



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus Bauru

PEDRO MARCELO PRADO

Algoritmos de Feedback Quântico: Estudo, Extensões e Aplicações

BAURU
2026

PEDRO MARCELO PRADO

Algoritmos de Feedback Quântico: Estudo, Extensões e Aplicações

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação em Ciência da Computação, da Faculdade de Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de Bauru.

Financiadora: FAPESP - Proc. 2023/12110-7

Orientador: Prof. Dr. Felipe Fernandes Fanchini

Coorientador: Dr. Felipe Rodrigues Perche Mahlow

BAURU
2026

P896a

Prado, Pedro Marcelo

Algoritmos de feedback quântico : estudo, extensões e aplicações /
Pedro Marcelo Prado. -- Bauru, 2026
90 p. : il.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP),
Faculdade de Ciências, Bauru

Orientador: Felipe Fernandes Fanchini

Coorientador: Felipe Rodrigues Perche Mahlow

1. Computação quântica. 2. Otimização quântica. 3. Algoritmo
quântico baseado em feedback. 4. Inicialização guiada por medições.
I. Título.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE PEDRO MARCELO PRADO, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO, DA FACULDADE DE CIÊNCIAS - CÂMPUS DE BAURU.

Aos 24 de julho de 2026, às 9h, por meio de Videoconferência, realizou-se a defesa de DISSERTAÇÃO DE MESTRADO de PEDRO MARCELO PRADO, intitulada "Algoritmos de Feedback Quântico: Estudo, Extensões e Aplicações". A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. FELIPE FERNANDES FANCHINI (Orientador(a) - Participação Virtual) do(a) Departamento de Física e Meteorologia / UNESP / Câmpus de Bauru - FC, Professor Doutor KELTON AUGUSTO PONTARA DA COSTA (Participação Virtual) do(a) Departamento de Computação / UNESP / Câmpus de Bauru - FC, Prof. Dr. IANN CUNHA (Participação Virtual) do(a) Computação / Instituto Israelita de Ensino e Pesquisa Albert Einstein, Após a exposição pelo mestrando e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, o discente recebeu o conceito final APROVADO. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.



Documento assinado digitalmente
FELIPE FERNANDES FANCHINI
Data: 24/07/2026 11:25:55-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. FELIPE FERNANDES FANCHINI

All models are wrong, but some are useful - George E.P. Box

Resumo

Esta dissertação investiga algoritmos quânticos baseados em feedback voltados à otimização combinatória, com ênfase no *Feedback-Based Algorithm for Quantum Optimization* (FALQON) e em estratégias para reduzir a profundidade de circuitos em dispositivos do regime NISQ. Inicialmente, são revisitados os fundamentos de otimização combinatória, as formulações QUBO/Ising e os conceitos essenciais de computação, informação e controle quânticos que sustentam a formulação desses métodos. Em seguida, são analisadas duas variantes relevantes da literatura, o TR-FALQON e o SO-FALQON, propostas para tornar a dinâmica do algoritmo mais eficiente em regimes de menor profundidade. Como principal contribuição do trabalho, é apresentada a estratégia *Measurement-Guided Initialization* aplicada ao FALQON, denominada MGI-FALQON. O método introduz um laço externo de refinamento no qual *bitstrings* obtidas por medições em execuções rasas são filtradas e utilizadas para estimar marginais de um qubit, a partir das quais se constrói um novo estado inicial enviesado para a execução seguinte. A avaliação experimental foi realizada no problema MaxCut, formulado em termos QUBO/Ising. Os resultados mostram que TR-FALQON e SO-FALQON aceleram a convergência em relação ao FALQON padrão, mas ainda demandam circuitos relativamente profundos. Em contraste, o MGI-FALQON, mesmo operando com profundidade interna fixa e rasa, apresentou melhora expressiva no desempenho, superando os métodos profundos na instância analisada, especialmente quando combinado a um filtro variável e ao TR-FALQON. Esses resultados indicam que a inicialização guiada por medições constitui uma estratégia promissora para tornar algoritmos quânticos baseados em feedback mais compatíveis com as restrições de hardware quântico de curto prazo.

Palavras-chave: computação quântica; otimização quântica; FALQON; feedback quântico; inicialização guiada por medições; NISQ.

