La geometria della Computazione Quantistica

Oscar Papini

29 luglio 2013

Sommario

Trovare algoritmi quantistici efficienti ha un'indubbia importanza, sia praticamente, che da un punto di vista teorico: avere dei parametri che misurino l'efficienza dell'algoritmo, e delle limitazioni su di esse, ci permette di sapere dove possiamo arrivare e quanto avanti possiamo spingerci. In questo seminario esporremo il lavoro di Nielsen *et al.* [3], analizzando in che modo abbiano ricondotto lo studio dell'efficienza di un algoritmo quantistico a problemi di geometria differenziale, e aprendo la strada all'uso di tali tecniche per valutare la potenza e i limiti dei calcolatori quantistici.

1 Panoramica della situazione

Già da qualche decennio si sta intravedendo la possibilità di sfruttare le peculiarità della Meccanica Quantistica (sovrapposizione di stati, *entanglement*) per aumentare le potenze di calcolo dei computer. Tra gli esempi più notevoli, ricordiamo l'algoritmo di Shor per la fattorizzazione di interi, che, se opportunamente implementato su un computer quantistico, permetterebbe di rompere alcuni tra i più diffusi sistemi crittografici, basati appunto sulla difficoltà del problema di fattorizzare.

Un algoritmo quantistico è solitamente descritto da una successione di porte quantistiche (*quantum gates*), che implementa un operatore unitario U. Eseguire l'algoritmo su un dato input corrisponde ad applicare U al vettore che descrive lo stato di input. L'efficienza di un algoritmo può essere misurata dal numero di porte necessarie alla sua implementazione.

Definizione 1.1. Diciamo che un algoritmo quantistico è *efficiente* se il numero di porte necessarie alla sua implementazione è $\mathcal{O}(n^k)$ per qualche $k \geqslant 1$, dove n rappresenta la dimensione del problema risolto dall'algoritmo (ad esempio, la quantità di bit necessaria per rappresentare il numero da fattorizzare).

Ma come stimare l'efficienza di un algoritmo? L'idea è quella di vedere U come punto di arrivo di un'evoluzione generata da una qualche hamiltoniana di controllo dipendente dal tempo $H(t)$ tramite l'equazione di Schrödinger

$$\frac{dU(t)}{dt} = -iH(t)U(t) \tag{1.1}$$

dove supporremo che, per un opportuno tempo finale t_f, $U(t_f) = U$.

A questo punto, vogliamo definire una funzione di *costo* F, che associ ad un'hamiltoniana H un numero reale $F[H]$. Tramite F sarà possibile definire una distanza sullo spazio $SU(2^n)$ degli operatori unitari su n qubit con determinante 1. Infatti, se $U(t) : [0, t_f] \to SU(2^n)$ è la curva ottenuta da $H(t)$ tramite l'equazione (1.1), è possibile definire la sua *lunghezza* come

$$\ell(U) := \int_0^{t_f} F[H(t)]dt.$$

Questa lunghezza è invariante rispetto a diverse parametrizzazioni della curva $U(t)$: possiamo allora riscalare l'hamiltoniana in modo che $F[H] = 1$, e quindi $\ell(U) = t_f$. Assumeremo d'ora in poi di lavorare con curve normalizzate in questo senso.

Infine, la nozione di lunghezza ci permette di definire la distanza tra l'identità I e l'operatore sintetizzato U come

$$d(I, U) := \inf_{U(t)} \{\ell(U(t))\}.$$

Per scegliere un'opportuna funzione di costo, in primo luogo scriviamo l'hamiltoniana in termini dell'espansione in operatori di Pauli:

$$H = \sum_{\sigma} h_\sigma \sigma,$$

dove gli h_σ sono coefficienti reali e $\sigma = \alpha_1 \otimes \cdots \otimes \alpha_n$ varia su tutti gli n-prodotti tensori tra le matrici di Pauli σ_x, σ_y e σ_z e l'identità I. Separiamo la sommatoria in due parti:

$$H = P(H) + Q(H)$$

dove in $P(H)$ compaiono i termini σ a uno e due corpi, ovvero quelli per cui tutti gli α_i sono I tranne al più due, mentre in $Q(H)$ ci sono i termini σ a tre o più corpi, in cui gli α_i diversi da I sono almeno tre. Scriveremo, con un piccolo abuso di notazione, "$\sigma \in P$" per indicare che il termine σ compare tra i termini di $P(H)$, e analogamente per $Q(H)$.

