



Algoritmos Quânticos em Geometria Molecular

Andrês Rodrigues Oliveira - IMECC/Unicamp

Carlile Lavor (Orientador) - IMECC/Unicamp

Palavras-chave: Computação Quântica, Oráculo, Geometria de Distâncias Moleculares.

1 Introdução

Se tivéssemos n pontos no espaço \mathbb{R}^3 , é computacionalmente simples calcular a distância dos pontos entre si. O Problema da Geometria de Distâncias é problema o inverso: temos algumas distâncias e queremos encontrar quais os pontos que satisfaçam essas distâncias. De um modo geral, a Geometria de Distâncias é um problema NP-Difícil (Saxe, 1979).

No caso da Geometria de Distancias Moleculares, em inglês *Molecular Distance Geometry Problem* (MDGP) (Crippen et al., 1988; Liberti et al., 2014), os pontos são átomos de uma molécula, conhecemos algumas distâncias entre os átomos desta molécula e queremos descobrir sua estrutura tridimensional. É ainda um problema NP-Difícil, mas com propriedades combinatórias que podem ajudar a encontrar uma solução.

Podemos usar da computação quântica para resolver esse problema, tendo que encontrar ao menos finitos candidatos à solução. Para isso temos a versão discreta do MDGP, chamado de *Discretizable Molecular Distance Geometry Problem* (DMDGP).

Em computação quântica temos os Algoritmos de Montanaro (2018, 2020), Campbell et al. (2019) e Grover (1996, 1997) que usam um oráculo, isto é, uma função que identifica qual a solução no espaço discreto de busca. E, após encontrarmos o oráculo (de modo teórico), precisamos encontrar uma forma de implementá-lo.

2 Geometria de Distâncias Moleculares

Definição 1 (*Molecular Distance Geometry Problem*). Seja $G = (V, E, d)$ um grafo simples, conectado, não direcionado e ponderado nas arestas por $d : E \rightarrow [0, \infty)$, em que os vértices V representam os átomos da molécula e as arestas E com peso d representam as distâncias conhecidas entre os átomos da molécula. Queremos encontrar uma função

$$\begin{aligned} \mathbf{x} : V &\rightarrow \mathbb{R}^3 \\ v &\mapsto \mathbf{x}(v) \end{aligned}$$

tal que

$$\forall (u, v) \in E, \|\mathbf{x}_u - \mathbf{x}_v\| = d_{u,v}, \quad (2.1)$$

em que $\|\cdot\|$ é a norma euclidiana e utilizaremos a notação \mathbf{x}_v para $\mathbf{x}(v)$ e $d_{u,v}$ para $d(u, v)$. Chamaromos \mathbf{x}_v de uma realização para o vértice v e a função \mathbf{x} apenas de realização.

Proposição 1. Dado um grafo $G = (V, E, d)$ de um MDGP, com $n = |V|$, se há uma ordem total em V tal que:

- Para todo $i \in \{4, \dots, n\}$ temos

$$\{(v_{i-3}, v_i), (v_{i-2}, v_i), (v_{i-1}, v_i)\} \subseteq E;$$

- Para todo $i \in \{3, \dots, n\}$ os vértices $\{v_{i-2}, v_{i-1}, v_i\}$ formam uma clique e vale a desigualdade

$$d_{v_{i-2}, v_i} < d_{v_{i-2}, v_{i-1}} + d_{v_{i-1}, v_i},$$

então o MDGP tem uma quantidade finita de soluções.

Para facilitar a notação, por vezes vamos denotar o vértice v_i simplesmente por i , bem como a realização de v_i por apenas \mathbf{x}_i , em que $i \in \{1, 2, \dots, n\}$. Podemos então entender uma realização \mathbf{x} como sendo uma sequência de n pontos no \mathbb{R}^3 , ou seja, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$.

Vamos definir o que chamaremos de Coordenadas Internas.

Definição 2. Denotaremos por $d_{i-1,i}$ a distância entre da ligação covalente entre os átomos $i - 1$ e i , para $i = 2, \dots, n$.

Definição 3. Denotaremos por $\theta_{i-2,i} \in (0, \pi)$ o ângulo plano, o ângulo de ligação entre os átomos $i - 2$, $i - 1$ e i , para $i = 3, \dots, n$.