Abstract

This dissertation investigates feedback-based quantum algorithms for combinatorial optimization, with emphasis on the *Feedback-Based Algorithm for Quantum Optimization* (FALQON) and on strategies to reduce circuit depth in NISQ devices. First, the work reviews the foundations of combinatorial optimization, QUBO/Ising formulations, and the essential concepts of quantum computation, quantum information, and quantum control that support the formulation of such methods. Next, two relevant variants from the literature are analyzed, namely TR-FALQON and SO-FALQON, both proposed to improve the efficiency of the algorithm dynamics in lower-depth regimes. As the main contribution of this dissertation, the *Measurement-Guided Initialization* strategy applied to FALQON, referred to as MGI-FALQON, is presented. The method introduces an outer refinement loop in which bitstrings obtained from measurements of shallow executions are filtered and used to estimate single-qubit marginals, from which a new biased initial state is constructed for the next execution. The experimental evaluation was carried out on the MaxCut problem formulated in the QUBO/Ising framework. The results show that TR-FALQON and SO-FALQON accelerate convergence when compared with standard FALQON, but still require relatively deep circuits. In contrast, MGI-FALQON, even under fixed shallow internal depth, achieved significant performance improvements, surpassing the deep methods in the analyzed instance, especially when combined with a variable filter and with TR-FALQON. These findings indicate that measurement-guided initialization is a promising strategy for making feedback-based quantum algorithms more compatible with the constraints of near-term quantum hardware.

Keywords: quantum computing; quantum optimization; FALQON; feedback-based quantum algorithms; measurement-guided initialization; NISQ.

Lista de ilustrações

Figura 1 – Representação esquemática do problema MaxCut em um grafo. . .	21
Figura 2 – Exemplo de uma iteração do MGI-FALQON	69
Figura 3 – Instância do MaxCut com 8 vértices	72
Figura 4 – Comparação entre o FALQON e suas variantes no problema MaxCut com 7 qubits. Foram usados $L = 200$ em todos os casos; para o FALQON, $\Delta t = 0.2$ e $\beta_1 = 0$; para o TR-FALQON, $\Delta t = 0.24$, $a = 1.5$, $\beta_1 = 0$, $t_f = L(a\Delta t)$ e $f'(\tau) = a - (a - 1) \cos\left(\frac{2\pi a}{t_f}\tau\right)$; para o SO-FALQON, $\Delta t = 0.4$ na versão híbrida. No painel a), energia em função de L ; no painel b), probabilidade da solução ótima. .	74
Figura 5 – Resultados do MGI-FALQON com profundidade interna $L = 2$, $R = 30$ iterações externas e 100 sementes. Foram comparados os filtros $n = 5$, $n = 30$ e n variável de $30 \rightarrow 5$. No painel a), energia média por iteração; no painel b), probabilidade média da solução ótima. As regiões sombreadas indicam a dispersão entre sementes. .	77
Figura 6 – Comparação entre as variantes acopladas ao MGI com filtro variável, profundidade interna $L = 2$, $R = 30$ iterações externas e 100 sementes. Foram usados $\Delta t = 0.4$ para o FALQON, $\Delta t = 0.2$ e $a = 2$ para o TR-FALQON, e $\Delta t = 0.42$ para o SO-FALQON. No painel a), energia média por iteração; no painel b), probabilidade média da solução ótima.	79

Lista de abreviaturas e siglas

VQA	Variational Quantum Algorithm
VQE	Variational Quantum Eigensolver
QAOA	Quantum Approximate Optimization Algorithm
FQA	Feedback-Based Quantum Algorithm
FALQON	Feedback-Based Algorithm for Quantum Optimization
TR-FALQON	Time-Rescaled FALQON
SO-FALQON	Second-Order FALQON
MGI	Measurement-Guided Initialization
NISQ	Noisy Intermediate-Scale Quantum
QUBO	Quadratic Unconstrained Binary Optimization
NP	Nondeterministic Polynomial

Sumário

1	INTRODUÇÃO	11
2	FUNDAMENTOS DE OTIMIZAÇÃO E FORMULAÇÕES	
	QUBO/ISING	15
2.1	Fundamentos de otimização combinatória	15
2.2	Otimização Binária Quadrática Irrestrita	17
2.2.1	Incorporação de restrições	17
2.3	Transcrição QUBO/Ising para Hamiltonianos de custo	18
2.4	O problema MaxCut	20
3	FUNDAMENTOS DE COMPUTAÇÃO E INFORMAÇÃO	
	QUÂNTICAS	23
3.1	Estados quânticos e notação bra-ket	23
3.2	Operadores unitários e evolução temporal	26
3.3	Medição quântica e valores esperados	28
3.4	Qubit, portas e circuitos quânticos	29
3.5	Formalismo de Hamiltonianos	31
3.6	Conceitos básicos de controle quântico	33
3.7	Introdução ao controle por Lyapunov	35
3.7.1	Escolha da função de Lyapunov	35
3.8	Relação entre controle quântico e algoritmos	37
3.9	Conexão com algoritmos variacionais e de feedback	40
4	ALGORITMOS DE FEEDBACK QUÂNTICO	43
4.1	Da formulação de controle ao algoritmo FALQON	43
4.2	Formulação do FALQON para otimização combinatória	44
4.2.1	Hamiltonianos H_p e H_d	44
4.3	Regra de feedback e monotonicidade da energia	47
4.3.1	Derivação da regra de feedback	47
4.4	Interpretação física e aspectos práticos de implementação	49

