alpha - \frac{i\theta}{2}\mathbf{n}\cdot\boldsymbol{\sigma}} = \mathrm{e}^{i\alpha}\left(\cos\frac{\theta}{2} - i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\theta}{2}\right)$, $\mathbf{n} \in \mathbb{R}^3$, $\|\mathbf{n}\| = 1$, $\theta \in [0, 4\pi]$ ed α fase globale. Il gate è completamente definito specificandone l'azione sulla base canonica. Una fondamentale operazione a un qubit è il *gate di Hadamard*, —$\boxed{H}$—, $H^2 = \mathbb{1}$, specificato dalla matrice

$$H = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \Rightarrow H|0\rangle = |+\rangle,\ H|1\rangle = |-\rangle.$$

Il *phase gate* si limita ad introdurre una fase relativa

$$P(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{i\phi} \end{pmatrix}.$$

Ricordiamo che, detto $\mathbf{S} = \frac{\boldsymbol{\sigma}}{2}$, $[S_i, S_j] = \epsilon_{ijk}S_k$; usando la decomposizione di Eulero $U = R_{\mathbf{n}}(\theta) = R_z(\alpha)R_y(\beta)R_z(\gamma)$, dove $R_z(\theta) = \begin{pmatrix} \mathrm{e}^{-i\frac{\theta}{2}} & 0 \\ 0 & \mathrm{e}^{i\frac{\theta}{2}} \end{pmatrix}$ è una rotazione attorno

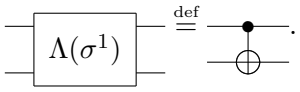
all'asse z e $R_y(\theta) = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$ è una rotazione attorno all'asse y, otteniamo così la più generica rotazione nella forma

$$R_z(\alpha)R_y(\beta)R_z(\gamma) = \begin{pmatrix} \cos\frac{\beta}{2}\,\mathrm{e}^{-i\frac{\alpha+\gamma}{2}} & -\sin\frac{\beta}{2}\,\mathrm{e}^{-i\frac{\alpha-\gamma}{2}} \\ \sin\frac{\beta}{2}\,\mathrm{e}^{i\frac{\alpha-\gamma}{2}} & \cos\frac{\beta}{2}\,\mathrm{e}^{i\frac{\alpha+\gamma}{2}} \end{pmatrix},$$

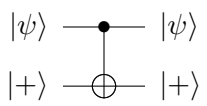
matrice di vettori ortogonali. Ponendo $A \stackrel{\text{def}}{=} R_z(\alpha)R_y\left(\frac{\beta}{2}\right)$, $B \stackrel{\text{def}}{=} R_y\left(-\frac{\beta}{2}\right)R_z\left(-\frac{\alpha+\gamma}{2}\right)$, $C \stackrel{\text{def}}{=} R_z\left(-\frac{\alpha-\gamma}{2}\right)$ è possibile scrivere la generica rotazione come $U = A\sigma^1 B\sigma^1 C$, $ABC = \mathbb{1}$, dove si è usato il fatto che $\sigma^1\,\mathrm{e}^{-i\frac{\theta}{2}\sigma^2}\,\sigma^1 = \mathrm{e}^{-i\frac{\theta}{2}\sigma^1\sigma^2\sigma^1} = \mathrm{e}^{i\frac{\theta}{2}\sigma^2}$.

5.2.2. Operazioni a due qubit

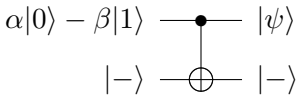
Le operazioni a due qubit sono rappresentate da matrici $U \in SU(4)$. La più importante operazione a due qubit è il gate CNOT,



Esso è tale che $\Lambda(\sigma^1)|00\rangle = |00\rangle$, $\Lambda(\sigma^1)|01\rangle = |01\rangle$, $\Lambda(\sigma^1)|10\rangle = |11\rangle$, $\Lambda(\sigma^1)|11\rangle = |10\rangle$. Inoltre osserviamo che, detto $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

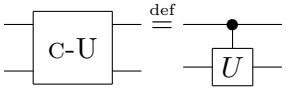


mentre

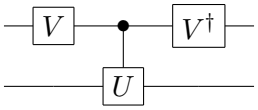


Questo evidenzia il fatto che non esiste unidirezionalità nell'accoppiamento e si può costruire un sistema in cui è il secondo bit a determinare le alterazioni del primo.

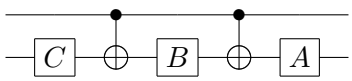
Il C-U (*control* U) consiste in un gate di controllo che applica U (1-qubit gate) ad uno dei due qubit:



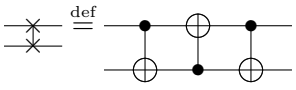
Per usare come base di controllo una base diversa da quella canonica basta alterare il gate introducendo un cambio di base V



Si può provare che un CNOT e possibili cambi di base come il precedente permettono qualsiasi C-U; infatti, detto $U = A\sigma^1 B\sigma^1 C$ è possibile scrivere un generico C-U:

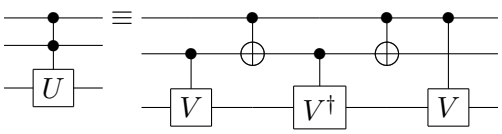


Analogamente uno *swap* può essere ottenuto come



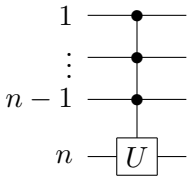
5.2.3. Operazioni ad n qubit ed universalità

Proviamo ora che ogni trasformazione $U \in SU(2^n)$ può essere ottenuta con CNOT e operazioni ad un qubit. Osserviamo preliminarmente che la seguente *operazione a 3 qubit* può essere scritta come segue:

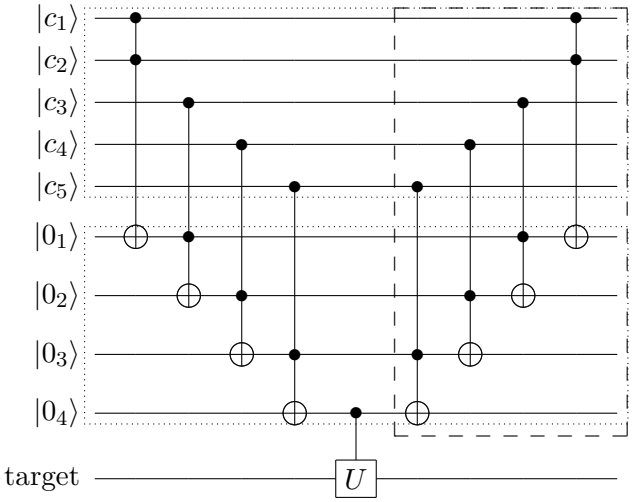


dove $V \stackrel{\text{def}}{=} \sqrt{U}$. Poiché il gate di Toffoli è realizzabile con operazioni a due bit[1] ed è il generatore della computazione reversibile classica, si può produrre tutta la computazione reversibile classica utilizzando solo gate a due bit.

Generalizziamo ora il gate di Toffoli e supponiamo di avere un'operazione ad n qubit in cui $n-1$ qubit controllano il qubit finale, come

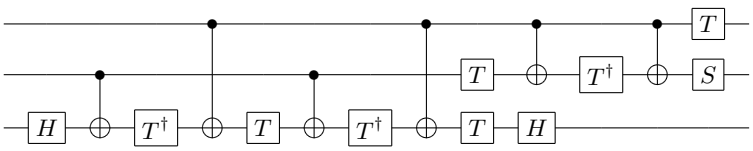


L'aggiunta di $n-2$ bit ausiliari permette di riprodurre il precedente gate con operazioni a tre qubit:



L'ultima parte del gate (nel box tratteggiato) è necessaria per eseguire l'*undo* sulle ancille, inizializzate tutte in $|0\rangle$. La spesa della precedente operazione è di $n-2$

[1] Infatti per il gate di Toffoli



dove $T \stackrel{\text{def}}{=} P\left(\frac{\pi}{4}\right)$ e $S \stackrel{\text{def}}{=} P\left(\frac{\pi}{2}\right)$.

ancille, $12(n-2)+2$ CNOT (ogni gate di Toffoli richiede 6 CNOT) e $\sim n$ operazioni ad un qubit; il costo computazionale dunque scala come n. Il precedente gate ad n qubit è banale su tutto lo spazio di Hilbert fuorché sul sottospazio $\mathrm{span}\{|1\cdots1\rangle|0\rangle, |1\cdots1\rangle|1\rangle\}$.

Data ora una trasformazione generica U sullo spazio $\mathcal{H}$ 2^n-dimensionale che descrive n qubit, essa sarà rappresentata da una matrice $m \times m$, $m = 2^n$. Allora si può scrivere $U = (u_{ij})_{ij} = W_k W_{k-1} \cdots W_1$, dove W_j operano solo su sottospazi bidimensionali e $k \sim 2^{2n}$. Se normalizziamo in modo tale che i vettori $\binom{u^*_{11}}{u_{21}}, \binom{u^*_{21}}{-u_{11}} \mapsto \binom{{u^{*(1)}}_{11}}{{u^{(1)}}_{21}}, \binom{{u^{*(1)}}_{21}}{-{u^{(1)}}_{11}}$ abbiano norma unitaria, la matrice unitaria

$$W_1^\dagger \stackrel{\text{def}}{=} \begin{pmatrix} {u^{*(1)}}_{11} & {u^{*(1)}}_{21} & \\ {u^{(1)}}_{21} & -{u^{(1)}}_{11} & \\ & & \mathbb{1}_{m-2} \end{pmatrix}$$

è tale che $U^{(2)} \stackrel{\text{def}}{=} W_1^\dagger U$ ha l'elemento nella posizione $(2,1)$ nullo. Definiamo ora

$$W_2^\dagger \stackrel{\text{def}}{=} \begin{pmatrix} {u^{*(2)}}_{11} & 0 & {u^{*(2)}}_{31} & \\ 0 & 1 & 0 & \\ {u^{(2)}}_{31} & 0 & -{u^{(2)}}_{11} & \\ & & & \mathbb{1}_{m-3} \end{pmatrix}$$

dove gli elementi non banali sono ottenuti dagli elementi di $U^{(2)}$ normalizzando in modo che W_2 sia unitaria. Così $W_2^\dagger U^{(2)} = W_2^\dagger W_1^\dagger U$ avrà termini nulli in $(2,1)$ e $(3,1)$. Continuando a procedere similmente, utilizzando matrici unitarie del tipo

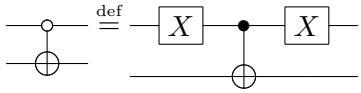
$$W_k^\dagger \stackrel{\text{def}}{=} \begin{pmatrix} {u^{*(k)}}_{11} & & {u^{*(k)}}_{k+1,1} & \\ & \mathbb{1}_{k-1} & & \\ {u^{(k)}}_{k+1,1} & & -{u^{(k)}}_{11} & \\ & & & \mathbb{1}_{m-k-1} \end{pmatrix}$$

si ha che gli elementi $(j,1)$, $j = 2, \ldots, m$, della matrice $W_m^\dagger W_{m-1}^\dagger \cdots W_1^\dagger U = W_0$ sono nulli per costruzione. Inoltre, l'elemento $(1,1)$ è del tipo $\mathrm{e}^{i\phi_1}$ essendo U unitario e dunque $W_0^\dagger W_0 = \mathbb{1}_m$. Infine dal fatto che $W_0 W_0^\dagger = \mathbb{1}_m$ si ha che la forma di W_0 è

$$W_0 = \begin{pmatrix} \mathrm{e}^{i\phi_1} & \\ & \tilde{W}_0 \end{pmatrix}.$$

È possibile a questo punto ripetere il procedimento sulla matrice $\tilde{W}_0$, in modo da ottenere, dopo $\sum_{j=2}^{m}(j-1) \sim m^2 = 4^n$ passaggi la sua riduzione alla forma diagonale $(\delta_{ij}\,\mathrm{e}^{i\phi_i})_{ij}$, corrispondente ad un insieme di *phase gate* di singolo qubit. La generica matrice W_k accoppia invece gli stati i e j corrispondenti agli indici dei termini non nulli fuori diagonale. Poiché tramite l'applicazione di NOT è possibile passare, invertendo un qubit alla volta, da uno stato ad un altro qualsiasi in un numero di passi al più pari al numero di elementi n della stringa che individua lo stato, possiamo scrivere $W_k = G_k^\dagger \Lambda(U_k) G_k$, dove G_k è detto *codice di Gray*. Perciò l'azione di un generico gate ad n qubit può essere ottenuto trasformando tramite il codice di Gray i due stati interessati in modo da farli differire di un solo qubit,

applicare un C-U sull'unico qubit diverso (condizionatamente all'eguaglianza di tutti gli altri bit) e quindi eseguire l'*undo* della sequenza di operazioni del codice di Gray. In questo tipo di circuito si usa spesso il gate



in cui l'inversione del secondo qubit avviene se il primo è nullo. Inoltre si indica con —[X]— il gate ottenuto applicando σ^1, con —[Y]— il gate ottenuto applicando σ^2 e con —[Z]— il gate ottenuto applicando σ^3. Dunque in totale un gate su n qubit può essere implementato utilizzando $\sim 4^n \text{poly}(n)$ gate.

5.2.4. Approssimazione di un generico circuito

Abbiamo visto dunque che qualunque gate può essere ottenuto combinando CNOT e operazioni ad un qubit. Poiché però quest'ultime sono infinite, ha senso chiedersi se è possibile approssimare un generico circuito usando un numero limitato di gate. In particolare si prova che è possibile approssimare un circuito generico utilizzando CNOT, gate di Hadamard e $T \equiv P\left(\frac{\pi}{4}\right)$ con approssimazione tanto migliore quanto più la rete è complessa. Se l'azione del circuito è espressa dalla matrice U e V è la matrice unitaria approssimante costruita con i tre gate anzidetti, sia $\|U - V\| < \epsilon$ e siano $\{E_k\}$ delle POVM assegnate. Prese $p_k(U) \stackrel{\text{def}}{=} \text{tr}\left[E_k U \rho U^\dagger\right]$ e $p_k(V) \stackrel{\text{def}}{=} \text{tr}\left[E_k V \rho V^\dagger\right]$, $|p_k(U) - p_k(V)| \xrightarrow{\epsilon \to 0} 0$. Infatti, se lo stato iniziale è uno stato puro[2] $\rho = |\psi\rangle\langle\psi|$

$$|p_k(U) - p_k(V)| = \left|\text{tr}\left[\rho\left(U^\dagger E_k U - V^\dagger E_k V\right)\right]\right| \leq \sup_{\||\psi\rangle\|=1} \left|\langle\psi|U^\dagger E_k U - V^\dagger E_k V|\psi\rangle\right|$$

$$\leq \sup_{\||\psi\rangle\|=1} \left[\left|\langle\psi|U^\dagger E_k (U - V)|\psi\rangle\right| + \left|\langle\psi|(U^\dagger - V^\dagger) E_k V|\psi\rangle\right|\right] \leq \sup_{\||\psi\rangle\|=1} 2\||\omega\rangle\| = 2\epsilon$$

avendo definito $|\omega\rangle = (U - V)|\psi\rangle$. Perciò in definitiva $|p_k(U) - p_k(V)| \leq \|U - V\|$.

