Alma Mater Studiorum · Università di Bologna

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

Tecniche di Tensor Networks per sistemi a molti corpi: il caso 1D con MPS e DMRG

Relatore:
Prof. Lorenzo Piroli

Presentata da:
Alberto Zaghini

Anno Accademico 2024/2025

Sommario

L'approccio al problema a molti corpi è in generale fortemente limitato dallo scaling esponenziale nel numero di componenti dei costi computazionali. Tuttavia, per un'ampia gamma di modelli a bassa dimensionalità vincoli sulla struttura di entanglement degli stati fondamentali rendono possibile uno studio efficiente, tramite algoritmi di rinormalizzazione basati su *ansatz* che catturano ottimamente tale struttura. Entrambi possono essere formulati naturalmente in un linguaggio diagrammatico, quello dei Tensor Network.
In questa trattazione si darà un'introduzione generale e si considererà poi in particolare il caso unidimensionale, introducendo i Matrix Product States e il Density Matrix Renormalization Group ed illustrandone una semplice applicazione.

Indice

Elenco delle figure

I would like to make a confession which may seem immoral: I do not believe in Hilbert space anymore.

John Von Neumann

Introduzione

La sfida fondamentale della fisica quantistica dei sistemi a molti corpi è quello della rappresentazione e manipolazione della vasta quantità di informazioni necessaria per caratterizzare lo stato di sistemi formati da un numero elevato di componenti (infinito nel limite termodinamico), da cui se ne possono calcolare le proprietà fisiche. In linea di principio si tratta di un problema molto arduo sia a livello teorico che numerico, a causa della *maledizione della dimensionalità*, ovvero la crescita esponenziale nel numero di componenti della quantità di valori da considerare per definire uno stato generico, dovuta al fatto che questo "vive" nello spazio di Hilbert dato dal prodotto *tensore* degli spazi associati alle singole unità, anziché nel semplice prodotto *cartesiano*.

Questa proprietà è all'origine del fenomeno dell'*entanglement* (lett. "intreccio"), che riveste un ruolo fondamentale soprattutto nei sistemi *fortemente correlati*, attualmente il maggior interesse di ricerca nella fisica degli stati condensati per via delle loro grandi potenzialità applicative: si parla in primo luogo di superconduttività ad alta temperatura, ma anche di altre fasi esotiche della materia come quelle topologiche, che permettono un controllo senza precedenti sulle proprietà dei materiali - e quindi la realizzazione di sensori estremamente accurati o piattaforme di computazione e simulazione quantistica facilmente manipolabili ma al contempo molto "robuste". Purtroppo, proprio in virtù dell'interrelazione che si stabilisce tra gli stati dei singoli componenti l'applicazione dei principali strumenti analitici e numerici ai modelli che astraggono le caratteristiche fondamentali di questi sistemi incontra grandi difficoltà: la diagonalizzazione esatta diviene impraticabile a scale ben lontane dal limite termodinamico (ove si manifestano le transizioni di fase), le tecniche perturbative e di campo medio falliscono in presenza di forti correlazioni soprattutto in basse dimensioni, i metodi Monte Carlo sono afflitti dal *sign problem*, ovvero l'emergere di probabilità negative.

Negli ultimi decenni, tuttavia, è emerso che per un'ampia classe di sistemi a bassa dimensionalità la struttura di entanglement e di correlazione tra osservabili è soggetta a importanti vincoli fisici, in virtù dei quali gli stati rilevanti a livello pratico, ovvero quelli fondamentali, sono localizzati in un "angolo" esponenzialmente piccolo dello spazio di Hilbert complessivo del sistema, la cui esplosione si rivela pertanto un problema fittizio. Grazie a ciò è stato possibile lo sviluppo di tecniche

di approssimazione efficienti, ossia che richiedono un impiego di risorse computazionali con *scaling* meramente polinomiale nel numero di componenti sia per la rappresentazione degli stati che per l'estrazione di informazioni.
Queste tecniche si basano sull'impiego di un "linguaggio" diagrammatico ideato originariamente dal fisico matematico Roger Penrose [1] - un grande *visual thinker* - che permette di operare intuitivamente con strutture numeriche complesse: i Tensor Networks. Si tratta di insiemi di tensori, ovvero essenzialmente matrici multidimensionali, rappresentati da forme; queste sono collegate da fili che stabiliscono il modo in cui i vari indici sono *contratti*, ovvero come i corrispondenti ingressi sono da moltiplicarsi e sommarsi per produrre nuovi tensori.
Costruendo istanze peculiari di tensor networks (di seguito abbreviati in **TN**) che riproducono la struttura di entanglement degli stati a bassa energia, sia per sistemi finiti che nel limite termodinamico, si ottengono degli *ansatz* ottimali per algoritmi di rinormalizzazione, che vanno proprio ad approssimare con bontà arbitraria (è sufficiente accrescere le dimensioni degli indici "di entanglement") il sottoinsieme degli stati fisicamente rilevanti. In particolare in questa trattazione ci si concentrerà sugli Stati Prodotto di Matrici (MPS), che modellano sistemi unidimensionali, e il corrispondente Gruppo di Rinormalizzazione della Matrice Densità (DMRG). Storicamente hanno ricoperto un ruolo pionieristico per l'approccio con TN al problema a molti corpi, rappresentando un significativo salto in avanti con la loro introduzione negli anni '90, ma costituiscono tuttora lo stato dell'arte nel proprio ambito.
Va in realtà notato che il quadro complessivo che si può tracciare oggi è stato intuito a posteriori: originariamente non si era compreso che il DMRG si basasse sugli MPS ed al contempo non si sapeva spiegare l'efficacia di questi nell'approssimare gli stati fondamentali. Come spesso nella ricerca, e l'analogia più immediata è forse proprio con la rinormalizzazione in teoria dei campi, prima è giunta la rivoluzione in termini di efficienza e precisione dei calcoli e solo successivamente è stata elaborata una concettualizzazione fisica che ha permesso di giustificarne il successo, illuminando le strutture profonde catturate dal formalismo.

Capitolo 1

Tensori e Tensor Networks

Si introduce innanzitutto il "linguaggio" diagrammatico che si adopererà per costruire le rappresentazioni degli stati quantistici e applicarvi le opportune manipolazioni, in modo efficiente e soprattutto intuitivo. Si descrive l'operazione fondamentale di contrazione, le possibili decomposizioni e si quantifica lo scaling delle risorse computazionali richieste.

1.1 Tensori

A livello pratico, è sufficiente la definizione dei tensori come generalizzazioni delle matrici, ovvero array numerici multidimensionali. Tuttavia, analogamente a quanto possibile con queste e gli operatori lineari su spazi vettoriali, si è in grado anche in questo caso di fornire una definizione intrinseca, che prescinde dalla scelta di specifiche basi [2, 3, 4]:

Definizione 1.1. *Sia V spazio vettoriale d-dimensionale su un generico campo $\mathbb{K}$*[1]*. Si definisce lo spazio duale V^* come spazio dei funzionali lineari su V a valori in K. Un* ***tensore di tipo*** *(p, q) è una mappa multilineare:*

$$T : \underbrace{V^* \times \cdots \times V^*}_{p} \times \underbrace{V \times \cdots \times V}_{q} \to \mathbb{K} \tag{1.1}$$

che associa quindi a q vettori e p covettori uno scalare. Equivalentemente può essere definito come un elemento del prodotto tensore:

$$T \in \underbrace{V \otimes \cdots \otimes V}_{p} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q} \tag{1.2}$$

[1] Di norma si considerano campi algebricamente chiusi. Specificamente, nella trattazione di tensori che descrivono stati di sistemi quantistici si ha sempre $K = \mathbb{C}$, in conseguenza del primo postulato (si faccia riferimento al capitolo 2).

Il prodotto tensore di due spazi V e W sul medesimo campo può essere definito, fissate due basi (canoniche) $\{\mathbf{v}_i\} \subseteq V$ e $\{\mathbf{w}_j\} \subseteq W$, come lo spazio generato da:

$$\{\mathbf{v}_i \otimes \mathbf{w}_j \,:\, 1 \leq i \leq \dim V \,,\, 1 \leq j \leq \dim W\} \tag{1.3}$$

O più formalmente come lo spazio $V \otimes W$ assieme ad una mappa $\phi : V \times W \to V \otimes W$ che soddisfa la **proprietà universale**: $\forall\, g : V \times W \to U$ (U spazio generico) $\exists g^*$ lineare t.c. il seguente diagramma commuta:

$$\begin{array}{ccc} V \times W & \xrightarrow{\;g\;} & U \\ & \searrow^{\phi} & \uparrow{\scriptstyle g^*} \\ & & V \otimes W \end{array} \tag{1.4}$$

L'equivalenza delle due definizioni di tensore si ha considerando l'isomorfismo naturale in dimensione finita tra uno spazio vettoriale ed il suo doppio duale V^{**}[2].

Per i tensori che saranno trattati la definizione è in generale da ampliarsi ammettendo che nelle varie posizioni possano essere collocati spazi differenti da una sola coppia V, V^* (ma sempre su medesimo campo) [4]. Fissando una base su ogni spazio (che ne identifica direttamente una sul rispettivo duale) si può quindi rappresentare il tensore come collezione di scalari $T_{j_1,\dots,j_k} \in \mathbb{K}^{d_1} \times \cdots \times \mathbb{K}^{d_k}$, ove d_k è la dimensione dello spazio corrispondente all'indice k-esimo, che può quindi assumere valori $1 \leq j_k \leq d_k$[3]. Facendo riferimento alla rappresentazione grafica che sarà introdotta a breve, gli indici sono anche definiti **gambe** del tensore; il loro numero totale definisce l'**ordine** k del tensore [2]. La **dimensione** del tensore è invece il prodotto delle dimensioni degli indici $d = \prod_j d_j$; la denominazione è appropriata in quanto corrisponde alla dimensione dello spazio vettoriale cui appartengono tensori dello stesso *tipo*[4]. Questi possono quindi essere combinati linearmente: per la rappresentazione in componenti il tensore risultante dalla somma e dal prodotto per scalare si ottiene effettuando le operazioni componente per componente. Si noti che è possibile reinterpretare le entità algebriche elementari come istanze particolari di tensori: gli scalari sono tensori di ordine 0, i vettori di ordine 1, le matrici (operatori lineari) di ordine 2 [5].

Di seguito si farà implicitamente uso della definizione "da lavoro" dei tensori, ma sarà fatto riferimento ove possibile a quella intrinseca in quanto permette una comprensione più chiara e profonda delle rappresentazioni che si andranno a costruire. In generale gli spazi associati alle gambe sono spazi di Hilbert complessi $\mathcal{H}_i$ finito

[2] L'azione di un vettore su un covettore è definita come la valutazione del covettore sul vettore.

[3] Nella letteratura sui TN non si distinguono normalmente indici covarianti e controvarianti.

[4] Per la definizione intrinseca la proprietà universale implica $\dim(V \otimes W) = \dim V \cdot \dim W$

dimensionali, dunque isomorfi a $\mathbb{C}^{d_i}$ [2]. Considerando tensori di ordine elevato e operazioni che ne combinino diversi per costruirne di nuovi, tenere conto dei vari indici in gioco e delle corrispondenti dimensioni può essere molto complesso. Per questo è di grandissima utilità introdurre una rappresentazione grafica dei tensori, come forme (identificate dalla denominazione) da cui si protendono le citate gambe corrispondenti agli indici [5].

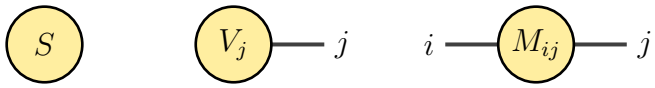


Figura 1.1: Rappresentazione di uno scalare S, un vettore V ed una matrice M

Sui tensori è possibile effettuare una serie di operazioni, che possono essere generalmente suddivise in due categorie: manipolazioni della forma, ovvero splitting, fusione e prodotto tensore, e operazioni di algebra lineare, ovvero contrazioni e decomposizioni. A livello computazionale è importante notare che solo le seconde richiedono operazioni con numeri a virgola mobile (FLOPs) per trattare gli ingressi in $\mathbb{C}$ e sono pertanto più complesse (quindi costose) e prone ad errori [6].

1.1.1 Manipolazione degli indici

Una conseguenza delle varie definizioni è che spazi di tensori con solamente la stessa d sono isodimensionali e pertanto isomorfi. A livello pratico, ciò significa che è possibile *ridimensionare* i tensori *fondendo* e *separando* gli indici [7].

Fusione Un generico tensore di ordine k si può sempre ridimensionare per ottenerne uno di ordine $k' < k$ tramite l'operazione di **fusione** degli indici.

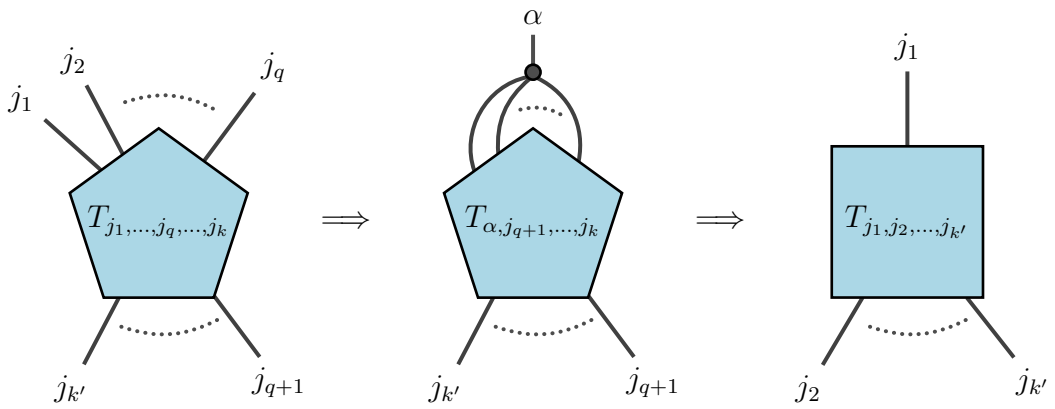


Figura 1.2: Fusione di indici

Per effettuarla si identificano $q \leq k$ indici da riunire in uno solo; ognuno corrisponde ad uno spazio vettoriale locale con rispettiva base.

Il singolo indice α risultante corrisponderà alla base complessiva del prodotto tensore di questi spazi, che è a sua volta uno spazio a dimensione finita:

$$\{|j_1, \cdots, j_q\rangle \equiv |\alpha\rangle \ : \ \alpha = 1, \ldots, d = \prod_{i=1}^{q} d_i\} \tag{1.5}$$

Si ha così un tensore di ordine $k' = k - q + 1$. L'operazione è reversibile: è sufficiente invertire la corrispondenza univoca tra α e le combinazioni degli indici originari.

Splitting Viceversa è possibile effettuare lo **splitting** di un indice, che è semplicemente l'inverso della fusione. Si noti che non essendo fissate le dimensioni degli spazi in cui si vuole decomporre quello di partenza, sottoposte solo al vincolo che il loro prodotto dia la dimensione dell'indice originario, l'operazione non è in generale definita univocamente.

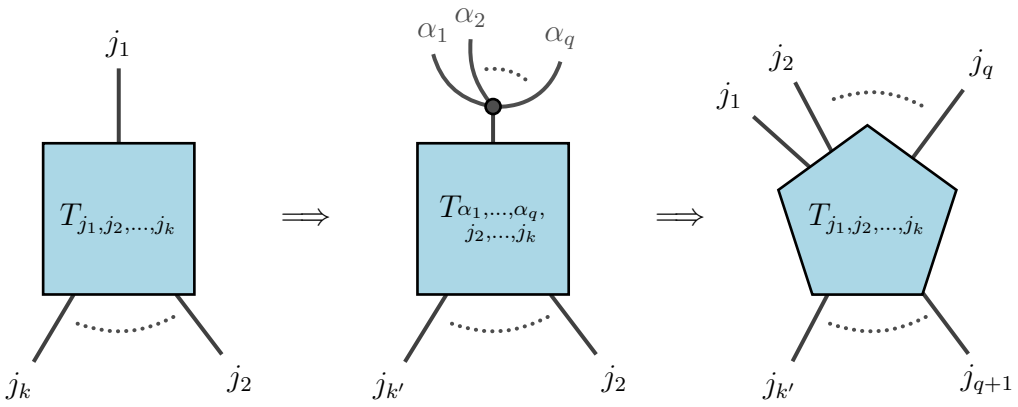


Figura 1.3: Splitting di indici

Prodotto tensore Infine, è possibile combinare tensori privi di connessioni fra loro per ottenerne uno di ordine maggiore tramite l'operazione di **prodotto tensore**; in generale in notazione indiciale il simbolo $\otimes$ viene omesso. A livello pratico l'operazione si compone di due parti: prima si fondono indipendentemente i due gruppi di indici coinvolti dei due tensori, quindi si fondono i due indici risultanti in un singolo indice del tensore finale [2].
Diagrammaticamente viene indicato accostando i tensori coinvolti o più praticamente inglobandoli in una sola forma che conserva le gambe preesistenti. In generale tensori disconnessi possono muoversi relativamente in modo libero all'interno di un diagramma, ovvero è possibile effettuarne una **deformazione piana** [4]. A livello simbolico ciò significa che vale:

$$(\mathbb{I} \otimes B)(A \otimes \mathbb{I}) = A \otimes B = (A \otimes \mathbb{I})(\mathbb{I} \otimes B) \tag{1.6}$$

Si tratta di una proprietà fondamentale in quanto rende le manipolazioni diagrammatiche molto più immediate di quelle algebriche.

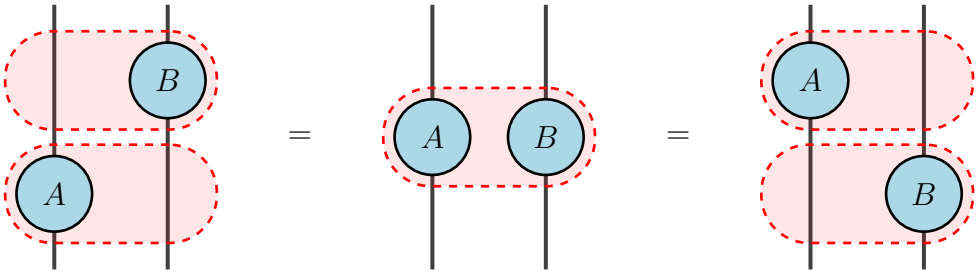


Figura 1.4: Prodotto tensore

1.2 Contrazione

L'operazione fondamentale di contrazione è alla base delle manipolazioni permesse dai TN e rappresenta la generalizzazione del prodotto matriciale [2]. Dati due tensori A e B di ordine k e q e individuati due sottoinsiemi di l indici di entrambi da contrarre fra loro, graficamente sono connesse le gambe corrispondenti e si produce un nuovo tensore C di ordine $k + q - 2l$ le componenti sono date secondo:

$$\sum_{j_1,\ldots,j_l} A_{m_1,\ldots,m_{k-l},j_1,\ldots,j_l} B_{j_1,\ldots,j_l,n_1,\ldots,n_{q-l}} = C_{m_1,\ldots,m_{k-l},n_1,\ldots,n_{q-l}} \tag{1.7}$$

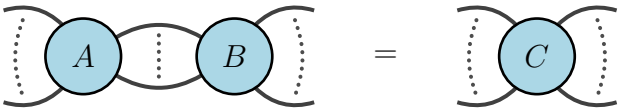


Figura 1.5: L'operazione di contrazione

Gli indici non contratti sono detti **liberi**. L'espressione delle contrazioni può essere semplificata, come si farà di seguito, facendo uso della convenzione di Einstein, per cui si somma implicitamente sugli indici ripetuti [4]. É essenziale che le gambe contratte siano riferite ai medesimi spazi [2].

1.2.1 Efficienza

Facendo uso delle manipolazioni viste in precedenza si può sempre esprimere una contrazione con ordini dei tensori e numero di gambe contratte arbitrari come un'operazione tra tensori di basso ordine, ovvero vettori e matrici: dopo averli ordinati opportunamente, si fondono fra loro tutti gli indici contratti e non contratti rispettivamente e si ottiene quindi l'ordinario prodotto matriciale (le matrici si riducono a vettori se non vi sono indici esclusi dalla contrazione):

$$A_{m,j} B_{j,n} = C_{m,n} \tag{1.8}$$

Splittando infine gli indici liberi residui, si recupera C. Lungi dall'essere utile solo a livello diagrammatico, questa procedura permette di efficientare il calcolo delle componenti del tensore risultante, grazie all'ottimizzazione del prodotto tra vettori e matrici possibile con le tecniche moderne, come quelle implementate nelle librerie LAPACK e BLAS. Infatti, se il costo della contrazione non dipende dal metodo scelto ed è comunque pari al prodotto delle dimensioni degli indici liberi per quelle degli indici contratti[5], ovvero [8]:

$$\text{cost}(A \cdot B) = \frac{d(A) \cdot d(B)}{d(A \cap B)} \tag{1.9}$$

(ove $d(\cdot)$ indica la dimensione totale del tensore e $A \cap B$ indica il tensore formato dai soli indici contratti), il tempo richiesto per il calcolo è significativamente minore se si fa uso delle routine ottimizzate [2, 8], grazie alla parallelizzazione [6]. Chiaramente per la fusione iniziale e lo splitting finale necessari per passare dalla forma originaria a quella matriciale e viceversa è aggiunto un *overhead*; però di norma questo non inficia significativamente il vantaggio computazionale in quanto i pacchetti specializzati nella manipolazione di tensori non modificano la posizione in memoria delle componenti, bensì solo il modo in cui sono indicizzate.
In realtà esistono algoritmi, come quello di Strassen, che permettono di ridurre anche il costo, passando per matrici quadrate $n \times n$ da uno scaling $\mathcal{O}(n^3)$ a $\mathcal{O}(n^{2.8})$ [9]; tuttavia non trovano ampio utilizzo pratico in quanto il vantaggio si ottiene per dimensioni molto elevate ed a fronte di una maggiore *instabilità numerica*, ovvero minore robustezza rispetto ad errori in input [10].

1.2.2 Traccia e norma

Dato un tensore con due indici di medesima dimensionalità, si può definire la **traccia (parziale)** rispetto ad essi come il tensore di ordine $k-2$ prodotto dalla loro contrazione [6]:

$$[\text{Tr}_{p,q}\, T]_{j_1,\dots,j_{p-1},j_{p+1},\dots,j_{q-1},j_{q+1},\dots,j_k} = T_{j_1,\dots,j_{p-1},\boldsymbol{\alpha},j_{p+1},\dots,j_{q-1},\boldsymbol{\alpha},j_{q+1},\dots,j_k} \tag{1.10}$$

Abusando della notazione, si dice **traccia** di due tensori A e B appartenenti allo stesso spazio lo scalare risultante dalla contrazione di tutti gli indici corrispondenti, ovvero dalla traccia parziale del prodotto su tutte le coppie di indici:

$$\text{Tr}(A, B) = A_{j_1,\dots,j_k} B_{j_1,\dots,j_k} \tag{1.11}$$

Dalla rappresentazione diagrammatica dell'operazione è evidente la proprietà di

[5] Intuitivamente: per ogni combinazione degli indici liberi del tensore risultante, bisogna moltiplicare tutte le coppie di ingressi per ogni combinazione di indici contratti e quindi sommare.

ciclicità, per cui vale $\mathrm{Tr}(A, B) = \mathrm{Tr}(B, A)$ [7]. Questa è facilmente estesa a permutazioni cicliche della traccia di più tensori (a questo punto non necessariamente dello stesso tipo) se questi formano un percorso chiuso [5].

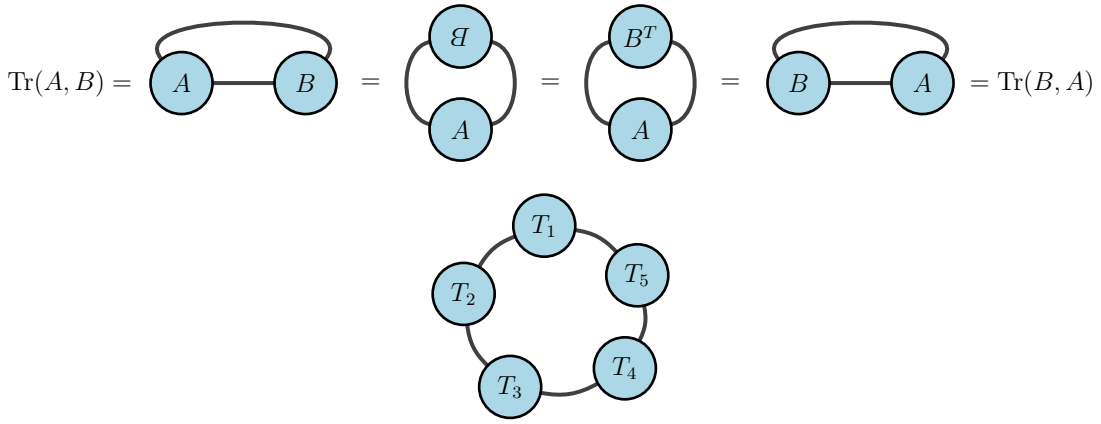


Figura 1.6: Traccia: proprietà ciclica per due tensori e generalizzazione

Tramite la traccia è possibile definire in ciascuno spazio di tensori a data dimensionalità un prodotto scalare (o interno), ovvero una forma sesquilineare hermitiana definita positiva, che generalizzando il caso matriciale viene definita **prodotto scalare di Frobenius**:

$$\langle A, B \rangle_F \equiv \mathrm{Tr}(A^*, B) \tag{1.12}$$

ove A^* indica il tensore complesso coniugato. Tale p.s. induce una norma, detta **norma di Frobenius** o **di Hilbert-Schmidt** [6]:

$$\|A\| = \sqrt{\mathrm{Tr}(A^*, A)} = \sqrt{A^*_{j_1,\dots,j_k} A_{j_1,\dots,j_k}} = \sqrt{\sum_{j_1,\dots,j_k} |A_{j_1,\dots,j_k}|^2} \tag{1.13}$$

Questa è particolarmente utile in quanto genera secondo $d(A, B) = \|A - B\|$ una metrica d che permette di quantificare la bontà delle approssimazioni. Si noti che la norma è invariante per splitting e fusione degli indici e che analogamente a quella matriciale soddisfa la proprietà di **sub-moltiplicatività** [11]:

Proposizione 1.1. *Sia C il tensore prodotto dalla contrazione di uno o più indici di una coppia di tensori A e B. Allora se A e B hanno norma finita anche C ha norma finita e specificamente vale*[6]

$$\|C\| \leq \|A\| \, \|B\| \tag{1.14}$$

[6] Si verifica facilmente applicando la diseguaglianza di Cauchy-Schwarz

Grazie alla sub-moltiplicatività e alla diseguaglianza triangolare l'errore di approssimazione sui vari tensori di una decomposizione controlla quello sull'intero tensore decomposto[7]. Considerando la contrazione di due tensori (la generalizzazione è immediata), se si approssima A con $\tilde{A}$ e B con $\tilde{B}$ di modo che l'errore di troncamento per ciascun tensore sia controllato dal medesimo ε, ovvero si abbia $\|A - \tilde{A}\| \leq \varepsilon$ e $\|B - \tilde{B}\| \leq \varepsilon$, allora detto $\tilde{C} = \tilde{A}\tilde{B}$ il tensore complessivo approssimato, per l'errore si ha:

$$\begin{aligned} \|C - \tilde{C}\| = \|AB - \tilde{A}\tilde{B}\| = \|AB - A\tilde{B} + A\tilde{B} - \tilde{A}\tilde{B}\| \leq \\ \leq \|A(B - \tilde{B})\| + \|(A - \tilde{A})\tilde{B}\| \leq \|A\|\|B - \tilde{B}\| + \|A - \tilde{A}\|\|\tilde{B}\| \leq \\ \leq \varepsilon\|A\| + \varepsilon\left(\|B\| + \varepsilon\right) = \left(\|A\| + \|B\|\right)\varepsilon + \varepsilon^2 \end{aligned} \tag{1.15}$$

che può essere a sua volta controllato da $C\varepsilon$, con una costante C opportuna dipendente dalle sole norme dei tensori di partenza, per $\varepsilon \to 0$. Come si vedrà di seguito, è tuttavia possibile migliorare questa stima applicando opportune decomposizioni ai tensori, di modo da rimuovere la dipendenza dalle norme.

1.3 Tensori rilevanti

Un classe di tensori triviali ma di particolare importanza è quella dei tensori **identità**, che esprimono l'azione del delta di Kronecker δ_{ij}; graficamente sono rappresentati come semplici fili [2][8]. La loro generalizzazione multidimensionale è il tensore copia (**copy tensor**), che impone l'uguaglianza di tutti gli indici associati alle gambe. Corrisponde pertanto al delta multidimensionale $\delta_{j_1 \cdots j_n}$, motivo per cui è definit anche tensore *superdiagonale*; può essere ridotto alla composizione dei delta ordinari per coppie di indici[9]:

$$\delta_{j_1 \cdots j_n} = \delta_{j_1 j_2}\delta_{j_2 j_3} \cdots \delta_{j_{n-1} j_n} \tag{1.16}$$

o graficamente alla concatenazione di tensori identità. Questa proprietà può essere espressa anche in senso inverso, ove è indicata come **fusion rule**: la contrazione di uno o più indici di due copy tensor (incluso il caso di ordine 1, quindi l'identità) produce un copy tensor sugli indici liberi residui [12].

[7] A patto che le norme siano finite, un assunto generalmente valido

[8] É importante notare che la rappresentazione non fornisce informazioni sugli specifici spazi vettoriali associati alle gambe: l'identità agisce come semplice connessione. Gli spazi divengono rilevanti una volta promosse le identità a operatori, operazione possibile contraendole con gli operatori $|i\rangle\langle j|$, ovvero ottenendo una risoluzione dell'identità [2]. Considerazioni analoghe valgono per copy e swap tensor.

[9] É equivalente una qualsiasi permutazione dell'ordine.

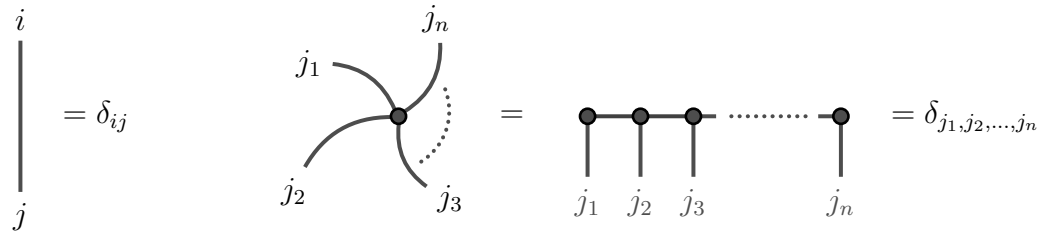


Figura 1.7: Tensore identità e copy tensor

Ciò dà un importante vantaggio computazionale nel caso della contrazione di più indici di due copy tensor, in quanto permette di ridurla a quella di un singolo indice senza alcuna approssimazione.
Con due tensori identità si può anche realizzare un tensore di scambio (**swap tensor**), che agendo su un sistema composto descritto dal prodotto tensore di spazi identici inverte l'ordine di due di questi. Qualsiasi permutazione più complessa è riducibile alla composizione di singoli scambi [2].

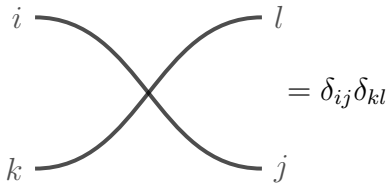


Figura 1.8: Swap tensor

1.4 Tensor Network

É ora possibile dare la definizione di **tensor network** [2, 7]:

Definizione 1.2. *Un tensor network è un insieme di tensori con una parte degli indici contratti sistematicamente. Il criterio di contrazione è specificato tramite una rappresentazione diagrammatica, sotto forma di rete.*

1.4.1 Efficienza: rappresentazione

Come detto in precedenza, i TN permettono in generale una *rappresentazione* efficiente dello stato di un sistema composto. Per definizione, ciò richiede che il numero totale di parametri sia polinomiale nel numero di componenti e dunque che siano controllati sia il numero di tensori che la loro dimensione.
Se N è il numero di componenti, si ha quindi per il numero di tensori $N_{TT} = \mathcal{O}(\text{poly}(N))$ (talvolta anche costante, $N = \mathcal{O}(1)$), mentre per il numero di parametri

necessari a definire ciascuno di essi vale[10]:

$$m(T) = \mathcal{O}\left(\prod_{i_T=1}^{\operatorname{rank} T} d_{i_T}\right) \tag{1.17}$$

e di conseguenza:

$$m(T) = \mathcal{O}\left(d_T^{\operatorname{rank} T}\right) \quad \text{ove} \quad d_T = \max_{1 \leq i_T \leq \operatorname{rank} T} d_{i_T} \tag{1.18}$$

Unendo il tutto, il numero di parametri totali scala come:

$$m_{tot} = \sum_{\{T\}} m(T) = \sum_{\{T\}} \mathcal{O}\left(d_T^{\operatorname{rank} T}\right) = \mathcal{O}\left(\operatorname{poly}(N)\operatorname{poly}(d)\right) \tag{1.19}$$

ove $d = \max_{\{T\}} d_T$ [5]. Quest'efficienza di rappresentazione è però ottenuta aggiungendo degli indici non liberi, ovvero nel caso di sistemi a molti corpi non riferiti agli spazi locali dei vari componenti, che vengono contratti all'interno del network, che vengono definiti indici di legame o ancillari (**bond** o **ancillary indices**). Analogamente si definisce la loro dimensione come **bond dimension**, termine che riferito all'intero TN indica il massimo valore tra i vari bond indices (si indica di norma con χ). Essendo in principio non vincolate, le bond dimension possono essere regolate a seconda del grado di approssimazione che si intende ottenere; incrementandole arbitrariamente si può in linea di principio rappresentare qualsiasi stato, anche quelli non soggetti ai vincoli fisici che rendono efficiente l'uso dei TN, chiaramente al prezzo di perdere la ratio del loro impiego. In ogni caso, di norma è proprio la bond dimension a determinare la d del TN [5, 6].
Nonostante non siano indici associati agli stati dei vari componenti concreti del sistema, i bond indices hanno un importante significato *fisico*: caratterizzano infatti la struttura di entanglement del sistema, e le loro dimensioni forniscono una misura *quantitativa* del grado di correlazione quantistica tra i vari siti [5].

1.4.2 Efficienza: contrazione

É essenziale notare che un'efficienza nella rappresentazione in memoria non si traduce, come nel caso di alcune tipologie di TN, in un'efficienza nella *manipolazione* e quindi estrazione di informazioni rilevanti: in primo luogo questo vale per le contrazioni. Un TN è detto **esattamente contraibile** se il costo della sua contrazione è polinomiale nella dimensione dei tensori elementari [12, 13].
L'approccio di "forza bruta" alla contrazione di un tensor network prevederebbe di effettuare tutte le contrazioni come una singola operazione, con una serie di loop iterativi (`for`) nidificati [8]. Chiaramente non è quello ottimale, come si può verificare da un semplice esempio.

[10] Una denominazione alternativa per l'ordine di un tensore è "rango" [12]. Non la si utilizza in questa trattazione per non confondere con il rango matriciale, con cui non ha relazione.

Esempio 1.1. Si consideri la contrazione di tre tensori, assunti non sparsi[11]:

$$F_{ijk} = A_{ljm}B_{iln}C_{nmk} \tag{1.20}$$

Siano inoltre tutte le dimensioni uguali ad un dato χ non triviale (ovvero > 1).

- Calcolando a forza bruta ogni elemento di F sono necessarie $\mathcal{O}(\chi^3)$ operazioni, e dunque in totale la contrazione ne richiede $\mathcal{O}(\chi^6)$;
- Se invece si calcola prima il tensore intermedio

$$D_{ijmn} = A_{ljm}B_{iln} \tag{1.21}$$

e quindi

$$F_{ijk} = D_{ijmn}C_{nmk} \tag{1.22}$$

allora sia il costo per ottenere D, essendo un solo indice contratto per ognuno dei $\mathcal{O}(\chi^4)$ elementi, che quello per ottenere infine F, essendo due indici contratti per ognuno dei $\mathcal{O}(\chi^3)$ elementi, è $\mathcal{O}(\chi^5)$.