4.4.1	Interpretação física e computacional	49
4.4.2	Estimativa de A_k por medições	49
4.5	Limitações práticas do FALQON	50
4.6	Variantes da literatura para redução de profundidade	52
4.6.1	TR-FALQON: reescalonamento temporal	52
4.6.2	SO-FALQON: correção de segunda ordem	54
4.7	Comparação conceitual	56
5	INICIALIZAÇÃO GUIADA POR MEDIÇÕES	59
5.1	Formulação geral do loop externo do MGI-FALQON	60
5.1.1	Seleção das <i>bitstrings</i> dominantes	62
5.1.2	Derivação explícita das marginais de um qubit	64
5.2	Construção do estado inicial enviesado	66
5.2.1	Estado de um qubit e correspondência com R_y	66
5.2.2	Construção do estado produto global	67
5.3	Interpretação conceitual e considerações sobre o método	68
6	APLICAÇÕES E RESULTADOS	71
6.1	Protocolo experimental	71
6.2	Resultados do FALQON e de suas variantes	73
6.3	Resultados do MGI-FALQON com diferentes filtros	75
6.4	Resultados do MGI aplicado às variantes do FALQON	78
6.5	Comparação quantitativa entre métodos	80
6.6	Limitações práticas e perspectivas	81
7	CONCLUSÃO	83
	REFERÊNCIAS	86

1 Introdução

A computação quântica surge como um novo paradigma computacional motivado pela necessidade de lidar com problemas cuja complexidade cresce de forma intratável para arquiteturas clássicas. Desde as propostas iniciais de Feynman e Deutsch, que destacaram a dificuldade fundamental de simular sistemas quânticos com computadores clássicos, a área evoluiu significativamente, consolidando-se como um campo interdisciplinar que conecta física, ciência da computação, matemática aplicada e engenharia (1). Ao longo das últimas décadas, avanços teóricos e experimentais permitiram a implementação de processadores quânticos programáveis, inaugurando uma fase em que algoritmos quânticos podem ser testados em dispositivos reais, ainda que com severas limitações (2).

Um dos principais atrativos da computação quântica reside em seu potencial para abordar problemas de otimização combinatória e de simulação de sistemas físicos complexos (3, 4, 5). Tais problemas são caracterizados por espaços de busca exponencialmente grandes, nos quais métodos clássicos frequentemente se tornam inviáveis à medida que o número de variáveis aumenta (6). Nesse contexto, algoritmos quânticos exploram superposição, interferência e emaranhamento como recursos computacionais capazes de explorar o espaço de soluções de maneira mais eficiente, oferecendo perspectivas promissoras para aplicações científicas e tecnológicas de grande impacto (2).

Entre os diversos modelos de computação quântica, o modelo baseado em portas quânticas destaca-se como o mais amplamente estudado e implementado (1). Nesse formalismo, algoritmos são expressos como circuitos compostos por sequências de operações unitárias que evoluem o estado quântico inicial até um estado final cuja distribuição de probabilidades codifica a solução do problema (1). A extração dessa informação é realizada por meio de medições repetidas, permitindo estimar observáveis de interesse. Algoritmos fundamentais desenvolvidos nesse paradigma evidenciam que, ao menos para classes específicas de problemas, vantagens assintóticas em relação a métodos clássicos podem ser alcançadas (4).

Apesar desses avanços conceituais, a implementação prática de algoritmos quânticos enfrenta desafios substanciais. Os dispositivos atualmente disponíveis enquadram-se no regime conhecido como *Noisy Intermediate-Scale Quantum* (NISQ), caracterizado por um número limitado de qubits, tempos de coerência finitos e operações sujeitas a diferentes fontes de erro, como erros de porta, ruído de medição, descoerência, relaxação e perda de fase (2, 7). Um fator crítico nesse cenário é a profundidade dos circuitos: circuitos mais profundos exigem sequências maiores de operações unitárias e, consequentemente, acumulam erros de forma mais significativa, comprometendo a fidelidade do estado preparado e limitando a confiabilidade dos resultados (7). Assim, mesmo algoritmos teoricamente eficientes podem tornar-se impraticáveis quando mapeados para hardware real, o que torna a redução de profundidade um requisito central para algoritmos quânticos de curto e médio prazo.