5.3. Modelli di computazione quantistica

I risultati precedenti possono essere utilizzati per implementare la computazione con diversi modelli. Una *macchina di Turing quantistica*, ad esempio, è una macchina di Turing in cui la stringa diventa una successione di qubit con associato un registro quantistico, mentre la macchina esegue trasformazioni unitarie come operazioni.

Un altro tipo di computazione adoperato è la *computazione quantistica adiabatica* (*adiabatic quantum computation*). In questo contesto si sfrutta il *teorema adiabatico*: il registro è composto da un insieme di sistemi interagenti la cui evoluzione è determinata da un operatore $H(t)$ dipendente esplicitamente da alcuni parametri di controllo di modo che si possa porre il sistema nell'istante iniziale nello stato fondamentale e che, al passare del tempo, sia possibile variare i parametri in modo da mantenere il sistema nello stato fondamentale. Il teorema adiabatico garantisce

[2] Analogamente per convessità si prova per stati misti.

che ciò è possibile purché non ci siano stati eccitati i cui autovalori incrocino l'autovalore corrispondente allo stato fondamentale e purché la trasformazione sia lenta. In questo modo un'eventuale interferenza con l'esterno diventa ininfluente.

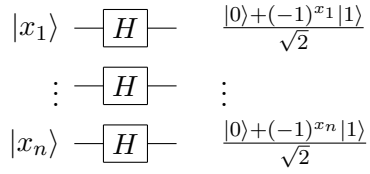
Nel caso del *cluster state quantum computation* la computazione introduce correlazioni fra i vari elementi del sistema partendo da uno stato iniziale fortemente entanglato (*cluster state*). A operazione ultimata, le misurazioni distruggono le correlazioni e i qubit superstiti conterranno l'informazione voluta.

Se non si è in grado di costruire forti interazioni fra i vari qubit, è possibile utilizzare gate *probabilistici* in grado di ottenere l'informazione sacrificando eventualmente più qubit (tramite la misurazione: si parla di *linear optics quantum computation*).

5.4. Algoritmi quantistici

5.4.1. Algoritmo di Hadamard

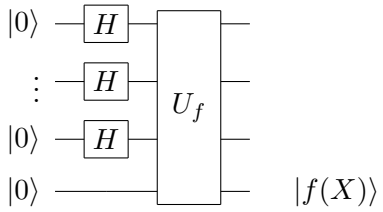
Il gate di Hadamard ha la forma $H|x\rangle = \frac{|0\rangle+(-1)^x|1\rangle}{\sqrt{2}}$, $x = 0, 1$. Consideriamo ora n qubit ed applichiamo a ciascuno un gate di Hadamard:



per cui $|X\rangle = |x_1\rangle\otimes\cdots\otimes|x_n\rangle \xrightarrow{H} \left(\frac{|0\rangle+(-1)^{x_1}|1\rangle}{\sqrt{2}}\right)\otimes\cdots\otimes\left(\frac{|0\rangle+(-1)^{x_n}|1\rangle}{\sqrt{2}}\right) = \frac{1}{2^{\frac{n}{2}}}\sum_Y(-1)^{X\cdot Y}|Y\rangle$, dove Y è una sequenza ordinata di simboli binari data dagli stati finali. Lo stato finale è ovviamente separabile. Se $|X\rangle = |\mathbf{0}\rangle \stackrel{\text{def}}{=} |0\rangle\otimes\cdots\otimes|0\rangle$, allora $H^{\otimes n}|\mathbf{0}\rangle = \frac{1}{2^{\frac{n}{2}}}\sum_Y|Y\rangle$ che è *una sovrapposizione coerente degli stati possibili.*

Quantum parallelism

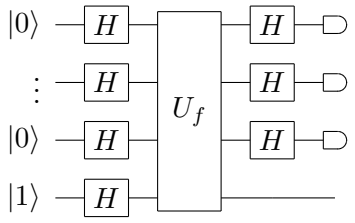
Sia ad esempio $f\colon \{0,1\}^n \to \{0,1\}$ una funzione assegnata e supponiamo di avere un oracolo che restituisca $|X\rangle \mapsto |f(X)\rangle$. Ciò non può essere implementato da un operatore unitario a meno di introdurre un'ancella in modo che $U_f\colon |X\rangle|y\rangle \mapsto |X\rangle|y \oplus f(X)\rangle$. La meccanica quantistica permette di calcolare, in una sola operazione, *tutti gli output* e selezionarne in seguito uno, utilizzando una serie di Hadamard:



Si ottiene dunque $|\mathbf{0}\rangle|0\rangle \xrightarrow{H} \frac{1}{2^{\frac{n}{2}}}\sum_X|X\rangle|0\rangle \xrightarrow{U_f} \frac{1}{2^{\frac{n}{2}}}\sum_X|X\rangle|f(X)\rangle$, che è una sovrapposizione coerente di tutti gli output. Osserviamo però che la misura del vettore $|X\rangle$ determina automaticamente un solo valore $|f(X)\rangle$ e non ci sono vantaggi rispetto al caso classico.

5.4.2. Algoritmo di Deutsch–Josza

Supponiamo di essere nel medesimo caso presentato sopra ma con l'informazione aggiuntiva che f definita su $\{0,1\}^n$ sia costante o *bilanciata* (ovvero f assume valore 0 su metà degli input, 1 sull'altra metà). Nel caso classico occorrono al più $2^{n-1}+1$ prove (ovvero occorre testare metà delle combinazioni possibili più una: se si accetta un approccio probabilistico, la complessità del problema è polinomiale). Nel caso quantistico, invece, basta una sola misura. Consideriamo il gate siffatto:



Esso corrisponde alla trasformazione

$$|\mathbf{0}\rangle|1\rangle \xrightarrow{H} \frac{1}{2^{\frac{n+1}{2}}}\sum_X |X\rangle\otimes(|0\rangle-|1\rangle)$$
$$\xrightarrow{U_f} \frac{1}{2^{\frac{n+1}{2}}}\sum_X |X\rangle\otimes(|f(X)\rangle-|1\oplus f(X)\rangle) = \frac{1}{2^{\frac{n+1}{2}}}\sum_X (-1)^{f(X)}|X\rangle\otimes(|0\rangle-|1\rangle)$$
$$\xrightarrow{H} \frac{1}{2^n}\left(\sum_X\sum_Y (-1)^{f(X)\oplus X\cdot Y}|Y\rangle\right)\otimes\frac{|0\rangle-|1\rangle}{\sqrt{2}}.$$

La probabilità di trovare $P(|Y\rangle=|\mathbf{0}\rangle) = \left|\frac{1}{2^n}\sum_X(-1)^{f(X)}\right|^2 = \begin{cases}1 & \text{se } f \text{ è costante}\\ 0 & \text{se } f \text{ è bilanciata}\end{cases}$.

5.4.3. Algoritmo di Bernstein–Vazirani

Sia ora $f_{\mathbf{a}}(X) = \bigoplus_{i=1}^n a_i x_i$; occorre trovare $\mathbf{a} = (a_1,\dots,a_n)$. Classicamente occorre testare gli n vettori di base ma, utilizzando lo stesso schema dell'algoritmo di Deutsch–Josza si ottiene in uscita

$$\left(\frac{1}{2^n}\sum_X\sum_Y(-1)^{X\cdot(\mathbf{a}\oplus Y)}|Y\rangle\right)\otimes\frac{|0\rangle-|1\rangle}{\sqrt{2}} = |\mathbf{a}\rangle\otimes\frac{|0\rangle-|1\rangle}{\sqrt{2}},$$

essendo $\sum_X(-1)^{X\cdot(Y\oplus\mathbf{a})} = 2^n\delta_{\mathbf{a},Y}$. Dunque l'output restituisce immediatamente il vettore cercato.

5.5. Quantum Fourier Transform

5.5.1. La trasformata di Fourier quantistica

La trasformata di Fourier quantistica (QFT) è una mappa $U_F\colon \{0,1\}^n \to \{0,1\}^n$ tale che, detto $N = 2^n$,

$$|x\rangle \mapsto U_F|x\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{N}}\sum_{y=0}^{N-1} \mathrm{e}^{i\frac{2\pi}{N}x\cdot y}\,|y\rangle,$$

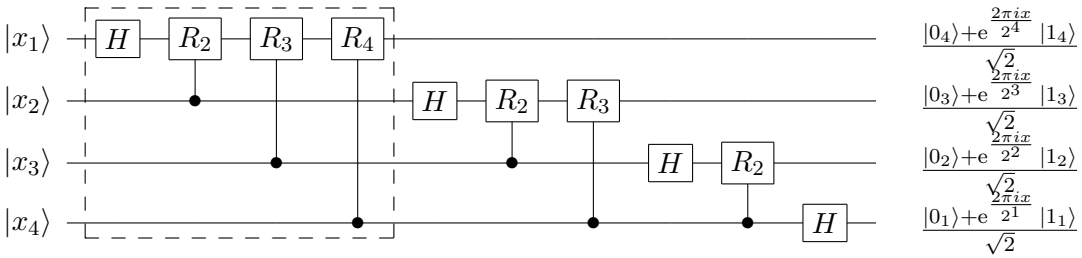
dove come solito abbiamo indicato con lo stesso simbolo uno stato con la sequenza binaria che lo identifica. Si verifica facilmente che la mappa è unitaria, potendosi scrivere

$$U_F = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \mathrm{e}^{i\frac{2\pi}{N} x\cdot y} \, |y\rangle\langle x|.$$

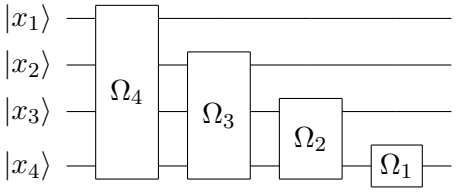
Classicamente calcolare il vettore trasformato richiede 2^n operazioni, mentre quantisticamente è sufficiente un numero polinomiale di gate e dunque il calcolo è molto più efficiente. Consideriamo anzitutto[3]

$$\begin{aligned} U_F|x\rangle = \frac{1}{2^{\frac{n}{2}}} \sum_{y=0}^{2^n-1} \mathrm{e}^{i\frac{2\pi}{2^n} x\cdot y} \, |y\rangle &= \frac{1}{2^{\frac{n}{2}}} \sum_{y_1=0}^{1} \cdots \sum_{y_n=0}^{1} \mathrm{e}^{\sum_l \frac{2i\pi x y_l}{2^l}} \, |y_1 y_2 \cdots y_n\rangle \\ &= \frac{1}{2^{\frac{n}{2}}} \sum_{y_1,\ldots,y_n=0}^{1} \bigotimes_{l=1}^{n} \mathrm{e}^{\frac{2i\pi x y_l}{2^l}} \, |y_l\rangle = \bigotimes_{l=1}^{n} \frac{|0_l\rangle + \mathrm{e}^{\frac{2\pi i x}{2^l}} \, |1_l\rangle}{\sqrt{2}}. \end{aligned}$$

Abbiamo così provato un fatto non banale: un vettore della base computazionale viene mappato in un vettore fattorizzabile e dunque, in questo caso, non viene creato entanglement. In particolare se $|x\rangle = |\mathbf{0}\rangle$ si ha $|\mathbf{0}\rangle \mapsto \frac{1}{2^{\frac{n}{2}}} \sum_y |y\rangle$ che è il risultato ottenuto utilizzando una serie di gate di Hadamard. In generale, tuttavia, la trasformata di Fourier quantistica *può* creare entanglement, dunque la trasformazione non è banalmente locale. Indicando con $R_k \stackrel{\text{def}}{=} P\left(\frac{2\pi}{2^k}\right)$ è possibile implementare, sulla base della scrittura precedente, il seguente circuito (per semplicità rappresentato per $n = 4$):



Per mostrare che il circuito precedente è corretto osserviamo che, $|x_1\rangle \xrightarrow{H} \frac{|0\rangle+(-1)^{x_1}|1\rangle}{\sqrt{2}} = \frac{|0\rangle+\mathrm{e}^{\frac{2\pi i x_1}{2}}|1\rangle}{\sqrt{2}} \xrightarrow{R_2 R_3 \cdots R_{n_1}} \frac{|0\rangle+\mathrm{e}^{2\pi i \sum_l \frac{x_l}{2^l}}|1\rangle}{\sqrt{2}} \equiv \frac{|0\rangle+\mathrm{e}^{\frac{2\pi x}{2^n}}|1\rangle}{\sqrt{2}}$ ed analogamente per gli altri qubit. Occorre tenere a mente che abbiamo ottenuto i qubit nell'ordine inverso rispetto a come voluto e dunque occorre eseguire una permutazione. Il numero di operazioni per l'i-esimo bit è $n-i$, per cui si ottiene un totale di $\frac{n(n+1)}{2}$ gate seguiti da al più $\frac{n}{2}$ operazioni di swap: dunque l'algoritmo è $\sim n^2$. Indichiamo con Ω_n il gate nel riquadro sopra (in particolare, nel nostro esempio si tratta di Ω_4). Il gate può complessivamente essere scritto nella forma

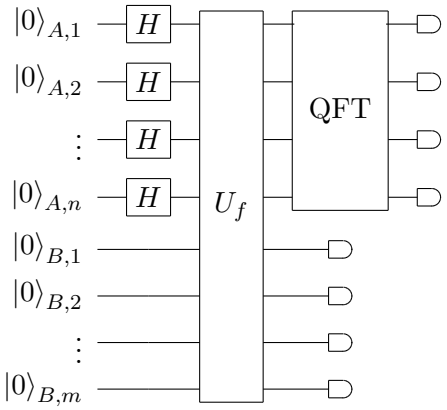


[3] Nell'espressione seguente si intende $x = \sum_{l=1}^{n} x_l 2^{n-l}$ e $y = \sum_{l=1}^{n} y_l 2^{n-l}$, dove x_l e y_l sono i valori della rappresentazione binaria.

5.5.2. *Period finding algorithm*

Sia $f\colon \{0,1\}^n \to \{0,1\}^m$ tale che $f(x \oplus_N r) = f(x)\ \forall x$, $N \stackrel{\text{def}}{=} 2^n$, dove $\oplus_N$ è la somma binaria modulo N. Vediamo come determinare il periodo r, che supponiamo divida N. Supponiamo inoltre[4] che per $x < y$ $f(x) = f(y) \Leftrightarrow y = x \oplus_N nr$, $n = 1, \ldots, \frac{N}{r}$. Una funzione di questo tipo, ad esempio, è la funzione a dente di sega, che è monotona a tratti. Classicamente questo problema è difficile da risolvere, viene affrontato con la *fast Fourier transform* ed ha un costo computazionale $O(n2^n)$, mentre quantisticamente ha un costo *polinomiale*!