Questo caso esemplifica una proprietà del tutto generale dei TN. Definendo l'efficienza in termini di costo [8]:

Teorema 1.1. *Per qualsiasi network di 3 o più tensori densi (non sparsi) è sempre almeno altrettanto efficiente suddividere la contrazione in una serie di contrazioni di coppie (pairwise contractions). Per bond dimension non triviali è in generale più efficiente.*

La riduzione a contrazioni di coppie porta con sè un altro vantaggio fondamentale: come visto in precedenza, queste possono essere riformulate come prodotti matriciali tramite manipolazione degli indici e quindi permettere l'impiego di routine altamente ottimizzate di algebra lineare. Dunque in generale consente un efficientamento anche in termini di complessità computazionale.
L'*ordine* di contrazione, detto **bubbling** [7], non è però di principio univoco e la scelta non è generalmente a sua volta indifferente a livello di costo e complessità, come si può verificare in semplici casi pratici [2, 14]:

Esempio 1.2. Si consideri la contrazione di tre tensori:

$$F_{ij} = A_{ikl}B_{km}C_{jlm} \tag{1.23}$$

Come in precedenza si assuma una dimensione uniforme non triviale degli indici $\chi > 1$. Si nota innanzitutto che il costo di forza bruta è $\mathcal{O}(\chi^5)$.

[11] Ovvero con un porzione importante di ingressi non nulli. Un controesempio sono i copy tensor.

- Se si contrae prima A con C, il primo step ha un costo $\mathcal{O}(\chi^5)$, quindi la contrazione di AC con B ha un costo $\mathcal{O}(\chi^4)$;
- Se si contrae invece prima C con B e quindi BC con A (o equivalentemente prima A con B e quindi AB con C), entrambi gli step hanno un costo $\mathcal{O}(\chi^4)$.

Il guadagno computazionale si può comprendere intuitivamente tramite la rappresentazione diagrammatica, contando il numero di gambe del tensore contratto ad ogni step (le gambe sono ordinate in senso orario, partendo dalle "ore 6").

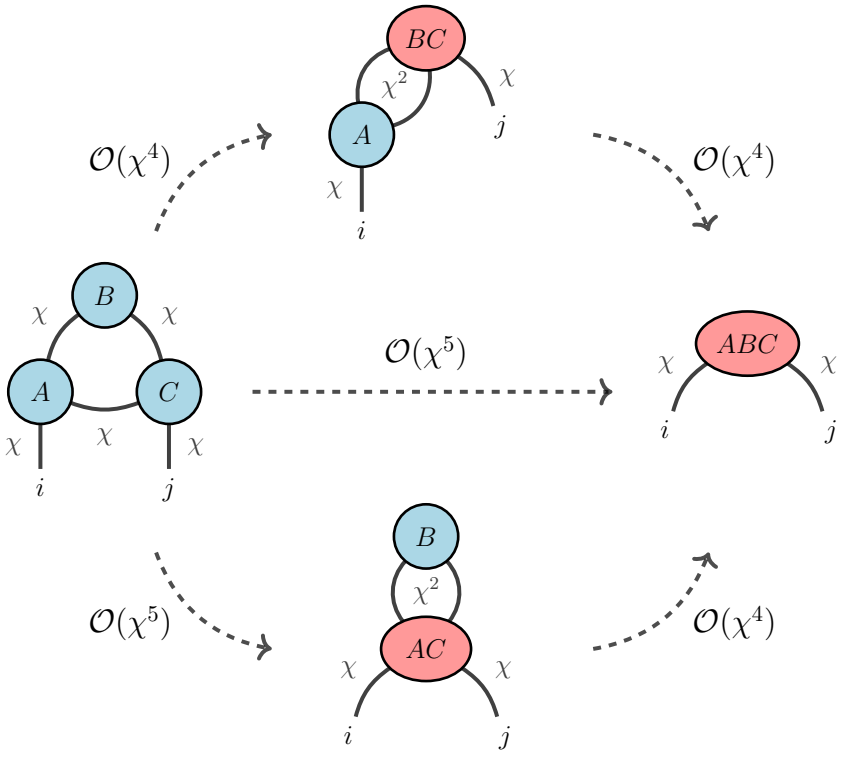


Figura 1.9: Importanza dell'ordine di contrazione

Al contempo, questo esempio illustra il principale vantaggio della rappresentazione diagrammatica: la possibilità di determinare il modo più efficiente in cui contrarre un network tramite l'ispezione visiva. Questo non vale solo per TN con pochi tensori, ma anche e soprattutto per quelli di grandi dimensioni (al limite infiniti) che esibiscono periodicità [7].

Esempio 1.3. Si consideri una doppia fila di tensori di una certa lunghezza, comunque ben maggiore di 2 tensori: è decisamente più conveniente contrarre uno "scalino" alla volta anziché procedere per il lungo, in quanto in questo modo non si dovranno tenere in memoria tensori di ordine superiore a 3, anziché arrivare ad un ordine pari alla lunghezza totale della fila. Scalando esponenzialmente il numero

di componenti nell'ordine (assumendo dimensionalità fissa o anche solo commensurabile degli indici) questo riduce notevolmente lo spazio in memoria necessario e mantiene pressoché costante il costo di ogni contrazione; di conseguenza anche il tempo richiesto per eseguire ogni step rimane controllato.

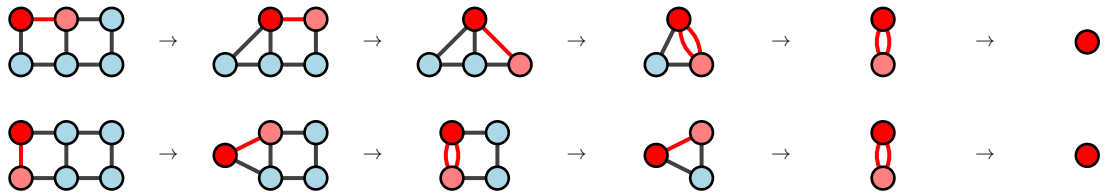


Figura 1.10: Esempio di bubbling inefficiente (sopra) ed efficiente (sotto) per la geometria a "ladder"

Esistono però anche geometrie per cui non è immediata l'individuazione di una strategia di contrazione ottimale; ad esempio talvolta può essere più efficace un approccio *multibubbling*. In generale sono state nel tempo sviluppate euristiche efficaci per network con meno di 20 tensori, che in genere sono sufficienti per gran parte dei casi pratici [8]. Purtroppo è difficile superare questa soglia, in quanto l'algoritmizzazione della ricerca soffre in generale di un'importante limitazione di scalabilità [15, 16, 17]:

Teorema 1.2. *La determinazione dell'ordine di contrazione ottimale a livello sia di costo che di complessità (di spazio o di tempo) è un problem NP-hard e in particolare NP-completo.*

Addirittura possono presentarsi delle situazioni in cui non esiste una sequenza ottimale che permetta di contenere su tutti gli step la quantità di risorse necessarie, nonostante in principio la *rappresentazione* potesse essere efficiente [7].

Esempio 1.4. Un esempio all'apparenza non tanto distante dal caso precedente è quello di una griglia piana con anche l'altra dimensione estesa: è inevitabile che col progredire delle contrazioni si ottenga un tensore di ordine paragonabile alle lunghezze dei lati del network.

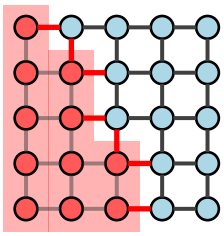


Figura 1.11: Geometria senza bubbling efficiente

1.5 Decomposizione

L'operazione inversa della contrazione è la **decomposizione** di un tensore in più tensori, che è alla base delle tecniche di TN in quanto permette di ridurre un tensore generico a tensori particolari per cui è più agevole la manipolazione e l'estrazione di informazioni rilevanti. Come si può intuire, analogamente allo splitting non è definita univocamente; tuttavia sono di particolari interesse alcune decomposizioni specifiche ottenute generalizzando quelle matriciali. Come per le operazioni precedenti si ottengono trasformando i tensori in matrici, applicandole e quindi invertendo la fusione degli indici [2]. Di seguito si assumerà che le matrici siano in generale complesse.
I costi delle decomposizioni sono espressi in termini delle dimensioni matriciali con $m > n$; considerando la fusione degli indici di un tensore $T_{(j_1,\dots,j_p),(j_{p+1},\dots,j_k)}$ è da assumersi [6]:

$$m = \max\{d_1 \cdot d_2 \cdots d_p, d_{p+1} \cdots d_k\} \qquad n = \min\{d_1 \cdot d_2 \cdots d_p, d_{p+1} \cdots d_k\} \quad (1.24)$$

1.5.1 Decomposizione spettrale (o autodecomposizione)

Una matrice quadrata diagonalizzabile può essere ridotta, tramite una similitudine, ad una diagonale, univocamente definita a meno di un riordino degli autovalori. Considerando l'applicazione dei TN ai sistemi quantistici le matrici diagonalizzabili rilevanti sono quelle hermitiane, per cui la matrice di cambio di base è unitaria:

$$M_{\alpha,\beta} = U_{\alpha,\gamma} D_{\gamma,\gamma} U^{\dagger}_{\gamma,\beta} \qquad U^{\dagger}U = UU^{\dagger} = \mathbb{I} \quad (1.25)$$

Se la dimensione delle matrici è $n \times n$, il costo della decomposizione spettrale scala come $\mathcal{O}(n^3)$ [2].

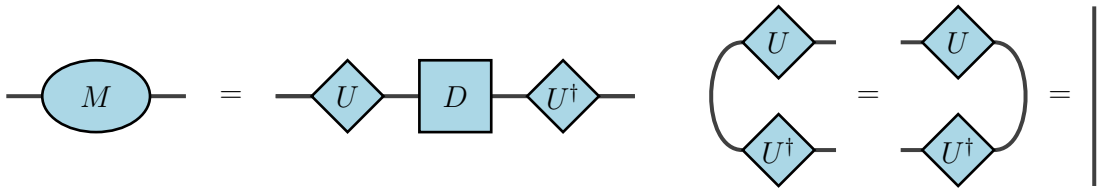


Figura 1.12: Decomposizione spettrale

1.5.2 Decomposizione ai valori singolari (Singular Value Decomposition)

Una decomposizione molto più generale, possibile per qualsiasi matrice di dimensione arbitraria (anche non diagonale) $m \times n$, permette di esprimerla come prodotto

di due matrici unitarie (quindi quadrate) $m \times m$ e $n \times n$, e una matrice diagonale[12] con medesima dimensione di quella di partenza e ingressi solo positivi, definiti **valori singolari**, determinati univocamente da M a meno di un riordino:

$$M_{\alpha,\beta} = U_{\alpha,\gamma} S_{\gamma,\gamma} V^{\dagger}_{\gamma,\beta} \qquad UU^{\dagger} = U^{\dagger}U = \mathbb{I}_m \quad VV^{\dagger} = V^{\dagger}V = \mathbb{I}_n \tag{1.26}$$

In generale il numero di valori singolari non nulli, uguale al rango di M, può essere minore di $\min(m, n)$: ciò permette di esprimere in modo più compatto la decomposizione:

$$M_{\alpha,\beta} = \sum_{\gamma=1}^{r} U_{\alpha,\gamma} S_{\gamma,\gamma} V^{\dagger}_{\gamma,\beta} \tag{1.27}$$

ove S è ora una matrice diagonale quadrata $r \times r$ con ingressi diagonali non nulli, U è $m \times r$ con solo *colonne* ortonormali ($U^{\dagger}U = \mathbb{I}_m$), dette normalizzate *a sinistra* e $V^{\dagger}$ è $r \times n$ con *righe* ortonormali ($VV^{\dagger} = \mathbb{I}_n$), dette normalizzate *a destra*. Assumendo $m > n$, il costo computazionale per la SVD è $\mathcal{O}(mn^2)$ [2].

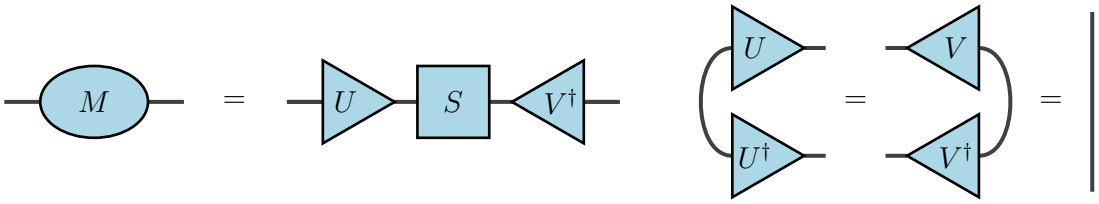


Figura 1.13: Decomposizione ai valori singolari

Si noti che i valori singolari sono le radici quadrate degli autovalori di $MM^{\dagger}$ e $M^{\dagger}M$ e in particolare quindi per una matrice hermitiana sono i valori assoluti degli autovalori. Se M è anche semi-definita positiva autovalori e valori singolari coincidono e le due decomposizioni sono equivalenti [8].
L'impiego della SVD è particolarmente importante in quanto agendo solo sulla matrice S si può ridurre il numero di parametri della rappresentazione tensoriale nel modo che conserva la migliore approssimazione possibile dello stato di partenza a parità di rango. Vale infatti un risultato fondamentale per l'approssimazione di matrici di dimensionalità $m \times n$ arbitraria, il **Teorema di Eckart-Young** [18]:

Teorema 1.3. *La miglior approssimazione di una matrice M data con una matrice di rango $r <$ rank M si ottiene considerandone la SVD e, ordinati in senso decrescente i valori singolari, ponendo a 0 gli ultimi* rank $M - r$.
Inoltre, l'errore di approssimazione (detto ***errore di troncamento*** *[12]) è pari*

[12] Definita in generale come una matrice con tutti gli ingressi fuori dalla diagonale nulli.

alla somma in quadratura dei valori singolari "tagliati"[13]*:*

$$\left\|M - M^{(r)}\right\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |M_{ij} - M_{ij}^{(r)}|^2} = \sqrt{\sum_{i=r+1}^{\operatorname{rank} M} \sigma_i^2(M)} \tag{1.28}$$

Come metrica si utilizza la norma di Frobenius, ma il risultato è stato poi generalizzato da Minsky a norme *unitariamente invarianti*, ovvero tali per cui

$$\|UMV\| = \|M\| \tag{1.29}$$

con qualsiasi U e V unitarie $m \times m$ e $n \times n$ rispettivamente [4].

1.5.3 Decomposizione QR

Una matrice di dimensione arbitraria $m \times n$ può essere decomposta nel prodotto di una matrice unitaria $m \times m$ e una triangolare superiore, ovvero con elementi non nulli solo per indice di colonna maggiore o uguale a quello di riga:

$$M_{\alpha,\beta} = Q_{\alpha,\gamma} R_{\gamma,\beta} \tag{1.30}$$

Se $m > n$, le ultime $m - n$ righe di R sono nulle e quindi si può riscrivere in forma compatta:

$$M_{\alpha,\beta} = \sum_{\gamma=1}^{n} Q'_{\alpha,\gamma} R'_{\gamma,\beta} \tag{1.31}$$

ove Q' è $m \times n$ normalizzata a sinistra e R' $n \times n$.

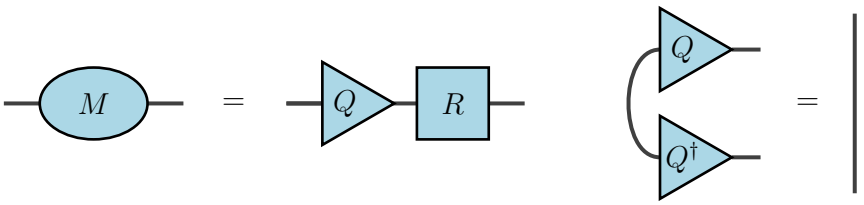


Figura 1.14: Decomposizione QR

Si possono equivalentemente definire la decomposizione RQ con ordine inverso e Q normalizzata a destra o la LQ con L triangolare inferiore [2].
Il costo della QR completa scala come $\mathcal{O}(m^2 n)$, mentre quello della versione compatta o *economica* (utilizzata più spesso negli algoritmi) solo come $\mathcal{O}(mn^2)$ [8].

[13]Dunque eliminando valori singolari nulli non si ha alcun errore, mentre se sono piccoli si mantiene comunque un'ottima approssimazione.

Il fatto che Q per la decomposizione ridotta sia un'isometria a sinistra, ovvero che determini una base ortonormale per lo spazio *delle colonne* della matrice di partenza[14], è di grande utilità a livello pratico: implica infatti che abbia norma unitaria [8].

1.6 Libertà di gauge

Gli indici di bond di un TN che descrive lo stato quantistico di un sistema a molti corpi rappresentano delle dimensioni ulteriori dei tensori non associate agli spazi locali dei vari componenti, ovvero dei gradi di libertà non fisici: dunque è possibile manipolarli senza alterare la fisica del sistema (essenzialmente, i valori di aspettazione delle osservabili) [12]. Questo può avvenire applicando ad un dato indice ausiliario una trasformazione invertibile e la sua inversa (che contraggono fra loro all'identità)[15], le quali sono quindi contratte separatamente con i tensori ai due estremi della gamba producendo i trasformati; trasformazioni più generali - ma sempre implementabili in modo efficiente - si ottengono quindi combinando quelle per singoli link differenti [6].
In analogia alle teorie di campo, le trasformazioni ammesse sono dette trasformazioni **di gauge** del TN. Spesso una variazione del *gauge fixing* al runtime è utile per efficientare gli algoritmi [6].

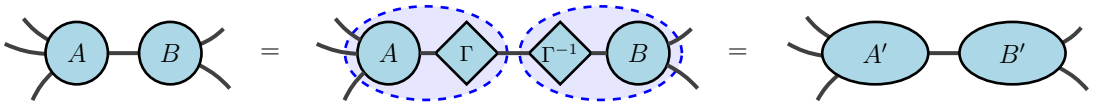


Figura 1.15: Trasformazione di gauge

1.6.1 TN loop-free e gauge locali

Alcune scelte di gauge particolarmente utili si ottengono applicando sistematicamente le decomposizioni introdotte in precedenza qualora il TN esibisca opportune proprietà generali. Una classe molto ampia, che include la tipologia di network su

[14]Le colonne di Q non sono che il risultato dell'applicazione della procedura di ortonormalizzazione di Gram-Schmidt a quelle di M. Si comprende facilmente in questo modo la generalità della decomposizione [19].

[15]In generale è sufficiente che sia un'isometria, in quanto può alterare la dimensione della gamba [7]. Sui link liberi o "fisici" le trasformazioni sono ristrette alle unitarie per mantenere una base ortonormale [6].

cui si concentrerà questa trattazione, è costituita dai TN privi di loop[16] [6]. Considerato la struttura di un TN definisce naturalmente un grafo (non orientato)[17], al netto di opportune tensorizzazioni si può identificare questa classe con i **Tree Tensor Network (TTN)**, in quanto un albero è un grafo connesso privo di loop, o *loop-free* [2, 8]. La struttura di grafo è utile in quanto permette di definire una metrica per i vertici, corrispondente al numero di spigoli che compongono il percorso (*path*) più breve che li connette; per network privi di cicli questo percorso è chiaramente unico. La caratteristica fondamentale dei TTN è che qualsiasi link contratto definisce univocamente una bipartizione del sistema [6].

Preliminarmente è utile dimostrare tramite la SVD un fatto fondamentale: per TN loop-free, se la struttura del network è vincolata le *sole* trasformazioni globali ammissibili sono quelle realizzate combinando risoluzioni dell'identità sui link virtuali. Tramite questa procedura si verifica anche che, per una data gamba virtuale, è possibile determinare chiaramente la corrispondente dimensione di bond minima per mantenere una rappresentazione esatta del network: è il rango della matrice ottenuta ridimensionando il network secondo la partizione degli indici liberi definita dalla gamba stessa [6].
Data una forma generica del network, ne si consideri l'espressione ottenuta fondendo gli indici liberi e contraendo quelli virtuali sui due lati della partizione:

$$T_{\alpha,\beta} = A_{\alpha,\gamma} B_{\gamma,\beta} \tag{1.32}$$

ove la somma è su tutta la dimensione D dell'indice d'interesse. Si operi ora invece la SVD ridotta di T (in questa forma), che è sempre possibile:

$$T_{\alpha,\beta} = \sum_{\gamma}^{\chi} \tilde{A}_{\alpha,\gamma} \tilde{B}_{\gamma,\beta} \tag{1.33}$$

assorbendo i valori singolari in una delle matrici risultanti. La somma è solo su un numero di termini dato dal rango χ, che è determinato dalla sola $T_{\alpha,\beta}$, ovvero dal tensore T e dalla partizione degli indici, e soddisfa $\chi \leq D$ per definizione (o equivalentemente per proprietà della SVD) - chiaramente a patto che la forma 1.32 sia esatta. Ora, se la dimensione di $T_{\alpha,\beta}$ è $m \times n$:

- $\tilde{A}$ è una matrice $m \times \chi$ di rango χ, dunque ha rango colonna massimo (le colonne sono l.i.). Di conseguenza ammette un'inversa a sinistra $\tilde{A}^{-1L}$;
- $\tilde{B}$ è una matrice $\chi \times n$ di rango χ, dunque ha rango riga massimo (le righe sono l.i.). Di conseguenza ammette un'inversa a destra $\tilde{B}^{-1R}$.

[16] Sotto opportune condizioni le manipolazioni illustrate si possono generalizzare anche a network che presentino cicli [8].

[17] Ovvero una coppia $G = (V, E)$, ove V è un insieme di vertici ed E di spigoli che li connettono.

Considerato che $AB = T = \tilde{A}\tilde{B}$, si ottiene quindi una coppia di trasformazioni la cui applicazione al link considerato trasforma il tensore dalla forma 1.32 alla 1.33:

$$\tilde{A} = A\big(B\tilde{B}^{-1R}\big) \equiv AX \qquad \tilde{B} = \big(\tilde{A}^{-1L}A\big)B \equiv YB \tag{1.34}$$

Queste sono $D \times \chi$ e $\chi \times D$ e contraggono all'identità, ma solo in un senso[18]:

$$YX = \mathbb{I}_\chi \qquad XY \neq \mathbb{I}_D \tag{1.35}$$

Iterando questa procedura per ogni gamba virtuale si ottiene quindi una forma equivalente a dimensione di bond minima. Non essendo stata fatta alcuna ipotesi sulla forma 1.32, si tratta di un risultato del tutto generale; inoltre scambiando i ruoli di X e Y si può recuperare la forma originaria dalla 1.33 (con Y in questo caso inversa a destra anziché a sinistra di X). Ma ciò implica che date due qualsiasi rappresentazioni del tensore a forma fissata è possibile determinare un insieme di trasformazioni locali che mappino l'una nell'altra [6].

1.6.2 Gauge centrale e canonica

Il vantaggio fondamentale dell'applicazione sistematica di decomposizioni QR o SVD è risiede nella possibilità di ridurre i nodi a delle isometrie, che semplificano notevolmente le contrazioni e di conseguenza permettono un computo immediato dell'errore di approssimazione complessivo su un network a partire da quello locale [8].

Gauge centrale o unitaria Si fissa innanzitutto un nodo come **centro di ortogonalità**. Tramite la metrica sul network, si possono definire dei **livelli** rispetto ad esso, ovvero insiemi di punti alla medesima distanza. Si applica quindi iterativamente una procedura detta di **pulling-through**, partendo dal livello a distanza massima:

1. Si applica la decomposizione QR ridotta a ciascun tensore del livello;
2. Lo si rimpiazza con l'isometria a sinistra Q e si assorbe la matrice triangolare R nel tensore adiacente nel livello inferiore.

Una volta completata, si ottiene un network in cui tutti i tensori salvo il centro sono isometrie a sinistra. Di conseguenza, sono isometrici gli interi rami che se ne dipartono, ovvero le porzioni connesse di network contratte con ogni sua gamba, che contratti con il proprio aggiunto danno il tensore identità. Ciò è di grande utilità in quanto implica che, detto T il network complessivo e C il centro [8]:

[18] In quanto in generale vale meramente $\chi \leq m, n, D$ e quindi le matrici sono solo isometriche.

- $\|T\| = \|C\|$;
- Più in generale, sostituendo $C \to C'$ (con la medesima forma) si ha $\langle T, T' \rangle = \langle C, C' \rangle$;
- Per l'errore in norma $\|T - T'\| = \|C - C'\|$.

Il costo computazionale del pulling through è dominato dalle decomposizioni; se l'ordine massimo di un tensore del network è K allora è $\mathcal{O}(\chi^{K+1})$[19]. Si noti che facendo uso della QR ridotta si ha il vantaggio di ridurre la dimensione di un link virtuale se questa è maggiore della dimensione complessiva delle altre gambe del tensore decomposto [6].
Esiste in realtà un altro modo, computazionalmente più efficiente ma al contempo più instabile numericamente, per creare un centro di ortogonalità. Anziché manipolare i singoli nodi, si può infatti rendere isometrici i rami nella loro interezza procedendo come di seguito [8]:

1. Si contrae ciascun ramo con il proprio aggiunto, ottenendo una matrice hermitiana ρ, che è semi-definita positiva[20];
2. La si diagonalizza: $\rho = UDU^\dagger$. Se ha autovalori nulli, si tronca la dimensione per eliminarli, di modo da renderla definita positiva;
3. Se ne prende la radice quadrata principale $X = \rho^{1/2} = UD^{1/2}U^\dagger$[21]. Per definizione $\rho = XX^\dagger$;
4. X è invertibile con $X^{-1} = UD^{-1}U^\dagger$; si applica allora come trasformazione di gauge sulla gamba tra il ramo e il centro, assorbendo X^{-1} nel primo e X nel secondo.

Si verifica quindi facilmente che:

$$\rho' = X^{-1}\rho(X^{-1})^\dagger = X^{-1}XX^\dagger(X^{-1})^\dagger = \mathbb{I} \tag{1.36}$$

Spesso negli algoritmi può essere utile *spostare* il centro di ortogonalità: questo può essere fatto semplicemente applicando una procedura analoga sul percorso che connette i due nodi, facendo uso eventualmente anche di decomposizioni LQ [2]. Chiaramente questo metodo, necessitando di decomposizioni ripetute, non è efficiente come l'applicazione di un'opportuna gauge che realizzi lo spostamento del tensore non isometrico; per poterlo esprimere nel secondo modo è tuttavia necessario che il centro, ovvero la parte non isometrica del network, ammetta a sua volta

[19] Suffice considerare il costo della ridotta con $m = \chi^{Z-1}$ e $n = \chi$.

[20] Ovvero ha autovalori non negativi.

[21] Definendo $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)^\alpha \equiv \mathrm{diag}(\lambda_1^\alpha, \ldots, \lambda_n^\alpha)$.

un'inversa - ovvero abbia rango massimo. Sebbene la gauge centrale possa permettere di ridurre le dimensioni dei link virtuali, non garantisce che siano minimali e quindi che righe o colonne eccedenti il rango siano completamente scartate: si può ovviare ricorrendo alla SVD ridotta [6].

Gauge canonica Si consideri un TTN in gauge unitaria; senza perdita di generalità si può assumere che il centro di ortogonalità sia una matrice, in quanto si può isolare su un singolo link applicando la QR ad un centro che avesse ordine diverso da 2. Applicandovi la SVD ridotta e scartando i valori singolari nulli, si ottiene una matrice S invertibile; se si opera in questo modo analogamente per tutti i possibili centri in forma matriciale ottenuti spostandolo su ogni link j-esimo del network, le isometrie $U^{[j]}$, $V^{[j]}$ corrispondenti danno le trasformazioni di gauge ricercate. Si noti che i tensori prodotti contraendole con quelli isometrici in gauge unitaria rimangono isometrici [6].
La loro applicazione permette così di ridurre la dimensione delle gambe al rango del corrispondente centro, che per quanto visto è la dimensione minima che mantiene una rappresentazione esatta del network. Lo spostamento del centro di ortogonalità nella gauge canonica diviene molto più efficiente in quanto per ogni link è espresso dalla matrice dei valori singolari corrispondente: a livello algebrico, per ottenere il tensore isometrico per un centro sul link i-esimo da quello per un centro sul link j-esimo è sufficiente moltiplicare per $S^{[j]}S^{[i]-1}$[22].
Si noti che anziché considerare simultaneamente tutte le gambe virtuali del network, si può più praticamente realizzare la gauge canonica procedendo a partire dal centro di ortogonalità come di seguito:

1. Si applica la SVD e si assorbe $S^{[j]}V^{[j]\dagger}$ o $U^{[j]}S^{[j]}$ nel tensore adiacente;

2. Si applica a questo la SVD rispetto ad ogni link verso il livello successivo, sostituendo ogni volta il tensore $U^{[j]}S^{[j]}$ prodotto al tensore di partenza salvo per l'ultima gamba, ove si mantiene solo $U^{[j]}$;

3. Si itera per il livello successivo, assorbendo le $S^{[j]}V^{[j]\dagger}$[23] nei tensori adiacenti;

Il network finale, al netto delle isometrie residue sugli indici liberi, è così in gauge canonica. Infatti, poiché i valori singolari su ogni link sono determinati - al netto dell'ordine - dal solo network iniziale, la gauge canonica è unica e non dipende dalla specifica procedura di implementazione, un fatto di grande utilità in quanto evita

[22] L'unico svantaggio è che per valori singolari piccoli l'inversione amplifica gli errori e pertanto comporta instabilità. La soluzione, anziché troncare solo sul rango delle matrici, è fissare un cutoff.

[23] Recuperabili per le gambe diverse dall'ultima trasformata tramite la semplice inversione dei valori singolari.

che gli errori dipendano dall'ordine in cui si effettuano le compressioni e permette di operare in parallelo sul network [6].

1.7 Differenziazione

Negli algoritmi variazionali si impone la condizione di estremalità su una certa funzione (tipicamente il valore di aspettazione dell'energia) rispetto ad un certo insieme di parametri; per un TN tipicamente si ottimizza rispetto alle componenti di un tensore del network il valore lo scalare prodotto dalla contrazione complessiva. La condizione può essere equivalentemente formulata come annullamento del differenziale della funzione rispetto al tensore; poiché la contrazione è un'operazione lineare, il gradiente rispetto al tensore in considerazione è calcolabile in modo immediato e corrisponde al network privato di esso[24]:

$$\frac{\partial}{\partial A}\left(A_{\boldsymbol{\alpha},\boldsymbol{\beta}} T_{\boldsymbol{\beta},\boldsymbol{\gamma}}\right) = T_{\boldsymbol{\beta},\boldsymbol{\gamma}} \tag{1.37}$$

[24]Per praticità si sono vettorizzati gli indici liberi e contratti.

Capitolo 2

Sistemi quantistici a molti corpi

L'oggetto cui si applicheranno gli strumenti descritti è lo stato dei sistemi a molti corpi. Si illustra il formalismo della Meccanica Quantistica, su cui si fonda la loro descrizione; tramite l'operatore densità si introducono quindi gli osservabili d'interesse e le misure di accuratezza delle approssimazioni. Si definisce infine la classe di sistemi d'interesse.

2.1 Postulati della MQ

Il framework alla base della fisica dei sistemi a molti corpi è la Meccanica Quantistica (di seguito **MQ**)[1]; per le finalità della trattazione è utile fornirne la formulazione più generale ed astratta, sviluppata principalmente ad opera di Paul A.M. Dirac e John Von Neumann. A Dirac si deve inoltre la notazione *bra-ket*, di cui si farà uso di seguito. Per i contenuti del capitolo si fa riferimento, ove non diversamente specificato, ai testi originali di Dirac [20] e Von Neumann [21], all'oramai altrettanto "classico" Nielsen-Chuang [22] e al già citato volume di Collura et al. [2].

2.1.1 Primi postulati

I primi postulati della MQ definiscono la natura matematica dello stato di un sistema quantistico *isolato*, ovvero l'oggetto che ne dà la descrizione completa [2], gli osservabili, ovvero le quantità fisiche misurabili, l'evoluzione temporale dello stato e le misure.

---

[1] A rigore non da considerarsi una *teoria* fisica, bensì un generale *paradigma* per lo sviluppo di teorie, che concernono una specifica classe di sistemi e le rispettive leggi - e hanno pertanto un dato campo di applicabilità. Ad esempio, la teoria *quantistica* dell'interazione EM è la QED.

[2] Ovvero da cui è possibile ricavare tutte le proprietà del sistema, in altri termini i risultati di ogni possibile misura. A differenza dello stato classico, il punto di fase, lo stato quantistico *non è direttamente osservabile*. La questione della sua realtà e in generale il problema dell'interpretazione della MQ esulano dagli scopi di questo lavoro.

Postulato 1 (Stato)**.** *Ad ogni sistema fisico isolato è associato uno spazio di Hilbert complesso $\mathcal{H}$ (separabile[3]). Il sistema è completamente descritto dal suo* ***vettore di stato****, ovvero un elemento $|\psi\rangle$ dello spazio di norma unitaria[4].*

Postulato 2 (Osservabili)**.** *Ad ogni osservabile del sistema è associato un operatore lineare autoaggiunto[5] su $\mathcal{H}$. La misura di un'osservabile su uno stato $|\psi\rangle$ dà come risultato uno tra i suoi autovalori λ_i , con probabilità:*

$$P_i = |\langle\psi|\lambda_i\rangle|^2 \tag{2.1}$$

ove $|\lambda_i\rangle$ è l'autostato associato[6], che corrisponde allo stato del sistema immediatamente dopo la misura.

In virtù del teorema spettrale, operatori autoaggiunti sono unitariamente diagonalizzabili e hanno spettro reale. Un'osservabile $\hat{A}$ ammette quindi sempre **decomposizione spettrale**[7]:

$$\hat{A} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \tag{2.2}$$

Il valore medio della sua misura su uno stato è definito **valore di aspettazione** e corrisponde al valore della forma quadratica associata:

$$\langle\hat{A}\rangle_\psi \equiv \sum_i P_i\lambda_i = \sum_i \lambda_i \langle\psi|\lambda_i\rangle\langle\lambda_i|\psi\rangle = \langle\psi|\Big(\sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|\Big)|\psi\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{2.3}$$

Non è necessariamente uguale a uno degli autovalori, bensì in ossequio alla legge dei grandi numeri al valore cui tende la media di misure identiche ripetute su stati identicamente preparati[8].
Ad esso è associata una **varianza** σ_A^2 e quindi una deviazione standard $\sigma_A = \sqrt{\sigma_A^2}$:

$$\sigma_A^2 = \sum_i P_i\big(\lambda_i - \langle\hat{A}\rangle\big)^2 = \sum_i P_i\lambda_i^2 - \langle\hat{A}\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2 \tag{2.4}$$

Può essere utile considerare un'estensione della definizione di misura:

---

[3]Una precisazione di poco riguardo nel caso di dimensione finita in quanto la separabilità vale trivialmente.

[4]Due elementi proporzionali descrivono lo stesso stato a meno di un fattore di normalizzazione. Dunque equivalentemente si può definire stato un raggio proiettivo dello spazio $\mathcal{H}$.

[5]Anche in questo caso, la dimensione finita è più "mansueta": non è necessario distinguere gli operatori autoaggiunti da quelli hermitiani. In dimensione infinita le due definizioni non sono equivalenti e il teorema spettrale vale solo per i primi.

[6]Si somma sui vari contributi se l'autovalore è degenere.

[7]Si fa uso di seguito di sommatorie ma chiaramente si può generalizzare a spettri continui o parzialmente continui tramite integrale.

[8]Per una definizione di preparazione, per quanto dichiaratamente operazionista, si può fare riferimento a Peres [23]

Postulato 3 (Misura)**.** *Una misura è descritta da una collezione di* ***operatori di misura*** $\{\hat{M}_m\}$ *su* $\mathcal{H}$ *che soddisfano l'****equazione di completezza****:*

$$\sum_m \hat{M}_m^\dagger \hat{M}_m = \mathbb{I} \tag{2.5}$$

La probabilità di ogni possibile esito $m \in \mathbb{R}$ *è* $p_m = \langle\psi| \hat{M}_m |\psi\rangle$ *e lo stato del sistema immediatamente dopo la misura è dato da:*

$$|m\rangle = \frac{1}{\sqrt{p_m}} \hat{M}_m |\psi\rangle \tag{2.6}$$

La misura di un'osservabile è quindi ridefinita come un tipo peculiare di misura, detta **proiettiva**, in cui gli operatori sono **proiettori ortogonali** - sugli autospazi[9]:

$$\hat{M}_m^\dagger = \hat{M}_m \qquad \hat{M}_m \hat{M}_n = \delta_{m,n} \hat{M}_n \tag{2.7}$$

Postulato 4 (Evoluzione temporale)**.** *L'evoluzione di un sistema quantistico chiuso è descritta da una trasformazione unitaria dipendente esclusivamente dalla differenza dei tempi:*

$$|\psi(t_2)\rangle = \hat{U}(t_2 - t_1) |\psi(t_1)\rangle \tag{2.8}$$

*Per sistemi che evolvono in tempo continuo l'evoluzione è equivalentemente descritta dall'****equazione di Schrödinger****:*

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi\rangle = \hat{H} |\psi\rangle \tag{2.9}$$

ove $\hbar$ *è la costante di Planck e* $\hat{H}$ *un'osservabile detta Hamiltoniana del sistema*[10]*.*

$\hat{H}$ è l'analogo della funzione $H(\mathbf{p}, \mathbf{q})$ per un sistema hamiltoniano classico[11]. Risolvendone l'equazione agli autovalori, detta **equazione di Schrödinger indipendente dal tempo**:

$$\hat{H} |\psi_n\rangle = E_n |\psi_n\rangle \tag{2.10}$$

se ne determinano gli autovalori, che sono le possibili **energie** del sistema $\{E_n\}$, e gli autostati, detti **stati stazionari**; o stato associato all'energia minima in particolare è detto **stato fondamentale (ground state)**. In generale lo spettro dell'Hamiltoniana può essere degenere - in particolare possono esserci più *stati fondamentali* - e pertanto è necessario considerare altre osservabili che commutano con essa (ovvero **compatibili**) per avere un insieme completo che rimuova tutte le degenerazioni; la base spettrale comune è quindi indicizzata tramite una stringa di autovalori.