Dentro desse panorama, os algoritmos quânticos variacionais, conhecidos pela sigla VQA (*Variational Quantum Algorithms*), emergiram como uma das abordagens mais exploradas para aplicações em otimização e simulação (5, 8). Métodos como o *Variational Quantum Eigensolver* (VQE) e o *Quantum Approximate Optimization Algorithm* (QAOA) utilizam circuitos parametrizados combinados com rotinas clássicas de otimização (5, 4). No entanto, à medida que o número de parâmetros, qubits e camadas cresce, esses algoritmos passam a enfrentar dificuldades estruturais, como a presença de *barren plateaus*. Nesses regimes, a superfície definida pelos valores da função de custo, ou *optimization landscape*, torna-se quase plana em grandes regiões do espaço de parâmetros, de modo que os gradientes se tornam exponencialmente pequenos e deixam de fornecer direções úteis para a atualização dos parâmetros (9, 8). Não há, contudo, um número universal de qubits, camadas ou portas a partir do qual esse fenômeno necessariamente ocorre, pois sua manifestação depende do *ansatz*, da função de custo, da conectividade e da estratégia de inicialização. Ainda assim, em *ansatz* suficientemente expressivos e profundos, sobretudo quando a profundidade cresce com o número de qubits, a otimização tende a se tornar progressivamente mais difícil, comprometendo a escalabilidade desses métodos em sistemas maiores (9, 8).

Como alternativa a esses esquemas híbridos, algoritmos baseados em *feedback* quântico foram propostos com o objetivo de eliminar a etapa de otimização clássica (10, 11). Nessa classe de métodos, os parâmetros do circuito são ajustados iterativamente

a partir de medições realizadas no próprio estado quântico durante a execução do algoritmo. Em particular, algoritmos de *feedback* inspirados em controle quântico utilizam regras determinísticas que garantem a diminuição monotônica de uma função custo associada ao problema, direcionando a evolução do sistema ao estado desejado de forma autônoma (12).

Entretanto, embora conceitualmente atrativos, algoritmos de *feedback* quântico também apresentam limitações práticas relevantes. Em sua forma padrão, a convergência para soluções de alta qualidade geralmente requer um número elevado de camadas, o que reintroduz o problema da profundidade excessiva em dispositivos NISQ (10, 13). Em particular, para instâncias não ponderadas de MaxCut em grafos 3-regulares, observa-se que a profundidade de circuito requerida pelo FALQON pode se tornar quase duas ordens de grandeza maior que a do QAOA para atingir níveis comparáveis de precisão (10, 14, 13). Diante disso, esforços recentes têm se concentrado no desenvolvimento de estratégias de redução de profundidade, que buscam acelerar a dinâmica do algoritmo ou aumentar a eficiência de cada camada aplicada. Essas estratégias incluem reparametrizações temporais da evolução, como no TR-FALQON (15), bem como aproximações de ordem superior para a regra de *feedback*, como no SO-FALQON (13), todas com o objetivo comum de atingir resultados comparáveis utilizando circuitos significativamente mais rasos.

Neste trabalho, dá-se especial atenção a uma estratégia adicional, denominada *Measurement-Guided Initialization* (MGI). A MGI consiste em uma abordagem baseada em inicialização guiada por medições, proposta para mitigar de forma ainda mais efetiva o custo em profundidade. Essa metodologia introduz um loop externo ao algoritmo de *feedback*, no qual execuções rasas são utilizadas para coletar estatísticas de medições no estado final (16). Neste contexto, o termo “rasas” refere-se a execuções em que a profundidade interna do circuito FALQON é previamente limitada e mantida fixa ao longo do processo, deslocando o refinamento do aumento do número de camadas para a atualização sucessiva do estado inicial. A partir das *bitstrings* medidas nessas execuções, constroem-se estimativas das marginais de um qubit, que são então utilizadas para preparar uma nova inicialização enviesada para a execução seguinte. Esse mecanismo concentra gradualmente a probabilidade em regiões promissoras do espaço de busca sem aumentar a profundidade interna do circuito, explorando diretamente os dados

gerados pelo próprio algoritmo e alinhando-se de maneira natural às restrições impostas por hardware NISQ.

Dessa forma, esta dissertação tem como objetivo geral investigar estratégias para reduzir a profundidade de circuitos em algoritmos quânticos baseados em *feedback*, com foco no FALQON e em extensões voltadas ao regime NISQ. Especificamente, busca-se estudar os fundamentos teóricos desses algoritmos, analisar variantes da literatura como TR-FALQON e SO-FALQON, desenvolver e avaliar a estratégia MGI-FALQON, baseada em inicialização guiada por medições, e comparar seu desempenho em termos de energia final, probabilidade de sucesso e número de camadas em problemas de otimização formulados no modelo QUBO/Ising. Com isso, pretende-se discutir os regimes de aplicabilidade, as limitações práticas e o potencial dessas abordagens para tornar algoritmos de *feedback* quântico mais compatíveis com dispositivos quânticos de curto prazo.