Nel nostro caso utilizziamo due registri, un registro di dimensione n qubit A e un registro di m qubit B, inizializzati tutti in $|0\rangle$.



dove U_f è un oracolo che fornisce l'azione della funzione f, $U_f|x\rangle \otimes |0\rangle = |x\rangle \otimes |f(x)\rangle$. Lo stato dopo l'applicazione di U_f è dato da $\frac{1}{\sqrt{N}} \sum_x |x\rangle_A \otimes |f(x)\rangle_B = \frac{1}{\sqrt{N}} \sum_{x'=0}^{r-1} \sum_{m=0}^{\frac{N}{r}-1} |x' + mr\rangle_A \otimes |f(x' + mr)\rangle_B = \frac{1}{\sqrt{N}} \sum_{x'=0}^{r-1} \sum_{m=0}^{\frac{N}{r}-1} |x' + mr\rangle_A \otimes |f(x')\rangle_B$. La misura del registro B restituisce con probabilità $\frac{1}{r}$ uno degli r valori assunti da f: il valore ottenuto $f(x)$ non ha rilevanza ma lo stato del registro A viene proiettato su $\sqrt{\frac{r}{N}} \sum_{m=0}^{\frac{N}{r}-1} |x + mr\rangle_A$. Se eseguiamo una trasformata di Fourier quantistica, otteniamo

$$\sqrt{\frac{r}{N}} \sum_{m=0}^{\frac{N}{r}-1} |x + mr\rangle_A \xrightarrow{\text{QFT}} \sqrt{\frac{r}{N}} \sum_{m=0}^{\frac{N}{r}-1} \sum_{y=0}^{N-1} \frac{\mathrm{e}^{\frac{2\pi i(x+mr)y}{N}}}{\sqrt{N}} |y\rangle_A.$$

Misuriamo ora il registro A; la probabilità di ottenere un certo $|y\rangle_A$ è dunque[5]

$$P(|y\rangle_A) = \frac{r}{N^2} \left| \sum_{m=0}^{\frac{N}{r}-1} \mathrm{e}^{\frac{2\pi iymr}{N^2}} \right|^2 = \frac{r}{N^2} \left| \frac{\sin(\pi y)}{\sin\left(\frac{\pi yr}{N}\right)} \right|^2.$$

Se consideriamo $y = \frac{lN}{r}$, $l \in \mathbb{N}_0$, si ottiene $P(|y\rangle_A) = \frac{1}{r}$. Poiché per $l = 0, \ldots, r-1$ si hanno r diversi valori di y, e dovendo essere P una probabilità, si ha che necessariamente la probabilità è nulla per $y \neq \frac{lN}{r}$, $l = 1, \ldots, r-1$. All'atto della

[4] Questo vincolo si aggira con un costo computazionale che va come poly(n).

[5] La fase globale $\mathrm{e}^{\frac{2\pi iyx}{N}}$ è irrilevante.

misura, dunque, otteniamo un valore di y tra quelli ammessi, ovvero un particolare rapporto $\frac{l}{r}$, con l come sopra ma non specificato. Dato r, ci sono circa $\frac{r}{2\ln r}$ numeri primi minori di r: inoltre vi saranno ancor più interi l coprimi con r, per i quali $\frac{l}{r}$ risulta ridotta ai minimi termini (in particolare, la probabilità di estrarne uno è $\sim \frac{1}{2\ln N}$). Ripetendo l'algoritmo $2\ln N$ volte, dunque, la probabilità di osservare frazioni ai minimi termini nella forma $\frac{l}{r}$ sarà molto alta. Sulla base della statistica ottenuta si può individuare r. La complessità di questa parte dell'algoritmo è di tipo polinomiale.

Se r non è un divisore di N bisogna ripetere i ragionamenti precedenti sostituendo $\frac{N}{r} \to \left\lceil \frac{N}{r} \right\rceil$ o $\frac{N}{r} \to \left\lfloor \frac{N}{r} \right\rfloor$: per $N \to +\infty$ l'approssimazione non comporta errori.

5.6. Algoritmo di Shor

5.6.1. L'algoritmo di crittografia a chiave pubblica RSA

Per esemplificare l'importanza dell'algoritmo di Shor, esposto nel paragrafo seguente, presentiamo qui l'algoritmo di crittografia a chiave pubblica RSA. Supponiamo che Alice intenda inviare un messaggio a Bob in maniera sicura e dunque che il messaggio vada crittografato. Alice e Bob devono stabilire prima della comunicazione due funzioni che permettano rispettivamente la codifica e la decodifica del messaggio M. L'algoritmo a chiave pubblica prevede l'esistenza di due chiavi, una chiave *pubblica*, nota a chiunque (compresa Alice) ed una *privata*, nota solo a Bob. L'idea è utilizzare un algoritmo che permetta una semplice codifica mediante chiave pubblica ma tale da essere quasi impossibile da invertire se non è nota la chiave privata.

L'*algoritmo RSA* prevede che Bob utilizzi come chiave pubblica un numero $N = pq$, con p, q primi scelti da Bob. Sia ora $1 < e < (p-1)(q-1)$ e sia inoltre $\mathrm{MCD}(e, (p-1)(q-1)) = 1$. Si crea dunque d tale che $1 < d < (p-1)(q-1)$ e $de = 1 \mod N$. Bob rende pubblici N ed e, mentre p, q e d rimangono privati. Alice codifica il messaggio M come $\tilde{M} = M^e \mod N$. La decodifica avviene considerando il fatto che $\tilde{M}^d \mod N = M$. La sicurezza del protocollo è legata alla difficoltà di trovare d dati N ed e.

5.6.2. Algoritmo di Shor

Peter Shor ha suggerito di mappare il problema della fattorizzazione di un intero in un problema di *period finding*. Il miglior algoritmo noto per affrontare il problema ha complessità $O(\mathrm{e}^{\sqrt[3]{n \ln^2 n}})$. Supponiamo di avere $N = pq$, dove p, q sono primi. Definiamo ora l'*esponenziale modulare* $f_{y,N}(x) \stackrel{\text{def}}{=} y^x \mod N$. L'esponenziale modulare è una funzione periodica: se $r = \min_x\{x \in \mathbb{N} : y^x \mod N = 1\}$, allora $f_{y,N}(x + r) = f_{y,N}(x)$. L'algoritmo di Shor procede come segue:

1. Generiamo un numero $y < N$ random. Se $\frac{N}{y}$ è intero abbiamo ovviamente individuato la fattorizzazione. Diversamente, si adopera l'algoritmo di Euclide per individuare il massimo comun divisore di N ed y. Se esso è maggiore di 1 il valore trovato costituisce uno dei fattori cercati; altrimenti N e y sono coprimi e si passa al secondo punto.

2. Se N ed y sono coprimi, allora risolviamo l'equazione $y^r - 1 = 0 \mod N$ utilizzando l'algoritmo per individuare la periodicità. Dunque, supponendo r pari[6] $(y^{\frac{r}{2}}+1)(y^{\frac{r}{2}}-1) = 0 \mod N$, che equivale a scrivere $(y^{\frac{r}{2}}+1)(y^{\frac{r}{2}}-1) = lN$ per un qualche $l \in \mathbb{N}$. Se uno dei due fattori almeno divide N abbiamo ottenuto la nostra fattorizzazione; altrimenti occorre ripartire dal primo punto.

L'efficienza dell'algoritmo è vincolata dall'efficienza dell'oracolo che esegue l'esponenziazione modulare. Si prova che la complessità è di tipo *polinomiale* ($O(n^2 \ln n \ln\ln n)$), quindi per un computer quantistico il problema della fattorizzazione risulta essere *semplice*.

Esempio

Sia $N = 15$ e supponiamo di avere estratto $y = 2$. Qui $r = 4$, essendo $2^4 = 1 \mod 15$. Si ha $2^{\frac{4}{2}} + 1 = 5$ e $2^{\frac{4}{2}} - 1 = 3$, il che risolve il nostro problema immediatamente.

5.7. Algoritmo di Grover

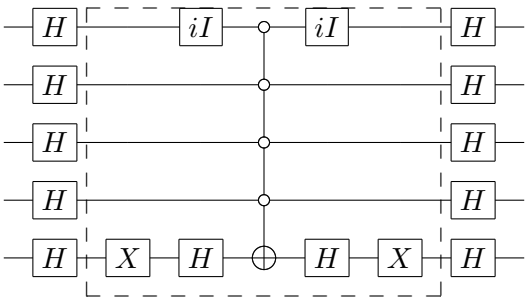
L'algoritmo di Grover non presenta guadagni significativi rispetto agli algoritmi classici che affrontano lo stesso problema; tuttavia la sua complessità è esattamente computabile e dunque si può dimostrare che non è possibile classicamente ottenere un algoritmo che risolve lo stesso problema con minore complessità computazionale. L'algoritmo è concepito per la ricerca in un database non strutturato (ovvero non indicizzato né ottimizzato in alcun modo per la ricerca). Classicamente occorre leggere al più tutti i bit di informazione 2^n, sicché se intendiamo cercare $x_0 \in \{0,1\}^n$ la nostra funzione f restituisce in x 1 se $x = x_0$, zero se diversamente. Vogliamo vedere quante volte occorre invocare f per individuare x_0, supponendo di disporre di un gate oracolare U_{x_0} tale che $U_{x_0}|x\rangle_A \otimes |y\rangle_B = |x\rangle_A \otimes |f(x) \oplus y\rangle_B$. In particolare, osserviamo che $U_{x_0}|x\rangle_A \otimes \frac{|0\rangle_B - |1\rangle_B}{\sqrt{2}} = (-1)^{f(x)}|x\rangle_A \otimes \frac{|0\rangle_B - |1\rangle_B}{\sqrt{2}}$. Tracciando via B, si ottiene per il registro A lo stato $(-1)^{\delta_{xx_0}}|x\rangle_A$, ovvero la funzione oracolare cambia la fase dell'input.

Dunque se consideriamo con la solita notazione

$$|s\rangle \stackrel{\text{def}}{=} \sum_x \frac{|x\rangle}{\sqrt{N}} \xrightarrow{U_{x_0}} -\frac{|x_0\rangle}{\sqrt{N}} + \sum_{x \neq x_0} \frac{|x\rangle}{\sqrt{N}} \equiv -\frac{|x_0\rangle}{\sqrt{N}} + \frac{\sqrt{N-1}}{\sqrt{N}}|r\rangle,$$

avendo definito $|r\rangle \stackrel{\text{def}}{=} \sum_{x \neq x_0} \frac{|x\rangle}{\sqrt{N-1}}$, con $\langle r|x_0\rangle = 0$. Definiamo ora l'operatore $D \stackrel{\text{def}}{=} 2|s\rangle\langle s| - \mathbb{1} \equiv H^{\otimes n}\,(2|0\rangle\langle 0| - \mathbb{1})\,H^{\otimes n} \stackrel{\text{def}}{=} H^{\otimes n} D_0 H^{\otimes n}$, trasformazione unitaria. L'operazione D_0 è una *control transformation* (in particolare una *control phase shift*) e il suo costo computazionale determina il costo di D (essendo il costo dei gate di Hadamard lineare in n). L'operatore D agisce come U_{x_0} ma sullo stato $|s\rangle$, dato che $D|s\rangle = |s\rangle$, mentre per $|x\rangle \perp |s\rangle$ $D|x\rangle = -|x\rangle$. Così sulla base di quanto detto D è del tipo

[6] Se r risulta dispari, occorre generare un altro valore di y e ricominciare.

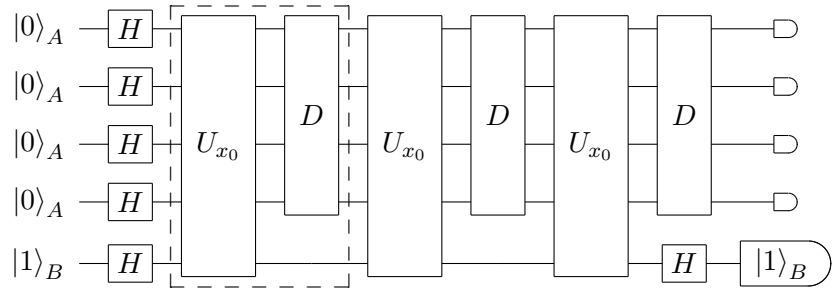


Come si vede, dunque, il costo computazionale del gate è polinomiale. Il gate nel riquadro coincide con l'azione di D_0. Osserviamo inoltre che, noto x_0, è possibile implementare U_{x_0} esattamente nella stessa maniera, potendo essere scritto (omettendo B) $U_{x_0} = \mathbb{1} - 2|x_0\rangle\langle x_0|$ e quindi anche il costo di U_{x_0} sarebbe polinomiale.

Dalla definizione $\langle x_0|s\rangle = \frac{1}{\sqrt{N}}$ e $\langle r|s\rangle = \sqrt{\frac{N-1}{N}}$, per cui nella base $\{|x_0\rangle, |r\rangle\}$ otteniamo la rappresentazione

$$U_{x_0} = \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{\text{riflessione}}, \quad D = \underbrace{\begin{pmatrix} -\frac{N-2}{N} & \frac{2\sqrt{N-1}}{N} \\ \frac{2\sqrt{N-1}}{N} & \frac{N-2}{N} \end{pmatrix}}_{\text{riflessione e rotazione}} \Rightarrow DU_{x_0} = \underbrace{\begin{pmatrix} \frac{N-2}{N} & -\frac{2\sqrt{N-1}}{N} \\ \frac{2\sqrt{N-1}}{N} & \frac{N-2}{N} \end{pmatrix}}_{\text{rotazione}} \stackrel{\text{def}}{=} G.$$

La matrice G ha la forma di una matrice di rotazione nello spazio bidimensionale generato da $\{|x_0\rangle, |r\rangle\}$ di un angolo $\theta = \arccos\frac{N-2}{N}$. Dato che per N molto grande $|s\rangle \approx |r\rangle$, è possibile tramite una rotazione di $\theta = \frac{\pi}{2}$ portare $|s\rangle \approx |x_0\rangle$: basta prendere, dunque, $\frac{\pi}{2} \approx \frac{2\sqrt{N-1}}{N} n \approx \frac{2n}{\sqrt{N}}$, ovvero applicare G $n \approx \frac{\pi}{2}\frac{\sqrt{N}}{2} = \sqrt{N}\frac{\pi}{4}$ volte.



Osserviamo infine che $|\langle x_0|\text{output}\rangle|^2 \approx 1 - \theta^2 \approx 1 - \frac{4}{N}$ e che la complessità dell'algoritmo è $\sim \text{poly}(n)2^{\frac{n}{2}}$. Si può provare che l'algoritmo di Grover è il miglior algoritmo possibile per risolvere questo problema. Ne segue che esistono problemi complessi anche in computazione quantistica, ma sono solitamente più efficacemente risolti rispetto alla computazione classica.

Capitolo 6

Error Correction

6.1. Classical error correction

Abbiamo finora supposto che fosse assente qualunque forma di disturbo. I computer reali, tuttavia, devono scontrarsi con errori di calcolo dovuti a imperfezioni di dispositivi o rumori esterni ed è necessario mettere a punto algoritmi di correzione per evitare che venga distrutta l'informazione.

Inoltre in computazione quantistica emergono aspetti computazionali sia di tipo digitale sia di tipo analogico. Tuttavia sono determinanti i primi e questo è di fondamentale importanza.

In teoria classica i metodi di controllo sono di vario tipo ma generalmente di semplice implementazione. Uno dei possibili errori, ad esempio, è un accidentale *bit flip*, ovvero l'inversione del valore di un bit da 0 a 1 o viceversa, oppure che l'informazione venga semplicemente cancellata (con una certa probabilità un bit viene azzerato, a prescindere dal suo valore iniziale). Una possibile soluzione è la *ridondanza*, ovvero associare ad ogni bit tre bit fisici con lo stesso valore: in caso di bit flip si assume che il valore corretto sia quello che compare più volte nella terna. Se ad esempio la probabilità di un bit flip è p, la probabilità di un errore è $p^2(3-2p) < p$. Si prova che in questa maniera la computazione è stabile. Osserviamo che un numero maggiore di bit comporta più operazioni e quindi più alta probabilità d'errore. Von Neumann ha provato che è comunque possibile mandare asintoticamente a zero la probabilità d'errore.