---

[9] In realtà una misura generica è sempre esprimibile come misura proiettiva, a meno di una trasformazione unitaria e dell'aggiunta di un altro sistema "ancillare".

[10] L'operatore di evoluzione si ottiene dall'esponenziale formale $\hat{U}(t_2 - t_1) = \exp\{-\frac{i}{\hbar}(t_2 - t_1)\hat{H}\}$

[11] Anche quantisticamente per un sistema non isolato $\hat{H}$ è dipendente dal tempo; quindi $\hat{U}$ non è invariante per traslazione temporale - in particolare i termini della serie esponenziale sono integrali multipli ordinati temporalmente, in quanto Hamiltoniane a tempi diversi non commutano.

2.1.2 Postulato per i sistemi composti

Di particolare interesse per la presente trattazione è il quinto postulato:

Postulato 5 (Sistemi composti)**.** *Lo spazio di Hilbert associato ad un sistema composto è dato dal prodotto tensore degli spazi associati ai suoi sottosistemi:*

$$\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i \tag{2.11}$$

Ove il prodotto interno è costruito secondo (in notazione si omette il simbolo $\otimes$):

$$\left(\langle \psi_1| \cdots \langle \psi_n| \, , |\phi_1\rangle \cdots |\phi_n\rangle \right) = \langle \psi_1, \ldots, \psi_n | \phi_1, \ldots, \phi_n \rangle = \prod_{i=1}^{n} \langle \psi_i | \phi_i \rangle \tag{2.12}$$

Dunque fissata una base $\{|k_i\rangle\}_{k_i=1}^{\dim \mathcal{H}_i}$ per ogni spazio, per convenienza ortonormale, una base per gli stati del sistema composto è:

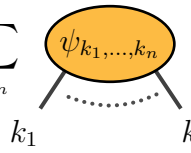
$$\{|k_1, k_2, \cdots, k_n\rangle = |k_1\rangle \, |k_2\rangle \cdots |k_n\rangle \, : \, 1 \leq k_i \leq \dim \mathcal{H}_i \, , \, 1 \leq i \leq n\} \tag{2.13}$$

a sua volta ortonormale e si ha $\dim \mathcal{H} = \dim \mathcal{H}_1 \cdot \dim \mathcal{H}_2 \cdots \dim \mathcal{H}_n$. Allo stesso modo si definisce l'algebra degli osservabili[12] del sistema composto come prodotto tensore di quelle dei singoli sottosistemi[13], ove ogni operatore "locale" agisce sullo spazio corrispondente:

$$(\hat{A}_1 \otimes \hat{A}_2) \, |\psi_1\rangle \, |\psi_2\rangle = (\hat{A}_1 \, |\psi_1\rangle) \otimes (\hat{A}_2 \, |\psi_2\rangle) \tag{2.14}$$

Un generico stato del sistema è un tensore di tipo $(n,0)$ e si può sviluppare sulla base, ottenendo un tensore scalare complesso di dimensione $\dim \mathcal{H}$:

$$|\psi\rangle = \sum_{k_1} \cdots \sum_{k_n} \psi_{k_1,\ldots,k_n} \, |k_1, \cdots, k_n\rangle = \sum_{k_1} \cdots \sum_{k_n} \psi_{k_1,\ldots,k_n} \, |k_1, \cdots, k_n\rangle$$



Essendo di norma nota l'azione degli osservabili sulle basi locali, $\psi_{k_1,\ldots,k_n}$ contiene equivalentemente tutta l'informazione sullo stato del sistema. Per uno stato generico, se la dimensione dei singoli spazi è controllata da un parametro finito d il numero di ingressi scala in quello dei sottosistemi n come $\mathcal{O}(d^n)$, ovvero esponenzialmente: in ciò consiste la **maledizione della dimensionalità (curse of dimensionality)** del problema a molti corpi.
Una tale quantità di valori è necessaria per descrivere la grande maggioranza degli stati dello spazio tensore, che esibiscono la proprietà di **entanglement**.

[12] Ovvero lo spazio $\mathcal{B}(\mathcal{H})$ degli operatori continui, che è uno s. di Hilbert con il prodotto interno di Hilbert-Schmidt, dotato della composizione come prodotto associativo [24].

[13] Che vi possono essere incluse in modo naturale tensorizzando con un'azione triviale su tutti gli altri spazi.

Definizione 2.1. *Un vettore di stato del sistema composto si dice* ***decomponibile*** *se può essere espresso come prodotto tensore di n vettori, ciascuno appartenente allo spazio di un sottosistema. In caso contrario lo stato è* ***entangled****.*

Dunque, dato un sistema composto in uno stato entangled, ciascun sottosistema di cui non è possibile "raccogliere a fattor comune" un vettore di stato individuale non è in uno stato definito; inoltre una misura su un sottosistema "intrecciato" ad un altro altera lo stato del sistema complessivo e di conseguenza comporta una determinazione (esatta o parziale) degli stati ammissibili del secondo, senza che in principio sia stato necessario interagirvi.

2.2 Operatore densità

Sistemi per cui la preparazione non determina univocamente uno stato, come quelli termici, sono descritti da una **miscela statistica** di vettori di stato[14] - definiti in questo contesto più specificamente **stati puri**, ovvero un insieme di più di essi con assegnate delle probabilità[15]:

$$\{p_i, |\psi_i\rangle\} \qquad p_i \in [0,1] \quad \textstyle\sum_i p_i = 1 \tag{2.15}$$

La descrizione di stati puri e **misti** può essere unificata introducendo l'**operatore densità**:

Definizione 2.2. *L'operatore densità (o* ***matrice densità****) associato ad un sistema descritto da uno stato puro è il proiettore ortogonale sul vettore di stato:*

$$\hat{\rho} \equiv |\psi\rangle\langle\psi| \tag{2.16}$$

L'operatore densità associato ad un sistema descritto da una miscela statistica è invece[16] *la somma convessa*[17] *degli operatori associati ai vari stati puri, con i pesi eguali alle rispettive probabilità:*

$$\hat{\rho} \equiv \textstyle\sum_i p_i \rho_i = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \tag{2.17}$$

[14] Non necessariamente ortogonali.

[15] É importante notare che non si tratta di una sovrapposizione, che è uno stato ben definito, ovvero puro. La differenza si può verificare matematicamente: essenzialmente, gli stati di una miscela non possono interferire.

[16] Chiaramente il caso puro si può assorbire in quello misto come miscela con un solo elemento a probabilità unitaria.

[17] Ovvero con coefficienti non negativi che sommano all'unità. Nello spazio operatoriale giace quindi internamente al simplesso definito dai termini della somma.

Si tratta di uno strumento estremamente utile in quanto è possibile riformulare integralmente i postulati della MQ facendo esclusivamente riferimento agli operatori densità, includendo così sia stati puri che ensemble nel medesimo formalismo. Preliminarmente è necessario definire la **traccia** di un operatore:

Definizione 2.3. *La traccia di un operatore $\hat{A}$ è definita come lo scalare:*

$$\mathrm{Tr}[\hat{A}] \equiv \sum_k \langle k|\hat{A}|k\rangle \tag{2.18}$$

ove $\{|k\rangle\}$ è una generica base ortonormale. La definizione è ben posta in quanto invariante per trasformazioni unitarie.

La denominazione è opportuna in quanto la traccia operatoriale possiede tutte le proprietà di quella matriciale (e tensoriale), incluse in particolare linearità e ciclicità. Innanzitutto si dà quindi un teorema di caratterizzazione degli operatori densità, che stabilisce una corrispondenza con possibili ensemble di stati puri:

Teorema 2.1. *Un operatore $\hat{\rho}$ è l'operatore densità associato ad un qualche ensemble di stati puri se e solo se soddisfa le seguenti due condizioni*[18]*:*

1. *Ha traccia unitaria:* $\mathrm{Tr}[\hat{\rho}] = 1$
2. *É positivo (o semi-definito positivo):* $\forall\psi \;\; \langle\psi|\hat{\rho}|\psi\rangle \in \mathbb{R} \wedge \langle\psi|\hat{\rho}|\psi\rangle \geq 0$

Non si riporta la dimostrazione ma si annota solamente che la positività è condizione sufficiente per l'hermitianità e che i proiettori ortogonali sono hermitiani e idempotenti (e quindi positivi). Questa corrispondenza è *biunivoca* solo considerando *classi di equivalenza* di ensemble, in quanto un medesimo stato misto ammette più espressioni come combinazione convessa di stati puri[19] - chiaramente non distinguibili a mezzo di misure o evoluzione temporale. In virtù di ciò è opportuno fare primariamente riferimento all'operatore densità, per cui come detto si può formulare consistentemente la MQ. Vale poi il seguente criterio:

Teorema 2.2. *Per un operatore densità $\hat{\rho}$ generico vale* $\mathrm{Tr}[\hat{\rho}^2] \leq 1$. *Si ha l'uguaglianza se e solo se $\hat{\rho}$ è uno stato puro.*

Si delineano quindi i punti essenziali della riformulazione dei postulati:

Post. 1. La descrizione completa del sistema è data da un operatore densità definito sullo spazio degli stati $\mathcal{H}$ associato;

Post. 2. Il valore di aspettazione[20] di un'osservabile è dato da $\langle\hat{A}\rangle_\rho = \mathrm{Tr}[\hat{\rho}\hat{A}]$

[18] Che sono di fatto le condizioni di buona definizione della probabilità.

[19] Equivalenti a meno di una trasformazione unitaria "pesata" $\sqrt{p_i}\,|\psi_i\rangle = \sum_j U_{ij}\sqrt{q_j}\,|\phi_j\rangle$

[20] Più precisamente, la media pesata sull'ensemble dei valori di aspettazione.

Post. 3. La probabilità di un risultato m di una misura è $p_m = \mathrm{Tr}[\hat{M}_m^\dagger M_m \rho]$ e lo stato corrispondente successivamente alla misura è:

$$\hat{\rho}_m = \frac{1}{p_m} \hat{M}_m \hat{\rho} \hat{M}_m^\dagger \tag{2.19}$$

Post. 4. Lo stato evolve unitariamente secondo $\hat{\rho}(t_2) = \hat{U}(t_2 - t_1)\hat{\rho}(t_1)\hat{U}^\dagger(t_2 - t_1)$; in tempo continuo vale l'**equazione di Von Neumann**:

$$i\hbar \frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = [\hat{H}, \hat{\rho}] \tag{2.20}$$

Post. 5. L'operatore densità del sistema composto agisce sul rispettivo spazio degli stati, quindi è espresso da una c.l. di termini nella forma $\hat{\rho}_1 \otimes \hat{\rho}_2 \otimes \cdots \otimes \hat{\rho}_n$

2.2.1 Heisenberg picture

Nella formulazione della MQ data, le osservabili sono operatori costanti e l'evoluzione temporale riguarda gli stati. In realtà questa è solo una delle diverse possibili rappresentazioni (*pictures*), detta *di Schrödinger*: equivalentemente - ovvero senza alcuna modifica a livello delle predizioni fisiche - si può considerare la rappresentazione detta *di Heisenberg*, in cui gli stati[21] sono fissi ed a evolvere sono gli operatori. Come si vedrà brevemente di seguito, questo cambiamento di prospettiva è utile in particolare per formulare una descrizione matematicamente coerente di sistemi infiniti.
I postulati non sono essenzialmente da modificarsi, salvo appunto per l'evoluzione temporale; a livello discreto l'unitaria, definita analogamente, evolve le osservabili:

$$\hat{A}(t_2) = \hat{U}^\dagger(t_1, t_2)\hat{A}(t_1)\hat{U}(t_1, t_2) \tag{2.21}$$

mentre in tempo continuo l'equazione di Schrödinger è rimpiazzata dall'**equazione di Heisenberg**:

$$\frac{\mathrm{d}\hat{A}(t)}{\mathrm{d}t} = \frac{i}{\hbar}\Big[\hat{H}(t), \hat{A}(t)\Big] + \frac{\partial \hat{A}(t)}{\partial t} \tag{2.22}$$

Si noti l'evidente corrispondenza con l'equazione per l'evoluzione di osservabili di un sistema hamiltoniano classico, che si ottiene semplicemente invertendo la quantizzazione canonica, rimpiazzando il commutatore con la parentesi di Poisson.

[21] Sia puri che misti.

2.2.2 Stati di Gibbs

Come anticipato, un caso rilevante descrivibile tramite l'operatore densità è quello dei sistemi *termici*. A rigore, ciò significa che quantisticamente si può definire un ensemble canonico per un sistema in contatto termico con un bagno a temperatura T[22] e che il suo stato è uno **stato di Gibbs** (**Gibbs state**), ovvero ha la forma:

$$\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{H}} = \frac{1}{Z} \sum_n e^{-\beta E_n} |\psi_n\rangle\langle\psi_n| \tag{2.23}$$

ove $\beta = {}^1\!/_{k_B T}$ è la temperatura inversa e $Z \equiv \mathrm{Tr}[e^{-\beta \hat{H}}] = \sum_n e^{-\beta E_n}$ la funzione di partizione canonica.

2.2.3 Operatore ridotto

Più in generale l'estensione del formalismo agli stati misti possibile tramite l'operatore densità è utile per la trattazione di sistemi composti, in quanto permette di definire uno stato anche per sottosistemi entangled. Per farlo è necessario introdurre l'**operatore ridotto**:

Definizione 2.4. *Sia $\hat{\rho}^{AB}$ l'operatore densità che descrive lo stato di un sistema composto da due sottosistemi A e B. L'operatore ridotto per il sistema A è definito secondo:*

$$\hat{\rho}^A \equiv \mathrm{Tr}_B[\hat{\rho}^{AB}] \tag{2.24}$$

*ove la **traccia parziale** Tr_B è l'operatore ottenuto tracciando rispetto ad una qualsiasi base ortonormale del solo spazio degli stati di B*[23]*.*

Anche in questo caso sono valide tutte le proprietà dell'analoga operazione tensoriale. Si può dimostrare che l'operatore ridotto è un operatore densità ben definito e che i valori e le probabilità dei risultati delle misure (e quindi in particolare delle misure di osservabili) definite sul solo spazio degli stati di A tramite $\hat{\rho}^A$ corrispondono a quelli delle misure associate nell'algebra tensore[24]. Per gli osservabili in particolare si ha quindi:

$$\langle \hat{O}_A \rangle_A = \mathrm{Tr}[\hat{\rho}^A \hat{O}_A] = \mathrm{Tr}[(\hat{O}_A \otimes \mathbb{I}_B)\hat{\rho}^{AB}] = \langle \hat{O}_A \otimes \mathbb{I}_B \rangle_{AB} \tag{2.25}$$

[22] Dunque aperto.

[23] Ovvero, nell'espressione dell'operatore totale, riducendo in ogni termine tensore l'operatore associato a B alla propria traccia. Il risultato è quindi effettivamente un'operatore sul solo spazio degli stati di A.

[24] Intuitivamente, la traccia parziale corrisponde a un'"integrazione" sui gradi di libertà del secondo sottosistema, non rilevanti per gli osservabili locali.

corrispondendo la traccia su AB a quella combinata su A e B (vedasi la definizione della base ortonormale del prodotto tensore). Viceversa si può verificare che l'operazione di traccia parziale è l'unica a soddisfare questa condizione.

É utile notare che tramite l'operatore ridotto si può formulare una descrizione dell'evoluzione anche per sistemi *aperti*. Infatti per un'ampia gamma di situazioni[25] si può costruire un più ampio sistema chiuso - per definizione - accoppiando un opportuno *ambiente* costituito dai gradi di libertà addizionali necessari, e rappresentare una generica trasformazione locale come restrizione di un'unitaria di tale sistema chiuso complessivo. Si definiscono così le **operazioni** sul sistema, una classe di trasformazioni che generalizza unitarie e misure:

Definizione 2.5. *Un'operazione $\hat{\mathcal{E}}$ è un super-operatore sugli operatori densità di un sistema che può essere espresso nella forma*

$$\hat{\mathcal{E}}(\hat{\rho}) = \mathrm{Tr}_E[\hat{U}(\hat{\rho} \otimes \hat{\rho}_E)\hat{U}^\dagger] \tag{2.26}$$

ove $\hat{\rho}_E$ è uno stato di un opportuno ambiente e $\hat{U}$ un'unitaria.

É possibile dare una definizione più ampia, espressa in termini di operatori sul solo spazio degli stati del sistema originario, che rende esplicita la generalizzazione:

Definizione 2.6. *Un'operazione è un super-operatore che ammette la seguente espressione, detta* ***Operator Sum Representation (OSR)***[26]*:*

$$\hat{\mathcal{E}}(\hat{\rho}) = \textstyle\sum_k \hat{E}_k \hat{\rho} \hat{E}_k^\dagger \tag{2.27}$$

ove gli ***operatori di Kraus*** *$\hat{E}_k$ soddisfano*[27]*: $\sum_k \hat{E}_k^\dagger \hat{E}_k \leq \mathbb{I}$*

La prima definizione corrisponde al sottoinsieme delle operazioni che preservano la traccia[28], per cui vale la completezza degli operatori di Kraus. É immediato osservare che le misure sono parte delle operazioni e lo stesso per l'evoluzione temporale, in cui si ha un unico operatore di Kraus che corrisponde all'unitaria.
Prima di proseguire è utile a fini di quanto si discuterà nel prossimo capitolo dare un'ultima caratterizzazione assiomatica delle operazioni, che ha la medesima generalità di quella appena riportata:

Definizione 2.7. *Un'operazione è una mappa $\hat{\mathcal{E}}$ sugli operatori densità tale che:*

---

[25] Quando il sistema non interagisce con i d.o.f. usati per la preparazione *dopo di essa*.

[26] La cui esistenza è di per sè condizione necessaria e sufficiente perché la mappa sia completamente positiva, secondo il Teorema di Choi [25].

[27] Operatori hermitiani possono essere ordinati definendo $A \geq B \Leftrightarrow A - B$ è positivo.

[28] Nel caso generale per ottenere uno stato bisogna propriamente riscalare su $\mathrm{Tr}[\hat{\mathcal{E}}(\hat{\rho})]$.

1. $0 \leq \mathrm{Tr}[\hat{\mathcal{E}}(\hat{\rho})] \leq 1$ *per ogni* $\hat{\rho}$*;*
2. *É lineare sulle combinazioni convesse;*
3. *É completamente positiva (**CP**), ovvero:*
   (a) *É positiva, cioè mappa operatori positivi in operatori positivi;*
   (b) *Introducendo un sistema ancillare A di dimensione arbitraria,* $(\mathbb{I}_A \otimes \hat{\mathcal{E}})(\hat{O})$ *è positivo per ogni* $\hat{O}$ *positivo sul sistema composto.*

Seguendo questa caratterizzazione le operazioni che preservano la traccia sono anche dette mappe **Completely Positive Trace-Preserving** (**CPTP**). Si noti che in generale la OSR di un'operazione non è univocamente determinata, ma i differenti insiemi possibili di operatori di Kraus sono unitariamente equivalenti. É utile ai fini della trattazione successiva riportare infine un teorema per le mappe CPTP che generalizza un risultato per matrici reali, il **Teorema di Perron-Frobenius** [26]:

Teorema 2.3. *Una mappa CPTP* $\hat{\mathcal{E}}$*:*

1. *Ha raggio spettrale unitario, ovvero* $|\lambda_i| \leq 1$ *per ogni suo autovalore* λ_i*;*
2. *Ha un autovalore dominante pari a* 1*;*
3. *L'autovettore corrispondente, che risulta per definizione un punto fisso di* $\hat{\mathcal{E}}$*, è un operatore positivo normalizzato.*

Si può ora procedere ad esporre la descrizione dell'entanglement in termini di operatore densità. Tuttavia è preliminarmente necessario riformularne la definizione, in quanto in generale uno stato misto non è necessariamente entangled:

Definizione 2.8. *Un operatore densità di un sistema composto si dice **separabile** se può essere espresso come combinazione convessa di prodotti tensori di operatori densità sui singoli sottosistemi. In caso contrario lo stato è **entangled**.*

Un operatore esprimibile come singolo prodotto tensore viene definito più specificamente **stato prodotto**. In generale la separabilità è difficile da caratterizzare ed esistono diverse condizioni necessarie in generale ma sufficienti solo in basse dimensioni. Per gli stati puri è invece più semplice determinare se un dato vettore è entangled e lo si può fare sfruttando una decomposizione legata strettamente all'operatore ridotto: la **decomposizione di Schmidt**.

2.2.4 Decomposizione di Schmidt

Vale il seguente risultato:

Teorema 2.4. *Un generico stato puro $|\psi\rangle$ di un sistema composto AB ammette sempre un'espressione nella forma:*

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \tag{2.28}$$

ove $\{|i_A\rangle\}$ e $\{|i_B\rangle\}$ sono insiemi ortonormali in $\mathcal{H}_A$ e $\mathcal{H}_B$[29] *e i λ_i sono numeri reali non negativi detti* ***coefficienti di Schmidt*** *che soddisfano $\sum_i \lambda_i^2 = 1$.*

Dalla decomposizione dello stato complessivo seguono le seguenti espressioni per le matrici ridotte:

$$\hat{\rho}^A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A| \qquad \hat{\rho}^B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B| \tag{2.29}$$

Sono dunque diagonalizzate con i medesimi autovalori, pari al quadrato dei coefficienti di Schmidt; ma tali autovalori non sono che le probabilità dei vari stati della miscela, che in questa particolare espressione risultano ortonormali. Il loro rango - si noti: limitato dalla dimensione dello spazio degli stati più piccolo - è quindi pari al numero di coefficienti non nulli, che viene detto **rango di Schmidt**.
La decomposizione di Schmidt non è che un caso di SVD, applicata alla matrice dei coefficienti dello stato composto su una generica base del prodotto tensore; operando diagrammaticamente si può verificare rapidamente il risultato per le matrici ridotte (si illustra la procedura per $\hat{\rho}^A$ ma chiaramente è analoga per B):

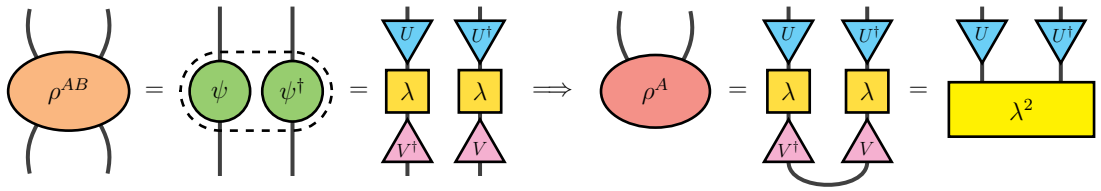


Figura 2.1: Decomposizione di Schmidt diagrammatica

Grazie alla decomposizione di Schmidt è possibile definire un'operazione "inversa" della traccia parziale, la **purificazione**: data un generico operatore densità, eventualmente misto, è sempre possibile introdurre un sistema aggiuntivo di dimensione almeno pari al numero di suoi autovalori non nulli, di modo che tale operatore sia la traccia parziale rispetto al sistema ausiliario di uno stato *puro* del sistema composto. Essenzialmente si costruisce di modo che la base di Schmidt per il sistema di partenza sia la base spettrale dell'operatore. La purificazione non è chiaramente unica, ma vi è una libertà limitata:

Teorema 2.5. *Tutte le possibili purificazioni di un dato stato misto sono equivalenti a meno di un'unitaria sul sistema ausiliario - o più generalmente un'isometria, ammettendo dimensioni differenti per il corrispondente spazio degli stati*[30].

[29] Senza perdita di generalità assumibili completi (i coefficienti possono essere nulli).

[30] Il rango di Schmidt determina solo un minimo.

Come anticipato, la decomposizione di Schmidt permette di formulare una condizione necessaria e sufficiente per l'entanglement di uno stato puro arbitrario:

Teorema 2.6. *$|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ è decomponibile se e solo se ha un coefficiente di Schmidt uguale a 1 e tutti gli altri nulli*[31]*, ossia un rango di S. pari ad 1*[32]*.*

2.3 Misure di entanglement

Al di là della distinzione binaria tra stati entangled e separabili, è possibile introdurre delle grandezze che permettono di *quantificare* l'entanglement di uno stato, al fine di definire una misura del grado di correlazione non classica tra i sottosistemi. Più tecnicamente, ad esserne descritto è il *contenuto informativo*, che pragmaticamente si può far corrispondere alle risorse necessarie per la specificazione ottimale dello stato. Infatti questi strumenti sono alla base della *teoria dell'informazione quantistica*, che estende la corrispondente teoria classica ai sistemi quantistici.
La grandezza fondamentale da introdurre è l'**entropia di entanglement**. Questa essenzialmente quantifica il grado di *incertezza* della definizione dello stato dei sottosistemi, o equivalentemente quello di informatività medio dei possibili valori delle variabili stocastiche discrete associate, le quali non sono che gli stati delle rispettive miscele, di cui le probabilità corrispondenti definiscono la distribuzione. É catturato così l'aspetto fondamentale, già evidenziato, del fenomeno: anche se lo stato complessivo è ben definito, ciò non implica lo sia anche quello dei sottosistemi, ed un maggior grado di entanglement corrisponde ad una maggiore indeterminatezza[33].

2.3.1 Entropia di Von Neumann

La quantità che in teoria dell'informazione classica misura l'incertezza di una distribuzione di probabilità è l'**entropia di Shannon**[34]:

$$H(X) = -\sum_{\{x\}} p(x) \log_2 p(x) \tag{2.30}$$

[31] Ovvero se e solo se gli operatori ridotti sono stati puri.

[32] Nel cap. 1 si è detto che la SVD è univoca a meno di riordino dei valori singolari; segue che i coefficienti di Schmidt sono univocamente determinati dallo stato $|\psi\rangle$, e quindi anche il rango. La condizione sul rango è equivalente poiché i coefficienti sono normalizzati.

[33] Nelle parole di Schrödinger [27]: *La conoscenza dei sistemi singoli può calare al minimo, proprio fino a zero, mentre quella del sistema complessivo resta costantemente massimale. La conoscenza migliore possibile di un tutto non include la conoscenza migliore possibile delle sue parti* [...].

[34] Il nome fu scelto da Shannon in analogia alla grandezza termodinamica, ed è in effetti possibile stabilire una corrispondenza in entrambi i sensi (si pensi al principio di Landauer).

ove si definisce $0 \log_2 0 = 0$. Se la distribuzione è quella degli stati di una miscela statistica quantistica, si ottiene l'**entropia di Von Neumann**[35]:

Definizione 2.9. *L'entropia di Von Neumann di uno stato $\hat{\rho}$ è la quantità:*

$$S(\hat{\rho}) \equiv -\operatorname{Tr}[\hat{\rho} \log_2 \hat{\rho}] \tag{2.31}$$

Chiaramente a meno di un fattore costante può essere ridefinita per qualsiasi altra base del logaritmo; la principale alternativa al binario è il logaritmo naturale[36]. Per proprietà della traccia, S dipende dal solo stato $\hat{\rho}$ e non da una sua particolare espressione; dunque diagonalizzando l'operatore si può esprimere più semplicemente in funzione dei suoi autovalori $\{\alpha_i\}$:

$$S(\hat{\rho}) = -\textstyle\sum_i \alpha_i \log \alpha_i \tag{2.32}$$

Poiché per uno stato puro del sistema composto gli autovalori della matrice ridotta di un sottosistema sono i quadrati dei coefficienti di Schmidt, l'entropia di Von Neumann è univocamente determinata dal solo stato complessivo. Tuttavia in generale i coefficienti e quindi l'entropia *dipendono dalla bipartizione scelta.*
Dall'espressione data segue immediatamente che uno stato puro ha $S = 0$, come da attendersi considerando che non vi è alcuna incertezza nella sua definizione; tale valore corrisponde inoltre al minimo di S in conseguenza della definizione delle probabilità. Si verifica poi che l'entropia ha anche un massimo, corrispondente allo stato **massimamente misto**, ovvero per cui la distribuzione non ha alcun bias: tutti gli elementi della miscela sono egualmente probabili. In tal caso la matrice densità diagonalizzata non ha che la semplice forma di una risoluzione dell'identità riscalata sulla probabilità uniforme.
In generale l'entropia di Von Neumann ha una serie di proprietà, anche non triviali, che la rendono una misura ben definita dell'entanglement. Prima di dare un teorema di caratterizzazione completo è utile introdurre una grandezza affine definita sempre in analogia al caso classico: l'**entropia relativa**.

Definizione 2.10. *L'entropia relativa di due operatori densità $\hat{\rho}$ e $\hat{\sigma}$ è la quantità:*

$$S(\hat{\rho}||\hat{\sigma}) \equiv \operatorname{Tr}[\hat{\rho} \log \hat{\rho}] - \operatorname{Tr}[\hat{\rho} \log \hat{\sigma}] = -S(\hat{\rho}) - \operatorname{Tr}[\hat{\rho} \log \hat{\sigma}] \tag{2.33}$$

Essa misura la distanza tra le due distribuzioni statistiche degli stati misti (per quanto non sia una vera metrica non essendo simmetrica). Più formalmente, quantifica la differenza tra l'entropia dello stato e l'informazione media che si può ottenere da un esito sulla distribuzione $\hat{\rho}$ *assumendo a priori che la distribuzione sia* $\hat{\sigma}$:

[35] Anch'egli scelse il nome in analogia alla quantità termodinamica, congetturando una piena corrispondenza. Questa è tutt'oggi argomento di dibattito, ma ad esempio i *Gibbs states* sono ottenuti in analogia all'estremizzazione vincolata classica nell'insieme canonico sostituendo all'energia media $\langle \hat{H} \rangle$ e all'entropia di Gibbs proprio quella di Von Neumann [23].

[36] Il logaritmo di un operatore è l'operatore che esponenziato lo produce.

l'informazione associata a ogni realizzazione è data da $\log \hat{\sigma}$ ma la sua probabilità *vera* è comunque descritta da $\hat{\rho}$ [28]. Chiaramente assumendo una qualsiasi distribuzione, in generale differente, l'informatività media sarà maggiorata dall'entropia della distribuzione *vera*: dunque l'entropia relativa è non negativa. Questo risultato si può dimostrare rigorosamente ed è detto **diseguaglianza di Klein**. All'estremo opposto, se la distribuzione ipotetica assegna probabilità nulla a degli esiti contemplati da quella vera, ovvero il nucleo di $\hat{\sigma}$ ha intersezione non triviale con il supporto di $\hat{\rho}$, il verificarsi di uno di questi sarà infinitamente meno informativo sotto l'ipotesi errata, che risulta del tutto inutile: si assegna il valore $S(\hat{\rho}||\hat{\sigma}) = +\infty$.
Infine, considerando un sistema composto la restrizione allo stato ridotto di un sottosistema rende necessariamente più difficile distinguere due stati complessivi: l'entropia relativa realizza matematicamente questa condizione tramite la proprietà di **monotonia**.

Teorema 2.7. *Per qualsiasi coppia di operatori densità per un sistema composto* $\hat{\rho}^{AB}$ *e* $\hat{\sigma}^{AB}$ *vale:*

$$S(\hat{\rho}^{A}||\hat{\sigma}^{A}) \leq S(\hat{\rho}^{AB}||\hat{\sigma}^{AB}) \tag{2.34}$$

Si enunciano quindi le proprietà di S:

Teorema 2.8. *L'entropia di Von Neumann soddisfa le seguenti proprietà:*

1. *É non negativa e si annulla se e solo se lo stato è puro;*
2. *Se lo spazio degli stati è d-dimensionale,* $S \leq \log d$[37]*. Il massimo è ottenuto se e solo se il sistema è nello stato massimamente misto, ovvero* $\hat{\rho} = \frac{1}{d}\mathbb{I}_d$*;*
3. *Per un operatore di rango* r*, vale* $S(\hat{\rho}) \leq \log r$*;*
4. *É una funzione continua*[38]*;*
5. *É invariante per trasformazioni unitarie* $S(\hat{\rho}) = S(\hat{U}\hat{\rho}\hat{U}^{\dagger})$[39]*;*
6. *Se un sistema composto* AB *è in uno stato puro,* $S(A) = S(B)$*;*

[37] Se si considera la base 2, l'entropia di Shannon non è che il numero medio di domande binarie necessarie per identificare un esito. Quantisticamente ciò corrisponde al numero di misure che distinguono singoli stati necessarie per identificare uno stato della miscela; è allora chiaro che deve valere $2^S \leq D$.

[38] Considerando per gli operatori la topologia indotta dalla metrica definita tramite la traccia.

[39] In quanto sono invarianti gli autovalori dell'operatore. Si noti che per un sistema composto si considerano unitarie sullo spazio del sottosistema, quindi un'evoluzione unitaria complessivamente ma non localmente (non decomponibile) può variare, ed in particolare aumentare, l'entanglement - si pensi alla porta CNOT in computazione quantistica. In generale l'entanglement non può aumentare per LOCC, ossia operazioni locali e comunicazione classica.

7. *Per una combinazione convessa di operatori con supporti ortogonali*[40] *vale:*

$$S\left(\sum_i p_i \hat{\rho}_i\right) = H(\{p_i\}) + \sum_i p_i S(\hat{\rho}_i) \tag{2.35}$$

8. ***Teorema dell'entropia congiunta (joint entropy)*** *Siano* $\{|i\rangle\}$ *stati ortogonali di un sistema* A *e* $\{\hat{\rho}_i\}$ *operatori densità di un altro sistema* B*; allora per una combinazione convessa vale:*

$$S\left(\sum_i p_i \, |i\rangle\langle i| \otimes \hat{\rho}_i\right) = H(\{p_i\}) + \sum_i p_i S(\hat{\rho}_i) \tag{2.36}$$

Dal teorema dell'entropia congiunta segue un utile risultato per stati prodotto, analogo a quanto vale classicamente per sistemi statisticamente indipendenti:

$$S(\hat{\rho} \otimes \hat{\sigma}) = S(\hat{\rho}) + S(\hat{\sigma}) \tag{2.37}$$

Più in generale, ammettendo la possibilità di correlazione tra i sottosistemi l'entropia dello stato totale è controllata dalle due diseguaglianze triangolari:

Teorema 2.9. *L'entropia di un sistema composto con due componenti* A *e* B*, indicata con* $S(A, B) = -\operatorname{Tr}[\hat{\rho}^{AB} \log \hat{\rho}^{AB}]$*, soddisfa le seguenti diseguaglianze:*

$$S(A, B) \leq S(A) + S(B) \tag{2.38}$$

$$S(A, B) \geq |S(A) - S(B)| \tag{2.39}$$

In particolare la prima definisce la proprietà di ***subadditività*** *di* S*.*

Grazie alla subadditività si può dimostrare che l'entropia è una funzione **concava**. Data una combinazione convessa di operatori densità vale infatti:

$$S\left(\sum_i p_i \hat{\rho}_i\right) \geq \sum_i p_i S(\hat{\rho}_i) \tag{2.40}$$

ossia l'incertezza sulla distribuzione complessiva è maggiore della media pesata delle incertezze sulle singole distribuzioni, in quanto è da considerarsi la stessa incertezza *su come queste sono pesate*; l'entropia relativa è invece convessa rispetto a ciascun argomento (mantenendo l'altro costante).
La concavità è una problematica per sistemi *complessivamente* in uno stato misto, per cui i sottosistemi - descritti da una combinazione convessa delle ridotte corrispondenti agli stati puri dell'ensemble - hanno un entanglement maggiore di quello che avrebbero se lo stato completo fosse puro, sebbene l'incertezza introdotta sia di origine classica e non quantistica. Di conseguenza S è più appropriata come misura per stati puri di sistemi composti; per quelli misti una prima soluzione può essere considerare l'estremo inferiore per le possibili decomposizioni[41]: si ottiene così l'**entanglement di formazione**.

[40] Ovvero che descrivono miscele di stati che generano sottospazi ortogonali.

[41] Ovvero le possibili preparazioni che danno miscele equivalenti.