A organização desta dissertação é a seguinte. No Capítulo 2, são apresentados os fundamentos de otimização combinatória e as formulações QUBO e Ising, que servem como base para os problemas estudados ao longo do trabalho. O Capítulo 3 revisa os conceitos essenciais de computação e informação quânticas, bem como noções de controle quântico relevantes para algoritmos baseados em *feedback*. No Capítulo 4, são apresentados o algoritmo FALQON e variantes relevantes da literatura, como TR-FALQON e SO-FALQON, destacando suas contribuições para a redução de profundidade e para o aprimoramento do método original. A estratégia de inicialização guiada por medições, denominada *Measurement-Guided Initialization* (MGI), e sua aplicação ao FALQON, denominada MGI-FALQON, são apresentadas no Capítulo 5. No Capítulo 6, são discutidas as aplicações consideradas, bem como os resultados numéricos obtidos e suas respectivas análises. Por fim, o Capítulo 7 sintetiza as conclusões do trabalho e apresenta perspectivas para investigações futuras.

2 Fundamentos de Otimização e Formulações QUBO/Ising

2.1 Fundamentos de otimização combinatória

Problemas de otimização combinatória envolvem encontrar, dentro de um conjunto discreto e finito de configurações possíveis, aquela que otimiza um critério previamente definido. Diferentemente da otimização contínua, na qual as variáveis de decisão assumem valores reais em domínios contínuos — como em problemas clássicos de minimização de funções diferenciáveis em $\mathbb{R}^n$ —, a otimização combinatória lida com variáveis discretas, frequentemente binárias ou inteiras, e com espaços de busca cujo tamanho cresce exponencialmente com o número de variáveis envolvidas (17, 18).

De forma geral, um problema de otimização pode ser expresso como

$$\min_{x \in \mathcal{X}} f(x) \quad \text{ou} \quad \max_{x \in \mathcal{X}} f(x), \quad (1)$$

onde x representa o vetor de variáveis de decisão, $\mathcal{X}$ denota o conjunto de soluções admissíveis e $f(x)$ é a função objetivo. No contexto combinatório, o conjunto $\mathcal{X}$ é discreto e, em muitos casos, definido implicitamente por restrições estruturais do problema, como condições de partição, alocação ou ordenamento, que restringem o conjunto de configurações viáveis.

Uma característica central desses problemas é a complexidade do espaço de busca. Para problemas definidos sobre n variáveis binárias, cada variável pode assumir dois valores possíveis, resultando em um total de 2^n configurações distintas. Esse crescimento exponencial do número de soluções candidatas torna abordagens baseadas em enumeração exaustiva inviáveis mesmo para valores moderados de n .

Esse comportamento está diretamente relacionado à dificuldade computacional de muitos problemas combinatórios. Em particular, diversos desses problemas pertencem à classe dos problemas NP-difíceis, isto é, problemas para os quais não se conhece, em geral, algoritmo determinístico capaz de resolvê-los em tempo polinomial. Essa

característica motiva o desenvolvimento e o uso de métodos aproximativos, heurísticos e metaheurísticos, capazes de produzir soluções de boa qualidade sem garantia de otimalidade global (19, 18).

Diversos problemas de relevância prática enquadram-se nessa classe, incluindo problemas de corte em grafos, roteamento, escalonamento e alocação de recursos. Um exemplo clássico é o problema MaxCut, que consiste em particionar os vértices de um grafo em dois subconjuntos de modo a maximizar o número ou o peso das arestas entre eles. O MaxCut é um problema NP-difícil e amplamente utilizado como benchmark teórico e computacional em estudos de otimização (20). Esses problemas compartilham o fato de que suas soluções correspondem a configurações discretas que maximizam ou minimizam uma função de custo global definida a partir de interações locais entre variáveis.

Uma observação fundamental é que muitos problemas de otimização combinatória podem ser reinterpretados como problemas de minimização de uma função de energia clássica. Nessa perspectiva, cada configuração admissível do problema corresponde a um estado do sistema, e a função objetivo desempenha o papel de uma energia associada a esse estado. Essa reformulação é particularmente relevante, pois permite estabelecer uma conexão direta entre problemas combinatórios e modelos físicos baseados em interações discretas, servindo como ponto de partida para formulações do tipo QUBO e Ising (6).