Definiamo quindi il costo per un'hamiltoniana H come

$$F[H] := \sqrt{\sum_{\sigma \in P} h_\sigma^2 + p^2 \sum_{\sigma \in Q} h_\sigma^2}$$

con p parametro che penalizzi i termini a tre o più corpi, che verrà scelto in modo da poter sopprimere tali termini.

Notiamo che una tale F può essere pensata come norma associata a una metrica riemanniana il cui tensore metrico ha componenti

$$g_{\sigma\tau} = \begin{cases} 0 & \text{se } \sigma \neq \tau \\ 1 & \text{se } \sigma = \tau \text{ e } \sigma \in P \\ p^2 & \text{se } \sigma = \tau \text{ e } \sigma \in Q. \end{cases} \tag{1.2}$$

2 Approssimazioni successive

Il risultato a cui vogliamo arrivare è il seguente: dato U, esiste un circuito quantistico che contiene un numero di porte polinomiale in $d(I, U)$, che approssimi U con grande precisione.

Il primo passo consiste nel mostrare che l'errore che nasce trascurando la parte $Q(H)$ dell'hamiltoniana può essere reso piccolo scegliendo opportunamente il parametro p.

Lemma 2.1. *Data un'hamiltoniana $H(t)$ che generi U, sia $H_P := P(H)$ come definita sopra, e sia U_P l'operatore unitario generato da $H_P(t)$. Allora*

$$\|U - U_P\| \leqslant \frac{2^n d(I, U)}{p}$$

dove $\|\cdot\|$ indica la norma operatoriale.

Dimostrazione. Innanzitutto notiamo che se H contiene solo termini a tre o più corpi, allora $F[H] = p\|H\|_2$, dove $\|\cdot\|_2$ indica la norma euclidea rispetto ai coefficienti dell'espansione di H in operatori di Pauli. Inoltre, per ogni H,

$$\|H\| = \left\| \sum_\sigma h_\sigma \sigma \right\| \leqslant \sum_\sigma |h_\sigma| \leqslant 2^n \|H\|_2$$

dove l'ultima è una disuguaglianza di Cauchy-Schwarz.

Infine, applicando la disuguaglianza triangolare e l'invarianza della norma operatoriale rispetto agli unitari, si ha che, se $H(t)$ e $K(t)$ sono hamiltoniane che generano al tempo t_f rispettivamente gli unitari U e V, allora

$$\|U - V\| \leqslant \int_0^{t_f} \|H(t) - K(t)\| dt.$$

Combinando questi risultati si ottiene

$$\begin{aligned}\ell(U) &= \int_0^{t_f} F[H(t)]dt \\ &\geqslant \int_0^{t_f} F[H(t) - H_P(t)]dt = \int_0^{t_f} p\|H - H_P\|_2 dt \\ &\geqslant \frac{p}{2^n}\int_0^{t_f} \|H - H_P\|dt \geqslant \frac{p}{2^n}\|U - U_P\|.\end{aligned}$$

Passando all'estremo inferiore si ottiene la tesi. □

Quindi, scegliendo p sufficientemente grande, ad esempio 4^n, ci possiamo assicurare che

$$\|U - U_P\| \leqslant \frac{d(I,U)}{2^n}.$$

Cerchiamo allora di sintetizzare U_P. Per fare ciò, spezziamo l'intervallo $[0, t_f]$, in cui agisce l'evoluzione tramite $H_P(t)$, in tanti piccoli intervalli, diciamo di lunghezza Δ. Il prossimo lemma ci mostra che su ciascun intervallino l'hamiltoniana $H_P(t)$ (dipendente dal tempo) può essere simulata con grande precisione da un'hamiltoniana media costante nel tempo, che indicheremo con $\overline{H_P}^{\Delta}$.

Lemma 2.2. *Sia U un operatore unitario generato da un'hamiltoniana $H(t)$, dipendente dal tempo, nell'intervallo $[0, \Delta]$. Supponiamo che $\|H(t)\| \leqslant c$ per una qualche costante $c > 0$. Detta*

$$\overline{H} := \frac{1}{\Delta}\int_0^{\Delta} H(t)dt$$

l'hamiltoniana media, si ha che

$$\|U - \exp(-i\overline{H}\Delta)\| \leqslant 2(e^{c\Delta} - 1 - c\Delta) \in \mathcal{O}(c^2\Delta^2).$$

Dimostrazione. U si può esprimere in *serie di Dyson* come

$$U = \sum_{m=0}^{\infty}(-i)^m \int_0^{\Delta}\int_0^{t_1}\dots\int_0^{t_{m-1}} H(t_1)H(t_2)\dots H(t_m)dt_m\dots dt_2 dt_1.$$