Definizione 2.11. *L'entanglement di formazione di uno stato misto di un sistema composto è definito secondo la seguente espressione, ove* $p_i \in [0,1]$ *e* $\sum_i p_i = 1$*:*

$$S_f(\hat{\rho}^{AB}) \equiv \inf \left\{ \sum_i p_i S\big(\operatorname{Tr}_B[\hat{\rho}^{AB}]\big) \ : \ \hat{\rho}^{AB} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \right\} \tag{2.41}$$

Alternativamente si possono introdurre altre grandezze, come la **negatività**[42]:

Definizione 2.12. *Dato uno stato* $\hat{\rho}$ *di un sistema composto* AB*, la negatività del sottosistema* A *è definita secondo:*

$$N(\hat{\rho}) \equiv \frac{\|\hat{\rho}^{T_B}\|_1 - 1}{2} \tag{2.42}$$

ove T_B *è la trasposizione parziale*[43] *rispetto al solo sottospazio di* B[44]*. Equivalentemente si può definire la* ***negatività logaritmica*** *secondo:*

$$E(\hat{\rho}) = \log \|\hat{\rho}^{T_B}\|_1 \tag{2.43}$$

Per le finalità di questa trattazione è comunque sufficiente limitarsi a considerare misure per stati puri. Grazie alla concavità si può estendere la proprietà 7 dell'entropia di Von Neumann a un caso misto generico $\hat{\rho} = \sum_i p_i \hat{\rho}_i$, ottenendo un limite superiore (oltre a quello inferiore che essa dà direttamente):

$$S(\hat{\rho}) \leq H(\{p_i\}) + \sum_i p_i S(\hat{\rho}_i) \tag{2.44}$$

Intuitivamente, se gli stati descritti dai vari operatori non sono pienamente distinguibili, diminuisce il contenuto informativo; il massimo è ottenuto per stati ortogonali, nel cui caso alla media pesata delle entropie si aggiunge il massimo contributo possibile di incertezza dalla distribuzione delle $\{p_i\}$, ovvero la sua entropia. Un'ultima proprietà interessante è l'estensione della subadditività a sistemi tripartiti:

Teorema 2.10. *Per un sistema composto di tre sottosistemi* A*,* B*,* C *valgono le seguenti diseguaglianze (in realtà equivalenti tramite una purificazione):*

$$S(A) + S(B) \leq S(A,C) + S(B,C) \tag{2.45}$$

$$S(A,B,C) + S(B) \leq S(A,B) + S(B,C) \tag{2.46}$$

e analoghe per permutazioni. Questa proprietà è detta di ***subadditività forte****.*

---

[42] In generale una misura valida è un qualsiasi **entanglement monotone**, ovvero una funzione che: (a) si annulla su stati separabili e (b) non aumenta per LOCC.

[43] La definizione viene dal criterio Peres-Horodecki o PPT.

[44] La 1-norma è $\|\hat{X}\|_1 \equiv \operatorname{Tr}[\sqrt{\hat{X}^\dagger \hat{X}}]$, ovvero per diagonalizzabili $\|\hat{X}\|_1 = \sum_i |\lambda_i|$.

Sempre seguendo l'analogia con il corrispettivo classico, si possono quindi definire a partire dall'entropia di Von Neumann delle altre misure di correlazione:

Definizione 2.13. *Dato un sistema composto con due componenti A e B, si definisce l'**entropia condizionale** di A (condizionata su/a B):*

$$S(A|B) \equiv S(A,B) - S(B) \tag{2.47}$$

*e l'**informazione mutua**, indicata alternativamente con $I(A:B)$:*

$$S(A:B) \equiv S(A) + S(B) - S(A,B) = \tag{2.48}$$
$$= S(A) - S(A|B) = S(B) - S(B|A) \tag{2.49}$$

L'entropia condizionale misura quanta informazione sullo stato di A possiamo ottenere osservando B (ed infatti è nulla per stati prodotto); l'informazione mutua (che è simmetrica) quantifica invece il grado di correlazione, ovvero quanta più incertezza è presente sugli stati individuali rispetto allo stato congiunto. Dalla subadditività dell'entropia di Von Neumann segue la subadditività dell'entropia condizionale, sia nei singoli argomenti che congiuntamente.
É importante notare che non vale una diseguaglianza elementare soddisfatta dall'entropia classica: $H(X) \leq H(X,Y)$ - o, equivalentemente, che l'entropia condizionale può essere negativa, come si può verificare nel caso di un qualsiasi stato puro entangled del sistema composto. Il significato di questa proprietà è chiaro alla luce dei ragionamenti fatti: l'incertezza sullo stato complessivo può essere minore (al limite nulla!) di quella sugli stati dei sottosistemi. Si tratta in realtà di una condizione necessaria e sufficiente, ovvero l'entropia è additiva se è solo se lo stato è prodotto:

Teorema 2.11. *Uno stato puro di un sistema composto AB è entangled se e solo se l'entropia condizionale per A (o equivalentemente B) è negativa.*

In generale tramite l'entropia condizionale e la mutua informazione si possono dare riformulazioni di proprietà dell'entropia di Von Neumann con un'interpretazione più immediata:

Teorema 2.12. *Per un sistema composto con tre sottosistemi ABC valgono le seguenti diseguaglianze:*

$$S(A|B,C) \leq S(A|B) \qquad S(A:B) \leq S(A:B,C) \tag{2.50}$$

Inoltre considerando l'informazione mutua prima e dopo l'applicazione di una qualsiasi operazione su B che preserva la traccia di $\hat{\rho}^B$ vale:

$$S(A:B) \leq S'(A:B) \tag{2.51}$$

Le prime due diseguaglianze possono essere interpretate come segue: considerando più sottosistemi si riduce l'entropia condizionale (la loro osservazione dà più

informazioni sullo stato del terzo) e restringendo invece i sistemi considerati non aumenta la mutua informazione (in quanto al più si rimuove una quantità nulla di correlazione, che non può essere negativa). Il significato della terza relazione è altrettanto immediato: agendo localmente con operazioni che preservano l'incertezza dello stato ridotto non si può aumentare la quantità di correlazione tra sottosistemi.

2.3.2 Entropie di Renyi

L'entropia di Von Neumann può essere sussunta in una famiglia più generale di entropie di entanglement, regolata da un singolo parametro scalare: le *entropie di Renyi*. Sono il diretto analogo delle grandezze definite in campo classico da Alfréd Rényi rilassando i postulati di Fadeev, che determinano univocamente la forma dell'entropia di Shannon come misura di incertezza per una distribuzione [29]:

Teorema 2.13. *L'entropia di Shannon di una distribuzione è univocamente determinata dalle seguenti proprietà:*

1. *É simmetrica per scambio di qualsiasi coppia di probabilità;*
2. $H(\{p, 1-p\})$ *è continua in* p*;*
3. *É normalizzata:* $H(\{\frac{1}{2}, \frac{1}{2}\}) = \log 2$*;*
4. *Per* $0 \leq t \leq 1$ *soddisfa:*

$$H(\{tp_1, (1-t)p_1, p_2, \ldots, p_n\}) = H(\{p_1, p_2, \ldots, p_n\}) + p_1 H(\{t, 1-t\}) \quad (2.52)$$

Renyi osservò che queste proprietà implicano l'*additività*, la quale però viceversa è più debole della *4.*. Dunque se la si sostituisce alla *4.* si ottiene una *famiglia* di entropie, che dimostrò avere la seguente forma:

$$H_\alpha(\{p_i\}) \equiv \frac{1}{1-\alpha} \log\left(\sum_i p_i^\alpha\right) \quad (2.53)$$

con parametro reale $\alpha > 0$ e $\alpha \neq 1$. Nel medesimo articolo in cui dava questa definizione Renyi formulò anche una caratterizzazione più immediata in termini di variabili stocastiche e distribuzioni *generalizzate*, che includono sia quelle ordinarie (ridefinite *complete*) sia quelle *incomplete*, ovvero non normalizzate o più precisamente definite su un sottoinsieme proprio degli eventi - a rigore, con misura inferiore a quella dell'insieme universo (unitaria per definizione) [29]. L'analogo quantistico delle distribuzioni di variabili incomplete sono operatori positivi non normalizzati, con traccia *minore* di 1 - dunque a rigore non stati. Si possono così riformulare gli assiomi per l'entropia di Von Neumann [30]:

Teorema 2.14. *Il funzionale sugli operatori densità entropia di Von Neumann è definito univocamente dalle seguenti proprietà:*

1. *É continuo;*
2. *É invariante per unitarie;*
3. *É normalizzato:* $S(\frac{1}{d}\mathbb{I}_d) = \log d$*;*
4. *É additivo:* $S(\hat{\rho} \otimes \hat{\sigma}) = S(\hat{\rho}) + S(\hat{\sigma})$*;*
5. ***Proprietà della media*** *Per la somma diretta di operatori positivi* $\hat{r}$ *e* $\hat{\sigma}$ *t.c.* $\mathrm{Tr}[\hat{\rho} + \hat{\sigma}] \leq 1$[45] *con supporti a intersezione triviale vale:*

$$S(\hat{\rho} \oplus \hat{\sigma}) = \frac{\mathrm{Tr}[\hat{\rho}]}{\mathrm{Tr}[\hat{\rho} + \hat{\sigma}]} S(\hat{\rho}) + \frac{\mathrm{Tr}[\hat{\sigma}]}{\mathrm{Tr}[\hat{\rho} + \hat{\sigma}]} S(\hat{\sigma}) \tag{2.54}$$

ove $\hat{\rho} \oplus \hat{\sigma}$ *è un operatore che valuta alla somma delle azioni di* $\hat{\rho}$ *e* $\hat{\sigma}$[46]*.*

L'assioma da rilassare, sempre in analogia con il ragionamento di Renyi, è il quinto. Anziché solo quella aritmetica, si ammette una media generica: è sufficiente che esista una funzione continua e strettamente monotona (dunque invertibile) g t.c.

$$S(\hat{\rho} \oplus \hat{\sigma}) = g^{-1}\Big(\frac{\mathrm{Tr}[\hat{\rho}]}{\mathrm{Tr}[\hat{\rho} + \hat{\sigma}]} g\big(S(\hat{\rho})\big) + \frac{\mathrm{Tr}[\hat{\sigma}]}{\mathrm{Tr}[\hat{\rho} + \hat{\sigma}]} g\big(S(\hat{\sigma})\big)\Big) \tag{2.55}$$

Scegliendo una g lineare si ottiene di nuovo l'entropia di Shannon; se invece si adotta una forma esponenziale, regolata da un parametro reale (la base è da variare a seconda del logaritmo che si sceglie di utilizzare):

$$g_\alpha(x) = 2^{(\alpha-1)x} \tag{2.56}$$

si ottengono le entropie di Renyi:

Definizione 2.14. *Dato un qualsiasi numero reale* $\alpha \in]0, +\infty[$*,* $\alpha \neq 1$*, l'entropia di Renyi di ordine* α *di uno stato* $\hat{\rho}$ *è la quantità:*

$$S_\alpha(\hat{\rho}) \equiv \frac{1}{1-\alpha} \log \mathrm{Tr}[\hat{\rho}^\alpha] \tag{2.57}$$

La definizione è estesa a $\alpha = 0, 1, +\infty$ *secondo* $S_\alpha(\hat{\rho}) \equiv \lim_{\gamma \to \alpha} S_\gamma(\hat{\rho})$*.*

Come per l'entropia di Shannon/Von Neumann, si pone convenzionalmente $f(0) = 0$ per funzioni divergenti all'origine. Alcune osservazioni [30]:

- Per $\alpha \to 1$ (il limite è univoco) si recupera proprio l'entropia di Von Neumann;

[45] Ovvero t.c. la loro somma possa essere interpretata a sua volta come "stato generalizzato."

[46] Si può pensare specificamente ai due operatori come proiettori pesati su stati ortogonali.

- Per $\alpha > 1$ viene dato maggior peso agli autovalori maggiori, ovvero agli stati più probabili della miscela ortonormale. Nel limite $\alpha \to +\infty$ si ottiene la **min-entropy** $S_\infty(\hat{\rho}) = S_{\min}(\hat{\rho}) = -\log\left(\max_{\{\alpha_i\}} \alpha_i\right) = -\log \|\hat{\rho}\|_\infty$;
- Per $\alpha < 1$ contribuiscono invece maggiormente gli stati più "soprendenti", ovvero meno probabili. Nel limite $\alpha \to 0$ si ottiene l'**entropia di Hartley**: $S_0(\hat{\rho}) = \log \operatorname{rank} \hat{\rho}$[47];
- Per qualunque stato, S_α è non crescente in α[48].

Le proprietà dell'entropia di Von Neumann ulteriori a quelle caratterizzanti in generale non si estendono alle entropie di Renyi di ordine arbitrario. Ad esempio, S_α è concava per $\alpha \leq 1$ ma per $\alpha > 1$ diviene convessa sopra una soglia che dipende dalla cardinalità della miscela. Si possono invece generalizzare naturalmente le definizioni delle grandezze derivate, ovvero l'entropia relativa, l'entropia condizionale e l'informazione mutua [30]. Per la prima si ha in particolare:

$$S_\alpha(\hat{\rho}||\hat{\sigma}) \equiv \frac{1}{1-\alpha} \operatorname{Tr}[\hat{\rho}^\alpha \hat{\sigma}^{1-\alpha}] \tag{2.58}$$

Si nota che rimane congiuntamente convessa solamente per $\alpha \leq 2$ [30].

2.4 Misure di distanza

Si è introdotta in precedenza l'entropia relativa come prima misura della distanza tra due stati quantistici. In generale si possono definire diverse grandezze che quantificano la vicinanza tra due stati e quindi la bontà dell'approssimazione di un vettore di stato o più generalmente un operatore densità.

2.4.1 Overlap

Per stati puri uno strumento rudimentale ma efficace è l'**overlap**, ovvero il prodotto scalare tra i vettori (assunti normalizzati), più frequentemente definito a rigore come il modulo quadro di questo:

$$|\langle\psi|\phi\rangle|^2 \tag{2.59}$$

Si ottiene così un numero reale non negativo con una chiara interpretazione: quantifica la probabilità di ottenere un risultato affermativo effettuando su uno stato una misura che distingua l'altro, ovvero più in generale quanto è difficile *statisticamente* distinguere i due stati tramite misure. Stati ortogonali sono chiaramente discernibili con una singola misura, opportunamente costruita. Si noti che si tratta di una misura di *vicinanza*: stati più *distanti* hanno un overlap *minore*.

[47] A rigore pertanto non soddisfa la continuità sugli operatori.

[48] Dunque per qualsiasi entropia di Renyi vale $S_\alpha \leq \log \operatorname{rank} \hat{\rho}$.

2.4.2 Fidelity

L'overlap può essere generalizzato (si chiarirà in che termini) agli operatori densità definendo la **fedeltà (fidelity)**, definita in analogia ad una grandezza classica omonima per distribuzioni di probabilità:

Definizione 2.15. *Dati due stati $\hat{\rho}$ e $\hat{\sigma}$ la loro fidelity è la quantità:*

$$F(\hat{\rho},\hat{\sigma}) \equiv \mathrm{Tr}\left[\sqrt{\hat{\rho}^{1/2}\hat{\sigma}\hat{\rho}^{1/2}}\right] \tag{2.60}$$

Si tratta di una funzione simmetrica e concava sia in ciascun argomento che congiuntamente; inoltre è invariante per unitarie (applicate a entrambi gli argomenti). Poiché operatori positivi ammettono sempre radice quadrata, è sempre ben definita e si può esprimere anche come $F(\hat{\rho},\hat{\sigma}) = \left\|\hat{\sigma}^{1/2}\hat{\rho}^{1/2}\right\|_1$; da questa forma è semplice verificare che il suo intervallo di valori è [0,1], con l'unità ottenuta se e solo se i due stati coincidono. É d'interesse esaminare alcuni casi particolari:

- Se i due stati commutano, e sono dunque simultaneamente diagonalizzabili[49], la fidelity si riduce all'espressione classica per le distribuzioni associate:

$$F(\{r_i\},\{s_i\}) = \textstyle\sum_i \sqrt{r_i s_i} \qquad \text{ove} \qquad \hat{\rho} = \sum_i r_i \left|i\right\rangle\!\left\langle i\right| \quad \hat{\sigma} = \sum_i s_i \left|i\right\rangle\!\left\langle i\right| \tag{2.61}$$

- Se uno stato è puro, corrisponde alla radice quadrata del valore di aspettazione dell'altro stato, ovvero della media pesata degli overlap:

$$F(\hat{\rho},\left|\psi\right\rangle\!\left\langle\psi\right|) = \sqrt{\left\langle\psi\right|\hat{\rho}\left|\psi\right\rangle} = \sqrt{\textstyle\sum_i p_i \left|\left\langle\phi_i|\psi\right\rangle\right|^2} \quad \text{ove} \quad \hat{\rho} = \textstyle\sum_i p_i \left|\phi_i\right\rangle\!\left\langle\phi_i\right| \tag{2.62}$$

- Se entrambi sono puri, è il modulo dell'overlap: $F(\left|\phi\right\rangle\!\left\langle\phi\right|, \left|\psi\right\rangle\!\left\langle\psi\right|) = \left|\left\langle\phi|\psi\right\rangle\right|$

In realtà ammette un'interpretazione in termini di overlap anche nel caso doppiamente misto, grazie al **Teorema di Uhlmann**:

Teorema 2.15. *Siano $\hat{\rho}$ e $\hat{\sigma}$ stati di un sistema S. Introducendo una copia $\tilde{S}$ di S, vale il seguente risultato:*

$$F(\hat{\rho},\hat{\sigma}) = \max_{\{\left|\psi_\rho\right\rangle,\left|\psi_\sigma\right\rangle\}} \left|\left\langle\psi_\rho|\psi_\sigma\right\rangle\right| \tag{2.63}$$

ove si massimizza sulle possibili purificazioni di $\hat{\rho}$ e $\hat{\sigma}$ nel sistema composto $S\tilde{S}$.

[49] La non commutatività, che segue dall'overlap tra gli stati - ovvero la non piena distinguibilità degli "esiti", è la differenza fondamentale del caso quantistico.

A rigore, definire la fidelity una "misura" della vicinanza di due stati è un abuso di terminologia, in quanto non si tratta di una metrica. Tuttavia è possibile ricavarne una distanza[50] secondo:

$$d(\hat{\rho}, \hat{\sigma}) \equiv \arccos F(\hat{\rho}, \hat{\sigma}) \tag{2.64}$$

questa non è che l'angolo (ovvero la lunghezza della geodetica) tra i due stati sulla sfera unitaria, ove si adotti per definirla la 1-norma.

2.4.3 Trace distance

Si è osservato per i tensori che il prodotto scalare di Hilbert-Schmidt induce naturalmente una norma omonima; questo vale quindi nel caso particolare dello spazio degli operatori. Non si tratta chiaramente dell'unica possibile: ad esempio si è già introdotta per definire la negatività la 1-norma. In piena analogia con gli spazi L^p è in realtà possibile definire una classe di norme - e annessi spazi normati di operatori limitati - regolata da un parametro: le **norme di Schatten** [31].

Definizione 2.16. *Per ogni operatore limitato tra due spazi di Hilbert $\hat{A}$ e numero reale $p \in [0, +\infty[$ si definisce la p-norma di Schatten di $\hat{A}$ secondo:*

$$\|\hat{A}\|_p \equiv \left(\mathrm{Tr}[(\hat{A}^\dagger \hat{A})^{p/2}]\right)^{1/p} \tag{2.65}$$

Per operatori diagonalizzabili la norma si riduce a $\left(\sum_i |\lambda_i|^p\right)^{1/p}$. É interessante notare che nel limite $p \to +\infty$ si ottiene la norma operatoriale; valgono inoltre tutte le altre proprietà delle p-norme per vettori, successioni e funzioni, incluse in particolare le diseguaglianze di Hölder e Minkowski. Per i prodotti tensori (a prescindere da differenze tra gli spazi locali) vale un'interessante proprietà [31]:

$$\|\hat{A}_1 \otimes \cdots \otimes \hat{A}_n\|_p = \|\hat{A}_1\|_p \cdots \|\hat{A}_n\|_p \tag{2.66}$$

Ogni norma genera una distanza che permette di esprimere la vicinanza di due operatori in termini di differenti proprietà; quella più utilizzata è proprio la 1-norma, che dà luogo alla cosiddetta **distanza di traccia (trace distance)**.

Definizione 2.17. *La trace distance tra due stati $\hat{\rho}$ e $\hat{\sigma}$ è definita secondo:*

$$D(\hat{\rho}, \hat{\sigma}) \equiv \frac{1}{2}\|\hat{\rho} - \hat{\sigma}\|_1 \tag{2.67}$$

Essa è l'analogo quantistico della **distanza di Kolmogorov** per distribuzioni:

$$D(\{r_i\}, \{s_i\}) \equiv \frac{1}{2}\sum_i |r_i - s_i| \tag{2.68}$$

[50] Che quindi soddisfa tutti gli assiomi della definizione tra cui in particolare la diseguaglianza triangolare.

cui si riduce sempre per operatori commutanti. Classicamente la trace distance è interpretabile come la massima differenza di probabilità cumulativa ottenibile per un sottoinsieme degli eventi opportunamente scelto; quantisticamente ha invece un significato geometrico intuitivo: è metà della distanza tra i vettori di Bloch associati ai due stati (vedasi dopo). A livello pratico il suo utilizzo è equivalente a quello della fidelity, in quanto vale la coppia di diseguaglianze:

$$1 - F(\hat{\rho}, \hat{\sigma}) \leq D(\hat{\rho}, \hat{\sigma}) \leq \sqrt{1 - F(\hat{\rho}, \hat{\sigma})^2} \tag{2.69}$$

A differenza di F, però, D è una metrica e quindi soddisfa in particolare la diseguaglianza triangolare. Inoltre è invariante per unitarie, monotona (nel senso utilizzato per l'entropia relativa), convessa in ciascun argomento e più generalmente **fortemente convessa**:

Teorema 2.16. *Considerando due combinazioni convesse di operatori densità con medesimo numero di termini vale:*

$$D\left(\sum_i p_i \hat{\rho}_i, \sum_i q_i \hat{\sigma}_i\right) \leq D\left(\{p_i\}, \{q_i\}\right) + \sum_i p_i D(\hat{r}_i, \hat{\sigma}_i) \tag{2.70}$$

La monotonia in realtà è un caso particolare di una proprietà più generale:

Teorema 2.17. *Dati due stati $\hat{\rho}$ e $\hat{\sigma}$ e un'operazione $\hat{\mathcal{E}}$ che preserva la traccia, vale:*

$$D(\hat{\mathcal{E}}(\hat{\rho}), \hat{\mathcal{E}}(\hat{\sigma})) \leq D(\hat{\rho}, \hat{\sigma}) \tag{2.71}$$

La traccia parziale è infatti un'operazione che preserva la traccia complessiva. A rigore, per dimostrare questo risultato è sufficiente assumere come ipotesi che $\hat{\mathcal{E}}$ sia una mappa *positiva* e non specificamente un'operazione.

2.5 Spin chains

Lo studio di sistemi a molti corpi, già arduo per la maledizione della dimensionalità, risulta del tutto proibitivo se si intende tenere conto della piena complessità degli oggetti fisici. Fortunatamente si tratta di complicazioni spesso superflue, in quanto è possibile definire sistemi ben più trattabili isolando i limitati gradi di libertà rilevanti nei contesti sperimentali d'interesse: si ottengono così modelli dalla definizione essenziale ma pur sempre all'origine di una fisica sufficientemente elaborata. Una classe generale di modelli con amplissimo campo di applicazione sono quelli su **reticolo (lattice)**, originati dalla discretizzazione dello spazio(-tempo) e/o dei gradi di libertà associati ai nodi del reticolo. Sono ampiamente utilizzati anche in meccanica statistica classica, ma per i fini della trattazione si considera specificamente il caso quantistico.
Di seguito si dà una breve introduzione ai concetti fondamentali per descrivere un sistema su reticolo, delineando le modifiche al formalismo necessarie per trattare il

caso infinito, che è rilevante in quanto descrive il **limite termodinamico**. Per la definizione rigorosa di un lattice system sono necessari tre elementi[51] [24, 32]:

Definizione 2.18. *Un modello su reticolo d-dimensionale è definito da:*

1. *Un reticolo, ovvero un insieme di punti $\Lambda \subset \mathbb{R}^d$ finito o al più numerabile[52]. Alternativamente si utilizza un grafo, ove gli spigoli definiscono le relazioni di vicinato. In entrambi i casi è naturalmente definita una metrica;*

2. *Un insieme di spazi di Hilbert $\mathcal{H}_j$ associati a ciascun punto $x_j \in \Lambda$ del reticolo. Il sistema complessivo ha quindi spazio degli stati $\mathcal{H}_\Lambda = \bigotimes_{x_j \in \Lambda} \mathcal{H}_j$[53];*

3. *Una mappa Φ sull'insieme delle parti del reticolo, detta **interazione**, t.c. $\forall X \in \mathcal{P}(\Lambda)\ \Phi(X) \in \mathcal{B}(\mathcal{H}_X) \wedge \Phi(X)^\dagger = \Phi(X)$;*

L'hamiltoniana per ciascun sottosistema definito da $Z \subset \Lambda$ è $H_Z = \sum_{X \in \mathcal{P}(Z)} \Phi(X)$

Se gli spazi locali sono di dimensione finita o al più numerabile, di principio la separabilità dello spazio degli stati complessivo è garantita; tuttavia la definizione immediata come prodotto tensore infinito è patologica in quanto il prodotto scalare può divergere [33]. Si considerano allora primariamente gli operatori [34], in particolare quelli **locali**, che sono ben definiti e hanno senso fisico [24, 33, 35]:

Definizione 2.19. *Il **supporto** di un operatore su un reticolo è il più piccolo insieme di siti su cui la sua azione è non triviale.*
*Un operatore su reticolo è detto **locale** se equivalentemente:*

1. *Ha supporto compatto;*

2. *Ha supporto su un numero finito di siti;*

3. *Detto **raggio** dell'operatore il diametro del supporto[54], se ha raggio finito.*

Specificamente se ha raggio k è detto k-locale. A rigore si può distinguere tra un operatore con supporto connesso di raggio k, detto propriamente k-locale, e un generico operatore con supporto su k siti eventualmente anche non limitrofi, detto **a k corpi** (**k-body**).
In quest'ottica è giustificata la costruzione dell'hamiltoniana tramite la mappa di interazione: la somma infinita è infatti definita solo formalmente [34, 33]. Per

[51] La definizione data richiede a ri

[52] Quindi in generale in biezione con un sottoinsieme di $\mathbb{Z}^d$.

[53] A rigore, la definizione dello spazio generale e di conseguenza le considerazioni di seguito richiedono alcune modifiche per il caso di sistemi di seconda quantizzazione; tuttavia non si ritiene necessario dettagliarle in quanto tali sistemi non saranno specificamente presi in esame.

[54] Ovvero la distanza massima tra due suoi punti.

costruire l'algebra del sistema complessivo si considera innanzitutto l'insieme degli osservabili locali, che è l'unione delle algebre per tutti i sottoinsiemi finiti[55]:

$$\mathcal{B}_{loc} \equiv \bigcup_{\substack{X\subset\Lambda \\ \#X<\infty}} \mathcal{B}_X \tag{2.72}$$

L'algebra complessiva è quindi ottenuta completando rispetto alla norma operatoriale[56] [34]. A questo punto si possono definire gli stati sfruttando essenzialmente le proprietà caratterizzanti dell'operatore densità (e della traccia) [34, 24]:

Definizione 2.20. *Uno stato è un funzionale ω su $\overline{\mathcal{B}_{loc}}$ tale per cui*[57]*:*

$$\omega(\mathbb{I}) = 1 \quad \wedge \quad \omega(\hat{O}^\dagger \hat{O}) \geq 0 \quad \forall \hat{A} \tag{2.73}$$

Si può dimostrare che sono generalizzate tutte le proprietà rilevanti degli operatori densità su spazi a dimensione finita. In questo formalismo, si può dare una definizione rigorosa anche per gli stati fondamentali, che saranno rilevanti nella trattazione successiva, per quanto non si possa diagonalizzare rigorosamente l'hamiltoniana infinita [24, 34]. Si noti innanzitutto che in termini di operatore densità:

Definizione 2.21. *Uno stato fondamentale è un operatore densità il cui supporto è incluso nell'autospazio dell'energia minima.*

Infatti se l'autovalore è degenere bisogna considerare anche possibili miscele. Per generalizzare questa definizione è necessario determinare una proprietà caratterizzante che si possa trasportare naturalmente nel caso infinito.
Si consideri uno stato fondamentale puro di un sistema finito, assunto normalizzato. Applicandovi una generica perturbazione:

$$|\psi_0\rangle \mapsto |\phi\rangle \equiv \frac{\hat{V}|\psi_0\rangle}{\|\hat{V}|\psi_0\rangle\|} = \frac{\hat{V}|\psi_0\rangle}{\sqrt{\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle}} \tag{2.74}$$

chiaramente la sua energia non può diminuire, in quanto al più lo stato trasformato ha componenti su autospazi a energia maggiore:

$$\langle\phi|\hat{H}|\phi\rangle = \frac{\langle\psi_0|\hat{V}^\dagger\hat{H}\hat{V}|\psi_0\rangle}{\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle} \geq \langle\psi_0|\hat{H}|\psi_0\rangle \tag{2.75}$$

[55] La procedura ha senso in quanto, generalizzando quanto discusso per il prodotto tensore di spazi di Hilbert, se un sottoinsieme ne contiene un altro allora vi è una naturale inclusione anche per le rispettive algebre [24].

[56] Intuitivamente, di modo da includere gli operatori approssimabili con precisione arbitraria da quelli locali, detti "quasi-locali"; formalmente è necessario per avere la proprietà di Banach.

[57] L'identità è l'unità dell'algebra.

Questa diseguaglianza si può riformulare sfruttando il fatto che il proiettore sullo stato fondamentale commuta con l'hamiltoniana, come si può verificare considerandone la decomposizione spettrale, e che $|\psi_0\rangle$ è normalizzato:

$$\langle\psi_0|\hat{V}^\dagger\hat{H}\hat{V}|\psi_0\rangle \geq \langle\psi_0|\hat{V}^\dagger\hat{V}\,|\psi_0\rangle\langle\psi_0|\,\hat{H}|\psi_0\rangle = \langle\psi_0|\hat{V}^\dagger\hat{V}\hat{H}|\psi_0\rangle \tag{2.76}$$

Ovvero:

$$\langle\psi_0|\hat{V}^\dagger[\hat{H},\hat{V}]\,|\psi_0\rangle \geq 0 \tag{2.77}$$

Si può notare poi che questa condizione, oltre ad essere necessaria, è anche sufficiente. Infatti considerando il contronominale dell'implicazione inversa, se $|\psi\rangle$ non è uno stato fondamentale, allora la diseguaglianza è violata per $\hat{V} = |\psi_0\rangle\langle\psi|$.
Questa proprietà può essere immediatamente generalizzata per linearità agli operatori densità:

Teorema 2.18. *Un'operatore $\hat{\rho}$ descrive uno stato fondamentale se e solo se per qualsiasi operatore $\hat{V}$ vale:*

$$\mathrm{Tr}[\hat{\rho}(\hat{V}^\dagger[\hat{H},\hat{V}]\,)] \geq 0 \tag{2.78}$$

Ora, sebbene l'hamiltoniana nel limite infinito - detto limite **termodinamico** - sia definita solo formalmente, per operatori locali si può ben definire il commutatore:

$$[\hat{H},\hat{O}] \equiv [\hat{H}_\Lambda,\hat{O}] \qquad \Lambda = \mathrm{supp}\,\hat{O} \tag{2.79}$$

Si può così dare la definizione generale [24, 34]:

Definizione 2.22. *Uno stato ω_0 è uno stato fondamentale di $\hat{H}$ se vale:*

$$\omega(\hat{O}^\dagger[\hat{H},\hat{O}]\,) \geq 0 \qquad \forall\hat{O} \in \mathcal{B}_{loc} \tag{2.80}$$

In questo lavoro si considereranno specificamente sistemi con spazi locali finito dimensionali, che sono definiti in generale **sistemi di spin (spin systems)** [24]. La denominazione è appropriata poiché osservabili e interazioni sono di norma costruiti sull'algebra dei momenti angolari, in conseguenza del fatto che fisicamente descrivono interazioni tra dipoli magnetici originati da momenti di cariche.

2.5.1 Spin

Fisicamente lo **spin** è il momento angolare intrinseco; è associato propriamente alle particelle elementari - per cui è introdotto rigorosamente tramite la teoria dei campi - ma per estensione indica anche quello degli oggetti composti, come gli atomi, dato dalla composizione dei singoli momenti intrinseci ed orbitali. Si tratta di una grandezza vettoriale quantizzata sia in modulo che in lunghezza delle

proiezioni sui vari assi di un qualsiasi sistema cartesiano [36].
Formalmente, uno spin s (o un *sistema di* spin s), con s intero o semintero - ovvero $s \in \frac{1}{2}\mathbb{N}$, è definito da una rappresentazione irriducibile[58] di dimensione finita $d = 2s + 1$[59] del gruppo unitario speciale $SU(2)$[60].
La rappresentazione è determinata fissando quella dell'algebra dei generatori del gruppo, che sono gli operatori corrispondenti alle proiezioni dello spin sui tre assi ortogonali $\hat{S}^x$, $\hat{S}^y$, $\hat{S}^z$ - o alternativamente $\hat{S}^1$, $\hat{S}^2$, $\hat{S}^3$; tramite questi e l'identità si costruisce l'algebra degli osservabili $(2s+1)^2$-dimensionale[61] [37]. Le componenti dello spin soddisfano le relazioni di commutazione (espresse concisamente tramite il tensore di Levi-Civita):

$$[\hat{S}^i, \hat{S}^j] = i\hbar\varepsilon_{ijk}\hat{S}^k \tag{2.81}$$

e possono essere racchiuse in forma compatta in un vettore $\hat{\mathbf{S}}$, che però in conseguenza della noncommutatività non è osservabile. Il modulo dello spin è determinato dal parametro dimensionale e vale $\hbar^2 s(s+1)$ - unico autovalore sullo spazio dell'osservabile associata $\hat{S}^2 = \hat{\mathbf{S}} \cdot \hat{\mathbf{S}} \equiv \hat{S}^i\hat{S}^i = \hbar^2 s(s+1)\mathbb{I}_{2s+1}$[62]. Questo commuta con le singole proiezioni e di conseguenza con qualsiasi hamiltoniana[63].
Le proiezioni hanno spettro non degenere e pertanto è possibile definire una base scegliendo autovettori di una di esse; di consueto si utilizza l'asse z (3). I ket sono denominati tramite un parametro s^z soggetto al vincolo $|s^z| \leq s$ e legato allo spettro secondo[64] [38]:

$$\hat{S}^z |s^z\rangle = \hbar s^z |s^z\rangle \tag{2.82}$$

É utile definire una coppia di operatori **di innalzamento** e **abbassamento**, detti generalmente **operatori scaletta (ladder operators)** che mappano gli autospazi

[58] O più in generale da una classe di rappresentazioni equivalenti a meno di un unitaria, ovvero essenzialmente una rotazione del sistema di riferimento.

[59] Lo spazio degli stati è identificabile con $\mathbb{C}^{2s+1}$.

[60] Che è il doppio ricoprimento del gruppo delle rotazioni nello spazio tridimensionale $SO(3)$.

[61] Sono una base per la rappresentazione fondamentale $s = 1/2$, per quelle di dimensione maggiore si considerano anche delle combinazioni di prodotti.

[62] L'azione non è triviale per sistemi composti, ove il prodotto tensore delle singole rappresentazioni ne dà una riducibile, ovvero decomponibile in somma diretta di rappresentazioni a diversi valori del modulo (tramite i coefficienti di Clebsch-Gordan). Per singolo spin vale invece il lemma di Schur, per cui un operatore che commuta con le tre proiezioni deve essere multiplo dell'identità [37].

[63] Trivialmente qualsiasi operatore commuta con l'identità.

[64] I suoi valori sono i cosiddetti pesi della rappresentazione; essendo la proiezione compatibile con l'hamiltoniana si dice che s^z è un *buon numero quantico*.

di $\hat{S}^z$ in quelli adiacenti, con s^z rispettivamente maggiore e minore [38]:

$$\hat{S}^+ |s^z\rangle = \hbar\sqrt{s(s+1) - s^z(s^z+1)}\, |s^z+1\rangle \tag{2.83}$$

$$\hat{S}^- |s^z\rangle = \hbar\sqrt{s(s+1) - s^z(s^z-1)}\, |s^z-1\rangle \tag{2.84}$$

Si può dare un'espressione in termini di questi per le proiezioni sugli altri assi:

$$\hat{S}^x = \frac{1}{2}\left(\hat{S}^+ + \hat{S}^-\right) \qquad \hat{S}^y = \frac{1}{2}\left(\hat{S}^+ - \hat{S}^-\right) \tag{2.85}$$

Per praticità le espressioni possono essere semplificate, come si farà di seguito, usando il sistema delle unità naturali, per cui $\hbar = 1$.