Do ponto de vista computacional, a dificuldade inerente dos problemas de otimização combinatória, associada ao crescimento exponencial do espaço de busca e à ausência de algoritmos clássicos eficientes para problemas NP-difíceis, motiva a investigação de paradigmas alternativos de computação. Nesse contexto, algoritmos quânticos de otimização surgem como candidatos naturais, uma vez que permitem codificar a função de custo como um Hamiltoniano de problema e explorar propriedades da dinâmica quântica para guiar a busca por soluções de baixa energia. Essa perspectiva tem impulsionado o desenvolvimento de algoritmos quânticos variacionais e de controle por realimentação, particularmente formulados sobre representações do tipo QUBO e Ising, que serão discutidas nas seções seguintes.

2.2 Otimização Binária Quadrática Irrestrita

A Otimização Binária Quadrática Irrestrita, conhecida pela sigla QUBO (*Quadratic Unconstrained Binary Optimization*), constitui uma classe fundamental de problemas de otimização combinatória na qual a função objetivo possui termos quadráticos e as variáveis de decisão são binárias, assumindo valores em $\{0, 1\}$ (21, 22). Essa formulação é particularmente relevante por sua expressividade, pois uma ampla variedade de problemas práticos pode ser reescrita nessa forma, incluindo aplicações em logística, engenharia e inteligência artificial (23, 24).

De maneira geral, um problema QUBO é definido pela minimização de uma função custo quadrática sobre variáveis binárias, da forma

$$\min_{x \in \{0,1\}^n} C(x), \quad C(x) = x^T Q x = \sum_{i,j} Q_{ij} x_i x_j, \quad (2)$$

onde $x = (x_1, x_2, \dots, x_n)^T$ representa o vetor de variáveis binárias e $Q \in \mathbb{R}^{n \times n}$ é a matriz de coeficientes do problema. Sem perda de generalidade, pode-se tomar Q como simétrica, uma vez que apenas sua parte simétrica contribui para o valor de $x^T Q x$. Termos lineares podem ser incorporados à diagonal de Q , pois, para variáveis binárias, vale $x_i^2 = x_i$. O objetivo consiste, portanto, em determinar a configuração binária x que minimiza o valor de $C(x)$.

2.2.1 Incorporação de restrições

Embora a formulação QUBO seja, por definição, irrestrita, problemas de otimização reais frequentemente envolvem restrições explícitas, como condições de balanceamento, conservação de recursos ou exigências de seleção exclusiva, por exemplo, “escolher exatamente um” dentre vários itens. Uma estratégia amplamente utilizada para lidar com esse aspecto consiste em incorporar as restrições diretamente à função objetivo por meio de termos de penalização quadráticos (6). Dessa forma, o problema original com restrições é transformado em um problema QUBO equivalente, no qual soluções inviáveis são penalizadas energeticamente.

Considere uma restrição de igualdade da forma

$$\sum_i a_i x_i = b. \quad (3)$$

Essa condição pode ser incorporada à função custo por meio de um termo de penalização quadrático,

$$P_{\text{eq}}(x) = \left(\sum_i a_i x_i - b \right)^2. \quad (4)$$

A escolha quadrática é conveniente porque $P_{\text{eq}}(x)$ é sempre não negativo e se anula apenas quando a restrição é satisfeita. Dessa forma, violações da igualdade em qualquer direção, isto é, tanto acima quanto abaixo do valor b , são penalizadas. Um termo linear proporcional a $\sum_i a_i x_i - b$, por outro lado, poderia assumir valores negativos e alterar a função objetivo favorecendo certas violações da restrição, em vez de penalizá-las de maneira simétrica. A função custo modificada passa então a ser

$$C'(x) = C(x) + \lambda P_{\text{eq}}(x), \quad (5)$$

onde $\lambda > 0$ controla a severidade da penalização associada à violação da restrição. Para que a formulação penalizada preserve as soluções ótimas viáveis do problema original, λ deve ser escolhido suficientemente grande em relação à escala dos valores assumidos pela função objetivo.

2.3 Transcrição QUBO/Ising para Hamiltonianos de custo

Para empregar algoritmos quânticos de otimização, é necessário reescrever o problema clássico de minimização em uma linguagem compatível com a dinâmica quântica. Em vez de operar diretamente sobre a função custo $C(x)$, busca-se construir um Hamiltoniano quântico cujo espectro de energia codifique os valores dessa função nas configurações da base computacional (6).

Nesse mapeamento, cada variável binária $x_i \in \{0, 1\}$ é associada a um qubit, de modo que os estados $|0\rangle$ e $|1\rangle$ representam os valores possíveis da variável. O objetivo é construir um Hamiltoniano H_f tal que, para cada estado computacional $|x\rangle$,

$$H_f |x\rangle = C(x) |x\rangle. \quad (6)$$

Assim, a minimização da função custo clássica passa a ser equivalente à busca pelo menor autovalor de H_f .