Scrivendo $\exp(-i\overline{H}\Delta)$ in serie di potenze, cancellando i termini $\mathcal{O}(1)$ e $\mathcal{O}(\Delta)$, e dalla disuguaglianza triangolare, si ha

$$\begin{aligned}&\|U - \exp(-i\overline{H}\Delta)\| \leqslant \\ &\leqslant \sum_{m=2}^{\infty}\frac{\|(-i\overline{H}\Delta)^m\|}{m!} + \int_0^{\Delta}\int_0^{t_1}\dots\int_0^{t_{m-1}}\|H(t_1)H(t_2)\dots H(t_m)\|dt_m\dots dt_2 dt_1 \\ &\leqslant 2\sum_{m=2}^{\infty}\frac{c^m\Delta^m}{m!} = 2(e^{c\Delta} - 1 - c\Delta),\end{aligned}$$

dove nella seconda riga abbiamo usato che $\|XY\| \leqslant \|X\|\|Y\|$, che $\|H(t)\| \leqslant c$ e che

$$\int_0^{\Delta} \int_0^{t_1} \dots \int_0^{t_{m-1}} dt_m \dots dt_2 dt_1 = \frac{\Delta^m}{m!}.$$

Questo è il risultato cercato. □

Per poter applicare il lemma osserviamo che ci sono al più $3n$ termini a un corpo, e al più $9n(n-1)/2$ termini a due corpi; quindi

$$\|H_P(t)\| \leqslant \left\| \sum_{\sigma \in P} h_\sigma \sigma \right\| \leqslant \sum_{\sigma \in P} |h_\sigma| \leqslant \frac{3}{\sqrt{2}} n \|H_P(t)\|_2 = \frac{3}{\sqrt{2}} n F[H_P].$$

In più, osservando che $F[H] = 1$ implica $F[H_P] \leqslant 1$, abbiamo

$$\|H_P(t)\| \leqslant \frac{3}{\sqrt{2}} n.$$

In particolare, su un intervallo ampio Δ, abbiamo

$$\|U_P^{\Delta} - \exp\left(-i\overline{H_P}^{\Delta}\Delta\right)\| \leqslant 2\left(e^{\frac{3}{\sqrt{2}}n\Delta} - 1 - \frac{3}{\sqrt{2}} n\Delta\right) \in \mathcal{O}(n^2\Delta^2)$$

dove U_P^{Δ} è l'unitario generato da $H_P(t)$ su un intervallo Δ.

L'ultimo lemma ci garantisce che un unitario generato da un'hamiltoniana indipendente dal tempo e contenente solo termini a uno e due corpi può essere implementato con un numero non troppo grande di porte quantistiche.

Lemma 2.3. *Sia H un'hamiltoniana indipendente dal tempo e contenente solo termini a uno e due corpi, tale che i suoi coefficienti dell'espansione in operatori di Pauli soddisfino $|h_\sigma| \leqslant 1$. Allora esiste un unitario U_A tale che*

$$\|e^{-iH\Delta} - U_A\| \leqslant c_2 n^4 \Delta^3$$

e che può essere implementato usando al più $c_1 n^2/\Delta$ porte a uno o due qubit, dove c_1 e c_2 sono costanti.

Dimostrazione. Dividiamo l'intervallo $[0, \Delta]$ in $N := 1/\Delta$ passi di taglia Δ^2. Se

$$H = \sum_{j=1}^{L} h_j \sigma_j$$

è la scrittura di H in termini di operatori di Pauli, con $L \in \mathcal{O}(n^2)$, definiamo

$$U_{\Delta^2} := e^{-ih_1\sigma_1\Delta^2} e^{-ih_2\sigma_2\Delta^2} \dots e^{-ih_L\sigma_L\Delta^2}.$$

Ciascun fattore del prodotto precedente è un porta a uno o due qubit. Sapendo che per A, B operatori hermitiani vale che

$$e^{i(A+B)\Delta^2} = e^{iA\Delta^2} e^{iB\Delta^2} + \mathcal{O}(\Delta^4)$$

(vedi, per esempio, [2, pag. 208]), si ottiene facilmente che

$$U_{\Delta^2} = e^{-iH\Delta^2} + \mathcal{O}(L^2\Delta^4).$$

Applicando la disuguaglianza triangolare per N passi, e sfruttando l'invarianza della norma operatoriale rispetto agli unitari, abbiamo

$$\begin{aligned}\|e^{-iH\Delta} - U_{\Delta^2}^N\| &\leqslant \|e^{-iH\Delta} - U_{\Delta^2}\| + \|U_{\Delta^2} - U_{\Delta^2}^2\| + \cdots + \|U_{\Delta^2}^{N-1} - U_{\Delta^2}^N\| \\ &\leqslant \widetilde{c_2} N L^2 \Delta^4 \leqslant c_2 n^4 \Delta^3\end{aligned}$$

con c_2 costante. Di conseguenza, è possibile approssimare $e^{-iH\Delta}$ con $NL = c_1 n^2/\Delta$ porte, dove c_1 è una costante. □