Spin $1/2$ Un caso particolarmente rilevante è quello della rappresentazione fondamentale, ovvero lo spin $1/2$. In questo caso si può sfruttare l'isomorfismo con $\mathbb{C}^2$ e rappresentare gli stati come vettori complessi bidimensionali. Scegliendo la base che diagonalizza la componente z, gli stati possono essere rappresentati secondo:

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.86}$$

Gli operatori di spin sono invece rappresentati tramite le **matrici di Pauli** [39]:

$$\hat{S}^i = \frac{1}{2}\sigma^i \qquad \sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.87}$$

Come annotato, assieme all'identità queste danno una base completa per le osservabili [36]. In particolare qualsiasi stato può essere espresso nella forma:

$$\hat{\rho} = \frac{1}{2}(\mathbb{I}_2 + \mathbf{r} \cdot \boldsymbol{\sigma}) \tag{2.88}$$

ove le matrici sono riunite in un vettore $\boldsymbol{\sigma}$[65].

2.5.2 Dimensione finita e infinita

La tipologia più semplice di sistemi di spin non triviali sono quelli unidimensionali, per cui $\Lambda \subset \mathbb{R}$ o equivalentemente $\Lambda \subseteq \mathbb{Z}$, che sono detti **catene di spin (spin chains)** [39]. Le spin chains sono classificate come:

[65] Il vettore $\mathbf{r}$ definisce un punto interno alla sfera unitaria in $\mathbb{R}^3$, che in questo contesto è detta *sfera di Bloch*. Si noti che, per la caratterizzazione degli operatori densità, uno stato è puro se è solo se il punto corrispondente è sulla *superficie* della sfera.

- **Finite** o **infinite**, in base alla cardinalità del reticolo. Il caso infinito è definito come limite termodinamico $n \to \infty$ di quello finito;
- **Aperte** o **chiuse**, a seconda delle condizioni al contorno specificate dall'interazione (indicate di norma come **Open Boundary Conditions, OBC** o **Periodic Boundary Conditions, PBC**).

Come anticipato, gli osservabili e le interazioni sono costruite a partire dalle algebre dei singoli siti, dunque dagli operatori di spin locali $\hat{S}_j^x$, $\hat{S}_j^y$, $\hat{S}_j^z$. Considerati come elementi dell'algebra complessiva, hanno supporti a intersezione triviale e quindi sono compatibili; si hanno così relazioni generalizzate di commutazione [37]:

$$[\hat{S}_\alpha^i, \hat{S}_\beta^j] = i\delta_{\alpha\beta}\varepsilon_{ijk}\hat{S}_\alpha^k \tag{2.89}$$

Si possono quindi definire gli operatori di spin per il sistema complessivo:

$$\hat{\mathbf{S}} = \sum_\alpha \hat{\mathbf{S}}_\alpha \tag{2.90}$$

Ogni componente commuta trivialmente con le singole componenti dei siti, e quindi vi è compatibile; questo non vale però per il modulo (la rappresentazione è riducibile). Di conseguenza $\hat{\mathbf{S}}$ commuta con ogni prodotto scalare di due spin [37]:

$$[\hat{\mathbf{S}}, \hat{\mathbf{S}}_\alpha \cdot \hat{\mathbf{S}}_\beta] = \mathbf{0} \quad \forall \alpha \neq \beta \qquad\qquad \hat{\mathbf{S}}_\alpha \cdot \hat{\mathbf{S}}_\beta \equiv \hat{S}_\alpha^i \hat{S}_\beta^i \tag{2.91}$$

2.6 Problema generale

Definito il formalismo e il tipo di sistema d'interesse, è ora possibile formulare il problema a molti corpi considerato, specificando quali siano i dati, la soluzione e le osservabili rilevanti.

1. Il problema è specificato definendo:
   (a) Il sistema, quindi lo spazio degli stati;
   (b) L'hamiltoniana associata.

   Come detto nel caso infinito la definizione è solo formale e si deve far uso del formalismo più generale per considerazioni rigorose;
2. Il problema è risolto determinando lo spettro $\{E_n\}$ dell'hamiltoniana e una corrispondente base di stati stazionari $\{|\psi_n\rangle\}$. In realtà è sufficiente determinare la porzione a bassa energia in quanto è la sola rilevante intorno all'equilibrio per temperature considerate ordinariamente in contesto sperimentale;
3. Gli stati danno una descrizione completa del sistema, quindi è possibile estrarre i valori di aspettazione per qualsiasi osservabile. In generale hanno rilevanza classi specifiche:

(a) Osservabili a singolo sito;

(b) Correlatori *connessi* a due siti: $\langle \hat{O}_i \hat{O}'_j \rangle - \langle \hat{O}i \rangle \langle \hat{O}'_j \rangle$. In particolare per sistemi uniformi si considerano in generale correlatori ad una data distanza:

$$C(r) \equiv \langle \hat{O}_i \hat{O}'_{i+r} \rangle - \langle \hat{O}_i \rangle \langle \hat{O}'_{i+r} \rangle \tag{2.92}$$

In generale lo scopo non è determinare una configurazione esatta del sistema, bensì proprietà locali e globali nel limite termodinamico. In questo senso correlatori sono utili per descrivere la struttura delle correlazioni (classiche e puramente quantistiche) degli stati e di conseguenza l'*ordine* su diverse scale - permettendo di definire *parametri d'ordine* che lo quantificano. Si noti che oltre alle proprietà statiche sono anche d'interesse dal punto di vista dinamico, in quanto caratterizzando l'accoppiamento tra diverse regioni permettono di descrivere la propagazione di perturbazioni[66]. Tramite i parametri d'ordine si possono così definire delle **fasi**, ovvero degli insiemi di stati che esibiscono le medesime proprietà. Sotto un cambiamento di parametri **di controllo** lo stato di equilibrio del sistema può spostarsi tra le fasi in una **transizione** [7].
Nella teoria classica di Landau, le transizioni di fase sono associate a fluttuazioni termiche e pertanto il parametro di controllo è di consueto la temperatura. Per i sistemi fortemente correlati in basse dimensioni, invece, si possono avere (a rigore per $T = 0$, dunque quando lo stato di equilibrio è lo stato fondamentale) transizioni di fase associate a fluttuazioni *quantistiche*, dette appunto **Transizioni di Fase Quantistiche (Quantum Phase Transitions, QPT)** - in cui il parametro di controllo è una quantità (energetica) che regola gli accoppiamenti dell'hamiltoniana [36, 40]. In virtù delle correlazioni, alcune di queste fasi possono esibire ordine a lungo raggio e di conseguenza degli invarianti di simmetria - ovvero, nel senso formale del termine, delle *cariche* - la cui conservazione determina la struttura energetica del sistema (si dice che le simmetrie *proteggono* certe proprietà). Per via di tali caratteristiche si parla di **fasi topologiche**, particolarmente rilevanti nel contesto della superconduttività ad alta temperatura o del fault-tolerant quantum computing [7].

[66] In generale proprietà di equilibrio e di risposta dinamica di un sistema sono legate, come noto dal teorema di fluttuazione-dissipazione.

Capitolo 3

Matrix Product States

Definito il problema e sviluppato il linguaggio dei TN con cui esprimerlo e risolverlo, si illustra la particolare categoria di network che realizza ansatz ottimali per il caso 1D. Se ne descrivono le proprietà, le gauge e le procedure diagrammatiche per estrarre le informazioni rilevanti; in primo luogo però si giustifica formalmente la loro efficacia per descrivere l'insieme degli stati rilevanti.

3.1 Area law

Per superare i limiti delle tecniche tradizionali di risoluzione del problema a molti corpi un approccio naturale e del tutto generale è quello di generalizzare il campo medio, ossia di applicare il principio variazionale su un particolare insieme - più precisamente, una *varietà* - di stati, identificato secondo ragioni *fisiche*, parametrizzabile in modo efficiente (rispetto al numero di siti). Il punto cruciale di questo lavoro è che per una classe estremamente ampia di modelli unidimensionali si può giustificare rigorosamente, fissato un errore arbitrariamente piccolo, l'esistenza di una varietà - ovvero equivalentemente di una struttura di ansatz - tramite cui si possono approssimare stati fondamentali, primi eccitati e termici[1] e che questa è ottimamente caratterizzata tramite una particolare forma di TN [41].
La dimostrazione - la quale suggerisce anche in modo naturale la costruzione di questa classe di stati - si basa sul fatto che, sebbene nominalmente si tratti di sistemi fortemente correlati, se le interazioni avvengono solo *localmente* e vi è una soglia finita di energia per eccitarli dallo stato fondamentale, allora negli stati rilevanti l'entanglement è egualmente *locale*, ovvero coinvolge solo siti limitrofi: formalmente, l'entropia di entanglement di blocchi di siti non è *estensiva*, bensì dipende essenzialmente solo dal numero di siti prossimi alla frontiera.

[1] In quest'ultimo caso chiaramente si parametrizza un operatore densità.

3.1.1 Località e gap

Di seguito si assume in generale che l'hamiltoniana abbia intensità finita, ovvero che la norma operatoriale di un qualsiasi termine sia controllata da una costante uniforme finita. Si considerano quindi specificamente sistemi con hamiltoniana **locale**:

Definizione 3.1. *Un sistema ha hamiltoniana locale (o interazione locale) se $\hat{H}$ è definita da una somma di operatori locali (secondo la definizione rigorosa).*

Si tratta chiaramente di un requisito poco restrittivo in quanto di norma giustificato sul piano fisico; le considerazioni di seguito possono comunque essere generalizzate con correzioni non sostanziali al caso di interazione quasi-locale con decadimento esponenziale, ovvero t.c. in norma operatoriale [42]:

$$\|\Phi(X)\|_\infty \leq e^{-\mathrm{diam}X/\xi} \qquad \xi > 0 \tag{3.1}$$

Una seconda proprietà riguarda lo spettro energetico in prossimità dello stato fondamentale: si considerano sistemi che presentano - a rigore, nel solo limite infinito[2] - un **gap** tra l'energia di ground e il primo livello eccitato, ovvero sono **gappati (gapped)** (in alternativa si dicono **gapless**) [41, 35]. In dimensione finita, sfruttando il fatto che è sempre possibile traslare lo spettro di un operatore aggiungendovi un termine proporzionale all'identità, si può porre in tutta generalità $E_0 = 0$ e definire il gap ΔE secondo [24, 42]:

$$\Delta E \equiv \sup\{\gamma > 0 \,:\, \mathrm{spec}\hat{H} \cap]0, \gamma[= \emptyset\} \tag{3.2}$$

Un primo modo di generalizzare a sistemi infiniti la definizione è analogo a quanto si è fatto per lo stato fondamentale; tuttavia in questo caso si ottiene più restrittivamente un ground state **localmente unico**, ovvero non trasformabile localmente in un altro stato fondamentale ortogonale [34].
Senza perdita di generalità, si consideri una trasformazione che mappa il ground $|\psi_0\rangle$ (normalizzato) in uno stato ad esso ortogonale:

$$|\psi_0\rangle \mapsto |\phi\rangle \equiv \frac{\hat{V}|\psi_0\rangle}{\|\hat{V}|\psi_0\rangle\|} = \frac{\hat{V}|\psi_0\rangle}{\sqrt{\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle}} \qquad \langle\psi_0|\phi\rangle \propto \langle\psi_0|\hat{V}|\psi_0\rangle = 0 \tag{3.3}$$

Allora per definizione di gap, se $0 < \gamma \leq \Delta E$:

$$\langle\phi|\hat{H}|\phi\rangle = \frac{\langle\psi_0|\hat{V}^\dagger\hat{H}\hat{V}|\psi_0\rangle}{\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle} \geq \gamma + \langle\psi_0|\hat{H}|\psi_0\rangle \tag{3.4}$$

[2] L'hamiltoniana di un sistema finito ha spettro discreto, quindi ha sempre un gap.

Sfruttando sempre la commutazione del proiettore con l'hamiltoniana:

$$\langle\psi_0|\hat{V}^\dagger\hat{H}\hat{V}|\psi_0\rangle \geq \gamma\,\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle + \langle\psi_0|\hat{V}^\dagger\hat{V}\,|\psi_0\rangle\langle\psi_0|\,\hat{H}|\psi_0\rangle = \\ = \gamma\,\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle + \langle\psi_0|\hat{V}^\dagger\hat{V}\hat{H}|\psi_0\rangle \quad (3.5)$$

ovvero:

$$\langle\psi_0|\hat{V}^\dagger[\hat{H},\hat{V}]\,|\psi_0\rangle \geq \gamma\,\langle\psi_0|\hat{V}^\dagger\hat{V}|\psi_0\rangle \quad (3.6)$$

La condizione ottenuta è anche sufficiente in quanto se uno stato $|\psi\rangle$ non la soddisfa si può, come per la definizione generale di stato fondamentale, prendere $\hat{V} = |\psi_0\rangle\langle\psi|$. Generalizzando al caso infinito [34]:

Definizione 3.2. *Uno stato fondamentale ω_0 è localmente unico e gappato se esiste una costante reale $\gamma > 0$ t.c.*

$$\omega(\hat{O}^\dagger[\hat{H},\hat{O}]\,) \geq \gamma\omega(\hat{V}^\dagger\hat{V}) \qquad \forall\hat{O}\in\mathcal{B}_{loc}\,:\,\omega(\hat{O}) = 0 \quad (3.7)$$

l'estremo superiore dei γ per cui vale questa proprietà è il gap ΔE.

Si noti che uno stato fondamentale unico è anche *localmente* unico, ma non vale in generale il viceversa: se si ha rottura spontanea di simmetria due stati fondamentali ortogonali, per quanto di medesima energia, possono non essere collegati tramite trasformazioni locali[3] [34].
Per avere una definizione generale di sistema gappato, anziché operare direttamente nel limite infinito si può considerare la sua costruzione tramite successioni di sistemi su reticoli finiti $\Lambda_n \to \Lambda$ [24, 34]. Per ciascuno di questi ha senso diagonalizzare l'hamiltoniana $\hat{H}_{\Lambda_n}$ e si può fissare uno stato fondamentale $|\psi_n\rangle$; è così costruita una successione per cui si può definire sensatamente un limite, sebbene non sia possibile fare lo stesso per gli spazi di Hilbert:

$$\omega(\hat{O}) \equiv \lim_{n\to\infty}\,\langle\psi_n|\hat{O}|\psi_n\rangle \qquad \forall\hat{A}\in\mathcal{B}_{loc} \quad (3.8)$$

In generale non è assicurato che il limite esista, ma si può ovviare a questo prendendo una sottosuccessione convergente $\omega_{n(k)} \equiv \omega_k$ [4]: in questo modo, sebbene per

[3] É interessante notare che una fase può essere definita come una classe di equivalenza di stati fondamentali dati dal limite infinito di sistemi le cui hamiltoniane al finito sono collegate da una trasformazione continua che non chiude il gap [42]. Questo ha anche un'interpretazione naturale nel contesto dei TN, che tuttavia non sarà trattata in questo elaborato. Si veda per approfondire ad esempio [41].

[4] Si applica il Teorema di Banach-Alaoglu, per cui una successione di funzionali lineari su uno spazio di Banach separabile ammette una sottosuccessione convergente puntualmente. Si noti che la convergenza puntuale non è che la convergenza debole per gli elementi del doppio duale identificabili naturalmente con quelli dello spazio originario (in dimensione infinita non si può assumere corrispondenza suriettiva).

un dato sistema il limite non sia necessariamente univoco, si è definito uno stato fondamentale su Λ, che soddisfa equivalentemente la definizione 2.22. Si può ora definire il gap [34, 42, 24]:

Definizione 3.3. *Se per ogni sistema finito il gap spettrale è $\Delta E^{(k)}$, allora il gap del sistema infinito è:*

$$\Delta E \equiv \liminf_{k\to\infty} \Delta E^{(k)} = \lim_{k\to\infty} \inf_{l\geq k} \Delta E^{(l)} \tag{3.9}$$

In pratica, il sistema nel limite infinito è gappato se si può determinare un elemento finito della successione Λ_k dopo il quale il gap non si "richiude" più. É interessante notare che la presenza di un gap entro una certa dimensione finita non implica che questo permanga anche nel limite. La definizione della classe rilevante di sistemi è ora completa e si può dare il primo risultato fondamentale [43, 12]:

Teorema 3.1. *In un sistema unidimensionale locale gappato i correlatori connessi decadono esponenzialmente*[5]*, ovvero esiste una lunghezza di correlazione finita:*

$$\exists \xi \in [0, \infty[\quad : \quad C(r) = \mathcal{O}(e^{-r/\xi}) \, \forall \hat{O}, \hat{O}' \tag{3.10}$$

Una dimostrazione rigorosa dell'implicazione esula dagli scopi di questa trattazione; tuttavia è possibile giustificarla in modo intuitivo tramite un'analogia [33]. Nelle teorie quantistiche di campo (QFT) che presentano un *mass gap* sopra lo stato di vuoto, i correlatori decadono esponenzialmente, in quanto essenzialmente i mediatori sono massivi. Ora, sebbene in principio si operi in contesto non relativistico, per spin systems con interazioni locali è stato dimostrato che esiste una velocità limite di propagazione dell'informazione[6], detta **velocità di Lieb-Robinson**. Dunque è legittimo attendersi che si abbia un vincolo simile al caso relativistico sull'andamento dei correlatori quando è presente un gap spettrale (che può essere intepretato come massa delle eccitazioni con proprietà quasiparticellari). Si può stimare la scala di lunghezza a cui si possono stabilire correlazioni in un tempo di evoluzione finito tramite un semplice argomento dimensionale:

$$\tau \sim \frac{1}{\Delta E} \implies \xi \sim \tau v \sim \frac{v}{\Delta E} \tag{3.11}$$

L'andamento dei correlatori suggerisce che complessivamente la correlazione tra gli stati dei singoli siti nei ground state in questi sistemi sia locale e di conseguenza i parametri di principio necessari per descrivere l'entanglement tra siti distanti siano superflui.

[5] Per operatori con norma limitata, un assunto del tutto generale.

[6] Ovvero, formalmente, di espansione del supporto di un operatore locale nella rappresentazione di Heisenberg.

3.1.2 Formalizzazione dell'area law

Considerando l'entanglement tra blocchi di siti, questa natura locale implica che a determinarlo siano essenzialmente i soli elementi collocati lungo la partizione, e non quelli nel *bulk*. Una tale struttura di entanglement area law [2]:

Definizione 3.4. *Uno stato a molti corpi soddisfa un'area law se il contributo dominante*[7] *all'entropia di entanglement per qualsiasi partizione è al più proporzionale alla dimensione*[8] *della frontiera tra i due sottosistemi che determina.*

Ovvero per l'entropia di una regione vale:

$$S(\hat{\rho}_B) = \mathcal{O}(|\partial B|) \tag{3.12}$$

In particolare, se il blocco ha una scala di lunghezza L[9] e il sistema è d-dimensionale si ha [12]:

$$S = \mathcal{O}(L^{d-1}) \tag{3.13}$$

In 1D un'area è quindi un punto: di conseguenza l'area law implica che l'entropia sia *costante* nella dimensione del blocco. Si tratta di una proprietà assolutamente non triviale: per uno stato generico l'entropia è estensiva, ovvero scala proporzionalmente al *volume* del blocco; gli stati soddisfacenti l'area law occupano quindi un "angolo" esponenzialmente piccolo[10] rispetto alla totalità dello spazio di Hilbert - il quale risulta quindi essere in contesti fisicamente ragionevoli una "conveniente illusione" [44, 2, 12].

Figura 3.1: Blocchi e frontiere in 1D e 2D. Immagine da [45].

Ciò ha profonde implicazioni fisiche: essenzialmente, indica che lo stato del *bulk* di un sistema può essere descritto osservandone solamente la *boundary*, diminuendo

[7] E quindi il solo rilevante nel limite di dimensione infinita.

[8] Espressa come numero di siti.

[9] Che ne controlla l'estensione in tutte le dimensioni

[10] Piuttosto che in relazione al numero di parametri, si può definire rigorosamente una misura di probabilità uniforme sugli stati. Per la precisione la misura, detta *m. di Haar*, è definita sul gruppo unitario [26]: si fissa allora uno stato normalizzato e si definisce la misura di un insieme di stati come la probabilità dell'insieme di unitarie che vi mappa lo stato di riferimento.

quindi di una dimensione i gradi di libertà *senza perdita di informazione*. Questa proprietà è in generale detta **principio olografico** ed ha grande importanza nello studio dei buchi neri e più in generale nelle teorie di quantizzazione della gravità, ove si congettura che una teoria di gravità quantistica in un volume possa essere descritta da una teoria di campo senza gravità sulla frontiera.
In realtà l'area law vale anche più in generale per entropie di Renyi, anche quelle più sensibili alla coda dei coefficienti di Schmidt[43]:

Teorema 3.2. *Per sistemi 1D con hamiltoniana locale il ground state soddisfa la seguente area law per le entropie di Renyi:*

$$S^{\alpha}(\hat{\rho}_L) \approx \frac{c}{6}\left(1+\frac{1}{\alpha}\right)\log \xi \tag{3.14}$$

ove c è la carica centrale[11] *e ξ la lunghezza di correlazione. Nel caso di sistemi critici vi è una lieve violazione in quanto è presente una correzione logaritmica, ovvero ξ è da sostituirsi con L*[12].

3.1.3 Parametrizzazione degli stati rilevanti

Gli stati fondamentali riproducono quindi in qualche modo la località dell'hamiltoniana e di conseguenza è ragionevole attendersi ammettano una parametrizzazione efficiente.
Infatti dato un sistema di n siti con interazione k-locale[13] ogni termine dell'hamiltoniana:

- Agisce su una combinazione di k siti[14];
- É operatore dell'algebra tensore d^{2k}-dimensionale.

Dunque in totale $\hat{H}$ è definita da $\binom{n}{k}d^{2k}$ parametri, ovvero una quantità polinomiale in n [46, 45]; in altri termini, è rappresentata su una base dell'algebra operatoriale complessiva da una matrice sparsa [41]. La località dell'interazione è di grande utilità anche in quanto comporta che il ground state contenga informazioni sullo spettro eccitato, quantomeno a bassa energia: si può infatti dimostrare che vincola i primi stati eccitati a essere dati da mere perturbazioni locali (di natura quasiparticellare) dello stato fondamentale, che possono essere caratterizzate quindi tramite

[11] Un parametro calcolabile tramite l'opportuna Teoria di Campo Conforme (CFT), essenzialmente quantifica i modi indipendenti di eccitazione.

[12] Rispetto ad uno scaling lineare in L l'entropia resta comunque esponenzialmente minore.

[13] Ovvero tale per cui $\Phi(X)$ è k-locale per ogni X.

[14] Si può assumere in generalità in quanto se dei siti non sono esplicitamente presenti in un termine l'azione su di essi è triviale.

la sua struttura di correlazione [41]. É importante notare tuttavia che la località dell'interazione non è *sufficiente* perché il sistema soddisfi un'area law; difatti nella derivazione è stato necessario introdurre ipotesi ulteriori [12].
Poiché la minimizzazione globale dell'energia di uno stato fondamentale, come osservato nella definizione, è una conseguenza di quella locale, quantomeno in assenza di frustrazione, la sua struttura completa può essere caratterizzata studiandone le proprietà locali [43]. Si consideri il caso di uno spin system con hamiltoniana locale gappato con un unico stato fondamentale $|\psi_0\rangle$ e un gap dipendente in generale dal numero di siti n. Si assuma specificamente he il gap non decada più rapidamente di un polinomio in n:

$$\frac{1}{\Delta E^{(n)}} = \mathcal{O}(\text{poly}(n)) \tag{3.15}$$

una proprietà verificata sempre nel caso non critico e in tutti i casi noti per quello non critico con invarianza traslazionale. Si assuma esista uno stato $|\psi_{app}\rangle$ che approssimi le proprietà *locali* del ground state entro una soglia globale δ, ove la distanza è misurata sugli operatori ridotti per coppie di siti vicini:

$$\|\hat{\rho}_0 - \hat{\rho}_{app}\| \leq \delta \tag{3.16}$$

Allora si può controllare l'errore globale secondo:

$$\| \, |\psi_{app}\rangle - |\psi_0\rangle \, \|^2 \leq \frac{n\delta}{\Delta E^{(n)}} \tag{3.17}$$

Questo indica che è sufficiente regolare l'approssimazione locale per minimizzare globalmente la distanza (o equivalentemente massimizzare la parte reale dell'overlap), ed in particolare per avere un errore globale costante aumentando il numero di spin è sufficiente che l'errore locale scali come il reciproco di un polinomio in n: ciò suggerisce la possibilità di sviluppare algoritmi iterativi efficienti [43].
Si può ora dare anche un argomento euristico che suggerisce naturalmente l'introduzione di un ansatz di TN per descrivere questi ground state [43, 12]. Considerato un tale sistema con lunghezza di correlazione finita ξ se ne definisca una tripartizione in regioni connesse A, B, C. Per lo stato ridotto $\hat{\rho}_{AC}$ nel limite di distanza tra i blocchi $L_{AC} \gg \xi$ varrà:

$$\hat{\rho}_{AC} \approx \hat{\rho}_A \otimes \hat{\rho}_C \tag{3.18}$$

al netto di correzioni esponenziali nei correlatori di operatori con supporto sui blocchi[15] [47]:

$$\text{Tr}[\hat{O}_A \hat{O}'_C \hat{\rho}_{AC}] = \text{Tr}[\hat{O}_A \hat{O}'_C \hat{\rho}_A \otimes \hat{\rho}_C] + \mathcal{O}(e^{-L_{AC}/\xi}) \tag{3.19}$$

[15] In realtà la 3.19 non implica la 3.18, e si è dimostrato che esistono stati ortogonali fra loro ma con andamenti esponenzialmente vicini dei correlatori. Pertanto l'argomento sviluppato non può essere reso rigoroso.

Segue che lo stato del sistema complessivo ABC possa considerarsi una purificazione di questo stato con la forma:

$$|\psi_{AB_L}\rangle \otimes |\psi_{B_RC}\rangle \tag{3.20}$$

ove B_L e B_R sono due blocchi di siti sinistro e destro che partiscono B. Nel capitolo precedente si è osservato che tutte le possibili purificazioni di uno stato misto sono equivalenti a meno di isometrie sul sistema ausiliario, o più restrittivamente unitarie; dunque è possibile "disintrecciare" A e C, o più precisamente le porzioni destra e sinistra del sistema, tramite un'unitaria su B:

$$\hat{\mathbb{I}}_A \otimes \hat{U}_B \otimes \hat{\mathbb{I}}_C \, |\psi_{ABC}\rangle = |\psi_{AB_L}\rangle \otimes |\psi_{B_RC}\rangle \tag{3.21}$$

Da ciò segue che lo stato del sistema possa essere espresso come:

$$|\psi_{ABC}\rangle \approx B_j^{\alpha,\gamma} \left|\psi_\alpha^A\right\rangle \left|\psi_j^B\right\rangle \left|\psi_\gamma^C\right\rangle \tag{3.22}$$

ove la dimensione di ciascuno degli indici sommati è al più pari alla dimensione dello spazio locale di B. Questa procedura può essere ora iterata aumentando la "risoluzione" dei blocchi e ottenendo così una decomposizione del tensore (in componenti) che descrive lo stato complessivo come contrazione di tensori di ordine 3 associati ai singoli siti.

3.2 MPS

Questa decomposizione definisce una classe di stati quantistici detti **stati prodotto di matrici (Matrix Product States)**, di seguito **MPS**. Il nome è dovuto al fatto che permettono di rappresentare il tensore come prodotto di una serie di matrici (sebbene si tratti di tensori di ordine 3, ma ci si concentra sugli indici non fisici contratti [7]) associate ai singoli spazi locali [41]. Esistono diversi modi equivalenti di definirli, ciascuno dei quali mette in evidenza differenti caratteristiche fondamentali.

3.2.1 Matrix Product Decomposition

Sfruttando la libertà di ridimensionare i tensori, si possono iterare le decomposizioni matriciali introdotte in precedenza per ridurre qualsiasi tensore (non triviale) ad un prodotto di tensori di ordine 3: si parla di **decomposizione in prodotto di matrici (Matrix Product Decomposition)**[16]. Di seguito si fa uso della SVD ma si può ottenere il medesimo risultato (ovviamente in generale con tensori finali

[16] Per i matematici anche Tensor Train Decomposition (TTD) [5, 12]

differenti) tramite le altre decomposizioni.

Dato T un tensore di ordine k, si procede come di seguito [2]:

1. Lo si ridimensiona in una matrice tramite fusione, si applica una prima SVD e poi si splittano gli indici per recuperare la forma iniziale:
$$T_{j_1,j_2,\dots,j_k} \rightarrow T_{j_1,(j_2,\dots,j_k)} = U_{j_1,c_1} S_{c_1,c_1} V^{\dagger}_{c_1,(j_2,\dots,j_k)} = A^{c_1}_{j_1} T_{c_1,j_2,\dots,j_k} \tag{3.23}$$
rinominando U come A e contraendo S e $V^{\dagger}$ per ridefinire T;

2. Si ripete la medesima procedura per quest'ultimo: fusione degli indici, SVD e quindi contrazione e splitting delle matrici risultanti:
$$T_{c_1,j_2,\dots,j_k} \rightarrow T_{(c_1,j_2),(j_3,\dots,j_k)} = U_{(c_1,j_2),c_2} S_{c_2,c_2} V^{\dagger}_{c_2,(j_3,\dots,j_k)} = A^{c_1,c_2}_{j_2} T_{c_2,j_3,\dots,j_k} \tag{3.24}$$

3. Si itera fino ad ottenere la decomposizione completa del tensore di partenza:
$$T_{j_1,j_2,\dots,j_k} = A^{c_1}_{j_1} A^{c_1,c_2}_{j_2} \cdots A^{c_{k-2}c_{k-1}}_{j_{k-1}} A^{c_{k-1}}_{j_k} \tag{3.25}$$

Quella ottenuta è propriamente una **Left Canonical Decomposition (LCD)**, in quanto i tensori risultanti (eccetto quello più a destra) sono normalizzati *a sinistra*:
$$A^{c_{p-1}c_p}_{j_p} \left[A^{c_{p-1}c'_p}_{j_p} \right]^{\dagger} = \delta^{c_p,c'_p} \tag{3.26}$$

Chiaramente non è univoca: si può egualmente partire fondendo i primi $k-1$ indici e procedendo da destra verso sinistra, ricavando la *Right* Canonical Decomposition (RCD), in cui tutti i tensori (eccetto quello più a sinistra) sono normalizzati *a destra*:
$$T_{j_1,j_2,\dots,j_k} = B^{c_1}_{j_1} B^{c_1,c_2}_{j_2} \cdots B^{c_{k-2}c_{k-1}}_{j_{k-1}} B^{c_{k-1}}_{j_k} \qquad B^{c_{p-1}c_p}_{j_p} \left[B^{c'_{p-1}c_p}_{j_p} \right]^{\dagger} = \delta^{c_{p-1},c'_{p-1}} \tag{3.27}$$

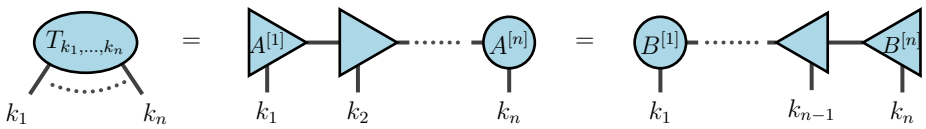
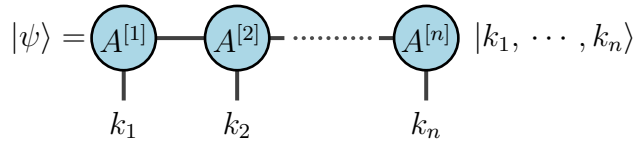


Figura 3.2: Left e Right Canonical Decompositions

L'applicazione di questa procedura allo stato quantistico di un sistema composto è esattamente equivalente a considerare iterativamente delle bipartizioni e applicare la decomposizione di Schmidt, per poi assorbire la matrice dei valori singolari a destra o a sinistra, a seconda della decomposizione scelta[17] [7]. Si ha così una prima definizione di MPS [2]:

[17]Si intuisce quindi che LCD e RCD sono equivalenti; si formalizzerà e generalizzerà questa proprietà tramite la libertà di gauge dei MPS.

Definizione 3.5. *Uno stato di un sistema a molti corpi è un MPS se è esprimibile come un TN decomposto in prodotto di matrici, ovvero ha la forma*[18]*:*



Gli indici di bond c_p hanno quindi una chiara interpretazione: descrivono l'entanglement attraverso la corrispondente bipartizione del sistema [7]. Per quanto visto in precedenza questo è un primo elemento fondamentale che contribuisce all'efficienza degli MPS [2, 7]:

- É sufficiente che la dimensione di bond χ_p sia pari al rango di Schmidt associato per avere una rappresentazione ottimale ma esatta dello stato[19];
- Riducendo χ_p di modo da mantenere gli indici associati ai maggiori $\chi < \chi_p$ coefficienti di Schmidt si ottiene la miglior compressione possibile a parità di numero di parametri - e la distanza dallo stato esatto è facilmente quantificabile con l'errore di troncamento[20]. Un criterio alternativo è la definizione di un *cutoff* per i coefficienti [4], che permette una stima più immediata dell'errore globale al prezzo di non avere un controllo uniforme sulle dimensioni - in casi estremi, potrebbe non essere addirittura "tagliato" alcun coefficiente[21].

[18] In generale da normalizzare [43].

[19] Viceversa il rango di Schmidt per la corrispondente partizione è minore o uguale a χ_p [4]

[20] Si noti che tuttavia non si ha una compressione ottimale effettuando il "taglio" dei coefficienti su più gambe virtuale contemporaneamente, in quanto si tratta di un'operazione non unitaria [6].

[21] Sia ξ il cutoff e si abbia, ordinando in modo crescente i valori singolari:

$$\lambda_1 \leq \ldots \leq \lambda_n \leq \xi \leq \lambda_{n+1} \leq \ldots \leq \lambda_\chi \qquad |\psi\rangle = \textstyle\sum_{i=1}^{\chi} \lambda_i |i_A\rangle |i_B\rangle \tag{3.28}$$

Si rimuovano quindi i primi n vettori di Schmidt, ottenendo l'approssimazione (non normalizzata) $|\psi'\rangle$ di $|\psi\rangle$. Per stimare l'errore si sfrutta la diseguaglianza di Cauchy-Schwarz considerando l'overlap con un generico stato normalizzato $|\phi\rangle$ (ciò permette una stima più generale per qualsiasi osservabile):

$$\big|\langle\phi|\left(|\psi\rangle - |\psi'\rangle\right)\big| \leq \textstyle\sum_{i=1}^{n} \lambda_i \big|\langle\phi|i_{AB}\rangle\big| \leq \sum_{i=1}^{n} \lambda_i \leq n\xi \tag{3.29}$$

Dunque in generale l'errore in modulo è lineare in ξ, ovvero $\mathcal{O}(\xi)$. Si può stimare anche il prodotto scalare tra $|\psi\rangle$ e l'approssimazione normalizzata $|\psi''\rangle \equiv \frac{1}{\||\psi'\rangle\|} |\psi'\rangle$:

$$\langle\psi|\psi''\rangle = \frac{1}{\sqrt{1 - \sum_{i=1}^{n} \lambda_i^2}} \Big(\textstyle\sum_{i=1}^{\chi} \lambda_i \langle i_{AB}|\Big)\Big(\sum_{i=1}^{n} \lambda_i |i_{AB}\rangle\Big) = \sqrt{1 - \textstyle\sum_{i=1}^{n} \lambda_i^2} \tag{3.30}$$

Da cui l'andamento dell'overlap in ξ:

$$\big|\langle\psi|\psi''\rangle\big|^2 = 1 - \textstyle\sum_{i=1}^{n} \lambda_i^2 \geq 1 - n\xi^2 = 1 + \mathcal{O}(\xi^2) \tag{3.31}$$

Tuttavia, non essendo necessarie ipotesi particolari per l'applicazione della SVD, in linea di principio qualsiasi stato composto ammette rappresentazione come MPS [7, 43]. Poiché in linea generale uno stato entangled presenta correlazioni significative tra siti arbitrariamente distanti, ciò significa che lo scaling esponenziale nel numero di componenti è meramente trasferito alle dimensioni degli indici di bond, o più limitatamente al rispettivo rango di Schmidt [36, 12]. Infatti considerando la LCD di un tensore generico di ordine k vista in precedenza (ma la stima è equivalente per le altre costruzioni) e assumendo una dimensione uniforme d per i j_i [2]:

- La dimensione dei tensori estremi è $d \times d = d^2$ [22];
- Procedendo verso destra:
  - per $p \leq \lfloor k/2 \rfloor$ la dimensione di $A_{j_p}^{c_{p-1},c_p}$ è $d \times d^{p-1} \times d^p = d^{2p}$;
  - per $p > \lfloor k/2 \rfloor$ la dimensione di $A_{j_p}^{c_{p-1},c_p}$ è $d \times d^{k-(p-1)} \times d^{k-p} = d^{2(k+1-p)}$.