A conexão entre a formulação QUBO e a formulação de Ising é dada pela mudança de variáveis

$$s_i = 1 - 2x_i,$$

em que $s_i \in \{-1, +1\}$ representa uma variável de spin. Equivalentemente,

$$x_i = \frac{1 - s_i}{2}.$$

No contexto quântico, a variável de spin s_i é representada pelo operador de Pauli σ_i^z , cujos autovalores na base computacional são

$$\sigma_i^z |0\rangle = +|0\rangle, \quad \sigma_i^z |1\rangle = -|1\rangle.$$

Portanto, a variável binária x_i é representada pelo operador

$$\hat{x}_i = \frac{\mathbb{1} - \sigma_i^z}{2},$$

que satisfaz

$$\hat{x}_i |x_i\rangle = x_i |x_i\rangle.$$

Esse operador é diagonal na base computacional e preserva, no espaço de Hilbert, os valores assumidos pela variável binária clássica.

Partindo da expressão clássica definida em (2), o Hamiltoniano de custo é obtido pela substituição $x_i \mapsto \hat{x}_i$, isto é,

$$H_f = \sum_{i,j} Q_{ij} \hat{x}_i \hat{x}_j, \quad (7)$$

onde os coeficientes Q_{ij} são aqueles definidos na formulação clássica do problema.

Essa construção permite reescrever o Hamiltoniano de custo na forma de Ising. De fato, para termos fora da diagonal, isto é, $i \neq j$, tem-se

$$\hat{x}_i \hat{x}_j = \frac{1}{4} (\mathbb{1} - \sigma_i^z - \sigma_j^z + \sigma_i^z \sigma_j^z).$$

Portanto, os termos do QUBO produzem contribuições constantes, termos locais em σ_i^z e acoplamentos entre pares de qubits em $\sigma_i^z \sigma_j^z$. Os termos diagonais, por sua vez, contribuem como termos locais, pois $\hat{x}_i^2 = \hat{x}_i$. Dessa forma, o Hamiltoniano H_f pode ser reorganizado como

$$H_f = \sum_i h_i \sigma_i^z + \sum_{i < j} J_{ij} \sigma_i^z \sigma_j^z + C_0 \mathbb{1},$$

onde h_i representam campos locais, J_{ij} representam acoplamentos de Ising e C_0 é uma constante aditiva determinada pelos elementos da matriz Q . Como C_0 apenas desloca todos os autovalores pela mesma quantidade, ele não altera a configuração que minimiza a energia.

A solução do problema pode então ser formulada como a minimização do valor esperado do Hamiltoniano de custo,

$$\min_{|\psi\rangle} \langle \psi | H_f | \psi \rangle. \quad (8)$$

Caso existam múltiplas configurações ótimas, o menor autovalor de H_f será degenerado, e o subespaço fundamental será gerado pelos estados computacionais associados a essas soluções. Assim, as formulações QUBO, Ising e Hamiltoniana descrevem o mesmo problema de otimização em linguagens distintas: variáveis binárias, variáveis de spin e operadores quânticos diagonais, respectivamente (6, 22).

Ao longo dos capítulos seguintes, o Hamiltoniano de custo também será denotado por H_p , por representar o Hamiltoniano de problema utilizado nos algoritmos quânticos de otimização.

2.4 O problema MaxCut

O problema MaxCut constitui um dos problemas clássicos de otimização combinatória definidos sobre grafos. Dado um grafo não direcionado $G = (V, E)$, com conjunto de vértices V e conjunto de arestas E , o objetivo consiste em particionar os vértices em dois subconjuntos disjuntos, S e T , de forma que o número de arestas entre S e T seja o maior possível (20). Existe também uma versão mais geral do problema, denominada *weighted MaxCut*, na qual cada aresta possui um peso associado. Nesse caso, o objetivo deixa de ser maximizar apenas o número de arestas cortadas e passa a ser maximizar o peso total das arestas que conectam vértices pertencentes a subconjuntos diferentes.

Formalmente, considera-se uma variável binária associada a cada vértice do grafo,

$$x_i \in \{0, 1\},$$

que indica a qual subconjunto o vértice i pertence. Uma aresta (i, j) contribui para o corte sempre que os vértices i e j pertencem a conjuntos diferentes, isto é, quando $x_i \neq x_j$.

A função objetivo do problema pode ser escrita como

$$C_{\text{MaxCut}}(x) = \sum_{(i,j) \in E} w_{ij}(x_i + x_j - 2x_i x_j), \quad (9)$$

onde w_{ij} representa o peso associado à aresta (i, j) . Essa expressão assume valor máximo quando o peso total das arestas conectando vértices em partições distintas é maximizado. No caso não ponderado, em que $w_{ij} = 1$ para todas as arestas, isso equivale a maximizar o número de arestas cortadas.