Vediamo come combinare i lemmi precedenti. Partiamo da un'hamiltoniana normalizzata e dipendente dal tempo $H(t)$, e supponiamo che tale hamiltoniana generi l'unitario U tramite una curva $U(t)$ di lunghezza minima $d(I, U)$. Sia $H_P(t)$ la corrispondente hamiltoniana, ottenuta come nel lemma 2.1, che generi U_P tale che

$$\|U - U_P\| \leqslant \frac{d(I, U)}{2^n}.$$

Dividiamo l'intervallo $[0, d(I, U)]$ in N intervalli ciascuno ampio $\Delta := d(I, U)/N$, con N scelto successivamente. Siano ora $U_P^{(j)}$ l'unitario generato da $H_P(t)$ durante il j-esimo intervallo, e $U_M^{(j)}$ l'unitario generato dalla corrispondente hamiltoniana media $\overline{H_P}$ sullo stesso intervallo. Per il lemma 2.2

$$\left\| U_P^{(j)} - U_M^{(j)} \right\| \leqslant 2\left(e^{\frac{3}{\sqrt{2}} n\Delta} - 1 - \frac{3}{\sqrt{2}} n\Delta \right).$$

Per il lemma 2.3, invece, possiamo ottenere un unitario $U_A^{(j)}$ tale che $\|U_M^{(j)} - U_A^{(j)}\| \leqslant c_2 n^4 \Delta^3$, usando al più $c_1 n^2/\Delta$ porte. In conclusione, sfruttando ripetutamente la disuguaglianza triangolare, otteniamo

$$\begin{aligned}\|U - U_A\| &\leqslant \|U - U_P\| + \|U_P - U_A\| \\ &\leqslant \frac{d(I, U)}{2^n} + \sum_{j=1}^{N} \|U_P^{(j)} - U_A^{(j)}\| \\ &\leqslant \frac{d(I, U)}{2^n} + \sum_{j=1}^{N} \|U_P^{(j)} - U_M^{(j)}\| + \|U_M^{(j)} - U_A^{(j)}\| \\ &\leqslant \frac{d(I, U)}{2^n} + 2\frac{d(I, U)}{\Delta}\left(e^{\frac{3}{\sqrt{2}} n\Delta} - 1 - \frac{3}{\sqrt{2}} n\Delta \right) + c_2 d(I, U) n^4 \Delta^2.\end{aligned}$$

A patto di scegliere $\Delta \sim 1/(n^2 d(I, U))$, possiamo assicurarci che $\|U - U_A\|$ sia piccola. Notiamo infine che il numero di porte richieste, in accordo con il lemma 2.3, è $\mathcal{O}(Nn^2/\Delta) = \mathcal{O}(n^6 d(I, U)^3)$.

3 Ulteriori sviluppi

Come possiamo scegliere allora l'hamiltoniana di controllo $H(t)$ in modo che generi U tramite un'evoluzione ottimale, cioè tale che la curva abbia effettivamente lunghezza $d(I, U)$?

Abbiamo visto che il costo definisce una metrica riemanniana tramite (1.2). Ebbene, l'evoluzione ottimale corrisponde al calcolo di una geodetica su $SU(2^n)$ dotato di questa metrica.

A questo punto, sarebbe interessante una classificazione delle geodetiche in questo spazio: questo ci fornirebbe informazioni importanti sull'effettiva potenza del calcolo quantistico. Tale classificazione esula dagli scopi di questo seminario; un'analisi più dettagliata, comprendente la costruzione di un'ampia classe di queste geodetiche, può essere trovata in [1].

Riferimenti bibliografici

[1] Michael A. Nielsen. A Geometric Approach to Quantum Circuit Lower Bounds. *Quantum Info. Comput.*, 6(3):213–262, Maggio 2006.

[2] Michael A. Nielsen e Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[3] Michael A. Nielsen, Mark R. Dowling, Mile Gu, e Andrew C. Doherty. Quantum Computation as Geometry. *Science*, 311(5764):1133–1135, 2006.