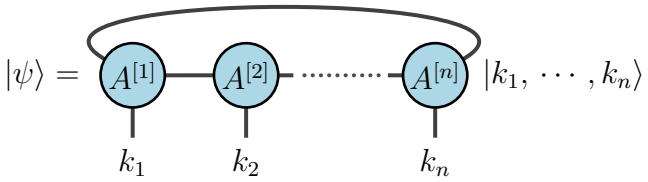
Di conseguenza il rango di Schmidt per possibili partizioni di un MPS di un sistema composto con n sottosistemi è $\mathcal{O}(d^{\lfloor n/2 \rfloor})$, come si può intuire euristicamente considerando che a priori tutti i siti dei due blocchi possono essere coinvolti nell'entanglement - e di conseguenza l'entropia di entanglement è $\mathcal{O}(n)$, ovvero estensiva. Per avere una crescita controllata dei parametri è necessario che l'entanglement dipenda più debolmente da n - o per nulla, ovvero che solo porzioni limitate dei blocchi siano correlate, in una proporzione fortemente decrescente all'aumentare della dimensione di questi: in altri termini, deve valere un'area law.
Una dimensione di bond triviale (ovvero la rispettiva matrice che si riduce a semplice scalare) corrisponde a uno stato prodotto sulla partizione corrispondente, dunque a entanglement nullo. Uno stato prodotto del sistema complessivo, ovvero un ansatz di campo medio, è così semplicemente rappresentabile come MPS - rimuovendo direttamente i link una volta normalizzato, secondo quanto visto per il prodotto tensore [2, 9]. Ciò permette un'intuizione grafica sul vantaggio di questa classe più generale: è costruita proprio aggiungendo allo stato fattorizzato dei link per catturare i gradi di libertà ulteriori dati dall'entanglement - in un modo ottimizzato per la sua struttura locale.

3.2.2 Tipologie di MPS

La costruzione che si è data definisce propriamente un MPS *finito aperto* e di principio *non uniforme*. Generalizzandola si possono introdurre MPS [5]:

- **Periodici**, in cui anche i tensori estremi sono matrici:

[22] Si può assorbire nei casi successivi in quanto $d^0 = 1$.



Algebricamente sono espressi nella forma compatta [7]:

$$|\psi\rangle = \mathrm{Tr}[A^{[1]}_{k_1} A^{[2]}_{k_2} \cdots A^{[n]}_{k_n}] \, |k_1, \ldots, k_n\rangle \tag{3.32}$$

Assumendo in generale matrici differenti per i vari siti;

- **Invarianti per traslazione (Translation-Invariant, TI)**, ove tutte le matrici sono identiche;
- **Infiniti**, che descrivono sistemi nel limite termodinamico[23].

3.2.3 Valence Bond e Projected Entangled Pair States

In alternativa a quella vista, è possibile una definizione degli MPS tramite una costruzione più formale e meno intuitiva ma che cattura più profondamente la natura di questi stati e soprattutto permette una generalizzazione a dimensioni maggiori [7, 43]. Dato un generico sistema a molti corpi con n siti, si associ a ciascuno di essi una coppia di spin virtuali di dimensione χ e si assuma che ogni coppia di spin adiacenti appartenenti a siti differenti si trovi in uno stato massimamente entangled. Nel caso di condizioni al contorno aperte, i singoli spin virtuali estremi sono posti in una sovrapposizione uniforme degli stati di base. Si applichi quindi a ogni sito una mappa:

$$\hat{\mathcal{A}}_i \equiv A^{[i]}_{j_i;\alpha_i,\beta_i} \, |j_i\rangle\langle\alpha_i, \beta_i| \tag{3.33}$$

ove l'indice latino è riferito al sito fisico e quelli greci ai siti virtuali. Questa mappa è un tensore di rango 3 (o 2 per i siti estremi in assenza di PBC), e gli scalari $A_{i,\alpha,\beta}$ ne definiscono la rappresentazione in componenti. Gli indici virtuali sono completamente contratti, quindi l'applicazione delle mappe $\mathcal{A}_j$ produce un nuovo stato del sistema, che è proprio un MPS con coefficienti dati dalla contrazione dei tensori corrispondenti. Questo si può verificare considerando l'azione delle mappe per i primi due siti fisici sulla coppia virtuale condivisa [7]:

$$\begin{aligned}(\hat{\mathcal{A}}_1 \otimes \hat{\mathcal{A}}_2) \, |\gamma, \gamma\rangle &= A^{[1]}_{j_1;\alpha_1,\beta_1} A^{[2]}_{j_2;\alpha_2,\beta_2} \, |j_1, j_2\rangle\langle\alpha_1, \beta_1, \alpha_2, \beta_2| \left(|\alpha\rangle\langle\alpha| \otimes |\gamma, \gamma\rangle \otimes |\beta\rangle\langle\beta| \right) = \\ &= A^{[1]}_{j_1;\alpha_1,\beta_1} A^{[2]}_{j_2;\alpha_2,\beta_2} \delta_{\alpha_1,\alpha} \delta_{\beta_1,\gamma} \delta_{\alpha_2,\gamma} \delta_{\beta_2,\beta} \, |j_1, j_2\rangle\langle\alpha, \beta| = A^{[1]}_{j_1;\alpha,\gamma} A^{[2]}_{j_2;\gamma,\beta} \, |j_1, j_2\rangle\langle\alpha, \beta|\end{aligned} \tag{3.34}$$

[23] Pertanto sovente sono TI

Tramite la rappresentazione a legami di valenza si può comprendere intuitivamente la struttura di entanglement degli MPS: le coppie virtuali sono massimamente "legate", quindi la correlazione tra i siti fisici dipende da quanto siano correlati ai siti virtuali associati, che è determinato dall'azione degli operatori $\mathcal{A}_i$ - ovvero equivalentemente dalla loro specificazione in componenti nei tensori A_{j_i,α_i,β_i}. Questi creano quindi entanglement tra siti non direttamente comunicanti tramite operazioni su quelli intermedi, in analogia ai cosiddetti protocolli di *entanglement swapping* in computazione quantistica [41]. L'applicazione di unitarie sugli indici di bond che permette di modificare la MPD dello stato corrisponde semplicemente ad un cambiamento di base per le coppie virtuali [7].
Questa costruzione è talvolta detta a legame di valenza (**Valence Bond, VB**), considerando le coppie massimamente entangled come legami; la configurazione iniziale degli spin virtuali corrisponde a uno stato di grande importanza in fisica della materia condensata, detto **Valence Bond Solid (VBS)** [37]; il suo vantaggio principale rispetto alla costruzione tramite MPD, come detto, è che si può generalizzare naturalmente a dimensioni maggiori (in particolare 2D e 3D) e geometrie arbitrarie, collocando una coppia virtuale in corrispondenza di ogni spigolo tra siti primi vicini: si ottengono così i **Projected Entangled Pair States (PEPS)**, una classe di TN molto utilizzata per lo studio di sistemi a bassa dimensionalità [7].
La costruzione degli MPS tramite PEPS risulta naturale applicando alle spin chains un risultato fondamentale di teoria dell'informazione quantistica: date N copie di un sistema bipartito con un certo quantitativo di entanglement[24] è possibile convertirle tramite LOCC in $N \cdot S$ stati di Bell, ovvero coppie massimamente entangled di sistemi bidimensionali, preservando[25] il quantitativo di entanglement [49]. Per transitività dell'equivalenza, si possono rimpiazzare le coppie di Bell con i *Pairs* definiti in precedenza e così le proiezioni per ogni sito fisico vanno a corrispondere alle operazioni locali [41].

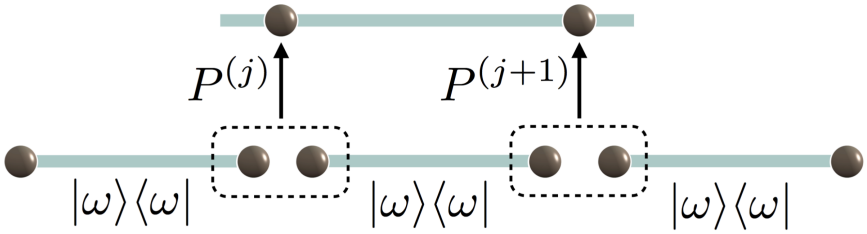


Figura 3.3: Costruzione di un PEPS. Immagine da [35].

[24] Quantificato nel caso puro dall'entropia di Von Neumann nel caso puro e in quello misto da una sua generalizzazione detta *Distillable Entanglement*, definita proprio tramite questa procedura.

[25] Esattamente nel caso puro e asintoticamente, ovvero per $N \to \infty$, in quello misto [48].

3.2.4 Proprietà di entanglement e approssimazione degli stati rilevanti

É immediato verificare che MPS con dimensione di bond costante (o comunque controllata da una costante) soddisfano l'area law unidimensionale. Considerando infatti un blocco contiguo di siti, l'entropia di entanglement per la partizione che definisce è $S(\hat{\rho}_B) = \mathcal{O}(\log \chi)$, in conseguenza del teorema 2.8[26]:

$$S(\hat{\rho}_B) \leq \log \chi^2 = 2 \log \chi \tag{3.35}$$

Dipendendo solamente dalla dimensione locale di bond tra i due blocchi, risulta così costante nel numero di siti [2, 5]. Più in generale si può facilmente verificare che PEPS a dimensione di bond costante soddisfano l'area law [5]. Considerando infatti un blocco di siti $L \times L$, l'entropia di entanglement del corrispondente operatore ridotto è controllata dal logaritmo della dimensione complessiva delle gambe virtuali "tagliate" dalla partizione:

$$S(\hat{\rho}_L) \leq \log d^{4L} = 4L \log d \tag{3.36}$$

Si noti che per proprietà del logaritmo si può interpretare equivalentemente considerando il limite complessivo come la somma sulle gambe del contributo massimo di ciascuna, ovvero $\log d$, e che concorrono allo scaling sia la dimensione di bond che la *geometria* dei legami: gli ansatz in forma di TN non sarebbero di principio utili se non riproducessero in modo ottimale la struttura fisica delle correlazioni dei sistemi [5].

Il fatto che i MPS a dimensione di bond *uniformemente limitata* (definita di solito, abusando della terminologia, come "finita") soddisfino l'area law per l'entropia di Von Neumann e di conseguenza le entropie di Renyi con $\alpha > 1$ non implica che, viceversa, stati che soddisfano l'area law per S siano approssimabili in modo efficiente con MPS; tuttavia è possibile dimostrare formalmente l'esistenza di tale approssimazione qualora sia soddisfatta l'area law, anche con correzioni logaritmiche, per entropie di Renyi con $\alpha < 1$, più sensibili alla coda dello spettro di entanglement. Chiaramente vale al contempo la contronominale: uno scaling lineare di S o in generale più che logaritmico (ossia a potenza con un qualsiasi esponente positivo) di S_α con $\alpha > 1$ è condizione sufficiente perché l'approssimazione non sia possibile [50]. I risultati riportati di seguito sono originariamente derivati in [51] ma si può fare riferimento anche alla discussione in [43].
Innanzitutto vale il seguente teorema, che segue direttamente dalla decomposizione in prodotto di matrici:

[26] La partizione è al più su due gambe virtuali.

Teorema 3.3. *Dato un generico stato $|\psi\rangle$ di un sistema a n siti, esiste un MPS $|\psi_{MPS}\rangle$ con dimensione di bond χ t.c.*

$$\| \, |\psi\rangle - |\psi_{MPS}\rangle \, \|^2 \leq 2 \sum_{i=1}^{n} \varepsilon_i(\chi) \tag{3.37}$$

ove $\varepsilon_i(\chi) = \sum_{k=\chi+1}^{\operatorname{rank} \hat{\rho}_i} \lambda_k^{[i]2}$ è l'errore di troncamento sulla partizione tra i primi i siti, con operatore ridotto $\hat{\rho}_i$, e gli ultimi $n-i$, con i coefficienti di Schmidt ordinati in senso decrescente.

Ciò indica che per sistemi i cui coefficienti di Schmidt sulle partizioni sono fortemente soppressi per indici crescenti, ovvero con un errore di troncamento che decade rapidamente in χ, esiste una rappresentazione approssimativa con MPS che cattura sia le proprietà locali, ovvero in particolare l'energia, che quelle globali, ossia le correlazioni su larga scala.
Sfruttando la loro definizione, l'errore di approssimazione può essere direttamente collegato allo scaling delle entropie di Renyi - e quindi all'area law:

Teorema 3.4. *Per $0 < \alpha < 1$, l'errore di troncamento ai primi χ autovalori di un operatore densità è controllato secondo:*

$$\log \varepsilon(\chi) \leq \frac{1-\alpha}{\alpha} \left(S^{\alpha}(\hat{\rho}) - \log \frac{\chi}{1-\alpha} \right) \tag{3.38}$$

Si noti tuttavia che l'esistenza di un'approssimazione efficiente non implica quella di un *algoritmo* efficiente per determinarla. É infatti possibile, specificamente qualora il gap non sia costante al variare della dimensione del sistema ma scali come l'inverso di un polinomio in N, che tale problema sia NP-hard. Nel caso di gap costante è stato invece derivato un algoritmo con scaling polinomiale; in generale non è stata ancora data una caratterizzazione completa della complessità del problema per un sistema qualsiasi [41].
Tramite questi risultati si può determinare un limite per la dimensione di bond necessaria per sistemi critici, che come detto in precedenza rappresentano il caso limite. Considerando un sistema finito di $2L$ e come blocco metà della catena, l'entropia di Renyi è data dalla 3.14 salvo per una correzione di dimensione finita che scala come $\frac{1}{L}$. Supponendo si voglia mantenere l'errore sullo stato complessivo sotto:

$$\| \, |\psi_\chi\rangle - |\psi_0\rangle \, \|^2 \leq \frac{\varepsilon}{L} \tag{3.39}$$

con ε costante indipendente da L si ha:

$$\chi_L \leq \text{cost} \cdot \left(\frac{L^2}{(1-\alpha)\varepsilon} \right)^{\frac{\alpha}{1-\alpha}} L^{\frac{c+c^*}{12} \frac{1+\alpha}{\alpha}} \tag{3.40}$$

ovvero χ deve scalare solo polinomialmente in L e quindi la rappresentazione approssimata tramite MPS è efficiente.

3.3 Manipolazioni ed estrazione di informazioni

Come già si è osservato in generale nel capitolo 2, l'utilizzo dei MPS non sarebbe giustificato se l'efficienza riguardasse meramente rappresentazione degli stati e non l'estrazione di informazioni da questi. In effetti, essi soddisfano anche la seconda condizione, in virtù della loro struttura.

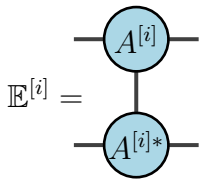
3.3.1 Transfer matrix

Un vantaggio cruciale dei MPS, che discende dalla loro costruzione, è la possibilità di determinare proprietà complessive tramite lo studio della struttura *locale*; questo è di particolare utilità nel caso di invarianza traslazionale e permette un approccio rigoroso ed efficiente al limite infinito. Questa proprietà fondamentale è del tutto generale per stati in forma di TN di sistemi locali: per il calcolo di prodotti scalari - quindi in particolare norme, overlap, valori di aspettazione ed elementi fuori diagonale - si adotta una strategia di *riduzione della dimensionalità* [52]. L'oggetto matematico fondamentale per caratterizzare la struttura locale è la **matrice di trasferimento (Transfer Matrix)** [2]:

Definizione 3.6. *Dato un MPS nella forma della definizione, la Transfer Matrix per il sito j-esimo è il super-operatore:*

$$\hat{\mathbb{E}}^{[i]} \equiv \hat{A}^{[i]}_{j_i} \otimes \hat{A}^{[i]\dagger}_{j_i} \tag{3.41}$$

ovvero in componenti il tensore:



La matrice di trasferimento può essere equivalentemente interpretata come funzionale multilineare, secondo la definizione di tensore[27]:

$$\hat{\mathbb{E}}^{[i]} \ : \ \mathbb{C}^{\chi_i} \otimes \mathbb{C}^{\chi_i} \otimes \mathbb{C}^{\chi_{i+1}} \otimes \mathbb{C}^{\chi_{i+1}} \ \rightarrow \ \mathbb{C} \tag{3.42}$$

o come super-operatore:

$$\hat{\mathbb{E}}^{[i]} \ : \ \mathcal{B}(\mathbb{C}^{\chi_i}) \cong \mathbb{C}^{\chi_i^2} \ \rightarrow \ \mathcal{B}(\mathbb{C}^{\chi_{i+1}}) \cong \mathbb{C}^{\chi_{i+1}^2} \tag{3.43}$$

[27] Sugli spazi associati agli indici virtuali.

In questa seconda forma non è che una mappa[28]:

$$\hat{\mathbb{E}}^{[i]}(\hat{X}) = \hat{A}^{[i]}_{j_i} \hat{X} \hat{A}^{[i]\dagger}_{j_i} \tag{3.44}$$

Le matrici del MPS sono quindi operatori di Kraus per l'OSR della matrice di trasferimento, che di conseguenza è una mappa completamente positiva [2, 43]. Senza perdita di generalità si può assumere che soddisfino la relazione di Kraus e pertanto la mappa sia un'operazione; se poi specificamente vale la completezza - ovvero $\hat{\mathbb{E}}$ preserva la traccia - confrontando la costruzione 3.22 con la def.2.5 si può notare che lo spazio ancillare necessario per costruire l'unitaria globale è proprio quello dell'indice fisico j_i. La libertà unitaria nella determinazione degli operatori di Kraus si traduce nella libertà di cambiare la base dell'indice fisico senza alterare l'azione di $\hat{\mathbb{E}}$ [43]. Assumendo che le dimensioni di bond siano uniformi, essendo un operatore positivo su spazio complesso la matrice di trasferimento è unitariamente diagonalizzabile. Il suo spettro contiene, come si vedrà di seguito, tutte le proprietà rilevanti del MPS. In particolare è utile ai fini dello studio spettrale l'applicabilità del teorema di Perron-Frobenius per mappe CPTP enunciato nel capitolo precedente.
Si noti che la denominazione della TM non è incidentale in quanto si può stabilire un'analogia con l'oggetto matematico omonimo che si introduce in la meccanica statistica[29] [36]. La differenza è che nel caso degli MPS si decompone uno stato anziché l'hamiltoniano e si ha una struttura a doppio anziché singolo "strato" [54].

3.3.2 Forme canoniche

Si è visto in generale che i TN godono di libertà di gauge; nel caso degli MPS è possibile sfruttarla sugli indici di bond senza alterare lo stato quantistico descritto[30]. A rigore si può dimostrare un'istanza specifica del risultato derivato nel primo capitolo, ovvero che si tratta dell'unica libertà nella parametrizzazione dello stato

[28] Analogamente si può considerare l'azione da destra, scambiando il ruolo delle $\hat{A}$ e $\hat{A}^\dagger$.

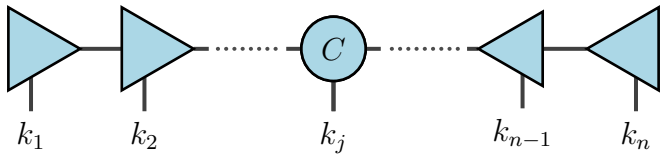
[29] Dato un sistema unidimensionale con interazioni solo tra primi vicini, si possono fattorizzare i fattori di Boltzmann ed esprimere la funzione di partizione come prodotto di matrici - sia nel caso classico che quantistico, con delle differenze formali che non si approfondiscono in questa sede. Per OBC e PBC rispettivamente:

$$Z = \mathcal{T}^{[1]}_{j_1} \mathcal{T}^{[2]}_{j_1 j_2} \cdots \mathcal{T}^{[n]}_{j_{n-1}} \qquad Z = \mathrm{Tr}[\mathcal{T}^{[1]} \cdots \mathcal{T}^{[n]}] \tag{3.45}$$

É interessante notare che l'analogia si estende al calcolo dei valori medi sull'ensemble statistico e dei valori di aspettazione sulla sovrapposizione in un MPS, in particolare con l'impiego delle proprietà spettrali nel limite termodinamico. Sfruttando questa affinità sono state sviluppate delle tecniche di TN per lo studio di sistemi classici, ad esempio un algoritmo di rinormalizzazione basato sull'estensione della DMRG, detto *Transfer Matrix Renormalization Group* (*TMRG*) [53].

[30] Che dunque non determina univocamente l'espressione come MPS, mentre vale il reciproco.

come MPS, ovvero che tutte le possibili decomposizioni sono equivalenti a meno di risoluzioni dell'identità sulle gambe virtuali [2]. Chiaramente questo a patto di mantenere invariata la forma del network, ovvero il numero di siti; altrimenti si può considerare come libertà addizionale quella di fondere gli indici fisici di più siti e contrarre le corrispondenti matrici [7].
Già con la costruzione MPD si sono introdotte delle forme canoniche (secondo il senso definito nel primo capitolo). É di grande utilità osservare che se la matrice di un sito è un'isometria a sinistra, allora la matrice di trasferimento ha come autovalore dominante 1 e come autovettore associato a sinistra l'identità; l'analogo vale per isometrie a destra considerando l'autovettore a destra [2].
Anziché procedere unicamente da destra o sinistra, in analogia al procedimento generale per creare centri di ortogonalità si può anche procedere alla isometrizzazione di un MPS contemporaneamente nelle due direzioni, fino a ottenere la **Mixed-Canonical Decomposition**:



ove i due rami per costruzione definiscono vettori ortonormali:

$$|\psi\rangle = C_{\alpha,k_j\beta}\,|\alpha\rangle\,|k_j\rangle\,|\beta\rangle \qquad \langle\alpha|\alpha'\rangle = \delta_{\alpha\alpha'} \quad \langle\beta|\beta'\rangle = \delta_{\beta\beta'} \tag{3.46}$$

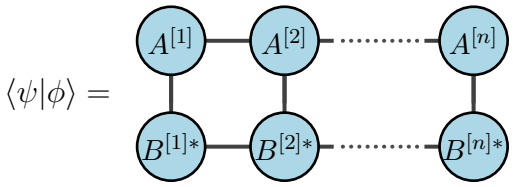
La procedura iterativa della MPD è semplice ed efficiente soprattutto nel caso di OBC [7]. In generale LCD, RCD e MCD sono utilizzate di norma negli algoritmi variazionali in quanto semplificano il computo delle contrazioni [7].
Si noti che tramite la matrice di trasferimento è possibile estendere la forma canonica al caso infinito, secondo una procedura detta **ortonormalizzazione degli indici** [52]. Si nota infatti che per definizione della SVD gli indici di bond corrispondono a destra ed a sinistra a basi ortonormali (in generale differenti); considerati gli autovettori dominanti destro e sinistro della matrice di trasferimento, che sono positivi per quanto visto in precedenza, se ne prendono le radici quadrate:

$$L_1 = Y^\dagger Y \qquad R_1 = XX^\dagger \tag{3.47}$$

X e Y^T definiscono quindi le trasformazioni di gauge da applicare agli indici di bond per portare il MPS infinito in forma canonica.

3.3.3 Modulo e overlap

La rappresentazione diagrammatica suggerisce immediatamente come calcolare l'overlap con un altro MPS dalla stessa forma, e quindi anche il modulo:



Per MPS con condizioni aperte al contorno, la procedura di contrazione ottimale è quella identificata tramite ispezione diagrammatica per la geometria a ladder (vd. capitolo 2) [2]. Assumendo una dimensione di bond uniforme, il costo del calcolo della norma è quindi $\mathcal{O}(n\chi^3)$ [2].

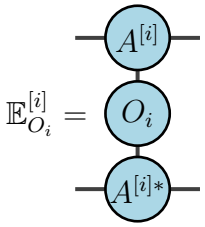
3.3.4 Valori di aspettazione

Si consideri innanzitutto il calcolo per un operatore decomponibile. La procedura ottimale per la contrazione è un'immediata generalizzazione di quella vista per l'overlap. Procedendo da sinistra, si possono raggruppare le contrazioni per ogni "scalino" in un unico step di aggiornamento della **boundary matrix**:

$$(\mathbb{L}^{[k]})^{\alpha\beta} = (A^{[k]})_i^{\gamma\alpha}(\mathbb{L}^{[k-1]})^{\gamma\delta}(O^{[k]})_{ij}(A^{[k]\dagger})_j^{\delta\beta} \tag{3.48}$$

Ove $\mathbb{L}^{[0]}$ è uno scalare e $\mathbb{L}^{[n]}$ il risultato finale. Per dimensioni uniformi si evince dall'espressione della contrazione che il costo di un singolo step è $\mathcal{O}(d^2\chi^3)$ [2]. Per condizioni al contorno aperte, in un sistema finito il costo è $\mathcal{O}(nd\chi^3)$, in un sistema infinito è $\mathcal{O}(d\chi^3)$. Nel secondo caso si opera infatti con una contrazione simultanea da entrambi i lati, assunte opportune condizioni al contorno - una procedura equivalente allo studio spettrale della matrice di trasferimento, ovvero un problema 0-dimensionale. Nel caso di condizioni periodiche il punto di partenza è l'espressione compatta[31] [43]:

$$\langle \bigotimes_{i=1}^n \hat{O}_i \rangle = \mathrm{Tr}\left[\prod_{i=1}^n \hat{\mathbb{E}}_{O_i}^{[i]}\right] \qquad \hat{\mathbb{E}}_{O_i}^{[i]} \equiv \langle j|\hat{O}_i|k\rangle \, \hat{A}_j^{[i]} \otimes \hat{A}_k^{[i]\dagger} \tag{3.49}$$



A livello computazionale, si sceglie un sito di partenza e si procede in analogia al caso aperto; tuttavia essendo tutti i tensori di medesima dimensionalità il costo è

[31] La matrice di trasferimento ordinaria può essere considerata il caso particolare di $\hat{O}_i = \hat{\mathbb{I}}$

maggiore, in particolare $\mathcal{O}(nd\chi^5)$.
Il vantaggio di operare direttamente a livello di matrici di trasferimento è evidente a livello di algoritmi di rinormalizzazione: ad ogni iterazione non risulta infatti necessario scartare gradi di libertà per mantenere una rappresentazione *efficace* efficiente dello stato del blocco di siti accumulato, in quanto la dimensione della matrice ottenuta dalla contrazione sequenziale - ovvero la transfer matrix del blocco - è costantemente χ^4 [7].

3.3.5 Correlatori

Con il formalismo appena introdotto si può ora studiare anche la struttura di correlazione dei MPS, oltre a quella di entanglement. Considerando un MPS uniforme, il calcolo del correlatore a due corpi:

$$C(r) \equiv \langle \hat{O}_i \hat{O}'_{i+r} \rangle - \langle \hat{O}_i \rangle \langle \hat{O}'_{i+r} \rangle \tag{3.50}$$

si può esprimere in termini della matrice di trasferimento [5]. Diagonalizzando e sfruttando le proprietà di ortogonalità degli autovettori:

$$\mathbb{E}^r = \left(\lambda_i R_i^T L_i\right)^r = \lambda_i^r R_i^T L_i = \lambda_1^r \left(R_1^T L_1 + \sum_{i>1} \left(\lambda_j/\lambda_1\right)^r R_i^T L_i \right) \tag{3.51}$$

ove si sono ordinati gli autovalori in senso decrescente e si è normalizzato su quello dominante, assunto non degenere. Ora per $r \gg 1$ la somma residua è dominata dal secondo autovalore, pertanto detta μ_2 la sua molteplicità:

$$\mathbb{E}^r \approx \lambda_1^r \left(R_1^T L_1 + \left(\lambda_j/\lambda_1\right)^r \sum_{i=1}^{\mu_2} R_i^T L_i \right) \tag{3.52}$$

il valore di aspettazione congiunto diviene, tenendo conto della normalizzazione:

$$\langle \hat{O}_i \hat{O}'_{i+r} \rangle \sim \frac{(L_1 \mathbb{E}_O R_1^T)(L_1 \mathbb{E}_{O'} R_1^T)}{\lambda_1^2} + \left(\lambda_2/\lambda_1\right)^{r-1} \sum_{i=1}^{\mu_2} \frac{(L_1 \mathbb{E}_O R_i^T)(L_i \mathbb{E}_{O'} R_1^T)}{\lambda_1^2} \tag{3.53}$$

Nel limite termodinamico si ha quindi:

$$\langle \hat{O}_i \hat{O}'_{i+r} \rangle \propto L^\dagger \hat{\mathbb{E}}^j R \tag{3.54}$$

con L e R autovettori dominanti sinistro e destro rispettivamente.
Poiché si può sempre normalizzare la matrice di trasferimento e di conseguenza applicare Perron-Frobenius, si può assumere senza perdita di generalità che l'autovalore dominante sia non degenere e che gli autovalori subdominanti giacciono entro il disco unitario. Di conseguenza i correlatori possono essere costanti (autovalore 1) o decadere esponenzialmente (altri autovalori): quindi gli MPS a dimensione di bond finita possono rappresentare solo stati con correlazioni che decadono esponenzialmente, ovvero sono sempre **finitamente correlati**, e non quelli a invarianza di

scala, ovvero con lunghezza di correlazione divergente [7, 5]. Ora, il primo termine nella 3.53 non è che $\langle \hat{O}_i \rangle \langle \hat{O}'_{i+r} \rangle$ e dunque:

$$C(r) \sim \left(\lambda_2/\lambda_1\right)^{r-1} \sum_{i=1}^{\mu_2} \frac{(L_1 \mathbb{E}_O R_i^T)(L_i \mathbb{E}_{O'} R_1^T)}{\lambda_1^2} = f(r) a e^{-r/\xi} \tag{3.55}$$

con $a = \mathcal{O}(\mu_2)$ costante reale, $f(r) = e^{i\phi(r)}$ fattore di fase che per operatori hermitiani può essere solo ± 1 e $\xi \equiv -\frac{1}{\log|\lambda_2/\lambda_1|}$ lunghezza di correlazione.
Finora si sono considerati valori di aspettazione di operatori, quindi elementi di matrice diagonali. Le considerazioni fatte si possono però generalizzare al caso fuori diagonale rimpiazzando il vettore od il covettore di stato con quello di un altro MPS [7].

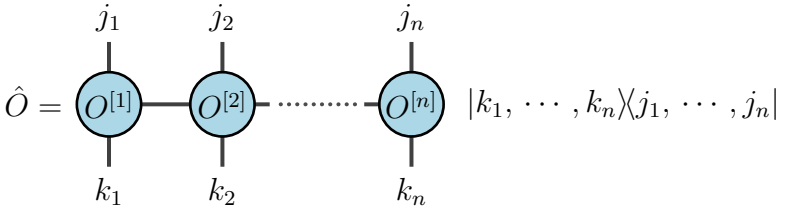
3.3.6 Matrix Product Operators

Il calcolo visto per operatori a singolo decomponibili può poi essere generalizzato a qualsiasi osservabile sul sistema complessivo; per renderlo efficiente è però necessario determinare una rappresentazione pratica anche per gli operatori. Questa è semplicemente ottenuta estendendo la MPD allo spazio degli operatori, che è il prodotto tensore:

$$\left(\bigotimes_{i=1}^{n} \mathcal{H}_i\right) \otimes \left(\bigotimes_{i=1}^{n} \mathcal{H}_i\right)^* \cong \bigotimes_{i=1}^{n} (\mathcal{H}_i \otimes \mathcal{H}_i^*) \tag{3.56}$$

Dato un operatore generico, lo si sviluppa su di una base ortonormale, costruita tensorizzando basi di ket e bra per gli spazi ed i rispettivi duali: Ci si riduce alla decomposizione per stati fondendo l'indice di ciascuno spazio locale con quello del suo duale; al termine della procedura si effettua quindi lo splitting e si ottiene la rappresentazione come **Matrix Product Operator (MPO)** [2]:

Definizione 3.7. *Un operatore su un sistema a molti corpi è un MPO se è esprimibile nella forma:*



In linea di principio, la decomposizione di un operatore generico è altamente inefficiente a causa dell'elevata dimensionalità; tuttavia si può determinare in modo efficiente qualora l'operatore non incrementi eccessivamente il rango di Schmidt

dello stato - ovvero non *crei* troppo entanglement [2].
Il caso triviale è quello di operatori di singolo sito e di loro prodotti tensore, ove in analogia al caso degli stati le matrici della MPD si riducono a scalari; più in generale però la proprietà è soddisfatta dagli operatori *locali*, che come si è detto in precedenza sono quelli di maggiore rilevanza fisica [7].
In realtà vi è anche un caso di operatore altamente non-locale di cui si può ottenere facilmente la rappresentazione come MPO: il proiettore su un MPS, ovvero l'operatore densità dello stato puro. Infatti per costruirla è sufficiente accostare il MPS al corrispondente covettore nella medesima forma[32] [2, 7]. L'operatore ridotto per un blocco di siti può essere facilmente ottenuto portando le porzioni esterne del MPS in forma canonica - destra o sinistra a seconda della posizione relativa al blocco, con le matrici non isometriche residue assorbite nei siti estremi - e quindi contraendo sulle gambe fisiche di queste. In tal modo non si ottiene che la contrazione delle matrici di trasferimento, la quale per le proprietà viste si riduce all'identità da ambo i lati. Le operazioni lineari tra MPO sono definite in modo del tutto analogo ai MPS [2]: in particolare, la somma produce un MPO con matrici diagonali a blocchi. Tuttavia, per operatori locali (dunque anche hamiltoniane locali), esiste una procedura per costruire sistematicamente rappresentazioni sempre esatte ma più efficienti, che evitino essenzialmente la ridondanza di matrici a blocchi in buona parte sparse [2].
Il caso triviale è quello di un'hamiltoniana non interagente, ovvero in cui ciascun termine ha supporto su un solo sito:

$$\hat{H} = \sum_j \hat{\mathbb{I}}_1 \otimes \ldots \hat{\mathbb{I}}_{j-1} \otimes \hat{H}_j \otimes \hat{\mathbb{I}}_{j+1} \otimes \ldots \otimes \hat{\mathbb{I}}_n \tag{3.57}$$

ossia è un operatore prodotto, rappresentabile da un MPO con dimensione di bond triviale. Per identificare i tensori della decomposizione completa, ci si concentri su quelli di *bulk* e li si consideri come matrici a valori operatoriali (nell'algebra degli spazi fisici) agenti sulle gambe virtuali; l'espressione dell'hamiltoniana è costruita dalla loro contrazione progressiva, che si può alternativamente concepire come l'iterazione di processi a stati finiti:

1. Di base, l'azione su un sito è triviale, quindi si inserisce l'identità;
2. Se sul sito agisce un termine diverso dall'identità, lo si inserisce e *se ne ha memoria*, ovvero successivamente vi può essere collocata solo l'identità.

Gli stati per ogni sito sono quindi due: uno di "idle", in cui l'azione di $\hat{H}$ è triviale, e uno in cui si passa se si incontra un termine non triviale - stato dal quale, per la memoria, non si può tornare indietro. Di conseguenza la dimensione di bond della rappresentazione efficiente è 2. Si può visualizzare in modo pratico tramite un diagramma:

[32] Ovvero effettuandone il prodotto tensore.

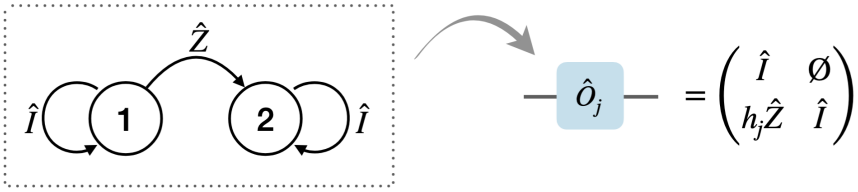


Figura 3.4: Macchina a stati finiti per la costruzione dei tensori del MPO hamiltoniano. Immagine da [2].

i vettori estremi sono semplicemente dati dall'ultima riga e dalla prima colonna.
Per un'hamiltoniana interagente, i termini in generale tensorizzano più operatori non triviali: di conseguenza anche in uno stato "attivo" si può aggiungere un operatore non triviale. In generale ciò significa che ad ogni incremento del raggio dell'interazione di un unità si fa corrispondere l'aggiunta di un layer di stati nel diagramma, in numero pari a quello dei termini a più siti; lo stato attivo finale rimane unico e si ha sempre la memoria. L'ampiezza del layer intermedio si somma quindi a 2 per dare la dimensione degli indici di bond.
Si noti che:

- Se l'interazione è solo tra siti contigui, allora non vi sono loop sugli stati dei layer intermedi, ovvero non vi si può aggiungere l'identità, bensì sono solo mappati in quelli successivi;
- Se l'interazione accoppia anche gruppi di siti non contigui, vi possono essere dei loop sui layer intermedi.