6.2. Quantum error correction

6.2.1. Bit flip e phase flip

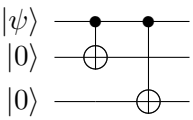
I codici di controllo classici sono basati sulla ridondanza, ovvero sulla copia multipla di informazioni. In computazione quantistica emergono delle sostanziali difficoltà:

1. il teorema di *no–cloning* non ci permette di utilizzare il metodo della ridondanza, in quanto non è possibile duplicare uno stato fisico;

2. il principio di misurare e correggere è inapplicabile, in quanto la misurazione distrugge l'entanglement;

3. al rumore discreto classico si contrappone un rumore continuo quantistico.

Queste difficoltà possono tuttavia essere sormontate. Supponiamo di avere un errore di *flip* con probabilità p, ovvero $\Phi[\rho] = (1-p)\rho + p\sigma^1\rho\sigma^1$. Siano ora $|0_L\rangle \stackrel{\text{def}}{=} |000\rangle \in \mathcal{H}^{\otimes 3}$, $|1_L\rangle \stackrel{\text{def}}{=} |111\rangle \in \mathcal{H}^{\otimes 3}$ e lo stato $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ da codificare come $|\psi_L\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$. Il gate per la codifica è



Il rumore può essere pensato come una correlazione con un *environment* $|e\rangle$ esterno; pertanto, ricorrendo alla forma di Stinespring, la mappa Φ (che opera indipendentemente su ciascun qubit) fornisce:

$$
\begin{aligned}
&|\psi_L\rangle|e\rangle \mapsto \sqrt{(1-p)^3}|\psi_L\rangle|e_0\rangle \\
&+\sqrt{p(1-p)^2}\left[\underbrace{(\alpha|100\rangle + \beta|011\rangle)}_{|\psi_1\rangle}|e_1\rangle + \underbrace{(\alpha|010\rangle + \beta|101\rangle)}_{|\psi_2\rangle}|e_2\rangle + \underbrace{(\alpha|001\rangle + \beta|110\rangle)}_{|\psi_3\rangle}|e_3\rangle\right] \\
&+\sqrt{p^2(1-p)}\left[(\alpha|110\rangle + \beta|001\rangle)|e_4\rangle + (\alpha|101\rangle + \beta|010\rangle)|e_5\rangle + (\alpha|011\rangle + \beta|100\rangle)|e_6\rangle\right] \\
&\qquad + \sqrt{p^3}\,(\alpha|111\rangle + \beta|000\rangle)|e_7\rangle,
\end{aligned}
$$

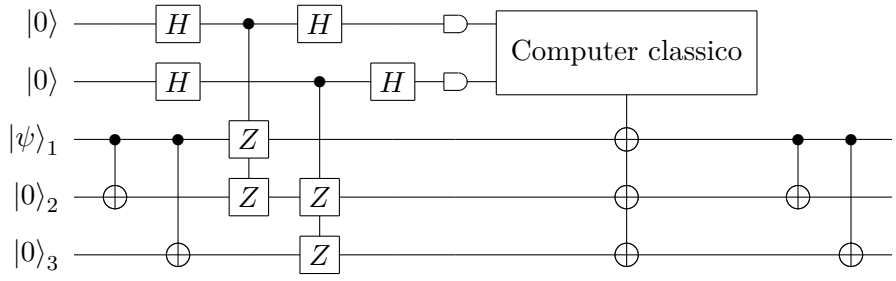
con $|e_i\rangle$ stati dell'*environment*. Indichiamo ora con X_i, Y_i, Z_i l'applicazione di $\sigma^1, \sigma^2, \sigma^3$ al qubit i-esimo dei tre di nostro interesse. Osserviamo che:

$$
\begin{aligned}
&Z_1Z_2|\psi_L\rangle = |\psi_L\rangle, \quad Z_1Z_2|\psi_1\rangle = -|\psi_1\rangle, \quad Z_1Z_2|\psi_2\rangle = -|\psi_2\rangle, \quad Z_1Z_2|\psi_3\rangle = |\psi_3\rangle \\
&Z_2Z_3|\psi_L\rangle = |\psi_L\rangle, \quad Z_2Z_3|\psi_1\rangle = |\psi_1\rangle, \quad Z_2Z_3|\psi_2\rangle = -|\psi_2\rangle, \quad Z_2Z_3|\psi_3\rangle = -|\psi_3\rangle.
\end{aligned}
$$

Dunque dopo le misure Z_1Z_2 e Z_2Z_3 lo stato acquisisce semplicemente un segno globale ma è possibile distinguere i quattro casi utili in cui c'è stato un singolo bit flip e correggerlo.

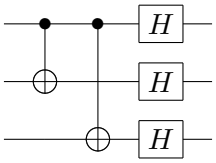
Se abbiamo due sistemi A e B è possibile provare che, dato lo stato $\sum_{ij} c_{ij}|\psi_i\rangle_A \otimes |\phi_j\rangle_B$, se si riesce a correggere gli errori su A, le correlazioni vengono preservate. In tutta la procedura, come si sarà notato, non viene letto lo stato e quindi l'informazione viene preservata. Inoltre, qualunque cosa succeda agli altri stati, essendo essi in partenza ortogonali a $|\psi_L\rangle$, una loro alterazione può solo far aumentare la *fidelity*.

È fondamentale che $[Z_1Z_2, Z_2Z_3] = 0$ e che gli stati da discriminare siano ortogonali tra loro. Il circuito ha la forma



Il computer classico (eventualmente sostituibile con uno quantistico) legge le misure ed esegue gli eventuali *flip*.

Un *phase flip* $\rho \mapsto \Phi(\rho) = (1-p)\rho + p\sigma^3\rho\sigma^3$ è unitariamente equivalente ad un bit flip. Osservando infatti che $\sigma^3|+\rangle = |-\rangle$ e $\sigma^3|-\rangle = |+\rangle$, possiamo studiare il phase flip utilizzando questa base per mezzo di un filtro di Hadamard:

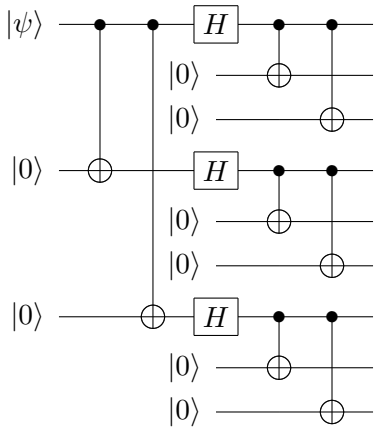


6.2.2. Algoritmo di Shor

Nel 1994 Shor ha proposto un algoritmo per la correzione di un qualsiasi errore su qubit utilizzando nove qubit

$$|0_L\rangle \stackrel{\text{def}}{=} \left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right)^{\otimes 3} \equiv |+\rangle^{\otimes 3}, \quad |1_L\rangle \stackrel{\text{def}}{=} \left(\frac{|000\rangle - |111\rangle}{\sqrt{2}}\right)^{\otimes 3} \equiv |-\rangle^{\otimes 3}.$$

Il circuito proposto da Shor per la codifica $|\psi\rangle \mapsto |\psi_L\rangle$ è il seguente:



Uno stato $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ pertanto viene rappresentato come

$$|\psi_L\rangle = \alpha\frac{(|0\rangle_1|0\rangle_2|0\rangle_3+|1\rangle_1|1\rangle_2|1\rangle_3)(|0\rangle_4|0\rangle_5|0\rangle_6+|1\rangle_4|1\rangle_5|1\rangle_6)(|0\rangle_7|0\rangle_8|0\rangle_9+|1\rangle_7|1\rangle_8|1\rangle_9)}{2\sqrt{2}} + \beta\frac{(|0\rangle_1|0\rangle_2|0\rangle_3-|1\rangle_1|1\rangle_2|1\rangle_3)(|0\rangle_4|0\rangle_5|0\rangle_6-|1\rangle_4|1\rangle_5|1\rangle_6)(|0\rangle_7|0\rangle_8|0\rangle_9-|1\rangle_7|1\rangle_8|1\rangle_9)}{2\sqrt{2}}$$

Come osservabili commutanti per constatare se sono avvenuti bit flip o phase flip adoperiamo

$$Z_1Z_2, \quad Z_2Z_3, \quad Z_4Z_5, \quad Z_5Z_6, \quad Z_7Z_8, \quad Z_8Z_9$$

per valutare i bit flip e

$$X_1X_2X_3X_4X_5X_6, \quad X_4X_5X_6X_7X_8X_9$$

per valutare i phase flip. Tutte queste osservabili commutano e quindi non distruggono entanglement.

Se per esempio avviene un bit flip $|0_L\rangle \mapsto \frac{|100\rangle+|011\rangle}{\sqrt{2}} \otimes |+\rangle^{\otimes 2}$ oppure $|1_L\rangle \mapsto \frac{|100\rangle-|011\rangle}{\sqrt{2}} \otimes |-\rangle^{\otimes 2}$ i risultati della misura sono $Z_1Z_2 \mapsto -1$ e tutte le altre $+1$ che è un esito che identifica univocamente un bit flip sul primo qubit. Se avviene un phase flip almeno una delle misure che coinvolge gli operatori X_i restituisce -1: è possibile inoltre identificare univocamente dove il processo è avvenuto e correggerlo appropriatamente. Un bit-phase flip determina misure con esiti -1 sia nel primo set di osservabili che nel secondo.

Per mostrare che il metodo funziona per ogni errore, scriviamo una generica CPT, che esprime la alterazione prodotta dal rumore su $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ come $U\left[(\alpha|0\rangle + \beta|1\rangle) \otimes |e\rangle\right] = \alpha\left(|0\rangle \otimes |e_{00}\rangle + |1\rangle \otimes |e_{01}\rangle\right) + \beta\left(|0\rangle \otimes |e_{10}\rangle + |1\rangle \otimes |e_{11}\rangle\right)$, ovvero

$$|\psi\rangle = |\psi\rangle \otimes \tfrac{|e_{00}\rangle+|e_{11}\rangle}{2} + \sigma^3|\psi\rangle \otimes \tfrac{|e_{00}\rangle-|e_{11}\rangle}{2} + \sigma^1|\psi\rangle \otimes \tfrac{|e_{01}\rangle+|e_{10}\rangle}{2} + \sigma^1\sigma^3|\psi\rangle \otimes \tfrac{|e_{01}\rangle-|e_{10}\rangle}{2}.$$

Ripetendo la stessa analisi per $|\psi_L\rangle$, con U che accoppia *solo il primo bit* con l'*environment*, si ha

$$|\psi_L\rangle = |\psi_L\rangle \otimes \tfrac{|e_{00}\rangle+|e_{11}\rangle}{2} + Z_1|\psi_L\rangle \otimes \tfrac{|e_{00}\rangle-|e_{11}\rangle}{2} + X_1|\psi_L\rangle \otimes \tfrac{|e_{01}\rangle+|e_{10}\rangle}{2} + X_1Z_1|\psi_L\rangle \otimes \tfrac{|e_{01}\rangle-|e_{10}\rangle}{2}.$$

Dunque una qualunque alterazione del primo bit è sempre scomponibile un bit flip, un phase flip e una composizione di questi. Aggiungiamo ora un'ancella A_1 su cui scriviamo il risultato dell'eventuale misura, ad esempio di bit flip sul primo qubit, e inizializziamola in $|0\rangle_{A_1}$:

$$\begin{aligned} &|\psi_L\rangle \otimes \frac{|e_{00}\rangle + |e_{11}\rangle}{2} \otimes |\text{no bit flip}\rangle_{A_1} + Z_1|\psi_L\rangle \otimes \frac{|e_{00}\rangle - |e_{11}\rangle}{2} \otimes |\text{bit flip}\rangle_{A_1} \\ &+ X_1|\psi_L\rangle \otimes \frac{|e_{01}\rangle + |e_{10}\rangle}{2} \otimes |\text{no bit flip}\rangle_{A_1} + X_1Z_1|\psi_L\rangle \otimes \frac{|e_{01}\rangle - |e_{10}\rangle}{2} \otimes |\text{bit flip}\rangle_{A_1}. \end{aligned}$$

L'esito della misura su A_1 permette di comprendere quale operazione eseguire per correggere l'errore. Dopodiché ripetiamo il ragionamento aggiungendo altre ancelle per altri tipi di errore e per gli altri qubit. Abbiamo così spostato le correlazioni fra sistema ed environment in correlazioni fra ancelle ed environment; alla fine abbiamo quindi $|\psi_L\rangle \otimes |\text{garbage state}\rangle$.

6.3. Teoria generale

Sia dato $\mathcal{H}^{\otimes n}$ e $C \subseteq \mathcal{H}^{\otimes n}$ sottospazio di codifica (nel caso dell'algoritmo di Shor $n = 9$ e $C = \text{span}\{|0_L\rangle, |1_L\rangle\}$). Sia inoltre il rumore una mappa $\Phi\colon \sigma(\mathcal{H}^{\otimes n}) \to \sigma(\mathcal{H}^{\otimes n})$, $\Phi[\rho] = \sum_k M_k \rho M_k^\dagger$, con M_k operatori di Kraus.

Definizione. Si dice che C corregge Φ se $\exists R$ mappa CPT tale che

$$R \circ \Phi(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|, \quad \forall |\psi\rangle\langle\psi| \in C.$$

Ricordiamo che le mappe CPT non sono in generale invertibili: qui però stiamo richiedendo che l'inversa esista solo su C.

Teorema. *Condizione necessaria e sufficiente affinché C corregga Φ è che valga la relazione*

$$PM_k^\dagger M_{k'} P = \alpha_{kk'} P, \tag{6.1}$$

dove gli operatori M_k sono gli operatori di Kraus relativi alla mappa Φ, P è il proiettore su C e $(\alpha_{kk'})_{kk'}$ è una matrice hermitiana[1].

Dimostrazione. Dimostriamo anzitutto che la (6.1) implica che C corregge Φ. Essendo $(\alpha_{kk'})_{kk'}$ definita positiva, esiste una matrice unitaria che la diagonalizza, ovvero $d_l \delta_{ll'} = \sum_{kk'} u_{lk}^\dagger \alpha_{kk'} u_{k'l}$. Definiamo ora $F_l \stackrel{\text{def}}{=} \sum_k u_{kl} M_k$: tali operatori sono anch'essi operatori di Kraus per Φ. Allora $PF_l^\dagger F_{l'} P = d_l \delta_{ll'} P$[2]. Definendo $N_l \stackrel{\text{def}}{=} \frac{PF_l^\dagger}{\sqrt{d_l}}$, allora $R(\Theta) = \sum_l N_l \Theta N_l^\dagger$ è la mappa che cerchiamo. Infatti $\forall |\psi\rangle \in C$:

$$R \circ \Phi(|\psi\rangle\langle\psi|) = R \circ \Phi(P|\psi\rangle\langle\psi|P) = \sum_{l,l'} N_l F_{l'} P |\psi\rangle\langle\psi| P F_{l'}^\dagger N_l^\dagger =$$

$$\sum_{l,l'} (PF_l^\dagger F_{l'} P) \frac{|\psi\rangle\langle\psi|}{d_l} (PF_{l'}^\dagger F_l P) = \sum_{l,l'} d_l \delta_{ll'} P \frac{|\psi\rangle\langle\psi|}{d_l} P d_l \delta_{ll'} = P|\psi\rangle\langle\psi|P \sum_l d_l = |\psi\rangle\langle\psi|.$$

Il fatto che $\sum_l d_l = 1$ segue dalla normalizzazione degli operatori di Kraus.