Definição 4. Denotaremos por $\omega_{i-3,i} \in [0, 2\pi)$ o ângulo de torção, o ângulo entre os vetores normais dos planos definidos pelos átomos $i - 3$, $i - 2$, $i - 1$ e pelos átomos $i - 2$, $i - 1$, i , para $i = 4, \dots, n$.

Geometricamente, estamos como na Figura 1.

Podemos calcular o ângulo plano $\theta_{i-2,i}$ usando:

$$\theta_{i-2,i} = \arccos \left(\frac{d_{i-2,i-1}^2 + d_{i-1,i}^2 - d_{i-2,i}^2}{2 d_{i-2,i-1} d_{i-1,i}} \right). \quad (2.2)$$

Podemos expressar o cosseno do ângulo de torção $\omega_{i-3,i}$ em função das distâncias conhecidas, dada por (2.3), em que

$$\begin{aligned} d_{i-3,i-2,i-1} &= d_{i-3,i-2}^2 + d_{i-2,i-1}^2 - d_{i-3,i-1}^2, \\ d_{i-2,i-1,i} &= d_{i-2,i-1}^2 + d_{i-1,i}^2 - d_{i-2,i}^2. \end{aligned}$$

Do valor de $\cos(\omega_{i-3,i})$, nós obtemos duas possibilidades para cada $\omega_{i-3,i}$. Sendo assim, podemos encontrar a estrutura da molécula escolhendo o sinal $+$ ou $-$ de $\sin(\omega_{i-3,i}) = \pm \sqrt{1 - \cos^2(\omega_{i-3,i})}$.

Sendo assim, tendo as coordenadas internas de uma solução, existe uma fórmula que retorna a solução em coordenadas cartesianas (Phillips et al., 1994).

3 Oráculo

Para problemas de busca em computação quântica, precisamos definir um oráculo, uma função que, dada um candidato à solução, nos diga se, de fato, aquele candidato é (ou não) solução do nosso problema.

Para isso, vamos precisar de dois registradores quânticos: o primeiro para armazenar os candidatos à solução e o segundo para armazenar a resposta do oráculo.

Vamos tomar $N = 2^{n-3}$, em que n é a quantidade de vértices no DMDGP; então teremos $n - 3$ *qubits* no primeiro registrador. Precisamos associar cada estado do primeiro registrador $\{|0\rangle, |1\rangle, \dots, |N - 1\rangle\}$ a uma solução do DMDGP. O segundo registrador terá apenas 1 *qubit*.

Para o q -ésimo *qubit* do primeiro registrador, com $q \in \{1, \dots, n - 3\}$,

- Vamos associar o estado $|0\rangle$ com

$$\sin(\omega_{q,q+3}) = +\sqrt{1 - \cos^2(\omega_{q,q+3})};$$

- Vamos associar o estado $|1\rangle$ com

$$\sin(\omega_{q,q+3}) = -\sqrt{1 - \cos^2(\omega_{q,q+3})}.$$

Definição 5. Seja a função $g: \mathbb{R}^{3n} \rightarrow [0, \infty)$ dada por

$$g(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{(u,v) \in E} \left(\|\mathbf{x}_u - \mathbf{x}_v\|^2 - d_{u,v}^2 \right)^2.$$

Vamos dizer que \mathbf{x} é solução do DMDGP se $g(\mathbf{x}) < \delta$, em que $\delta > 0$ é uma tolerância de erro.

Definição 6. Seja a função $h: \{0, 1, \dots, N - 1\} \rightarrow \mathbb{R}^{3n}$ de tal forma que h recebe um candidato à solução k e associa a uma realização

$$h(k) = \mathbf{x}^k = (\mathbf{x}_1^k, \mathbf{x}_2^k, \dots, \mathbf{x}_n^k);$$

realização essa com o valor dos senos e cossenos associados ao valor do estado $|k\rangle$, com $k \in \{0, 1, \dots, N - 1\}$.

Sendo assim, temos $g \circ h: \{0, 1, \dots, N - 1\} \rightarrow [0, \infty)$ dada por

$$g(h(k)) = \sum_{(u,v) \in E} \left(\|\mathbf{x}_u^k - \mathbf{x}_v^k\|^2 - d_{u,v}^2 \right)^2. \quad (3.1)$$

O resultado principal do trabalho é o Teorema 1, que garante a existência de um Oráculo para o DMDGP.