In analogia a quanto possibile per gli stati la costruzione degli MPO può essere estesa a dimensioni maggiori tramite i **Projected Entangled Pair Operators (PEPO)** [12].

3.3.7 Finitely Correlated States (FCS)

É interessante aggiungere in conclusione una digressione su un'altra caratterizzazione dei MPS, maggiormente legata alla struttura di *correlazione* che si è appena descritto. Come notato nell'introduzione, il quadro unificato che si può presentare oggi è in realtà frutto di una serie di sviluppi indipendenti le cui interconnessioni naturali sono state rivelate solo in un secondo tempo. Ciò vale anche per gli stessi MPS, che non furono introdotti in origine utilizzando il linguaggio tensoriale, bensì quello puramente algebrico: la prima forma è infatti quella dei **Finitely Correlated States (FCS)**, a posteriori identificabili con gli stati infiniti TI, la cui costruzione generalizzava VBS e AKLT e stabiliva soprattutto un primo legame esplicito tra sistemi su reticolo e teoria dell'informazione quantistica, tramite la caratterizzazione della struttura di correlazione locale in termini di canali quantistici [55]. La costruzione è infatti la seguente [43]: data una spin chain infinita, si associa a ogni spazio di Hilbert locale $\mathcal{H}_S$ uno sistema ancillare $\mathcal{H}_A$ e si definisce una mappa $\hat{\mathcal{E}} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_S)$ completamente positiva e che preserva la traccia. Un FCS per la catena è uno stato tale per cui la matrice ridotta per un blocco di n siti è ottenuta come traccia parziale sull'ancella:

$$\hat{\rho}_n = \mathrm{Tr}_A[\hat{\mathcal{E}}^n(\Lambda)] \tag{3.58}$$

ove Λ soddisfa $\Lambda = \mathrm{Tr}_B[\mathbb{E}(\Lambda)]$, ossia è un punto fisso della mappa indotta sugli stati di A. É immediato identificare questa con la matrice di trasferimento e gli

spazi ancillari con gli spin virtuali negli MPS; la costruzione tramite l'iterazione della mappa $\mathcal{E}$ è detta difatti anche generazione sequenziale dei MPS [46].

Capitolo 4

Il Density Matrix Renormalization Group

Si definisce la tecnica di riferimento per il calcolo dello spettro e degli stati a bassa energia di sistemi 1D. La si introduce come perfezionamento del gruppo di rinormalizzazione numerico e si mostra la naturale riformulazione come algoritmo variazionale sui MPS. Se ne dà quindi un'applicazione tramite la libreria ITensor per il linguaggio Julia.

4.1 Rinormalizzazione

Il concetto di *rinormalizzazione* trae le proprie origini dalla teoria quantistica dei campi: venne infatti introdotto negli anni '40 per descrivere una procedura, di principio solo formale, che permetteva di eliminare le divergenze negli sviluppi perturbativi dovute ai diagrammi di auto-interazione introducendo una dipendenza dell'accoppiamento dalla scala di energia in grado di sopprimere i contributi nell'*ultravioletto*, ossia ad alte energie [56].
Fu però solo negli anni '70 che venne formulata una teoria generale della rinormalizzazione in grado di darne una giustificazione rigorosa sul piano matematico ed al contempo illuminare le profonde ragioni fisiche della sua efficacia. A svilupparla fu Kenneth Wilson, che definì il **Gruppo di Rinormalizzazione** (**Renormalization Group, RG**) con lo scopo iniziale di risolvere un problema non di fisica delle interazioni fondamentali, bensì degli stati condensati[1] - inaugurando un proficuo dialogo tra i due ambiti che prosegue ancora oggi e che coinvolge anche le tecniche di TN [57].
Il gruppo di rinormalizzazione è formato da trasformazioni di scala spaziale od

[1] Il cosiddetto *problema di Kondo*, che riguarda l'andamento della resistività di metalli in presenza di impurità magnetiche.

energetica, regolate da un parametro di lunghezza o, appunto, energia; queste sono applicate incrementando progressivamente la scala del sistema ed effettuando per ogni step una riduzione (detta **decimazione**) dei gradi di libertà, ottenendo quindi una rappresentazione *efficace* dello stato e delle interazioni [58]. Il principio fondamentale dietro una tale procedura è legato alle transizioni di fase del secondo ordine[2]: queste sono infatti caratterizzate dalla presenza di fluttuazioni rilevanti su tutte le scale, ovvero da una lunghezza di correlazione divergente, e pertanto esibiscono invarianza di scala[3] - in altri termini, i punti critici sono punti fissi del flusso di rinormalizzazione [59].
É interessante notare che una delle conseguenze principali dell'introduzione del RG è la possibilità di definire delle **classi di universalità**, ovvero degli insiemi di sistemi che in virtù di alcune proprietà comuni del tutto generali, prescindenti dalla specifica struttura microscopica - essenzialmente la dimensionalità, le simmetrie del parametro d'ordine e l'ampiezza del range delle interazioni - presentano i medesimi **esponenti critici** nello scaling delle grandezze termodinamiche in prossimità delle transizioni. Questo fornisce la vera giustificazione formale per lo studio di modelli semplificati su reticolo: per quanto concerne le proprietà nel limite termodinamico, sono infatti ininfluenti tutti quei gradi di libertà che divengono irrilevanti nella definizione delle interazioni efficaci col procedere della rinormalizzazione - ossia le strutture addizionali di un sistema che non ne discriminano la classe di universalità da quella del modello semplificato [59].

4.1.1 Numerical Renormalization Group

Per gli scopi di questo lavoro, ci si concentrerà sulle applicazioni del RG a spin system, evitando la caratterizzazione formale generale della rinormalizzazione, che non è necessaria per definirle e comprenderle. Per sistemi su reticolo si ha una discretizzazione naturale del flusso, in quanto si può definire uno step di riscalamento come l'aggiunta di uno o più siti, e corrispondentemente anche della scala energetica [60]. Si noti che a livello computazionale la soglia di decimazione determina l'efficienza dell'algoritmo, in quanto limita la quantità di risorse necessarie alla rappresentazione di stati ed hamiltoniane, ma che l'efficacia, ossia in particolare la convergenza, dipendono dal *criterio* con cui sono scelti gli stati da mantenere.
La prima tecnica di rinormalizzazione non perturbativa per lo studio di sistemi su reticolo fu introdotta da Wilson stesso proprio per la soluzione del problema di Kondo e prende il nome di **Gruppo di Rinormalizzazione Numerico** (**Numerical**

[2]Ovvero quelle in cui, secondo la classificazione di Sommerfeld, le quantità termodinamiche rimangono continue ma le loro derivate presentano discontinuità.

[3]Una lunghezza di correlazione divergente implica un decadimento dei correlatori a legge di potenza, (*power law*), che è invariante per riscalamento a meno di un fattore moltiplicativo dipendente dal solo fattore di riscalamento.

RG, NRG) [4]. Nel NRG il criterio di troncamento è energetico: si mantengono gli stati a energia minore [58].
L'algoritmo procede come di seguito. Si fissa preliminarmente una soglia dimensionale D (di norma nell'ordine delle centinaia). Quindi si procede iterativamente [58, 9]:

1. Si considera un blocco di lunghezza L t.c. $Ld = D$, ove d è la dimensione dello spazio locale. Si ottiene la rappresentazione dell'hamiltoniana su una base e la si diagonalizza (esattamente o tramite Lanczos);

2. Si raddoppia la dimensione del blocco. Utilizzando per la rappresentazione la base prodotto di quelle ottenute allo step precedente, si diagonalizza l'hamiltoniana. Si scelgono quindi i primi D autostati a energia minore come base per il blocco $2L$, ovvero si proiettano lo spazio degli stati e le osservabili sul sottospazio generato da $\{|\psi_i\rangle\}_{i=1}^{D}$;

3. Si ripete lo step 2 fino al raggiungimento della dimensione del sistema originario (nel caso finito) oppure un punto fisso, ovvero alla convergenza (nel caso infinito).

Sebbene efficace in certe specifiche situazioni, come il problema di Kondo, in generale l'applicazione del NRG agli spin system dà risultati del tutto errati [58]. La ragione fondamentale è che quando si raddoppia il blocco accostandovi una sua copia non si tiene in conto degli effetti di bordo, ovvero dell'interazione tra i due: se questa non è trascurabile, come avviene in generale[5], allora cade l'assunzione fondamentale del NRG, ovvero che gli stati a minore energia del blocco $2L$ in generale siano dati da combinazioni dei soli stati a minore energia dei blocchi L [9]. Lo si può verificare con un esempio semplice ma molto eloquente [58].

Esempio 4.1. Si consideri una particella in una buca rettangolare ("particle in a box"). Questa può essere semplicemente descritta come una versione a singolo sito di un modello di *tight binding*[6]:

$$\hat{H} = -t\sum_i \big(|i\rangle\langle i+1| + |i+1\rangle\langle i| \big) \tag{4.1}$$

Tale hamiltoniana non è che una discretizzazione del laplaciano [61], quindi è diagonale nella base dei momenti: la base spettrale è data dalle onde piane con k che

[4] Sebbene nel caso specifico si tratti di un sistema fermionico - in cui peraltro i siti sono dati dalle impurezze e dai livelli discretizzati logaritmicamente della banda di conduzione - la tecnica si può applicare senza alcuna modifica sostanziale al caso di spin systems.

[5] Basti pensare ad un sistema omogeneo con interazioni tra primi vicini.

[6] Ovvero l'approssimazione di una struttura elettronica come sovrapposizione di stati discreti localizzati tra cui gli elettroni possono saltare.

soddisfa alle condizioni di annullamento al contorno. Ma ciò implica che lo stato fondamentale per il sistema di lunghezza $2L$ non può essere in alcun modo approssimato da combinazioni di stati a bassa energia dei due sottosistemi di lunghezza L, in quanto queste presenteranno sempre un nodo al centro, ove esso ha invece un antinodo.

Una prima soluzione è modificare opportunamente le condizioni al contorno per risolvere questo problema di bordo, ma si tratta di un accorgimento ad hoc e di scarsa praticità. Una via d'uscita molto più efficace si può invece trovare sfruttando gli strumenti introdotti nel capitolo 3, in particolare l'applicazione del formalismo dell'operatore densità a sistemi composti entangled: anziché considerare combinazioni di stati puri dei blocchi, si può fare riferimento alle miscele degli stati ridotti corrispondenti alle basse energie del *superblocco* di dimensione doppia, ottenuto accoppiando al sistema un *ambiente* che descrive i gradi di libertà al contorno [62]. Fu seguendo questo principio che White introdusse il **Gruppo di Rinormalizzazione della Matrice Densità** (**Density Matrix Renormalization Group, DMRG**) [63]. Essenzialmente, anziché utilizzare un mero criterio energetico per la decimazione, si sfrutta la decomposizione di Schmidt per determinare gli stati dei blocchi con maggiore rilevanza per la descrizione del sistema *complessivo*.

4.2 Density Matrix Renormalization Group

Il DMRG può essere formulato in due modi simili ma ciascuno più adatto ad una diversa situazione: il DMRG finito e il DMRG infinito, a seconda della natura del sistema in esame [58]. Per comprendere il principio generale è utile partire dal caso infinito, che è anche quello storicamente precedente [9]. Si noti che si assumono OBC.

4.2.1 DMRG infinito

Lo scopo è determinare lo spettro a basse energie nel limite termodinamico. Si fissa preliminarmente la soglia di decimazione D (di norma anche nell'ordine delle migliaia) e una base per i singoli siti. Si considerano quindi due blocchi A (sistema) e B (ambiente) di uguale lunghezza, inizialmente composti di un singolo sito ciascuno. Si procede come di seguito [58]:

1. Si aggiunge un sito ad entrambi i blocchi in posizione intermedia, ottenendo il sistema $A \bullet \bullet B$;

2. Rappresentandola sulla base prodotto, si diagonalizza l'hamiltoniana, determinando in particolare gli stati a bassa energia. Di norma non si utilizza la diagonalizzazione esatta bensì dei metodi che sfruttano la sparsità delle hamiltoniane locali e determinano solo parte degli autovettori e autovalori, come quello di Lanczos o di Davidson [64]. Si seleziona quindi lo stato fondamentale;

3. Si effettua la decomposizione di Schmidt di quest'ultimo, tramite la diagonalizzazione delle matrici ridotte. Se la dimensione dello spazio degli stati di $A\bullet$ e $\bullet B$ è minore o uguale a D, si procede allo step successivo, altrimenti si effettua la decimazione selezionando come base per ogni blocco i primi D elementi della base di Schmidt corrispondente;

4. Si ripetono i punti 1. - 3. fino a convergenza.

Si noti che, al netto dell'argomento euristico sulla rilevanza degli stati, l'utilizzo della decomposizione di Schmidt - ovvero della SVD - è giustificato sul piano matematico in virtù del teorema di Eckart-Young [58]. La procedura descritta si può in realtà estendere ai primi stati eccitati, approssimati iterativamente in parallelo a quello fondamentale sotto il vincolo di ortogonalità; tuttavia il loro numero è fortemente limitato dai requisiti di efficienza degli algoritmi di diagonalizzazione approssimativa [62].

4.2.2 DMRG finito

Se si intende studiare un sistema di dimensione finita anziché stimare proprietà nel limite termodinamico, si può naturalmente pensare di interrompere l'algoritmo infinito alla scala opportuna. Tuttavia questo può portare a predizioni anche qualitativamente errate in presenza di disomogeneità (ad esempio impurezze) e in generale strutture non triviali, per cui certi effetti si possono osservare solo considerando l'hamiltoniana completa: il troncamento sulla base dell'hamiltoniana costruita iterativamente produce in questi casi degli stati locali che offrono una pessima approssimazione di quelli effettivamente rilevanti nel quadro complessivo. Formalmente, in questo caso l'errore di approssimazione *non* è controllato meramente dall'errore di troncamento dei coefficienti di Schmidt e pertanto si dice che l'algoritmo non è **quasi-esatto**, come avviene invece nel caso propriamente infinito. In conseguenza di ciò, una volta raggiunta la dimensione completa è opportuno applicare una variante finita del DMRG, che si descrive di seguito [62, 64].
Il sistema è dato nella forma prodotta dall'algoritmo infinito, due superblocchi con una coppia di siti intermedia. Si fissa sempre una soglia dimensionale D e si prosegue quindi secondo degli *sweep*:

1. Si aggiunge al blocco-ambiente un sito e ne si sottrae uno al sistema;

2. Si diagonalizza l'hamiltoniana complessiva e si applica il troncamento della base al solo blocco dell'ambiente, aggiornando quindi la base e la rappresentazione efficace degli operatori;

3. Si prosegue fino a raggiungere una dimensione per cui lo spazio di Hilbert del blocco-sistema ha dimensione minore od uguale a D, e quindi è descritto esattamente. A questo punto si inverte la direzione dello sweep e si procede come in precedenza.

Si ripetono quindi gli *sweep* fino a raggiungere la convergenza.

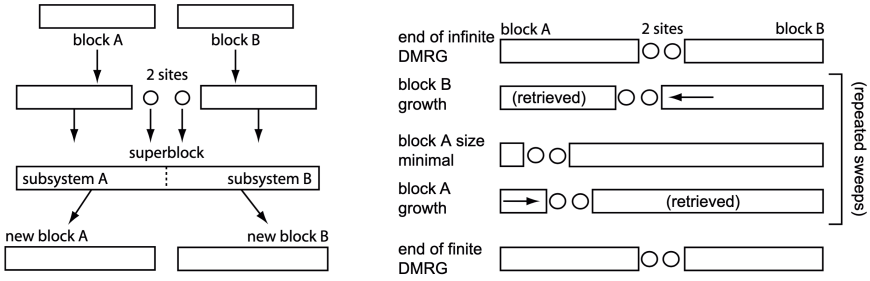


Figura 4.1: DMRG finito e infinito. Immagine da [62]

4.2.3 Legame con i MPS

La ragione profonda dell'efficacia del DMRG diviene evidente una volta applicativi i concetti di teoria quantistica dell'informazione e il formalismo dei TN. Si ottengono infatti due risultati fondamentali:

1. La procedura di rinormalizzazione definisce in generale naturalmente un MPS;
2. Il criterio di decimazione sui coefficienti di Schmidt, ovvero sull'entanglement, produce specificamente l'approssimazione *ottimale* a parità di soglia D, ossia un MPS con *dimensione di bond* D.

Poiché gli stati a bassa energia che mira a ricostruire giacciono, in virtù dell'area law, proprio nell'angolo dello spazio di Hilbert efficientemente parametrizzabile con MPS a dimensione di bond finita, il DMRG converge molto rapidamente e fornisce predizioni verificate con grande precisione. Si noti tuttavia che in virtù di quanto osservato nel capitolo precedente questo vale rigorosamente in presenza di area law per le entropie di Renyi con $\alpha < 1$, che esclude la possibilità di spettri di Schmidt patologici, cui non è invece sufficientemente sensibile l'entropia di Von Neumann. Purtroppo in generale la seconda è molto più facilmente stimabile delle prime (che richiedono l'applicazione di CFT) e pertanto l'argomento ha una valenza euristica, per quanto sia verificato in pressoché la totalità dei casi pratici.
Si dimostra il primo fatto[7], tramite una riformulazione dell'algoritmo da cui emerge naturalmente la struttura di prodotto matrice dell'ansatz [46, 65]. Sia $\mathcal{H}_1$ lo spazio degli stati di uno spin e d_1 la sua dimensione. Procedendo a partire da due spin:

[7]Sfruttabile per migliorare l'efficacia della rinormalizzazione generica ad alte energie includendo un approccio variazionale.

1. Si effettua la prima decimazione, considerando un sottospazio $\mathcal{H}_2 \subseteq \mathcal{H}_1 \otimes \mathcal{H}_1$ con $d_2 \leq d_1^2$;

2. Si aggiunge un terzo spin e nuovamente si restringe al sottospazio $\mathcal{H}_3 \subseteq \mathcal{H}_2 \otimes \mathcal{H}_1$ con $d_2 \leq d_1 d_2$;

3. Si itera fino ad ottenere $\mathcal{H}_n \subseteq \mathcal{H}_{n-1} \otimes \mathcal{H}_1$ con $d_n \leq d_{n-1} d_1$. A questo punto si approssima l'hamiltoniana tramite una proiezione da $\mathcal{H}_1^{\otimes n}$ a $\mathcal{H}_n$:

$$\hat{H} \mapsto \hat{P}_n \hat{H} \hat{P}_n \tag{4.2}$$

In generale, se ad ogni step vale l'eguaglianza dimensionale[8] non si ha alcuna approssimazione e pertanto d_n cresce esponenzialmente in n; si assuma pertanto la decimazione sopra una dimensione fissata D.
Si studi quindi la struttura degli autostati approssimati. Formando un insieme ortonormale nello spazio tensore $\mathcal{H}_1^{\otimes n}$ si possono esprimere in termini di una base prodotto, o più specificamente, di un suo sottoinsieme incluso in $\mathcal{H}_n$: sfruttando questo fatto, anziché operare direttamente nello spazio finale se ne può ricavare più praticamente procedendo iterativamente. Innanzitutto, si sviluppa una base di $\mathcal{H}_2$ su quella di $\mathcal{H}_1 \otimes \mathcal{H}_1$:

$$|\beta\rangle_2 = B^{\beta}_{k_1,k_2} |k_1\rangle_1 |k_2\rangle_1 \tag{4.3}$$

Il tensore B può essere sempre espresso come:

$$B^{\beta}_{k_1,k_2} = A^{[1]\alpha}_{k_1} A^{[2]\alpha,\beta}_{k_2} \qquad A^{[1]\alpha}_{k_1} = \delta^{\alpha}_{k_1} \qquad A^{[2]\alpha,\beta}_{k_2} = B^{\beta}_{\alpha,k_2} \tag{4.4}$$

In conseguenza dell'ortonormalità della base su $\mathcal{H}_1$ e della propria definizione questi tensori soddisfano la completezza (per contrazione sull'indice fisico k_i). Proseguendo allo step successivo, si sviluppa la base $\{|\beta\rangle_3\}$ di $\mathcal{H}_3$ su $\{|\alpha\rangle_2 |k_3\rangle_1\}$; si itera quindi fino a $\mathcal{H}_n$, ottenendo:

$$|\beta\rangle_n = A^{[n]\alpha,\beta}_{k_n} |\alpha\rangle_{n-1} |k_n\rangle_1 \qquad 1 \leq \alpha \leq n-1 \tag{4.5}$$

ove gli $A^{[n]}_{k_n}$ soddisfano a loro volta la completezza. Sostituendo ora le espressioni ricavate ricorsivamente si giunge alla forma finale:

$$|\beta\rangle_n = A^{[1]\alpha_1}_{k_1} A^{[2]\alpha_1,\alpha_2}_{k_2} \cdots A^{[n]\alpha_{n-1},\beta}_{k_n} |k_1, k_2, \ldots, k_n\rangle \tag{4.6}$$

che altro non è se non un MPS con dimensione di bond minore o uguale D. Si noti che ora l'intera procedura di rinormalizzazione è equivalentemente specificata dalle matrici $A^{[n]}$.

[8] E dunque per teorema del completamento gli spazi coincidono.

4.2.4 Formulazione variazionale

Poiché lo stato prodotto da una rinormalizzazione con soglia di decimazione D è localizzato all'interno dell'insieme dei MPS con al più dimensione di bond $\chi = D$ e la procedura è completamente parametrizzata dalle matrici, il problema della determinazione dello stato fondamentale di $\mathcal{H}_n$ può essere riformulato in termini variazionali su tale insieme. Si noti che in questo modo non è vincolato il criterio di decimazione, che è espresso proprio dalla struttura delle matrici: pertanto è esso stesso ad essere ottimizzato a parità di D [46, 65].

Si richiama preliminarmente il **teorema variazionale**, che riformula semplicemente la caratterizzazione data in precedenza per il ground state:

Teorema 4.1. *Uno stato $|\psi\rangle$ è un ground state di un hamiltoniana $\hat{H}$ se e solo se è un minimo del funzionale*

$$E[|\psi\rangle] = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{4.7}$$

Proiettando lo stato su una varietà la migliore approssimazione è data dal minimo[9] vincolato del funzionale.

Si considera inizialmente il caso finito con OBC. Data un'hamiltoniana k-locale generica:

$$\hat{H} = \sum_{i=1}^{n} \sum_{\alpha} \hat{O}_i^{\alpha} \otimes \cdots \otimes \hat{O}_{i+k-1}^{\alpha} \tag{4.8}$$

Il valore di aspettazione dell'energia su un MPS è:

$$\langle\psi|\hat{H}|\psi\rangle = \sum_{i=1}^{n} \sum_{\alpha} \mathbb{E}^{[1]} \cdots \mathbb{E}^{[i-1]} \mathbb{E}_{O_i^{\alpha}}^{[i]} \cdots \mathbb{E}_{O_{i+k-1}^{\alpha}}^{[i+k-1]} \cdots \mathbb{E}^{[n]} \tag{4.9}$$

Analogamente si può formulare in termini di matrici di trasferimento il vincolo di normalizzazione:

$$\langle\psi|\psi\rangle = \prod_{i=1}^{n} \mathbb{E}^{[i]} = 1 \tag{4.10}$$

É immediato verificare che sia l'energia che la normalizzazione sono espresse da un polinomio quadratico nelle $A^{[i]}$. Una strategia efficiente può quindi essere quella di minimizzare indipendentemente rispetto alle singole matrici, mantenendo fisse le altre [46, 65, 7]; questa può essere adottata anche nel caso di MPS TI rompendo

[9] A rigore, dall'*estremo*.

la simmetria traslazionale di modo da rendere il problema più trattabile [43]. In virtù del teorema variazionale, ogni minimizzazione individuale non può aumentare l'energia: pertanto l'algoritmo deve convergere ad un minimo, che di principio può essere locale o globale [66]. Il funzionale energia si può esprimere rispetto ad ogni $A^{[i]}$ come quoziente di Rayleigh[10] [67]:

$$E[A^{[i]}] = \frac{A^{[i]\dagger}_{k_i} H^{[i]}_{k_i,j_i} A^{[i]}_{j_i}}{A^{[i]\dagger}_{k_i} N^{[i]}_{k_i,j_i} A^{[i]}_{j_i}} \tag{4.11}$$

ove i tensori $H^{[i]}$ e $N^{[i]}$ sono definiti tramite la 4.9 e la 4.10 rispettivamente. Ora, poiché $\hat{H}$ è hermitiana lo è anche $H^{[i]}$ e analogamente per le proprietà delle matrici di trasferimento $N^{[i]}$ è positiva; dunque $E \in \mathbb{R}$ e si può minimizzare determinando il λ minimo che risolve l'equazione agli autovalori generalizzata:

$$H^{[i]}_{k_i,j_i} A^{[i]}_{j_i} = \lambda N^{[i]}_{k_i,j_i} A^{[i]}_{j_i} \tag{4.12}$$

In questa forma generale possono presentarsi delle problematicità sul piano numerico se lo spettro di $N^{[i]}$ presenta larghe differenze di scala, in quanto ciò comporta una forte sensibilità a minime variazioni del tensore incognito; nel caso di OBC si può ovviare adottando la gauge centrale, in cui diviene l'identità e quindi la 4.12 si riduce all'equazione agli autovalori ordinaria [45, 7]. Una volta ottimizzato $A^{[i]}$, con un costo $\mathcal{O}(d^2D^4)$ [43], lo si isometrizza e si passa ad un tensore adiacente, in cui è assorbita la componente non isometrica; si procede quindi fino all'estremità della catena e da questa in verso opposto, per poi ripetere nuovamente lo *sweep* fino a raggiungere la convergenza [46, 45]. Si noti che l'utilizzo della gauge centrale corrisponde a rinormalizzare procedendo dalle due estremità della catena verso il centro, anziché solo in un senso - procedura corrispondente invece alla gauge canonica a destra o a sinistra.
Quella che si è ottenuta è, al netto di dettagli implementativi, una formulazione variazionale equivalente del DMRG finito, che quindi risulta ottimale per costruzione [46, 63]. Per sistemi gappati in conseguenza dell'area law si osserva una convergenza esponenzialmente veloce, con un tempo medio di rilassamento proporzionale all'inverso del gap; poiché nel caso limite di sistemi critici il gap si chiude al più con un andamento polinomiale il DMRG variazionale è efficiente in generale [43]. Tuttavia, poiché il problema non è convesso non si può escludere che l'algoritmo conduca ad un minimo meramente locale e, come anticipato nel capitolo precedente, si è dimostrato che la determinazione del minimo globale, ovvero del MPS ottimale, è in generale NP-hard [68].
Si noti che lo sweep ripetuto in entrambe le direzioni è un elemento fondamentale:

[10] A rigore generalizzato, anche se può essere intepretato secondo la definizione ordinaria ridimensionando ciascun tensore in vettore.

il NRG corrisponde infatti solo ad un singolo sweep in una direzione, che quindi non permette di avere *feedback* tra le varie scale (spaziali e quindi energetiche) [65].

4.2.5 Condizioni al contorno periodiche

Finora si è discusso il caso di OBC. Tuttavia, in linea di principio sarebbe preferibile operare con condizioni periodiche, di modo da evitare effetti di bordo; purtroppo applicando direttamente il DMRG al caso di PBC si ottiene una precisione molto minore in quanto lo spettro degli operatori ridotti decade molto meno rapidamente [62].
Fortunatamente, si tratta essenzialmente di un artefatto della specifica struttura dell'algoritmo. Una prima soluzione è modificarlo (nel caso infinito) mantenendo le condizioni aperte ma incrementando i blocchi non con una coppia di siti intermedi, bensì aggiungendo sia al sistema che all'ambiente un sito *a destra*. La formulazione variazionale suggerisce però un modo molto più efficace, che risolve il problema alla radice: si può estendere la procedura variazionale al caso di MPS periodici, procedendo non per sweep avanti e indietro ma in senso orario ed antiorario, ponendo lo stato in forma canonica non mista ma sinistra o destra a seconda del senso di iterazione [66].

4.3 Esempio di applicazione

Si può dare un semplice esempio dell'applicazione del DMRG a un sistema di spin unidimensionale sfruttando **ITensor** [69], una libreria dedicata alla manipolazione efficiente dei TN per il linguaggio Julia [70].

4.3.1 ITensor

ITensor - abbreviazione di *Intelligent Tensor* - è una libreria basata direttamente sul linguaggio tensoriale diagrammatico; in questo modo permette di definire gli oggetti, effettuare decomposizioni e contrazioni ed infinite estrarre le informazioni in un modo intuitivo ed ottimizzato. La caratteristica principale dell'implementazione è che gli oggetti di base in ITensor non sono i tensori bensì gli indici, che contengono al proprio le informazioni addizionali necessarie per realizzare correttamente le contrazioni; inoltre sono identificati in modo permanente e pertanto non è necessario conservare in memoria alcuna struttura addizionale per poter riprodurre correttamente un network. In generale l'ordinamento e la verifica della corrette corrispondenze di forma dei tensori sono gestiti implicitamente dall'implementazione, ma è comunque possibile per l'utente effettuare manipolazioni manuali.
Di seguito si è fatto uso anche del pacchetto specializzato ITensorMPS, che permette di creare e manipolare facilmente MPS e MPO (implementando in particolare la compressione ottimale degli operatori non decomponibili) ed applicare il DMRG.

4.3.2 Problema

Si considera il caso finito dimensionale con condizioni al contorno aperte e si calcola esclusivamente il ground state. L'hamiltoniana è quella del cosiddetto **modello di Heisenberg** (isotropo):

$$\hat{H} = \sum_{j=1}^{N-1} \hat{\vec{S}}_j \cdot \hat{\vec{S}}_{j+1} = \sum_{j=1}^{N-1} \hat{S}_j^z \hat{S}_{j+1}^z + \frac{1}{2}\hat{S}_j^+ \hat{S}_{j+1}^- + \hat{S}_j^- \hat{S}_{j+1}^+ \tag{4.13}$$

Specificamente si assume un numero di siti è $N = 100$ e lo spin unitario ($S = 1$).

4.3.3 Risultati

I parametri della simulazione sono i seguenti:

- Si effettuano 5 sweep;
- Le soglie di decimazione sono in ordine: 10,20,100,100,200;
- Il cutoff sui coefficienti di Schmidt è 10^{-10}.

Al termine dell'esecuzione si ottiene come stima dell'energia di ground il seguente valore (in unità naturali):

$$E_0 = -138.94 \tag{4.14}$$

Si riporta quindi innanzitutto in figura 4.2 la convergenza dell'energia.
Si può poi verificare che lo stato ottenuto soddisfi le proprietà che sono state caratterizzate in precedenza. In primis si grafica in figura 4.3 l'andamento dell'entropia di Von Neumann per ogni bond virtuale. Si osserva che, trascurando i bordi in cui il valore è minore in virtù delle OBC, nel bulk si raggiunge un valore costante di:

$$S \approx 0.855 \tag{4.15}$$

che è nell'ordine dell'unità, come da attendersi considerando la dimensione degli spazi locali (si è utilizzato il logaritmo naturale).
Infine si calcola il correlatore connesso di spin su z:

$$C(r) = \langle \hat{S}_i^z \hat{S}_j^z \rangle - \langle \hat{S}_i^z \rangle \langle \hat{S}_j^z \rangle \qquad |i - j| = r \tag{4.16}$$

mediato sulle possibili coppie di siti, riportando l'andamento in figura 4.4. L'andamento è, come da attese, esponenziale, e dal fit sull'inviluppo si ottiene una stima della lunghezza di correlazione:

$$\xi \approx 3.82 \tag{4.17}$$

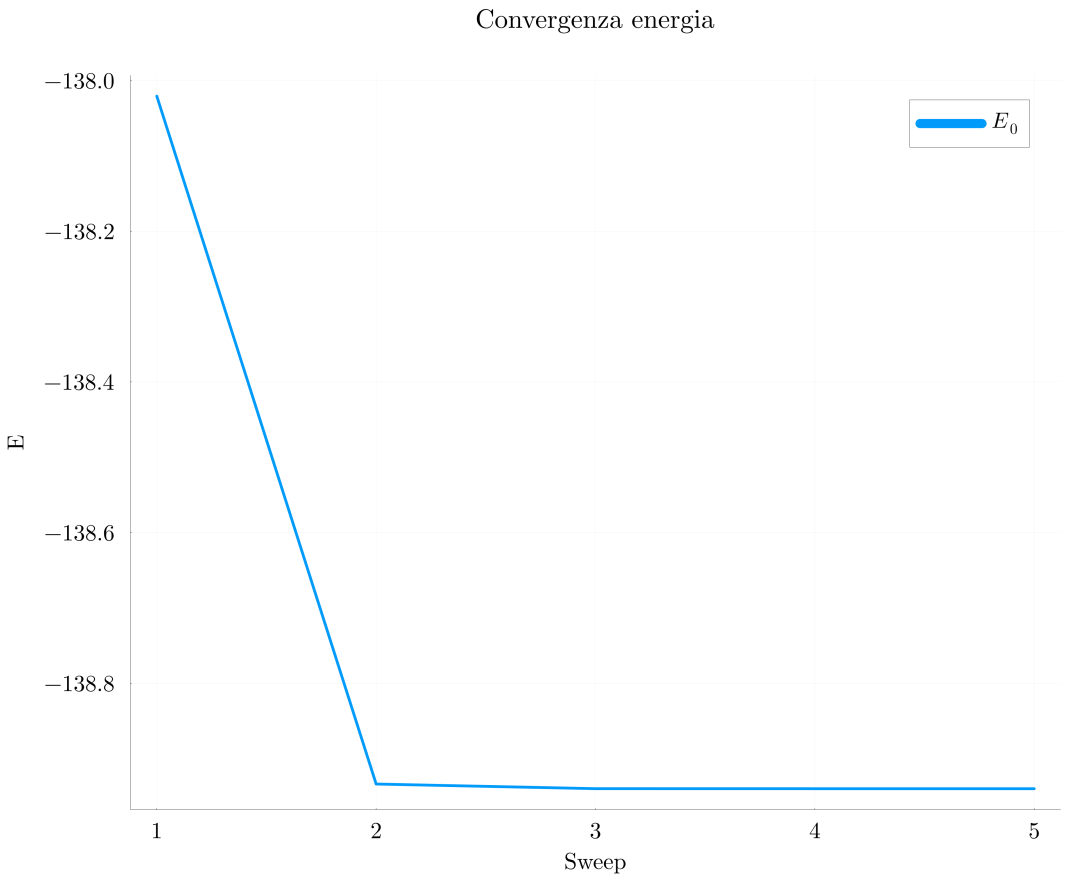


Figura 4.2: Convergenza in energia del DMRG

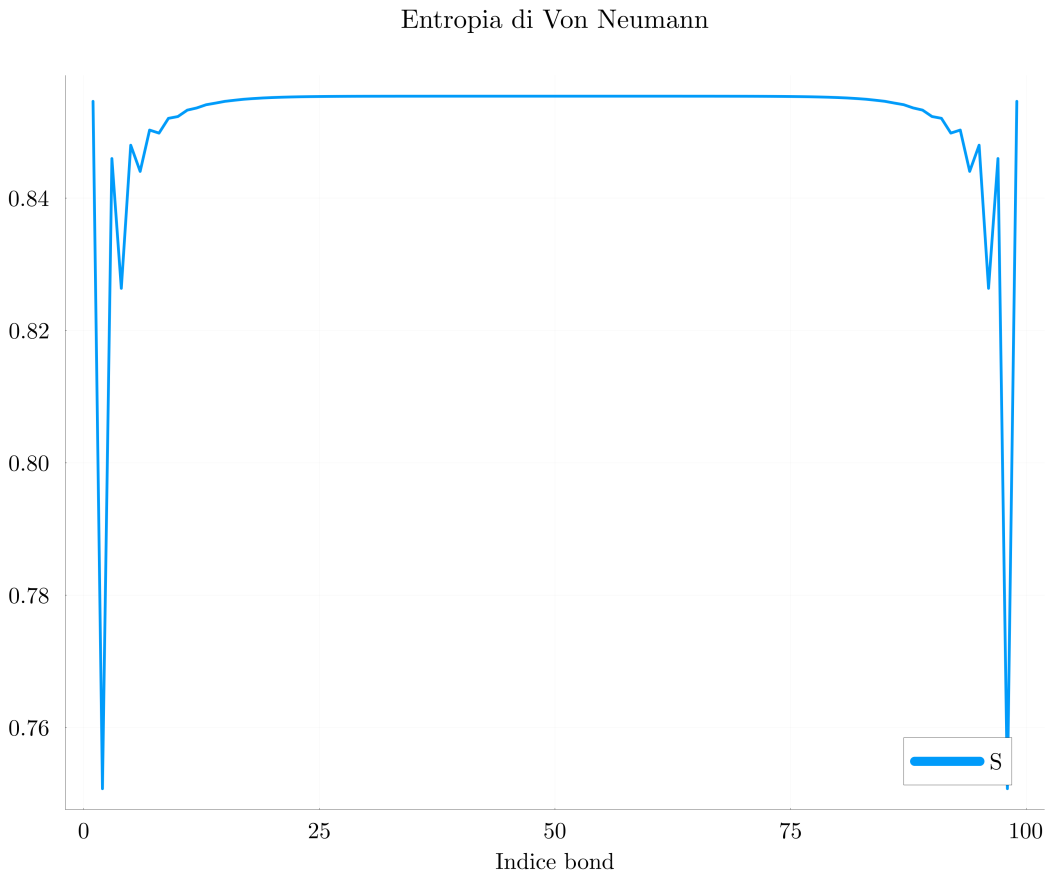


Figura 4.3: Entropia bipartita

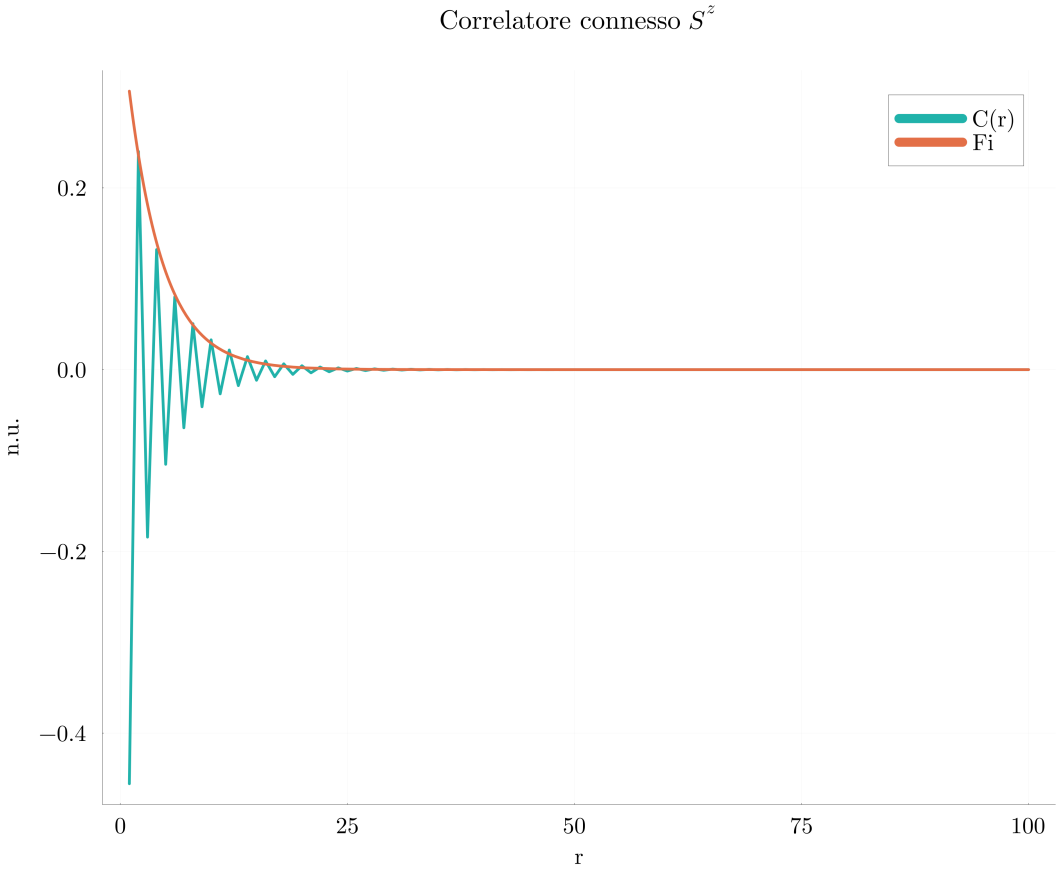


Figura 4.4: Andamento del correlatore

Conclusioni

Quello che si è illustrato è solamente l'esempio più elementare delle tecniche di Tensor Networks; come si è già detto nell'introduzione, i TN rappresentano un paradigma, un *framework* di estrema versatilità ed efficacia, che ha permesso sviluppi importanti su una vastissima gamma di problemi, tanto sul piano analitico quanto su quello numerico. Tra le varie applicazioni ulteriori si possono infatti menzionare: sempre per i sistemi a molti corpi, lo studio spettrale in dimensioni maggiore di 1, la determinazione degli stati termici, l'evoluzione temporale; nel quantum computing, la simulazione di circuiti e la progettazione di algoritmi; al di fuori dell'alveo della fisica, la parametrizzazione di modelli di Machine Learning. Specificamente, gli MPS possono essere generalizzati tramite i PEPS a dimensioni maggiori o i **Multiscale Entanglement Renormalization Ansatz** (**MERA**) ai sistemi critici, mente il DMRG propriamente detto si può estendere a sistemi termici e fuori equilibrio (rendendolo dipendente dal tempo), oppure applicare allo spazio dei momenti, con particolare utilità per i sistemi fermionici. Per la dinamica dei sistemi in particolare applicando gli stessi principi di decomposizione agli operatori di evoluzione si ottiene la **Time-Evolving Block Decimation** (**TEBD**) [7, 14].