Dimostriamo ora l'implicazione inversa. Sia R tale che $R \circ \Phi[\rho] = \rho \ \forall \rho \in C$. Allora se N_l ed M_k sono operatori di Kraus per R e Φ rispettivamente, $\rho = \sum_{k,l} N_l M_k \rho M_k^\dagger N_l^\dagger \Rightarrow P\rho P = \sum_{k,l} N_l M_k P \rho P M_k^\dagger N_l^\dagger$, dove questa volta ρ può considerarsi generico (non necessariamente in C). A questo punto è utile ricordare che due rappresentazioni di Kraus $\{A_k\}_k$ e $\{B_{k'}\}_{k'}$ di Φ sono legate dalla relazione $A_{k'} = \sum_k W_{k'k} B_k$, dove $W_{k'k}$ è unitaria. Nel nostro caso, dunque, l'identità è verificata se $N_l M_k P \rho = c_{kl} P \rho$ con $\sum_{kl} |c_{kl}|^2 = 1$: questa relazione dev'essere necessariamente verificata oppure è possibile trovare una trasformazione unitaria che permetta di ottenere questa rappresentazione. Dunque

$$\sum_l PM_k^\dagger N_l^\dagger N_l M_{k'} P = \sum_l c_{k'l} c_{kl}^* P \Leftrightarrow PM_k^\dagger M_{k'} P = \alpha_{kk'} P, \quad \alpha_{kk'} \stackrel{\text{def}}{=} \sum_l c_{k'l} c_{kl}^*.$$

□

Siano $|\psi\rangle, |\phi\rangle \in C$. Dato che $\Phi(|\psi\rangle\langle\psi|) = \sum_k M_k |\psi\rangle\langle\psi| M_k^\dagger$ la mappa "disperde" $|\psi\rangle$ in un sottospazio più ampio. Inoltre $\langle\psi|M_k^\dagger M_{k'}|\phi\rangle = \langle\psi|PM_k^\dagger M_{k'} P|\phi\rangle = \alpha_{kk'}\langle\psi|\phi\rangle$; se $\langle\psi|\phi\rangle = 0$ l'azione degli operatori M_k *preserva l'ortogonalità*. Se invece $|\psi\rangle = |\phi\rangle$ allora $\langle\psi|M_k^\dagger M_{k'}|\psi\rangle = \alpha_{kk'}$, ovvero l'angolo tra i vettori trasformati *non dipende dallo stato di partenza*.

Esistono protocolli di correzione diversi da quello di Shor che utilizzano meno qubit e che riescono a proteggere dagli errori singoli (protocollo di Steane a 7 qubit; protocollo di Laflamme a 5 qubit, minimo numero possibile). È possibile proteggere k qubit logici da al più e errori usando n qubit fisici solo se vale la relazione

$$n \geq 4e + k.$$

[1] L'ipotesi di hermitianità non è in realtà necessaria, in quanto la matrice è senz'altro definita positiva, essendo $\sum_{kk'} v_k^* \alpha_{kk'} v_{k'} P = \sum_{kk'} P v_k^* M_k^\dagger M_{k'} v_{k'} P \equiv P\Omega^\dagger \Omega P \geq 0$, con $\Omega \stackrel{\text{def}}{=} \sum_k M_k v_k$.

[2] La *polar decomposition* garantisce che $F_l P = U_l \sqrt{(F_l P)^\dagger F_l P} = \sqrt{d_l} U_l P$, con U_l è un operatore unitario: dunque l'azione di F_k può essere espressa come una rotazione dell'intero sottospazio C sullo spazio in cui proietta $P_k = U_k P U_k^\dagger$, con $P_k P_l = 0$ per $l \neq k$.

6.3.1. *Fault tolerant Quantum Computation*

La *Fault tolerant Quantum Computation* studia i limiti di correzione di errore da parte di componenti aggiuntive a loro volta fonti di rumore. In questo contesto vale il fondamentale teorema:

Teorema. *Esiste una soglia p_{th} per la probabilità di errore per singolo gate tale che qualunque circuito quantistico ad N componenti ideali possa essere simulato con probabilità $1-\epsilon$ da $M \sim \frac{poly(N)}{\epsilon}$ componenti rumorosi.*

Il valore di soglia citato dipende dall'architettura dell'algoritmo utilizzato e vale, secondo le stime più recenti, $p_{\text{th}} \approx 10^{-5} \div 10^{-6}$.

Capitolo 7

Quantum Cryptography

7.1. Algoritmi a chiave pubblica e algoritmi a chiave privata

È interessante valutare se è possibile sfruttare la meccanica quantistica per elaborare procedure efficienti di crittografia, ovvero algoritmi che permettano che Alice e Bob comunichino senza che Eve (pur in ascolto) riesca a capire cosa si stanno dicendo.

Esistono due tipi di algoritmi:

- *algoritmi a chiave pubblica*, in cui tutti hanno a disposizione una chiave per criptare il messaggio ma solo Bob ha la chiave per decriptare;
- *algoritmi a chiave privata*, in cui Alice e Bob si sono scambiati in precedenza una chiave segreta.

Non esiste una vera e propria dimostrazione della sicurezza degli algoritmi a chiave pubblica. Si dice che essi sono *computationally secure* in quanto dipendono dalla nostra conoscenza (ancora imperfetta) delle classi di complessità computazionale. L'esistenza di un computer quantistico, ad esempio, inficerebbe la sicurezza di algoritmi come RSA.

Un esempio di algoritmo a chiave privata è il seguente. Supponiamo che Alice e Bob si siano scambiati in passato una chiave k, data da una stringa di cifre binarie. Poiché $k \oplus k = 0$, Alice può crittografare il suo messaggio m con $m' = m \oplus k$, mentre Bob potrà recuperare il messaggio con $m = m' \oplus k$. In questa maniera Eve intercetterà un messaggio non decodificabile. Tuttavia usando la stessa chiave due volte Eve può recuperare informazioni sui messaggi scambiati, ad esempio sommando le informazioni intercettate, $m'_1 \oplus m'_2 = m_1 \oplus m_2$.

7.2. Algoritmi quantistici

La *crittografia quantistica* si occupa di metodi per la creazione di chiavi private per mezzo di comunicazione su canali pubblici. Alice e Bob dovranno essere collegati da un canale classico e da un canale quantistico (tipo fibra ottica per il trasferimento di fotoni). L'osservazione da parte di Eve provoca alterazioni negli stati e ciò permette ad Alice e Bob di rilevare l'intercettazione.

a_i	b_i	$\|\psi_i\rangle$
0	0	$\|0\rangle$
1	0	$\|1\rangle$
0	1	$\|+\rangle$
1	1	$\|-\rangle$

TABELLA 7.1.: Protocollo BB84. Come si vede a_i determina il valore del qubit e b_i la base in cui scrivere.

7.2.1. PROTOCOLLO BB84

Nel protocollo BB84 Alice crea due sequenze binarie random **a** e **b** di dimensione n; in seguito crea n qubit secondo la regola in tabella 7.1. Alice invia dunque la sequenza di qubit a Bob, che quindi ha la forma $|0\rangle|+\rangle|-\rangle|-\rangle|1\rangle\cdots$. Bob genera una sequenza casuale **c** della stessa dimensione e procede eseguendo sull'i-esimo qubit una misura su $\{|0\rangle, |1\rangle\}$ se $c_i = 0$, una misura su $\{|+\rangle, |-\rangle\}$ se $c_i = 1$. A questo punto Bob annuncia i tipi di misure fatte **c** e Alice i tipi di codifica **b**. Potendo quindi Bob comprendere su quali qubit ha eseguito una misura secondo una base corretta e su quali invece no, elimina questi ultimi conservando i restanti che sono circa la metà del totale. In questo modo Alice e Bob hanno una stringa comune di $\frac{n}{2}$ qubit. Se Eve interferisce sostituendosi a Bob, egli potrebbe non ricevere la stringa corretta. Il protocollo prevede di sacrificare ad esempio $\frac{n}{4}$ qubit da confrontare pubblicamente, in modo da verificare la presenza di eventuali alterazioni. Si può provare che qualunque tentativo di Eve di acquisire informazione può essere rilevato, proprio per effetto del teorema di *no–cloning*. In effetti, supponendo di voler costruire una macchina tale che, sfruttando un'ancella, permetta di ottenere

$$\begin{cases} |\psi\rangle_S \otimes |0\rangle_A \\ |\phi\rangle_S \otimes |0\rangle_A \end{cases} \xrightarrow{U} \begin{cases} |\psi\rangle_S \otimes |u\rangle_A \\ |\phi\rangle_S \otimes |v\rangle_A \end{cases},$$

per unitarietà dovrà essere ${}_S\langle\psi|\phi\rangle_{SA}\langle 0|0\rangle_A = {}_S\langle\psi|\phi\rangle_{SA}\langle u|v\rangle_A \Rightarrow {}_A\langle u|v\rangle_A = 1$, ovvero gli stati sull'ancella sono indistinguibili. Se la linea è intrinsecamente rumorosa, la presenza di Eve viene rilevata come rumore aggiuntivo.

7.2.2. PROTOCOLLO B92

Nel protocollo B92 Alice genera una sequenza random di n cifre binarie **a** e invia per ciascuna un qubit nello stato $0 \mapsto |0\rangle$ e $1 \mapsto |+\rangle$. Bob genera a sua volta una sequenza random **c** che, come nel protocollo BB84, utilizza per scegliere la base in cui misurare. Se Bob ottiene come risultato di una misura $|0\rangle$ o $|+\rangle$ non può ricostruire con certezza il corretto valore del qubit iniziale. Se invece ottiene $|1\rangle$ o $|-\rangle$ c'è un solo qubit possibile per ciascuno (si veda la tabella 7.2). Probabilisticamente, i qubit utili che vengono conservati sono $\frac{N}{4}$, mentre i restanti sono da eliminare. Come nel protocollo BB84, si rileva la presenza di Eve sacrificando una porzione della serie utile ottenuta

a_i	c_i	$\|\psi_i\rangle$
0	0	$\|0\rangle$
	1	$\|+\rangle, \|-\rangle$
1	0	$\|0\rangle, \|1\rangle$
	1	$\|+\rangle$

Tabella 7.2.: Protocollo B92.

7.2.3. Protocollo di Eckert (EPR)

Il protocollo di Eckert sfrutta l'entanglement. È possibile che Alice spedisca a Bob dei qubit massimamente entanglati con i propri in numero n e che poi si proceda alle misure. La presenza di Eve viene rilevata testando le disuguaglianze di Bell oppure sfruttando l'*entanglement monogamy*, secondo cui la presenza di una certa quantità di entanglement tra A e B (massimizzata per distillazione) esclude correlazioni tra A ed E o tra E e B.

Capitolo 8

Teoria dell'informazione

8.1. Teoria classica dell'informazione

8.1.1. Entropia di Shannon

Cerchiamo ora di introdurre una quantità che esprima il quantitativo di informazione contenuto in un messaggio. Supponiamo che esistano d simboli dell'alfabeto, ciascuno prodotto con probabilità p_i. Ad esempio, nel caso uniforme si ha $p_i = \frac{1}{d}$. Ci aspettiamo che la variabile cercata abbia massimo valore, essendoci massima informazione, nel caso di una distribuzione uniforme, mentre se $p_i = \delta_{ij}$ allora l'informazione è assente. Il funzionale che si introduce è

$$H = -\sum_i p_i \log_2 p_i, \quad 0 \log_2 0 \stackrel{\text{def}}{=} 0,$$

detta *entropia di Shannon*. Il funzionale è estremamente utile in quanto, se si associa all'evento j con probabilità p_j un funzionale informazione continuo $I(p_j)$ dovrà essere $I(p_j) \geq 0$ e $I(p_j) = 0 \Leftrightarrow p_j = 1$; vogliamo inoltre che $I(p_j = q_j w_j) = I(q_j) + I(w_j)$ (additività dell'informazione). Sotto queste ipotesi si prova che l'unico funzionale ammesso è del tipo $-k \log_2 p_j$, con k costante positiva. Dunque l'entropia di Shannon non è altro che una media su tutti i possibili valori di p_j, $H = \langle I(p_j) \rangle$.

Consideriamo un esempio che mostra come identificare H quale misura dell'informazione. Sia $X = \{A, B, C, D\}$ un alfabeto (assimilabile a tutti gli effetti ad una variabile casuale) in cui i vari simboli hanno diversa probabilità $p_A = \frac{1}{2}$, $p_B = \frac{1}{4}$, $p_C = p_D = \frac{1}{8}$. Le stringhe prodotte con un qualche processo stocastico di n lettere constano di 4^n possibili sequenze codificabili con $m = \log_2 4^n = 2n$ bit. Tuttavia l'entropia di Shannon $H(X) = \frac{7}{4}$ permette di ottenere una stima del numero minimo di bit necessari: esso è $m = nH(x) = \frac{7}{4}n < 2n$. Infatti possiamo utilizzare, al posto della codifica $A = 00$, $B = 01$, $C = 10$, $D = 11$, la codifica $A = 0$, $B = 10$, $C = 110$, $D = 111$: in questo modo, il numero di bit da impiegare è $m(n) = n_A + 2n_B + 3(n_C + n_D) \Rightarrow \frac{m(n)}{n} \xrightarrow{n \to \infty} p_A + 2p_B + 3(p_C + p_D) = \frac{7}{4}$.

Definizione (Sequenza tipica). Una sequenza di simboli $\mathbf{x} = (x_{j_1}, \ldots, x_{j_n}) \in \mathcal{L}^{(n)}$ ottenuta da una variabile stocastica $X = \{x_j\}_{j=1,\ldots,d}$ è detta *sequenza tipica* se il processo con cui viene generata è stocastico e non autocontrollato.

Primo teorema di Shannon. *Sia dato un insieme di simboli* $X = \{x_j\}_{j=1,\dots,d}$ *cui è associata una probabilità* $p_j \stackrel{def}{=} P(x_j) \geq 0$, $\sum_j p_j = 1$. *Nel limite* $n \to \infty$ *il numero di bit necessari per codificare una sequenza* $\mathbf{x} = (x_{j_1}, \dots, x_{j_n}) \in \mathcal{L}^{(n)}$ *è* $nH(X)$, *dove*

$$H(X) = -\sum_j p_j \log_2 p_j$$

è l'entropia di Shannon.