Teorema 1. Dado $\varepsilon \in (0, 1)$ e seja $\varphi: I \subseteq \mathbb{R} \rightarrow [0, 1]$, em que $\text{Im}(g \circ h) \subseteq I$, com as propriedades:

- φ é estritamente crescente;
- $\varphi(\delta) = 1 - \varepsilon$.

A função

$$f(k) = 1 - \lfloor \varphi(g(h(k))) + \varepsilon \rfloor \quad (3.2)$$

satisfaz a seguinte propriedade: $f(k) = 1 \Leftrightarrow k$ está associado a uma solução do DMDGP ($g(h(k)) < \delta$).

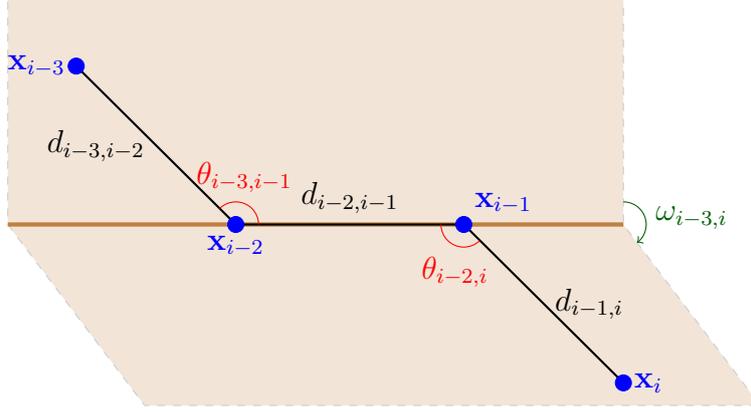


Figura 1: Posição dos Átomos.

$$\cos(\omega_{i-3,i}) = \frac{2d_{i-2,i-1}^2 (d_{i-3,i-2}^2 + d_{i-2,i}^2 - d_{i-3,i}^2) - (d_{i-3,i-2,i-1})(d_{i-2,i-1,i})}{\sqrt{4d_{i-3,i-2}^2 d_{i-2,i-1}^2 - (d_{i-3,i-2,i-1})^2} \sqrt{4d_{i-2,i-1}^2 d_{i-2,i}^2 - (d_{i-2,i-1,i})^2}}. \quad (2.3)$$

Para cada uma das instâncias do DMDGP (que é uma lista de distâncias), temos um oráculo f que identifica qual a solução. Para cada elemento em $\{0, 1, \dots, N-1\}$, associamos um valor em $\{0, 1\}$. Dessa forma, podemos interpretar esse oráculo como sendo uma tabela de verdade.

Exemplo 1. Na Tabela 1, temos um exemplo de uma instância com $n = 7$. Sendo assim, pelo Teorema 1, existe uma função que identifica sua solução, sendo esta função dada pela Tabela 2. Nesta tabela, estamos usando simetrias para reduzir o espaço de busca (Mucherino et al., 2011; Lavor et al., 2021).

Para o oráculo da Tabela 2, temos a porta quântica da Figura 2, uma porta Toffoli generalizada, em que $|q_1\rangle, |q_2\rangle, |q_3\rangle$ são os *qubits* do primeiro registrador (que serão os *qubits* controle) e $|t\rangle$ é o segundo registrador (que é o alvo).

4 Conclusão

Apesar da solução poder ser detectada em simulações computacionais, quando executado em um computador quântico real, nem sempre é o que acontece. Mesmo com poucos *qubits*, o erro nas aplicações ainda é muito alto, fazendo com que nem sempre consigamos o resultado esperado (em muitas vezes, a distribuição da saída é uniforme de tão grande o erro!) (Wang and Krstic, 2020). A principal “culpada” por esse erro elevado é a porta CNOT (que no nosso caso, é usada para compor a porta Toffoli generalizada que usamos no oráculo). Quanto mais portas CNOT são usadas,

maior o erro. Dessa forma, precisamos encontrar formas ótimas para implementar o oráculo. Um dos caminhos é usar a Porta de Margolus (Song and Klappenecker, 2003; BEIT, 2020).

Outro problema enfrentado é que estamos assumindo que conhecemos o oráculo, o que, de fato, em computação quântica, ainda não é o caso. O argumento usado é: como em computação clássica esse oráculo é facilmente implementado, logo existe sua versão usando computação reversível e podemos usá-la em computação quântica (Dalcumune et al., 2021). Apesar dessa afirmação ser verdadeira, essa implementação não é simples: seria necessário uma enorme quantidade de *qubits* (o que torna inviável a simulação usando computação clássica), corretores de erro bastante avançados, lidar com pontos fixos (ou flutuantes). Todos esses fatores são empecilhos para a implementação real da função $g \circ h$. De qualquer forma, algumas maneiras de como fazer essa implementação (mesmo que como algo a longo prazo) usando QRAM, por exemplo, estão na nossa mira.