Alla radice dell'efficacia dei TN vi è sicuramente l'intuitività del metodo diagrammatico, ma anche il dialogo proficuo che permettono di stabilire tra diversi ambiti di studio: si può dire in particolare che rappresentino uno degli elementi fondamentali dell'approccio *information-theoretic* alla fisica, in cui si incontrano la fisica degli stati condensati, quella delle interazioni fondamentali e la teoria dell'informazione quantistica. Questa generalità e fecondità dei TN suggeriscono che, lungi dal costituire una semplice rappresentazione pratica per operazioni di algebra lineare, questi catturino delle proprietà fondamentali comuni alle varie strutture matematiche che vi trovano una riformulazione naturale. Ciò è effettivamente vero, e per comprenderlo è necessario osservarli da un punto di vista estremamente astratto, che generalizza persino l'approccio dello stesso Penrose nel suo paper fondamentale: quello della **teoria delle categorie**. Questa considera classi estremamente generali di oggetti e mappe fra di essi, entro cui sono unificate strutture definite in varie branche della matematica - in particolare di quella applicata alla fisica, ed è formulabile in modo naturale tramite un linguaggio diagrammatico di cui i TN

costituiscono semplicemente un'istanza specifica [71].

Bibliografia

[1] Roger Penrose. "Applications of Negative Dimensional Tensors". In: (1971). URL: https://www.mscs.dal.ca/~selinger/papers/graphical-bib/public/Penrose-applications-of-negative-dimensional-tensors.pdf.

[2] Mario Collura et al. *Tensor Network Techniques for Quantum Computation*. arXiv:2503.04423 [quant-ph]. Dic. 2024. URL: http://arxiv.org/abs/2503.04423.

[3] Seymour Lipschutz. *Linear algebra*. eng. New York : McGraw-Hill, 2009. ISBN: 978-0-07-154352-1.

[4] Jacob Biamonte e Ville Bergholm. *Tensor Networks in a Nutshell*. en. arXiv:1708.00006 [quant-ph]. Lug. 2017. DOI: 10.48550/arXiv.1708.00006. URL: http://arxiv.org/abs/1708.00006.

[5] Roman Orus. "A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States". In: *Annals of Physics* 349 (ott. 2014). arXiv:1306.2164 [cond-mat], pp. 117–158. ISSN: 00034916. DOI: 10.1016/j.aop.2014.06.013. URL: http://arxiv.org/abs/1306.2164.

[6] Pietro Silvi et al. "The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems". In: *SciPost Phys. Lect. Notes* (mar. 2019). arXiv:1710.03733 [quant-ph], p. 8. ISSN: 2590-1990. DOI: 10.21468/SciPostPhysLectNotes.8. URL: http://arxiv.org/abs/1710.03733.

[7] Jacob C. Bridgeman e Christopher T. Chubb. "Hand-waving and Interpretive Dance: An Introductory Course on Tensor Networks". In: *J. Phys. A: Math. Theor.* 50.22 (giu. 2017). arXiv:1603.03039 [quant-ph], p. 223001. ISSN: 1751-8113, 1751-8121. DOI: 10.1088/1751-8121/aa6dc3. URL: http://arxiv.org/abs/1603.03039.

[8] Glen Evenbly. "A Practical Guide to the Numerical Implementation of Tensor Networks I: Contractions, Decompositions, and Gauge Freedom". en. In: *Front. Appl. Math. Stat.* 8 (giu. 2022), p. 806549. ISSN: 2297-4687. DOI: 10.3389/fams.2022.806549. URL: https://www.frontiersin.org/articles/10.3389/fams.2022.806549/full.

[9] Simone Montangero. *Introduction to Tensor Network Methods: Numerical simulations of low-dimensional many-body quantum systems*. en. Cham: Springer International Publishing, 2018. ISBN: 978-3-030-01408-7 978-3-030-01409-4. DOI: 10.1007/978-3-030-01409-4. URL: http://link.springer.com/10.1007/978-3-030-01409-4.

[10] Himeshi De Silva, John L. Gustafson e Weng-Fai Wong. "Making Strassen Matrix Multiplication Safe". en. In: *2018 IEEE 25th International Conference on High Performance Computing (HiPC)*. Bengaluru, India: IEEE, dic. 2018, pp. 173–182. ISBN: 978-1-5386-8386-6. DOI: 10.1109/HiPC.2018.00028. URL: https://ieeexplore.ieee.org/document/8638088/.

[11] Tom Kennedy e Slava Rychkov. "Tensor Renormalization Group at Low Temperatures: Discontinuity Fixed Point". en. In: *Ann. Henri Poincaré* 25.1 (gen. 2024). arXiv:2210.06669 [math-ph], pp. 773–841. ISSN: 1424-0637, 1424-0661. DOI: 10.1007/s00023-023-01289-y. URL: http://arxiv.org/abs/2210.06669.

[12] Shi-Ju Ran et al. *Lecture Notes of Tensor Network Contractions*. en. Vol. 964. arXiv:1708.09213 [physics]. 2020. DOI: 10.1007/978-3-030-34489-4. URL: http://arxiv.org/abs/1708.09213.

[13] Shi-Ju Ran et al. *Tensor Network Contractions: Methods and Applications to Quantum Many-Body Systems*. en. Lecture Notes in Physics. ISSN: 0075-8450, 1616-6361. Cham: Springer International Publishing, 2020. ISBN: 978-3-030-34488-7 978-3-030-34489-4. DOI: 10.1007/978-3-030-34489-4. URL: http://link.springer.com/10.1007/978-3-030-34489-4.

[14] G. Evenbly e G. Vidal. "Tensor network states and geometry". en. In: *J Stat Phys* 145.4 (nov. 2011). arXiv:1106.1082 [quant-ph], pp. 891–918. ISSN: 0022-4715, 1572-9613. DOI: 10.1007/s10955-011-0237-4. URL: http://arxiv.org/abs/1106.1082.

[15] Lam Chi-Chung, P. Sadayappan e Rephael Wenger. "On Optimizing a Class of Multi-Dimensional Loops with Reduction for Parallel Execution". In: *Parallel Process. Lett.* 07.02 (giu. 1997). Publisher: World Scientific Publishing Co., pp. 157–168. ISSN: 0129-6264. DOI: 10.1142/S0129626497000176. URL: https://www.worldscientific.com/doi/10.1142/S0129626497000176.

[16] Jianyu Xu et al. "Towards a polynomial algorithm for optimal contraction sequence of tensor networks from trees". en. In: *Phys. Rev. E* 100.4 (ott. 2019). Publisher: American Physical Society (APS). ISSN: 2470-0045, 2470-0053. DOI: 10.1103/physreve.100.043309. URL: https://link.aps.org/doi/10.1103/PhysRevE.100.043309.

[17] Jianyu Xu et al. *NP-Hardness of Tensor Network Contraction Ordering*. en. arXiv:2310.06140 [cs]. Ott. 2023. DOI: 10.48550/arXiv.2310.06140. URL: http://arxiv.org/abs/2310.06140.

[18] Carl Eckart e Gale Young. "The Approximation of One Matrix by Another of Lower Rank". en. In: *Psychometrika* 1.3 (set. 1936), pp. 211–218. ISSN: 0033-3123, 1860-0980. DOI: 10.1007/BF02288367. URL: https://www.cambridge.org/core/product/identifier/S0033312300051085/type/journal_article.

[19] Lloyd N. Trefethen. *Numerical linear algebra*. eng. Philadelphia, Pa: Society for Industrial e Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1997. ISBN: 978-0-89871-361-9 978-0-89871-957-4.

[20] P. A. M. Dirac. *I principi della meccanica quantistica*. ita. Rist. OCLC: 955652160. Torino: Bollati Boringhieri, 2009. ISBN: 978-88-339-5161-4.

[21] John Von Neumann. *Mathematical foundations of quantum mechanics*. eng. A cura di Nicholas A. Wheeler. New edition. Princeton: Princeton University Press, 2018. ISBN: 978-0-691-17856-1.

[22] Michael A. Nielsen e Isaac L. Chuang. *Quantum computation and quantum information*. 10th anniversary ed. Cambridge ; New York: Cambridge University Press, 2010. ISBN: 978-1-107-00217-3.

[23] Asher Peres. *Quantum theory: concepts and methods*. eng. Nachdr. Fundamental theories of physics vol. 72. Dordrecht: Kluwer Acad. Publ, 2010. ISBN: 978-0-7923-3632-7.

[24] Bruno Nachtergaele e Robert Sims. *An Introduction to Quantum Spin Systems*. 2016. URL: https://nextcloud.tfk.ph.tum.de/etn/wp-content/uploads/2022/09/JvN_lecture_notes_S2016_abcde-1.pdf.

[25] Man-Duen Choi. "Completely positive linear maps on complex matrices". In: *Linear Algebra and its Applications* 10.3 (giu. 1975), pp. 285–290. ISSN: 0024-3795. DOI: 10.1016/0024-3795(75)90075-0. URL: https://www.sciencedirect.com/science/article/pii/0024379575900750.

[26] W Bruzda e V Cappellini. *Institute of Physics, Jagiellonian University, Cracow, Poland*. en. Nov. 2012.

[27] E. Schrödinger. "Die gegenwärtige Situation in der Quantenmechanik". de. In: *Naturwissenschaften* 23.48 (nov. 1935), pp. 807–812. ISSN: 1432-1904. DOI: 10.1007/BF01491891. URL: https://doi.org/10.1007/BF01491891.

[28] V. Vedral. "The Role of Relative Entropy in Quantum Information Theory". en. In: *Rev. Mod. Phys.* 74.1 (mar. 2002). arXiv:quant-ph/0102094, pp. 197–234. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.74.197. URL: http://arxiv.org/abs/quant-ph/0102094.

[29] Alfréd Rényi. "On Measures of Entropy and Information". In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. Vol. 4.1. University of California Press, gen. 1961, pp. 547–562. URL: https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fourth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/On-Measures-of-Entropy-and-Information/bsmsp/1200512181.

[30] Martin Müller-Lennert et al. "On quantum Renyi entropies: a new generalization and some properties". en. In: *Journal of Mathematical Physics* 54.12 (dic. 2013). arXiv:1306.3142 [quant-ph], p. 122203. ISSN: 0022-2488, 1089-7658. DOI: 10.1063/1.4838856. URL: http://arxiv.org/abs/1306.3142.

[31] John Watrous. *The theory of quantum information*. eng. Cambridge New York Port Melbourne New Delhi Singapore: Cambridge University Press, 2018. ISBN: 978-1-316-84814-2. DOI: 10.1017/9781316848142.

[32] J. Eisert, M. Cramer e M. B. Plenio. "Area laws for the entanglement entropy - a review". en. In: *Rev. Mod. Phys.* 82.1 (feb. 2010). arXiv:0808.3773 [quant-ph], pp. 277–306. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.82.277. URL: http://arxiv.org/abs/0808.3773.

[33] Pieter Naaijkens. *Quantum spin systems on infinite lattices*. en. Vol. 933. arXiv:1311.2717 [math-ph]. 2017. DOI: 10.1007/978-3-319-51458-1. URL: http://arxiv.org/abs/1311.2717.

[34] Hal Tasaki. *The Lieb-Schultz-Mattis Theorem: A Topological Point of View*. arXiv:2202.06243 [cond-mat]. Ago. 2022. DOI: 10.48550/arXiv.2202.06243. URL: http://arxiv.org/abs/2202.06243.

[35] J. Eisert. *Entanglement and tensor network states*. arXiv:1308.3318 [quant-ph]. Set. 2013. DOI: 10.48550/arXiv.1308.3318. URL: http://arxiv.org/abs/1308.3318.

[36] Guglielmo Lami, Mario Collura e Nishan Ranabhat. *Beginner's Lecture Notes on Quantum Spin Chains, Exact Diagonalization and Tensor Networks*. en. arXiv:2503.03564 [cond-mat]. Mar. 2025. DOI: 10.48550/arXiv.2503.03564. URL: http://arxiv.org/abs/2503.03564.

[37] Stephen Blundell. "Advanced Quantum Mechanics for Condensed Matter - Lecture Notes". 2002. URL: https://users.ox.ac.uk/~phys1116/msm3.pdf.

[38] Matthias Troyer. "Computational Quantum Physics". en. In: (set. 2015).

[39] Rafael I. Nepomechie. "A SPIN CHAIN PRIMER". en. In: *Int. J. Mod. Phys. B* 13.24n25 (ott. 1999), pp. 2973–2985. ISSN: 0217-9792, 1793-6578. DOI: 10.1142/S0217979299002800. URL: https://www.worldscientific.com/doi/abs/10.1142/S0217979299002800.

[40] Anders W. Sandvik. "Computational Studies of Quantum Spin Systems". en. In: arXiv:1101.3281 [cond-mat]. 2010, pp. 135–338. DOI: 10.1063/1.3518900. URL: http://arxiv.org/abs/1101.3281.

[41] J. Ignacio Cirac et al. "Matrix product states and projected entangled pair states: Concepts, symmetries, theorems". In: *Rev. Mod. Phys.* 93.4 (dic. 2021). Publisher: American Physical Society, p. 045003. DOI: 10.1103/RevModPhys.93.045003. URL: https://link.aps.org/doi/10.1103/RevModPhys.93.045003.

[42] Angelo Lucia. *Locality and spectral gaps in quantum spin systems*. en. Dic. 2024.

[43] D. Perez-Garcia et al. *Matrix Product State Representations*. arXiv:quant-ph/0608197. Mag. 2007. DOI: 10.48550/arXiv.quant-ph/0608197. URL: http://arxiv.org/abs/quant-ph/0608197.

[44] David Poulin et al. "Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space". In: *Phys. Rev. Lett.* 106.17 (apr. 2011). Publisher: American Physical Society, p. 170501. DOI: 10.1103/PhysRevLett.106.170501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.170501.

[45] Norbert Schuch. *Condensed Matter Applications of Entanglement Theory*. arXiv:1306.5551 [quant-ph]. Giu. 2013. DOI: 10.48550/arXiv.1306.5551. URL: http://arxiv.org/abs/1306.5551.

[46] J. I. Cirac e F. Verstraete. "Renormalization and tensor product states in spin chains and lattices". en. In: *J. Phys. A: Math. Theor.* 42.50 (dic. 2009). arXiv:0910.1130 [cond-mat], p. 504004. ISSN: 1751-8113, 1751-8121. DOI: 10.1088/1751-8113/42/50/504004. URL: http://arxiv.org/abs/0910.1130.

[47] M. B. Hastings. "Lieb-Schultz-Mattis in Higher Dimensions". In: *Phys. Rev. B* 69.10 (mar. 2004). arXiv:cond-mat/0305505, p. 104431. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.69.104431. URL: http://arxiv.org/abs/cond-mat/0305505.

[48] Fangxin Li. "An Introduction to Entanglement Measures". en. In: (2020).

[49] Charles H. Bennett et al. "Mixed State Entanglement and Quantum Error Correction". In: *Phys. Rev. A* 54.5 (nov. 1996). arXiv:quant-ph/9604024, pp. 3824–3851. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.54.3824. URL: http://arxiv.org/abs/quant-ph/9604024.

[50] Norbert Schuch et al. "Entropy scaling and simulability by Matrix Product States". en. In: *Phys. Rev. Lett.* 100.3 (gen. 2008). arXiv:0705.0292 [quant-ph], p. 030504. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.100.030504. URL: http://arxiv.org/abs/0705.0292.

[51] F. Verstraete e J. I. Cirac. "Matrix product states represent ground states faithfully". In: *Phys. Rev. B* 73.9 (mar. 2006). arXiv:cond-mat/0505140, p. 094423. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.73.094423. URL: http://arxiv.org/abs/cond-mat/0505140.

[52] Roman Orus. "Advances on Tensor Network Theory: Symmetries, Fermions, Entanglement, and Holography". en. In: *Eur. Phys. J. B* 87.11 (nov. 2014). arXiv:1407.6552 [cond-mat], p. 280. ISSN: 1434-6028, 1434-6036. DOI: 10.1140/epjb/e2014-50502-9. URL: http://arxiv.org/abs/1407.6552.

[53] Tao Xiang. *Density matrix and tensor network renormalization*. eng. Cambridge New York Melbourne New Delhi Singapore: Cambridge University Press, 2023. ISBN: 978-1-009-39867-1 978-1-009-39870-1. DOI: 10.1017/9781009398671.

[54] Frank Verstraete et al. "Matrix Product Operators: Algebras and Applications". en. In: ().

[55] M. Fannes, B. Nachtergaele e R. F. Werner. "Finitely correlated states on quantum spin chains". en. In: *Commun.Math. Phys.* 144.3 (mar. 1992), pp. 443–490. ISSN: 1432-0916. DOI: 10.1007/BF02099178. URL: https://doi.org/10.1007/BF02099178.

[56] J Zinn-Justin. *RENORMALIZATION GROUP: AN INTRODUCTION*. en. 2009.

[57] Kenneth G. Wilson. "The renormalization group: Critical phenomena and the Kondo problem". en. In: *Rev. Mod. Phys.* 47.4 (ott. 1975), pp. 773–840. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.47.773. URL: https://link.aps.org/doi/10.1103/RevModPhys.47.773.

[58] Yu-An Chen e Hung-I Yang. *Density matrix renormalization group*. en. Mag. 2014.

[59] Leo P Kadanoff. *More is the Same Less is the Same, too; Mean Field Theories and Renormalization*. en.

[60] Theo Costi. "Wilson's numerical renormalization group". en. In: *Density-Matrix Renormalization*. A cura di Ingo Peschel et al. Berlin, Heidelberg: Springer, 1999, pp. 3–25. ISBN: 978-3-540-48750-0. DOI: 10.1007/BFb0106063.

[61] Eric Jeckelmann. *density-matrix renormalization group*. en.

[62] Ulrich Schollwoeck. "The density-matrix renormalization group". In: *Rev. Mod. Phys.* 77.1 (apr. 2005). arXiv:cond-mat/0409292, pp. 259–315. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.77.259. URL: http://arxiv.org/abs/cond-mat/0409292.

[63] Steven R. White. "Density matrix formulation for quantum renormalization groups". en. In: *Phys. Rev. Lett.* 69.19 (nov. 1992), pp. 2863–2866. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.69.2863. URL: https://link.aps.org/doi/10.1103/PhysRevLett.69.2863.

[64] Ulrich Schollw?ck. "The density-matrix renormalization group: a short introduction". In: *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 369.1946 (2011). Publisher: The Royal Society, pp. 2643–2661. ISSN: 1364-503X. URL: https://www.jstor.org/stable/23035851.

[65] A. Weichselbaum et al. "Variational matrix product state approach to quantum impurity models". en. In: *Phys. Rev. B* 80.16 (ott. 2009). arXiv:cond-mat/0504305, p. 165117. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.80.165117. URL: http://arxiv.org/abs/cond-mat/0504305.

[66] F. Verstraete, D. Porras e J. I. Cirac. "Density Matrix Renormalization Group and Periodic Boundary Conditions: A Quantum Information Perspective". en. In: *Phys. Rev. Lett.* 93.22 (nov. 2004), p. 227205. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.93.227205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.227205.

[67] T. Huckle, K. Waldherr e T. Schulte-Herbrueggen. "Computations in Quantum Tensor Networks". In: *Linear Algebra and its Applications* 438.2 (gen. 2013). arXiv:1212.5005 [quant-ph], pp. 750–781. ISSN: 00243795. DOI: 10.1016/j.laa.2011.12.019. URL: http://arxiv.org/abs/1212.5005.

[68] J. Eisert. "Computational Difficulty of Global Variations in the Density Matrix Renormalization Group". en. In: *Phys. Rev. Lett.* 97.26 (dic. 2006). arXiv:quant-ph/0609051, p. 260501. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.97.260501. URL: http://arxiv.org/abs/quant-ph/0609051.

[69] Matthew Fishman, Steven White e Edwin Miles Stoudenmire. "The ITensor Software Library for Tensor Network Calculations". en. In: *SciPost Physics Codebases* (ago. 2022), p. 004. ISSN: 2949-804X. DOI: 10.21468/SciPostPhysCodeb.4. URL: https://www.scipost.org/SciPostPhysCodeb.4.

[70] *The Julia Language*. Lug. 2025. URL: https://raw.githubusercontent.com/JuliaLang/docs.julialang.org/assets/julia-1.11.6.pdf.

[71] John C. Baez e Aaron Lauda. "A Prehistory of n-Categorical Physics". en. In: arXiv:0908.2469 [hep-th]. Apr. 2011, pp. 13–128. DOI: 10.1017/CBO9780511976971.003. URL: http://arxiv.org/abs/0908.2469.

[72] Jacob Biamonte. *Lectures on Quantum Tensor Networks*. arXiv:1912.10049 [quant-ph]. Gen. 2020. DOI: 10.48550/arXiv.1912.10049. URL: http://arxiv.org/abs/1912.10049.

[73] Aleksej Ju Kitaev, Aleksandr Ch Šen e Michail N. Vjalyj. *Classical and quantum computation*. eng. Trad. da Lester J. Senechal. Graduate studies in mathematics volume 47. Providence, Rhode Island: American Mathematical Society, 2002. ISBN: 978-0-8218-3229-5 978-1-4704-1800-7.

Appendice A

Complessità computazionale

La teoria della complessità computazionale si occupa di classificare i problemi in base alla loro difficoltà, quantificata in termini delle risorse necessarie per determinare una soluzione. Piuttosto che considerare specifiche istanze, si studiano tipologie generali di problemi e la dipendenza delle quantità richieste per la risoluzione (nel caso peggiore) dalla lunghezza dell'input di definizione; sono così definite delle **classi** di problemi di difficoltà simile, tra cui si possono anche stabilire delle relazioni di *riducibilità* di grande importanza.
Di seguito si specificheranno le misure di efficienza utilizzate in questo lavoro e si definiranno le principali classi di complessità. Ove non diversamente specificato, si può fare riferimento al Nielsen-Chuang [22].

A.1 Complessità computazionale

Per dare una nozione precisa di efficienza si classificano le possibile dipendenze funzionali della **complessità computazionale**:

Definizione A.1. *La complessità computazionale di un algoritmo è la quantità di risorse computazionali $f(n)$ necessarie per eseguirlo su un input di lunghezza n*[1]*. La complessità di un problema è la complessità minima degli algoritmi che lo risolvono.*

Come detto:

- $f(n)$ corrisponde al caso peggiore per un dato n, definito come estremo superiore per i valori sui possibili input di lunghezza determinata;

[1] Può definire equivalentemente lo specifico valore o la funzione; di seguito si assume il secondo significato.

- Interessa l'andamento asintotico, ovvero nel limite $n \to \infty$. Si utilizza pertanto la notazione "O-grande", definita secondo:

$$f(n) = \mathcal{O}(g(n)) \Leftrightarrow \exists C > 0, n_0 \geq 0 \, : \, \forall n \geq n_0 \, f(n) \leq Cg(n) \tag{A.1}$$

  A parole si dice che f è "dell'ordine" di g.

In generale la complessità si classifica come:

- **Polinomiale** quando $f(n)$ è un polinomio di grado finito k, ovvero più semplicemente asintoticamente $f(n) = \mathcal{O}(n^k)$ (se $k = 0$ in particolare f è costante);
- **Esponenziale** se $f(n) = \mathcal{O}(\alpha^n)$ per qualche $\alpha > 1$.

Si noti che la presenza di eventuali correzioni logaritmiche non incide sulla classificazione. Si può ora formalizzare l'efficienza:

Definizione A.2. *Un algoritmo è efficiente se ha complessità polinomiale.*

A.2 Risorse

Le principali risorse considerate per definire la complessità sono universali e prescindono dallo specifico supporto fisico della computazione in quanto in si rifanno al modello teorico della macchina di Turing. Esse sono:

- Il **tempo**, ovvero equivalentemente il numero di operazioni;
- Lo **spazio** in memoria.

Si noti che per definire l'efficienza delle manipolazioni sui TN si introduce anche il **costo**, definito come numero di FLOPs, che è una misura altrettanto valida ove non sia rilevante considerare anche l'overhead della manipolazione degli indici e in generale delle operazioni di ricollocamento in memoria, che si basano sull'aritmetica degli interi. É importante avere presente questa distinzione per capire come gli algoritmi di moltiplicazione matriciale ottimizzati possano avere lo stesso costo dei cicli nidificati ma essere più efficienti.

A.3 Classi di complessità

Le principali classi di complessità sono definite rispetto al *tempo* per algoritmi **deterministici**, ossia che producono sempre il medesimo output per un dato input, che risolvono problemi **di decisione**, ovvero che possono essere formulati come una serie di domande binarie sull'input:

Definizione A.3. *Si definiscono le seguenti classi:*

- ***P*** *è la classe di problemi che ammettono una soluzione efficiente;*
- ***NP*** *è la classe di problemi che ammettono soluzioni le quali possono essere verificate efficientemente.*

Chiaramente P è incluso in NP, mentre l'inclusione inversa (e quindi l'uguaglianza tra le classi) resta il principale problema aperto della teoria della complessità computazionale. Dunque in generale è possibile che esistano problemi non risolvibili efficientemente, anche se è difficile dimostrare in generale l'inammissibilità di un algoritmo efficiente: questi sono i problemi considerati "difficili". In particolare è d'interesse la sottoclasse dei problemi **NP-completi**:

Definizione A.4. *Un problema è NP-completo se:*

1. *É in NP;*
2. *Qualsiasi algoritmo che lo risolve può essere adattato con un overhead al più polinomiale per risolvere qualsiasi altro problema in NP.*

La classe NP-completo può essere anche definita come l'intersezione tra la classe NP è la classe **NP-hard**, ottenuta semplicemente rimuovendo l'assioma 1:

Definizione A.5. *Un problema è NP-hard se qualsiasi algoritmo che lo risolve può essere adattato con un overhead al più polinomiale per risolvere qualsiasi altro problema in NP.*

Gli NP-hard, come indica il nome, possono essere pertanto considerati ancora più ardui degli NP-completi in quanto non è neppure noto se ammettano la verifica efficiente delle soluzioni.
Si possono naturalmente definire classi anche rispetto allo *spazio*, anche se hanno minore utilizzo; in particolare **PSPACE** è l'equivalente di P (si ritiene che includa sia P che NP). Le definizioni si possono poi estendere al caso di algoritmi probabilistici e di macchine di Turing quantistiche; per le seconde si hanno in particolare due classi in diretta corrispondenza con il caso classico deterministico:

Definizione A.6. *Per problemi risolvibili da algoritmi quantistici (ovvero eseguiti su computer quantistici) sono definite le seguenti classi:*

- ***Bounded-error Quantum P, BQP*** *(analoga di P) è la classe dei problemi risolvibili efficientemente con probabilità almeno* $2/3$*;*
- ***Quantum Merlin Arthur, QMA*** *(analoga di NP) è la classe dei problemi con soluzioni verificabili efficientemente con la medesima probabilità.*

La relazione con le classi classiche non è del tutto definita, ma è noto che BQP include P ed è incluso in PSPACE. É interessante notare che i principali problemi considerati in questo testo rimangono difficili anche per macchine quantistiche:

- La contrazione anche solo approssimata di un TN è in generale QMA-hard. Un modo interessante di interpretare questa proprietà è considerare che è possibile mappare nella contrazione di un TN opportuno il problema della colorazione di un grafo, un noto problema hard [7, 72];
- Il problema della determinazione del ground state di un'hamiltoniana k-locale generica (detto *Problema dell'hamiltoniana locale*) è QMA-completo per $k \geq 2$. Se ridotto ad un problema classico è NP-completo [73, 7].

Appendice B

Codice simulazione

Si riporta di seguito il codice utilizzato per l'esperimento numerico.

```
# PACCHETTI

using ITensors, ITensorMPS, Dates, LaTeXStrings, Measures, Peaks, LsqFit
import Plots
const plt = Plots

# Setup generale dei plot
plt.gr()
plt.default(fontfamily="Computer Modern",
    titlefont=plt.font(50, "Computer Modern"),
    guidefont=plt.font(40, "Computer Modern"),
    tickfont=plt.font(40, "Computer Modern"),
    legendfont=plt.font(40, "Computer Modern"),
)

let
    # CONDIZIONI INIZIALI

    # Spin chain
    N = 100
    chain = siteinds("S=1",N)

    # MPO hamiltoniana
    ham = OpSum()
    for j=1:N-1
        ham += 0.5,"S+",j,"S-",j+1
        ham += 0.5,"S-",j,"S+",j+1
        ham += "Sz",j,"Sz",j+1
    end
    H = MPO(ham,chain)
```

```
# Parametri DMRG
nsweeps = 5 # sweeps
maxdim = [10,20,100,100,200] # soglia di decimazione
cutoff = [1E-10] # soglia di troncamento

# Ansatz
psi_0 = randomMPS(chain,2)

# Observer per estrarre dati durante l'esecuzione
obs = DMRGObserver()

# ESECUZIONE ALGORITMO

# Sono prodotti energia e ground state finale. Il livello di output
↪  specifica la verbosità delle informazioni riportate ne terminale
energy,psi = @time dmrg(H,psi_0;nsweeps,maxdim,cutoff,observer=obs,
↪  outputlevel = 1)

# ANALISI

# Estrazione energie per vari sweep
energies = ITensorMPS.energies(obs)

# Calcolo correlatore connesso
exp_Sz = expect(psi, "Sz")
corr_Sz = correlation_matrix(psi, "Sz", "Sz")
corr = zeros(N)
for r=1:N
    for i=1:(N-r)
        corr[r] += (corr_Sz[i,i+r] - exp_Sz[i]*exp_Sz[i+r]) / (N-r)
    end
end

# Calcolo entropia di entanglement
ent = zeros(N-1)
for b=1:N-1
    psi = orthogonalize(psi, b)
    U,S,V = svd(psi[b], (linkinds(psi, b-1), siteinds(psi, b)))
    for n=1:dim(S, 1)
        p = S[n,n]^2
        ent[b] -= p * log(p)
    end
    if b == N/2
        println("Entropia di bulk: ", ent[b])
    end
end
```

```
    # Fit esponenziale
    positions, values = findmaxima(corr)
    @. model(r,par) = par[1]*exp(-r*par[2])
    par0 = [1.,10.]
    fit = curve_fit(model, positions, values, par0)
    corf(r) = fit.param[1]exp(-r*fit.param[2])
    println("Lunghezza di correlazione: ", 1/fit.param[2])

    # Data per indicizzazione dei grafici
    now_str = Dates.format(now(), "yyyy-mm-dd_HHMMSS")

    # GRAFICI

    conv = plt.plot(1:length(energies), energies, size=(4000,3200),
    ↪  title="Convergenza energia" , margin=40mm, w=10, xlabel="Sweep",
    ↪  ylabel="E", label=L"$E_0$")
    plt.savefig("fig/conv_" * now_str * ".png")

    corr = plt.plot([1:N], corr, size=(4000,3200), title=L"Correlatore
    ↪  connesso $S^z$" , margin=40mm, w=10, xlabel="r", label="C(r)",
    ↪  ylabel="n.u.", color=:lightseagreen)
    plt.plot!(r -> corf(r), 1, N, w=10, label="Fi")
    plt.savefig("fig/corr_" * now_str * ".png")

    entr = plt.plot([1:N-1], ent, size=(4000,3200), title="Entropia di
    ↪  Von Neumann" , margin=40mm, w=10, xlabel="Indice bond",
    ↪  label="S")
    plt.savefig("fig/entr_" * now_str * ".png")

    return
end
```