Dimostrazione. Diamo qui un abbozzo di dimostrazione. In una sequenza tipica di n elementi il simbolo x_j apparirà $n_j \approx p_j n$ volte circa. Il numero di sequenza tipiche di lunghezza n è dato da $\left|\mathcal{L}^{(n)}_{\text{typ}}\right| = \frac{n!}{\prod_{j=1}^d n_j!}$. Usando la relazione di Stirling $n! \sim n^n \, \mathrm{e}^{-n} \sqrt{2\pi n} \sim n^n$ (tenendo solo il termine dominante) si ha

$$\left|\mathcal{L}^{(n)}_{\text{typ}}\right| \approx \frac{n^n}{\prod_{j=1}^d n_j^{n_j}} \approx \frac{n^n}{\prod_{j=1}^d (p_j n)^{p_j n}} = \frac{1}{\prod_j p_j^{n p_j}} \Rightarrow$$
$$\log_2 \left|\mathcal{L}^{(n)}_{\text{typ}}\right| = -n \sum_j p_j \log_2 p_j \stackrel{\text{def}}{=} nH(X).$$

Nei passaggi precedenti abbiamo considerato solo sequenze tipiche. Esse sono tuttavia, nel limite $n \to \infty$, le uniche sequenze a dare un contributo. □

Sia ora la probabilità che venga generata una precisa sequenza $\mathbf{x} = (x_{j_1}, \dots, x_{j_n})$ $P(\mathbf{x}) = \prod_i p_{j_i}$.

Definizione (Sequenza δ-tipica). Dato $\delta > 0$, la sequenza $\mathbf{x}$ è δ-tipica, $\mathbf{x} \in \mathcal{L}^{(n)}_{\text{typ}}(\delta)$, se $2^{-n(H(X)+\delta)} \leq P(\mathbf{x}) \leq 2^{-n(H(X)-\delta)}$.

La precedente definizione è compatibile con la definizione di sequenza tipica, in cui $\left|\mathcal{L}^{(n)}_{\text{typ}}\right| = 2^{nH(X)}$, avendo qui ammesso un piccolo errore.

Primo teorema di Shannon (o teorema delle sequenze δ-tipiche). *Sia fissato* $\delta > 0$; $\forall \epsilon > 0 \ \exists n_0 \colon \forall n \geq n_0$ *si ha:*

1. $P\left(\mathbf{x} \in \mathcal{L}^{(n)}_{typ}(\delta)\right) \geq 1 - \epsilon$;
2. $(1-\epsilon)2^{n(H(X)-\delta)} \leq \left|\mathcal{L}^{(n)}_{typ}(\delta)\right| \leq 2^{n(H(X)+\delta)}$;
3. $\forall R < H(X)$ *e* $\forall \mathcal{C}^{(n)} \subseteq \mathcal{L}^{(n)}_{typ}(\delta)$ *tale che* $\left|\mathcal{C}^{(n)}\right| \leq 2^{nR}$ *si ha* $P\left(\mathbf{x} \in \mathcal{C}^{(n)}\right) < \epsilon$, *ovvero non c'è un sottoinsieme più piccolo di quello delle sequenze* δ*-tipiche che soddisfa le relazioni richieste.*

8.1.2. Proprietà dell'entropia di Shannon

Usando le proprietà delle funzioni convesse sulla funzione $-\log_2 x$ si provano le seguenti relazioni.

1. Per costruzione $0 \leq H(X) \leq \log_2 d$.

2. Date due variabili stocastiche X e Y non correlate è possibile definire la probabilità congiunta $p(x,y)$, dove x e y sono due realizzazioni di X e Y rispettivamente e le probabilità marginali $p(x) = \sum_y p(x,y)$ e $p(y) = \sum_x p(x,y)$. Vale la relazione di *subadditività*:

$$H(X,Y) = -\sum_{x,y} p(x,y)\log_2 p(x,y) \leq H(X) + H(Y),$$

che è saturata se $p(x,y) = p(x)p(y)$.

3. Classicamente vale la relazione $H(X,Y) \geq H(X)$ e $H(X,Y) \geq H(Y)$. Quantisticamente queste relazioni possono essere violate. In teoria classica si definisce dunque l'*entropia condizionale* $H(X|Y) = H(X,Y) - H(Y) \geq 0$ e $H(X|Y) = 0$ se e solo se la variabile X è funzione della variabile Y. L'entropia condizionata permette di definire la *mutua informazione* come la quantità di informazione acquisibile su una variabile leggendone un'altra:

$$\begin{aligned} I(X:Y) &\stackrel{\text{def}}{=} H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) = H(Y) - H(Y|X) \Rightarrow \\ &\qquad I(X:Y) \leq H(X),\ I(X:Y) \leq H(Y). \end{aligned}$$

4. In teoria classica vale la *subadditività forte*

$$H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Z,Y),$$

che vale anche in meccanica quantistica sebbene vada ulteriormente precisata.

5. Oltre ad un limite superiore, $I(X:Y)$ ha un limite inferiore fornita dalla *disuguaglianza di Fano*. Supponiamo di voler inferire il valore della variabile X conoscendo Y, secondo la associazione $x_i \mapsto y_i$. La probabilità di errore è $p_E = \sum_{j \neq j'} p(x_j)p(y_{j'}|x_j)$. La disuguaglianza di Fano afferma che

$$I(X:Y) \geq H(X) - (\ln_2 d - 1)p_E - H_2(p_E),$$

dove $H_2(x) = -x\log_2 x - (1-x)\log_2(1-x)$ è l'*entropia dicotomica*.

6. Siano dati tre processi stocastici X, Y, Z ed introduciamo la catena di Markov $X \to Y \to Z$ in modo che $p(z|x,y) \stackrel{\text{def}}{=} \frac{p(x,y,z)}{p(x,y)} \equiv p(z|y)$, ovvero la variabile Z è correlata con X solo tramite Y. Allora vale la *data processing inequality* $I(Z:X) \leq I(Y:X)$, ovvero l'informazione si deteriora.

7. La mutua informazione è subadditiva. Infatti siano $\mathbf{X} = (X^{(1)}, X^{(2)}, \cdots, X^{(N)})$ e $\mathbf{Y} = (Y^{(1)}, Y^{(2)}, \cdots, Y^{(N)})$ due vettori di variabili stocastiche. Se il processo stocastico che manda $\mathbf{X}$ in $\mathbf{Y}$ è del tipo $X^{(i)} \mapsto Y^{(i)}$ si ha $I(\mathbf{X}:\mathbf{Y}) \leq \sum_i I(X^{(i)}:Y^{(i)})$.

8.2. Trasmissione su canali rumorosi

Supponiamo che Alice tenti di inviare a Bob un segnale su un canale in cui è presente del rumore. Sia x ciò che Alice invia e y ciò che riceve Bob. Bob può a priori ricevere qualunque messaggio y con probabilità $p(y|x)$. Possiamo pensare di rimediare a questo problema utilizzando n volte il canale, supponendo che il rumore sia sempre lo stesso e che il canale sia "senza memoria". Allora $p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y^{(i)}|x^{(i)})$. Se Alice dispone di un certo numero di stringhe che hanno immagini disgiunte nello spazio di Bob, allora la comunicazione può avvenire. Tuttavia, operando in questo modo, il numero di stringhe effettivamente utili è drasticamente ridotto dal numero iniziale. Se B è l'insieme delle *codewords* che Alice può usare si può introdurre il *rate* $R = \frac{1}{n}\log_2|B|$. Perché si lavori in condizioni ottimali (pochi usi del canale e molte codewords) bisogna massimizzare R. È possibile ammettere piccoli errori (*overlaps* tra le immagini delle codewords di B) purché si annullino all'aumentare di n. Allora si prova che

$$R_{\text{ottimale}} \equiv C = \max_{p(x)} I(X:Y), \tag{8.1}$$

dove si massimizza su tutti i possibili *random processes* in ingresso. La quantità C è detta *capacità del canale*.

Definizioni. Diamo le seguenti definizioni:

1. indichiamo con $\mathbf{x} = (x_1, \ldots, x_n)$ una codeword di n elementi;
2. l'insieme B delle codewords è detto *codebook* ed indichiamo con $M \stackrel{\text{def}}{=} |B|$;
3. il *codice* $\text{Codice}(M,n)$ è l'insieme dell'applicazione di codifica $E\colon \{1,\ldots,M\} \to B$ e decodifica $D\colon \{\mathbf{y}\} \to \{1,\ldots,M\}$;
4. il *rate* associato al codice è $R = \frac{\log_2 M}{n}$;
5. la *probabilità di errore del codice* è $p_E(\text{Codice}) = \frac{1}{M}\sum_{m=1}^{M} P(D(\mathbf{y}_m) \neq m|\mathbf{x}_m)$;
6. un rate R è *achievable* se $\lim_n p_E(\text{Codice}) = 0$;
7. la *capacità* di un canale è data da

$$\begin{aligned} C &\stackrel{\text{def}}{=} \max\{R\colon R \text{ achievable}\} \\ &= \lim_{\epsilon\to 0} \limsup_n \left\{ \frac{\log_2 M}{n} \,\middle|\, \exists \text{Codice}(M,n)\colon p_E(\text{Codice}) < \epsilon \right\}. \end{aligned}$$

Shannon ha provato che la precedente definizione di capacità di un canale coincide con la forma (8.1) (*secondo teorema di Shannon*). Si dimostra in maniera relativamente semplice che $C \leq \max_{p(x)} I(X:Y)$. Sia data infatti una $p(x)$ tale che B sia il suo insieme di sequenze tipiche. Data una codeword, essa è associata per il primo teorema di Shannon ad un insieme di dimensione $2^{nH(Y|X)}$, in modo che di conseguenza l'intero codebook ha immagine di dimensione $2^{nH(Y|X)}M$. Questo dev'essere minore del numero delle sequenze generate dal peso stocastico $p(y)$, $2^{nH(Y|X)}M \leq 2^{nH(Y)} \Rightarrow M \leq 2^{n(H(Y)-H(Y|X))} = 2^{nI(X:Y)} \Rightarrow R \leq I(X:Y)$.

8.2.1. *Noisy typewriter model*

Utilizzando l'alfabeto inglese a ventisei simboli, supponiamo che, digitando su una tastiera difettosa, con probabilità $\frac{1}{2}$ venga prodotta la lettera digitata x, con probabilità $\frac{1}{2}$ la successiva $y = x \oplus 1$. Allora $H(Y|X) = -\sum_{x,y} p(x,y) \log_2 p(y|x) = 1$ e $I(X:Y) = H(Y) - H(Y|X) = H(Y) - 1$. Dunque $C = \max_{p(x)}(H(Y) - 1)$, che può essere massimizzata massimizzando $H(Y)$, che ha valore massimo $H_{\max}(Y) = \log_2 26$, ovvero $C = \log_2 26 - 1 = \log_2 13$. La codifica si costruisce facilmente tralasciando le lettere "pari" all'atto dell'invio: se Bob riceve A o B decodifica come A, eccetera. La possibilità di comunicare senza errori su un canale con rumore è stata una rivoluzione di notevole portata.

8.3. Entropia di von Neumann

8.3.1. Definizione di entropia di von Neumann

L'entropia di von Neumann è un funzionale associato allo stato di un sistema definito come $S(\rho) \stackrel{\text{def}}{=} H(\lambda_k) = -\sum_k \lambda_k \log_2 \lambda_k$, dove λ_k sono gli autovalori della matrice densità. Storicamente, l'entropia di von Neumann è stata introdotta prima di quella di Shannon e non era chiaro il suo collegamento con la teoria dell'informazione. Le sue proprietà fondamentali sono le seguenti.

1. $0 \leq S(\rho) \leq \log_2 d$, $d = \dim \mathcal{H}$, $S(|\psi\rangle\langle\psi|) = 0$, $S\left(\frac{\mathbb{1}}{d}\right) = \log_2 d$;

2. Definendo la *relative entropy* $S(\rho\|\sigma) \stackrel{\text{def}}{=} \operatorname{tr}[\rho \log_2 \rho] - \operatorname{tr}[\rho \log_2 \sigma]$, si ottiene un funzionale dalle numerose proprietà:
   a) vale la *disuguaglianza di Klein* $S(\rho\|\sigma) \geq 0$, e $S(\rho\|\sigma) = 0 \Leftrightarrow \rho = \sigma$ per cui può essere utilizzata come misura della distanza tra stati, sebbene non sia una distanza in senso proprio, dato che $S(\rho\|\sigma) \neq S(\sigma\|\rho)$, e, se $\operatorname{Ker}(\sigma) \cap \operatorname{Supp}(\rho) \neq 0$, allora $S(\rho\|\sigma) = +\infty$ (qui $\operatorname{Supp}(\rho) = \operatorname{span}\{\text{autovettori non nulli}\}$);
   b) $S(\Phi[\rho]\|\Phi[\sigma]) \leq S(\rho\|\sigma)$ per Φ CPT.

3. Supponiamo di avere ρ_{AB} per $\mathcal{H}_A \otimes \mathcal{H}_B$. Dette ρ_A e ρ_B le tracce parziali, l'entropia di von Neumann gode della proprietà di *subadditività*, $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$. Inoltre $S(\rho_{AB}\|\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \geq 0$ come si vede per calcolo diretto (l'uguaglianza vale se e solo se $\rho_{AB} = \rho_A \otimes \rho_B$).

4. Sia $|\psi\rangle_{AB} = \sum_k \sqrt{\lambda_k} |\psi_k\rangle_A \otimes |\phi_k\rangle_B$ (decomposizione di Schmidt). Allora, essendo $\rho_A = \sum_k \lambda_k |\psi_k\rangle_A \langle\psi_k|$ e $\rho_B = \sum_k \lambda_k |\phi_k\rangle_B \langle\phi_k|$, ne discende che *le entropie locali sono uguali*;

5. $S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$. Classicamente la relazione è banale essendo $H(AB) \geq H(A), H(B)$; ma tale proprietà non vale in meccanica quantistica.

   Dimostrazione. Costruiamo una purificazione $|\psi\rangle_{ABR}$ di ρ_{AB} e poniamo $\rho_{ABR} = |\psi\rangle_{ABR}\langle\psi|$, allora $S(\rho_{AR}) \leq S(\rho_A) + S(\rho_R)$. Essendo $S(\rho_{ABR}) = 0$, dalla proprietà 4 si ha $S(\rho_R) = S(\rho_{AB})$ e $S(\rho_B) = S(\rho_{AR})$, per cui $S(\rho_B) \leq S(\rho_A) +$

$S(\rho_{AB})$ ed analogamente (invertendo $A \leftrightarrow B$) $S(\rho_A) \leq S(\rho_B) + S(\rho_{AB})$, da cui si ha la tesi. □

6. $S(\rho_{AB}) \not\geq S(\rho_A)$ (basta considerare uno stato massimamente entanglato: questa proprietà è tipicamente quantistica).

7. L'entropia di von Neumann è *concava*, $S\left(\sum_l p_l \rho_l\right) \geq \sum_l p_l S(\rho_l)$. Inoltre è possibile provare che $S(\sum_l p_l \rho_l) \leq H(p_l) + \sum_l p_l S(\rho_l)$. In definitiva non esiste nessuna precisa relazione d'ordine tra $H(p_l)$ ed $S\left(\sum_l p_l \rho_l\right)$.

8. Vale la *subadditività forte*, $S(\rho_{ABC}) + S(\rho_C) \leq S(\rho_{AC}) + S(\rho_{BC})$ (la dimostrazione è piuttosto complessa).