Como fruto desse trabalho, há um artigo em elaboração, *A Quantum Approach to the Discretizable Molecular Distance Geometry Problem*, escrito também com Franklin Marquezino da Coppe/UFRJ e Renato Portugal do LNCC.

Como projetos futuros, espero continuar estudando Computação Quântica, Computação Reversível e otimização de portas quânticas.

Finalmente, agradeço ao CNPq pelo financiamento da pesquisa.

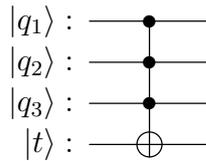


Figura 2: Porta Quântica da Tabela 2.

Referências

- BEIT (2020). Ibm quantum challenge – adam’s write-up. Disponível em <https://beit.tech/blog/adsz-iqc-writeup.html>. Acesso em Julho 2021.
- Campbell, E., Khurana, A., and Montanaro, A. (2019). Applying quantum algorithms to constraint satisfaction problems. *Quantum*, 3:167.
- Crippen, G. M., Havel, T. F., et al. (1988). *Distance geometry and molecular conformation*, volume 74. Research Studies Press Taunton.
- Dalcumune, E., Kowada, L. A. B., da Cunha Ribeiro, A., de Figueiredo, C. M. H., and de Lima Marquezino, F. (2021). A reversible circuit synthesis algorithm with progressive increase of controls in generalized toffoli gates. *JUCS - Journal of Universal Computer Science*, 27(6):544–563.
- Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC 96*. ACM Press.
- Grover, L. K. (1997). Quantum mechanics helps in searching for a needle in a haystack. *Physical Review Letters*, 79(2):325–328.
- Lavor, C., Oliveira, A., Rocha, W., and Souza, M. (2021). On the optimality of finding DMDGP symmetries. *Computational and Applied Mathematics*, 40(3).
- Liberti, L., Lavor, C., Maculan, N., and Mucherino, A. (2014). Euclidean distance geometry and applications. *Siam Review*, 56(1):3–69.
- Montanaro, A. (2018). Quantum-walk speedup of backtracking algorithms. *Theory of Computing*, 14(1):1–24.
- Montanaro, A. (2020). Quantum speedup of branch-and-bound algorithms. *Physical Review Research*, 2(1):013056.
- Mucherino, A., Lavor, C., and Liberti, L. (2011). A symmetry-driven bp algorithm for the discretizable molecular distance geometry problem. In *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, pages 390–395. IEEE.
- Phillips, A. T., Rosen, J. B., and Walke, V. H. (1994). Molecular structure determination by convex global underestimation of local energy minima. In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 181–198. American Mathematical Society.
- Saxe, J. B. (1979). Embeddability of weighted graphs in k-space is strongly np-hard. In *Proc. of 17th Allerton Conference in Communications, Control and Computing, Monticello, IL*, pages 480–489.
- Song, G. and Klappenecker, A. (2003). The simplified toffoli gate implementation by margolus is optimal. *arXiv*. Disponível em <https://arxiv.org/pdf/quant-ph/0312225.pdf>. Acesso em Julho de 2021.
- Wang, Y. and Krstic, P. S. (2020). Prospect of using grover’s search in the noisy-intermediate-scale quantum-computer era. *Physical Review A*, 102(4):042609.

u	v	$d_{u,v}$
2	7	4.813768718
4	7	3.4095310172
1	3	2.4913895358
3	7	4.7110300453
3	6	3.8349217603
1	5	2.444520579
2	4	2.4913895358
6	7	1.526
1	2	1.526
1	6	3.8016257964
5	7	2.4913895358
3	5	2.4913895358
4	5	1.526
3	4	1.526
1	4	2.9704371064
4	6	2.4913895358
2	5	2.5644614938
1	7	4.7642917197
2	6	4.0903451356
2	3	1.526
5	6	1.526

Tabela 1: Lista de Distâncias

k	$f(k)$
000	0
001	0
010	0
011	0
100	0
101	0
110	0
111	1

Tabela 2: Oráculo da Instância da Tabela 1.