Teorema (Schumacher)**.** *Sia data una sorgente che emette stati $|\psi_j\rangle$ puri con probabilità p_j e sia $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$. Se $d = \dim \mathcal{H}$, i qubit necessari per rappresentare una stringa di n caratteri generata da questa sorgente sono $M = nS(\rho)$ (classicamente erano $n \log_2 d$).*

8.3.2. Informazione accessibile ed Holevo bound

Supponiamo che Alice disponga di un alfabeto di simboli $\{j\}$ e abbia prodotto per ciascuno uno stato ρ_j con probabilità p_j. In seguito Alice invia a Bob tali stati. Bob conosce l'elenco dei simboli, delle matrici densità e delle probabilità associate. Ci si chiede se Bob può effettivamente recuperare il messaggio. Se Bob esegue una POVM $\{E_{j'}\}$ si ha $p(j'|j) = \operatorname{tr}\left[E_{j'}\rho_j\right]$. L'informazione recuperabile è dunque fornita dalla quantità classica $I_{\text{POVM}}(j' : j)$, mutua informazione. Tuttavia, come abbiamo evidenziato, c'è una dipendenza dalla POVM: si definisce *informazione accessibile* $I_{\text{acc}}(\{p_j, \rho_j\}) = \max_{\text{POVM}} I_{\text{POVM}}(j' : j)$: questa quantità solitamente non è calcolabile direttamente ma esiste un limite superiore, detto *Holevo bound*:

$$I_{\text{acc}}(\{p_j, \rho_j\}) \leq \chi(\{p_j, \rho_j\}), \quad \chi(\{p_j, \rho_j\}) \stackrel{\text{def}}{=} S\left(\sum_j p_j \rho_j\right) - \sum_j p_j S(\rho_j) \geq 0.$$

La funzione χ è detta *Holevo quantity*. Dunque $I_{\text{acc}}(\{p_j, \rho_j\}) \leq S\left(\sum_j p_j \rho_j\right) \leq \log_2 d$.

Dimostrazione. Introduciamo quattro insiemi, Q, X, Y, Z, dove Q è l'insieme degli stati ρ_j mentre i restanti sono insiemi di ancelle. Sia dunque

$$\rho_{QXYZ} = \sum_j p_j \rho_j \otimes |j\rangle_X\langle j| \otimes |0\rangle_Y\langle 0| \otimes |0\rangle_Z\langle 0|,$$

correlando così classicamente Q e X. Rappresentiamo una generica POVM $\{E_j\}$ come

$$U_{QYZ}|\psi\rangle_Q|0\rangle_Y|0\rangle_Z = \sum_j \left(\sqrt{E_j}|\psi\rangle_Q\right)|j\rangle_Y|j\rangle_Z.$$

Definendo ora

$$\rho'_{QXYZ} \stackrel{\text{def}}{=} U_{QYZ}\rho_{QXYZ}U^\dagger_{QYZ} = \sum_{j,k,k'} p_j \sqrt{E_k}\rho_j\sqrt{E_{k'}} \otimes |j\rangle_X\langle j| \otimes |k\rangle_Y\langle k'| \otimes |k\rangle_Z\langle k'|.$$

Dunque eseguendo le opportune tracce parziali

$$\rho'_{XY} = \sum_{j,k} p_j \operatorname{tr}\left[E_k \rho_j\right] |j\rangle_X\langle j| \otimes |k\rangle_Y\langle k| \equiv \sum_{j,k} p(j,k) |jk\rangle_{XY}\langle jk| \Rightarrow S(\rho'_{XY}) = H(X,Y),$$
$$\rho'_Y = \sum_k \operatorname{tr}\left[E_k \rho\right] |k\rangle_Y\langle k| \Rightarrow S(\rho'_Y) = H(Y).$$

Invece $S(\rho'_{QYZ}) = S(\rho_{QYZ}) = S(\rho)$ (essendo U_{QYZ} unitaria per questo stato) e $S(\rho'_{QXYZ}) = S(\rho_{QXYZ}) = S(\rho_{QX}) = -\sum_j \operatorname{tr}\left[p_j \rho_j \log_2 p_j \rho_j\right] = H(X) + \sum_j p_j S(\rho_j)$. Ricorrendo ora alla subadditività forte si ha

$$S(\rho'_{QXYZ}) + S(\rho'_Y) \leq S(\rho'_{XY}) + S(\rho'_{QYZ})$$
$$\Rightarrow H(X) + \sum_j p_j S(\rho_j) + H(Y) \leq H(X,Y) + S(\rho),$$

da cui segue la tesi

$$\boxed{I(X,Y) \stackrel{\text{def}}{=} H(X) + H(Y) - H(X,Y) \leq S(\rho) - \sum_j p_j S(\rho_j) \stackrel{\text{def}}{=} \chi(\{p_j, \rho_j\}).} \quad \square$$

8.3.3. Trasferimento di informazione su canali quantistici

Se Alice e Bob sono collegati da un canale quantistico ma Alice vuole inviare a Bob un'informazione classica, ci aspettiamo che il rate di comunicazione classica sia nettamente incrementato dalle proprietà quantistiche del canale. Supponiamo che il canale sia privo di memoria, ma associato ad una trasformazione CPT Φ che costituisce il rumore. La capacità C del canale, esprimibile come $C(\Phi) = \lim_n \left.\frac{\text{numero di bit con } n \text{ usi}}{n}\right|_{p_E \to 0}$, può essere espressa come segue:

$$C(\Phi) = \lim_n \frac{1}{n} \max_{\{p_j, \rho_j\}} \chi(\{p_j, \Phi^{\otimes n}[\rho_j]\}).$$

La precedente è detta *formula di Holevo–Schumacher–Westermoreland*. Come si vede χ svolge il ruolo della mutua informazione classica. Fino al 2008 si è dibattuto sull'utilità della regolarizzazione $\lim_n \frac{1}{n}$. Ci si chiedeva in particolare se

$$\max_{\{p_j, \rho_j\}} \chi(\{p_j, \Phi^{\otimes n}[\rho_j]\}) \stackrel{?}{\leq} n \max_{\{p_j, \rho_j\}} \chi(\{p_j, \Phi[\rho_j]\}).$$

Hastings ha successivamente provato che la relazione d'ordine non è sempre valida e la regolarizzazione è necessaria. Inoltre, in teoria classica un canale di feedback non aumenta la capacità della comunicazione. Quantisticamente invece questo avviene (consentendo, ad esempio, protocolli di distillazione).

L'entanglement, infine, può aumentare la capacità del canale (ad esempio tramite superdense coding): si parla in tal caso di *entanglement assisted classical capacity*.

Se usiamo il canale per trasmettere stati quantistici, la capacità quantistica è

$$Q(\Phi) \stackrel{\text{def}}{=} \lim_n \frac{\#\text{ qubit con } n \text{ usi}}{n} \leq C(\Phi).$$

Esistono canali per cui $Q(\Phi) = 0$ e $C(\Phi) \neq 0$. Nel caso di entanglement assisted quantum capacity (messa in atto tramite ad esempio teletrasporto) $\frac{C_E(\Phi)}{2} = Q_E(\Phi) \geq Q(\Phi)$.

Appendice A

Nota matematica

A.1. *Polar decomposition* e *singular value decomposition*

Sia Θ un operatore lineare su uno spazio di Hilbert $\mathcal{H}$ di dimensione finita n. Se $\Theta = \Omega^2$, Ω operatore lineare su $\mathcal{H}$, allora Ω è detto *radice* di Θ. Si dice che Ω è *radice positiva* di Θ se è semidefinita positiva, $\Omega \geq 0$. Se inoltre Θ è un operatore semidefinito positivo, si prova che la sua radice $\Omega \stackrel{\text{def}}{=} \sqrt{\Theta} \geq 0$ *esiste ed è unica.* In tali ipotesi è sempre possibile trovare una base $\{|j\rangle\}$ in cui Θ può essere scritto come $\Theta = \sum_{j=1}^{n} \lambda_j |j\rangle\langle j|$, $\lambda_j \geq 0$; in tale base $\Omega = \sum_{j=1}^{n} \sqrt{\lambda_j}|j\rangle\langle j|$.

Teorema (*Polar decomposition*)**.** *Sia Θ un operatore lineare; allora $\exists! J \geq 0$, $\exists! K \geq 0$ ed $\exists U$ unitario tali che $\Theta = UJ = KU$, dove*

$$J \stackrel{def}{=} \sqrt{\Theta^\dagger \Theta}, \quad K \stackrel{def}{=} \sqrt{\Theta \Theta^\dagger} \tag{A.1}$$

Dimostrazione. Dalla definizione data di J e K si deduce che esse sono matrici semidefinite positive e dunque hermitiane. Pertanto è possibile diagonalizzare $J = \sum_{l=1}^{n} j_l |l\rangle\langle l|$, dove $\{|l\rangle\}$ è un set ortonormale completo di $\mathcal{H}$. Se dunque $\Theta|l\rangle = |\psi_l\rangle \Rightarrow \langle\psi_l|\psi_l\rangle = \langle l|\Theta^\dagger\Theta|l\rangle = \langle l|J^2|l\rangle = j_l^2$. Se $j_l \neq 0$ allora $|\psi_l\rangle$ è diverso dal vettore nullo e può essere normalizzato come $|e_l\rangle = \frac{1}{j_l}|\psi_l\rangle$. I vettori così ottenuti sono ortonormali ($\langle e_l|e_k\rangle = \frac{\langle\psi_l|\psi_k\rangle}{j_l j_k} = \frac{\langle l|J^2|k\rangle}{j_l j_k} = \delta_{lk}$). Se $\dim \text{span}\{|e_l\rangle\} < n$ tale base può essere completata in maniera arbitraria. Introducendo ora $U = \sum_{l=1}^{n} |e_l\rangle\langle l|$, $UJ|l\rangle = j_l|e_l\rangle = |\psi_l\rangle = \Theta|l\rangle$, ovvero $\Theta = UJ$. Analogamente si prova che $\Theta = KU$. L'arbitrarietà nel completamento della base evidenzia l'eventuale non unicità di U, a meno che J (o K) non sia invertibile: infatti le matrici J e K hanno lo stesso spettro, essendo $UJ = KU \Rightarrow J = U^\dagger KU$. □

Corollario (*Singular value decomposition*)**.** *Sia Θ la matrice di rappresentazione di un operatore lineare come sopra. Allora $\exists U, V$ unitarie e $\exists D$ diagonale tale che $\Theta = UDV$. Inoltre D ha lo stesso spettro di J e K. Tali autovalori sono detti singolari.*

Dimostrazione. Basta osservare che $J = V^\dagger DV \Rightarrow \Theta = \tilde{U}J = UDV$, con $U = \tilde{U}V$. □

A.2. Identità notevoli

Segnaliamo di seguito dei risultati notevoli utili per lo studio dell'evoluzione dei sistemi quantistici e per la chiusura delle algebre che intervengono nella trattazione di un certo sistema[1].

Se $[A,[A,B]] = [B,[A,B]] = 0$, vale la seguente *identità di Baker–Campbell–Hausdorff*:

$$\mathrm{e}^{A+B} = \mathrm{e}^{A}\,\mathrm{e}^{B}\,\mathrm{e}^{-\frac{[A,B]}{2}};$$

la precedente vale in maniera approssimata se si ha un parametro piccolo ϵ, come

$$\mathrm{e}^{\epsilon(A+B)} \approx \mathrm{e}^{\epsilon A}\,\mathrm{e}^{\epsilon B}\,\mathrm{e}^{-\epsilon^2\frac{[A,B]}{2}} + O(\epsilon^3)$$

o se $[A,B]$ è piccolo. Da tale formula si ha, ad esempio, che per due hamiltoniane H_0, H_1

$$\mathrm{e}^{iH_0\Delta t}\,\mathrm{e}^{iH_1\Delta t}\,\mathrm{e}^{-iH_0\Delta t}\,\mathrm{e}^{-iH_1\Delta t} = \mathrm{e}^{[H_0,H_1]\Delta t^2} + O(\Delta t^2).$$

Vale anche l'*espansione di Lie–Trotter*:

$$\mathrm{e}^{A+B} = \left(\mathrm{e}^{\frac{A}{n}}\,\mathrm{e}^{\frac{B}{n}}\right)^n + O\left(\frac{1}{n}\right).$$

[1] La teoria del controllo quantistico (*quantum control theory*) si occupa ad esempio dello sfruttamento delle proprietà fisico-matematiche del sistema per ottenere informazioni non accessibili direttamente.

Appendice B

Implementazioni fisiche

B.1. Criteri di diVincenzo

DiVincenzo ha formulato un insieme di criteri da soddisfare per rendere possibile la computazione quantistica. In particolare è necessario avere

- registri di sistemi a due livelli ad n qubit;
- inizializzazione del registro;
- manipolazione dello stato (tramite trasformazioni unitarie);
- *read–out*, lettura finale del risultato;
- lungo tempo di coerenza.

I più ovvi sistemi a due livelli sono

1. *spin elettronico*;
2. *spin nucleari nei semiconduttori*;
3. *impurezze in solidi*;
4. *circuiti superconduttivi.*

Le interazioni fra spin come nei punti 1, 2, 3 sono del tipo $H = \sum_{i=1}^{n} \mathbf{h}_i \cdot \mathbf{s}_i + \frac{1}{2}\sum_{ij} J_{ij}\mathbf{s}_i \cdot \mathbf{s}_j$ (hamiltoniana di Heisenberg), dove le costanti di accoppiamento dipendono dalle caratteristiche microscopiche del sistema. Per il singolo qubit, la generica hamiltoniana quantistica ha la forma

$$H = -\frac{1}{2}\left(\epsilon\sigma^3 + \Delta\sigma^1\right) = -\frac{1}{2}\begin{pmatrix} \epsilon & \Delta \\ \Delta & -\epsilon \end{pmatrix}. \tag{B.1}$$

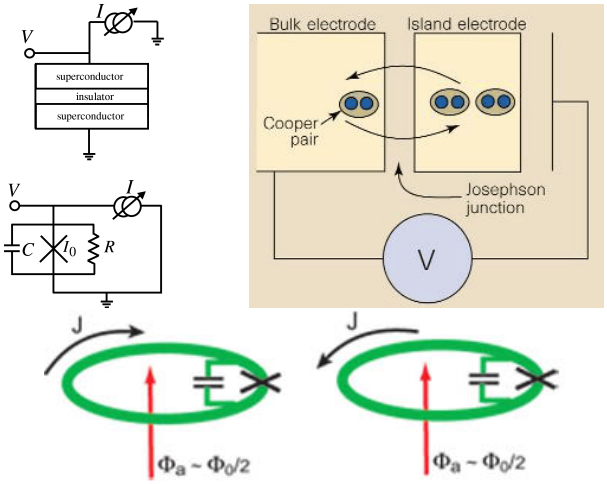


FIGURA B.1.: Schema di un *phase qubit*, di un SCB *qubit* e di un *rf*-SQUID.

B.2. Circuiti superconduttivi

B.2.1. Superconduttore

Un superconduttore ha la caratteristica di avere resistività ρ nulla al di sotto di una certa temperatura critica T_c. La funzione d'onda globale degli elettroni $\psi(\mathbf{r}) = |\psi(\mathbf{r})|\, \mathrm{e}^{i\phi(\mathbf{r})}$ ha la caratteristica di avere la fase $\phi(\mathbf{r})$ correlata macroscopicamente. Gli elettroni interagiscono tramite interazioni a due corpi e, per $T < T_c$, formano delle *coppie di Cooper* e la loro funzione d'onda diventa $|\psi\rangle_{\mathrm{BCS}} = \prod_{\mathbf{k}} \left(n_{\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{\mathbf{k}\uparrow} a^\dagger_{-\mathbf{k}\downarrow}\right) |0\rangle$.

B.2.2. Giunzione Josephson

Supponiamo di avere due superconduttori separati da una barriera. Su ciascun superconduttore vi sarà una fase differente φ_1 e φ_2 rispettivamente. Il fatto notevole è che, anche in assenza di differenza di potenziale, è possibile che scorra una corrente (per effetto tunnel attraverso la barriera) $I_s = I_c \sin\phi$, $\phi = \varphi_1 - \varphi_2$, con I_c corrente critica (massima corrente non dissipativa possibile). Applicando una differenza di potenziale V tra i due capi della giunzione, si ha[1]

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{2eV}{\hbar} = 2\pi \frac{V}{\Phi_0},$$

dove $\Phi_0 \stackrel{\mathrm{def}}{=} \frac{h}{2e}$. Utilizzando questo tipo di giunzione possiamo costruire diversi dispositivi.

B.2.3. Phase qubit (Current biased Josephson junction)

Un *phase qubit* è realizzabile utilizzando una giunzione di Josephson reale, modellizzabile come una giunzione ideale *in parallelo* con un capacitore C ed un resistore

[1] In realtà per avere una fase gauge invariante occorrerebbe considerare come differenza la quantità $\phi + \frac{2\pi}{\Phi_0} \int \mathbf{A} \cdot \mathrm{d}\,\mathbf{s}$, dove $\Phi_0 = \frac{h}{2e}$ è il quanto di flusso.

R. Con chiaro significato di notazione, la corrente di polarizzazione (*bias current*) è data da:

$$I_b = I_J + I_C + I_R \equiv I_c \sin\phi + C\frac{\mathrm{d}V}{\mathrm{d}t} + \frac{V}{R}, \quad \phi(t) = \frac{2e}{\hbar}\int^t V(\tau)\,\mathrm{d}\,\tau.$$

Dunque

$$\frac{\hbar}{2e}C\ddot{\phi} + \frac{\hbar}{2eR}\dot{\phi} + I_c \sin\phi = I_b$$

che è l'equazione di un oscillatore non lineare smorzato. Nel limite $R \to \infty$ (assenza di dissipazione) questa equazione può essere ottenuta da una "lagrangiana"; definendo $E_J \stackrel{\text{def}}{=} \frac{\hbar I_c}{2e}$ *energia della giunzione Josephson* e $E_C \stackrel{\text{def}}{=} \frac{(2e)^2}{2C}$ *energia di carica del capacitore per singola coppia di Cooper* la lagrangiana assume la forma

$$L(\dot{\phi}, \phi) = \frac{\hbar^2\dot{\phi}^2}{4E_C} - E_J(1 - \cos\phi) + \frac{\hbar I_b}{2e}\phi,$$

In assenza di I_b, l'equazione del moto è quella di un oscillatore puro, $\ddot{\phi} + \omega_J^2 \sin\phi = 0$, dove $\omega_J^2 = \frac{2E_C E_J}{\hbar^2}$ è detta *frequenza di plasma*. L'andamento del potenziale $U(\phi) = E_J(1 - \cos\phi) - \frac{\hbar I_b}{2e}\phi$ dipende da I_b ed è tanto più oscillante in funzione di $\frac{I_c}{I_b}$.

Passiamo ora alla formulazione hamiltoniana del problema introducendo $p = \frac{\partial L}{\partial \dot{\phi}} = \left(\frac{\hbar}{2e}\right)^2 C\dot{\phi} = \frac{\hbar}{2e}CV \equiv \hbar n$, dove n è il numero di coppie di Cooper sul condensatore. Dunque

$$H(n, \phi) = E_C n^2 - E_J \cos\phi - \frac{\hbar}{2e}I_b\phi$$

per cui possiamo associare un termine di "massa". Quantizzando $\hat{p} = -i\hbar\frac{\partial}{\partial\phi}$, $\hat{n} = -i\frac{\partial}{\partial\phi}$, da cui $[\hat{\phi}, \hat{n}] = i$.

A questo punto vogliamo ottenere da un'ansa del potenziale $U(\phi)$ due livelli energetici per la costruzione del nostro qubit. In particolare, occorre trovare una relazione tra i parametri di U e quelli che compaiono in (B.1). Da quanto detto sopra $U(\phi) = -E_J\left(\cos\phi + \frac{I_b}{I_c}\phi\right)$, per cui nel punto di minimo $U'(\phi) = E_J\left(\sin\phi - \frac{I_b}{I_c}\right) = 0$. Possiamo risolvere sviluppando la funzione seno attorno a $\frac{\pi}{2}$, da cui si ottengono le radici $\phi_0^{\pm} = \frac{\pi}{2} \pm \sqrt{2}\sqrt{1 - \frac{I_b}{I_c}}$. Il minimo corrisponde al valore ϕ_0^-, da cui $U(\phi) \approx U(\phi_0^-) + \frac{E_J\cos\phi_0^-}{2}(\phi - \phi_0^-)^2$. Possiamo introdurre una frequenza $\omega_p \approx \frac{E_J\cos\phi_0^-}{m} \approx \sqrt[4]{2}\omega_J\sqrt[4]{1 - \frac{I_b}{I_c}}$ in funzione della quale possono essere espressi i livelli prossimi al fondamentale $E_n = \hbar\omega_p\left(n + \frac{1}{2}\right)$. Limitandoci ai primi due livelli possiamo scrivere $H = -\frac{1}{2}\epsilon\sigma^3$, $\epsilon = \hbar\omega_p$, a meno di un termine costante ininfluente. Tenendo anche i termini cubici si può vedere che $E_1 - E_0 = \hbar\omega_p\left(1 - \frac{5}{36}\frac{\hbar\omega_p}{\Delta u}\right)$ e $E_2 - E_1 = \hbar\omega_p\left(1 - \frac{10}{36}\frac{\hbar\omega_p}{\Delta u}\right)$, con $\Delta u \stackrel{\text{def}}{=} U(\phi_0^+) - U(\phi_0^-) = E_J 2\sqrt{2}\left(1 - \frac{I_b}{I_c}\right)^{\frac{3}{2}}$. Se la quantità $\frac{\hbar\omega_p}{\Delta u}$ è dell'ordine di 5 o 7 è possibile inviare segnali di frequenza fissata ed eccitare lo stato $|0\rangle$ nello stato $|1\rangle$ senza rischiare di eccitare fino al secondo livello, che richiederebbe diversa frequenza.

La lettura avviene abbassando la barriera (manipolando I_b) in modo che, se il sistema si trova nel primo stato eccitato, esso scivola via. Occorre però segnalare che questo sistema non ha tempi di coerenza particolarmente lunghi.

Infine, un termine del tipo $\Delta\sigma^1$ nell'hamiltoniana può essere introdotto introducendo una corrente di bias alternata, $I_b = I_{\mathrm{DC}} + I(t)\cos(\omega_p t + \varphi)$.

B.2.4. Qubit di carica (Single Cooper pair Box)

Consideriamo ora una giunzione in cui una piccola isola superconduttiva è collegata ad un *bulk* per mezzo di una giunzione tunnel, il cui potenziale è controllato da un voltaggio V_g. Alla giunzione di Josephson sono associati una resistenza R ed una capacità C, mentre C_g è la capacità tra l'isola superconduttiva e la rete. Si parla di *regime di Coulomb blockade* quando $R \gg R_q = \frac{h}{2e^2}$, resistenza quantica, e $k_B T \ll E_C$: in queste circostanze gli elettroni possono essere portati sull'isola uno alla volta e controllati dal potenziale V_g (*single Cooper pair box*, SCB). L'energia elettrostatica ha la forma

$$\frac{CV^2}{2} + C_g\frac{(V_g - V)^2}{2},$$

Dunque, omettendo il termine costante e definendo $C_\Sigma = C + C_g$, la lagrangiana ha la forma

$$L(\dot\phi, \phi) = \frac{C_\Sigma}{2}\left(\frac{\hbar\dot\phi}{2e} - \frac{C_g}{C_\Sigma}V_g\right)^2 - E_J(1-\cos\phi) \Rightarrow p = \frac{\hbar C_\Sigma}{2e}\left(\frac{\hbar}{2e}\dot\phi - \frac{C_g}{C_\Sigma}V_g\right).$$

Passando in formalismo hamiltoniano

$$H = E_C(n - n_g)^2 - E_J\cos\phi,$$

dove $n = \frac{p}{\hbar}$ è il numero di coppie di Cooper, $E_C = \frac{(2e)^2}{2C_\Sigma}$ e $n_g = -\frac{C_g V_g}{2e}$ è la carica sul capacitore controllabile dall'esterno in unità di coppie di Cooper. Supponendo sempre $E_C \ll \Delta$, dove Δ è il gap del superconduttore tra condensato e livello eccitato, è possibile ignorare effetti di singolo elettrone nell'isola superconduttiva (corrispondenti ad uno stato eccitato). Se $E_J = 0$ allora lo spettro è dato da $E_n = E_C(n - n_g)^2$.

Per $E_J \neq 0$ occorre calcolare $\langle n|\cos\hat\phi|m\rangle$, dove $|n\rangle$ l'autostato dell'operatore $\hat n$ di autovalore n. In analogia con le variabili posizione–impulso, da $\hat\phi|\phi\rangle = \phi|\phi\rangle$ possiamo scrivere $\langle n|\phi\rangle = \frac{\mathrm{e}^{in\phi}}{\sqrt{2\pi}}$, per cui $\langle n|\,\mathrm{e}^{i\hat\phi}\,|m\rangle = \int_0^{2\pi}\frac{\mathrm{d}\,\phi}{2\pi}\,\mathrm{e}^{i\phi(n-m+1)} = \delta_{n,m-1}$. Se ci limitiamo ora a considerare gli stati $|0\rangle$ e $|1\rangle$, la matrice hamiltoniana che si ottiene ha la forma di una hamiltoniana per qubit, $H = -\frac{1}{2}(\epsilon\sigma^3 + E_J\sigma^1)$, dove $\epsilon = E_C\left[(1-n_g)^2 - n_g^2\right] = E_C(1-2n_g)$ è il gap tra i due livelli relativo al termine "cinetico". Lo spettro è dato dunque dagli autovalori $E_\pm = \pm\frac{1}{2}\sqrt{E_C^2(1-2n_g)^2 + E_J^2}$ della matrice hamiltoniana, mentre gli autostati sono $|E_\pm\rangle = |\pm\rangle$.

B.2.5. Qubit di flusso (rf-SQUID)

Un rf-SQUID è costituito da un loop superconduttivo su cui è presente una giunzione di Josephson, attraversato trasversalmente da un campo magnetico con un flusso

Φ_e. Il circuito può essere modellizzato come una resistenza R, una giunzione di Josephson, un capacitore C ed una induttanza L da considerarsi tutti in parallelo. Come sopra $\phi(t) = \frac{2e}{\hbar}\int^t V(\tau)\,\mathrm{d}\,\tau + \phi_e$, $\phi_e \stackrel{\text{def}}{=} 2\pi\frac{\Phi_e}{\Phi_0}$ fase indipendente dal tempo legata al flusso di campo magnetico Φ_e. La corrente associata all'induttanza è $I_L = \frac{\hbar}{2eL}(\phi - \phi_e)$ (ricordando che $V = L\frac{\mathrm{d}I_L}{\mathrm{d}t}$), mentre l'equazione di Kirchhoff è $\frac{\hbar}{2e}C\ddot{\phi} + \frac{\hbar}{2eR}\dot{\phi} + I_c\sin\phi + \frac{\hbar}{2eL}(\phi - \phi_e) = 0$. Tramite il flusso esterno ϕ_e è possibile controllare la corrente che scorre nel loop superconduttivo. Possiamo identificare due stati fisici: $|R\rangle$, associato a corrente destrorsa, e $|L\rangle$, associato a corrente sinistrorsa ($I_c = 0$ corrisponde al caso di un semplice oscillatore). Come sopra, possiamo scrivere una lagrangiana *ignorando il contributo dissipativo, $R \to \infty$,*

$$L(\phi, \dot{\phi}) = \frac{\hbar^2}{4E_C}\dot{\phi}^2 - E_J(1 - \cos\phi) - E_L(\phi - \phi_e)^2,$$

dove $E_L \stackrel{\text{def}}{=} \frac{\Phi_0^2}{4\pi^2 L}$.

Se $\Phi_e = m\Phi_0$, $m \in \mathbb{N}$, il potenziale ha un minimo in $\phi = \phi_e$, mentre si hanno due minimi distinti, simmetrici rispetto a ϕ_e, se Φ è multiplo semintero di Φ_0, corrispondenti a correnti nelle due direzioni. In presenza di tunneling i due stati separati (uno per buca) si sovrappongono e forniscono due stati ibridi globali. Alterando ϕ_e (ad esempio $\phi_e = (1 - 0.3)\pi$) si può rendere una buca molto più profonda dell'altra ed inizializzare lo stato.

B.3. Accoppiamenti

Esistono diversi possibili modi per accoppiare i qubit precedentemente introdotti.

Accoppiamento induttivo di due qubit di flusso: due qubit di flusso vengono accoppiati e interagiscono inducendo un flusso ciascuno nell'altro con la relazione $\Phi_i = \sum_k L_{ik}\Phi_k$, dove i termini fuori diagonale $L_{12} = L_{21}$ governano l'interazione (questo accoppiamento è tuttavia poco flessibile).

Accoppiamento capacitivo di qubit di carica: due qubit di carica vengono accoppiati tramite un capacitore C_3 tra le isole di carica. Il contributo all'hamiltoniana del capacitore è $\delta H = \frac{C_3 V_3^2}{2} = \frac{C_3}{2}\left(\frac{\hbar}{2e}\right)^2(\dot{\phi}_1 - \dot{\phi}_2)^2$. Indicando con $C_{12} \equiv C_3$ e $C_{ii} = C_{\Sigma,i} + C_3$ possiamo scrivere la lagrangiana come $\frac{1}{2}\left(\frac{\hbar}{2e}\right)^2\sum_{i,j=1}^2 C_{ij}\dot{\phi}_i\dot{\phi}_j - \frac{\hbar}{2e}\sum_{i=1}^2 C_{gi}V_{gi}\dot{\phi}_i - \sum_{i=1}^2 E_{J_i}(1 - \cos\phi_i)$. L'hamiltoniana allora ha la forma

$$H = E_{C_1}\underbrace{(n_1 - n_{g_1})^2}_{\sigma_1^3\otimes\mathbb{1}_2} + E_{C_2}\underbrace{(n_2 - n_{g_2})^2}_{\mathbb{1}_1\otimes\sigma_2^3} + E_C\underbrace{(n_1 - n_{g1})(n_2 - n_{g2})}_{\sigma_1^3\otimes\sigma_2^3} + \sum_{i=1}^2 E_{J_i}\cos\phi_i.$$

Accoppiamento LC tra qubit di carica: esistono sistemazioni più flessibili in cui due qubit di carica sono in parallelo con un capacitore ed un'induttanza.