Como a formulação QUBO foi apresentada como um problema de minimização, o MaxCut pode ser incorporado a essa estrutura minimizando-se $-C_{\text{MaxCut}}(x)$, ou, de forma equivalente, escolhendo uma convenção de sinal apropriada para o Hamiltoniano de custo. Essa formulação torna o MaxCut um dos exemplos mais utilizados em estudos de algoritmos de otimização quântica (6, 4).

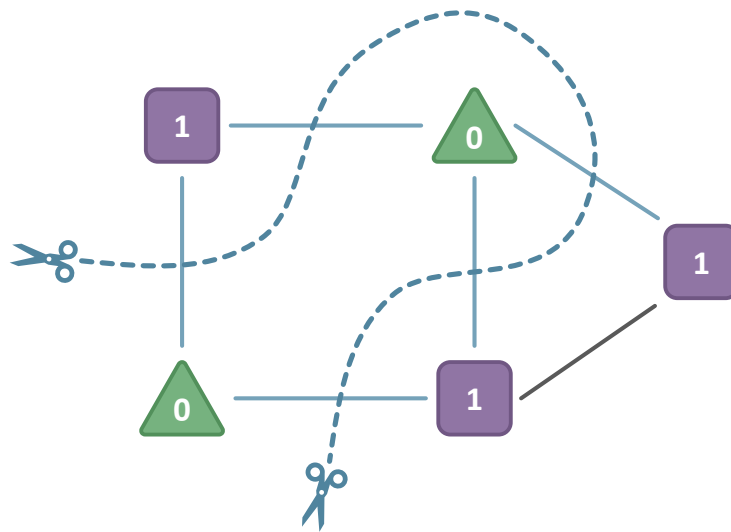


Figura 1 – Representação esquemática do problema MaxCut em um grafo.

Fonte: extraída de (16).

Na Figura 1, apresenta-se uma ilustração esquemática do problema MaxCut. Os vértices são distribuídos em duas partições, e as arestas que cruzam essa separação correspondem às contribuições efetivas para o valor do corte.

Uma propriedade importante do MaxCut é a presença de simetrias e degenerescências. Se uma partição dos vértices em dois subconjuntos, S e T , define um corte ótimo, então a troca global dessas partições, isto é, $S \leftrightarrow T$, produz uma solução equivalente com o mesmo valor da função objetivo. Assim, cada solução possui naturalmente uma configuração simétrica associada. Do ponto de vista Hamiltoniano, essa propriedade se manifesta como uma degenerescência do estado fundamental, pois diferentes estados computacionais podem estar associados ao mesmo menor autovalor. Essa característica é relevante para algoritmos quânticos de otimização, uma vez que o objetivo não é necessariamente preparar uma configuração ótima específica, mas qualquer estado pertencente ao subespaço de menor energia.

Além de sua relevância teórica, o MaxCut possui aplicações em diversas áreas, incluindo clustering, visão computacional, design de circuitos e problemas de segmentação em grafos. Sua importância também se deve ao fato de ser um problema NP-difícil, amplamente utilizado como benchmark no desenvolvimento e avaliação de algoritmos clássicos e quânticos de otimização.

3 Fundamentos de Computação e Informação Quânticas

No Capítulo 2, foram apresentados os fundamentos de otimização combinatória e as formulações QUBO/Ising, com ênfase na reescrita de problemas clássicos em termos de Hamiltonianos de custo. Para compreender como tais Hamiltonianos podem ser explorados por algoritmos quânticos de otimização, torna-se necessário introduzir os conceitos fundamentais da computação e da informação quânticas que servirão de base para os capítulos seguintes (1, 2). Em particular, os algoritmos estudados nesta dissertação dependem diretamente do formalismo de estados quânticos, da evolução unitária, da medição de observáveis e da representação de circuitos parametrizados, elementos centrais na literatura de computação quântica (1, 2). Além disso, a formulação desses métodos também requer noções de controle quântico, especialmente controle por Lyapunov, que constitui a base conceitual do FALQON e de suas variantes (10, 12).

Este capítulo tem, portanto, o objetivo de apenas retomar os principais conceitos de computação e informação quânticas que serão utilizados ao longo do trabalho. Em vez de apresentar uma exposição exaustiva da mecânica quântica, busca-se reunir, de forma organizada, os elementos mais relevantes para a compreensão dos algoritmos variacionais (9, 8, 25) e de feedback estudados nesta dissertação.

3.1 Estados quânticos e notação bra-ket

A formulação matemática da mecânica quântica descreve o estado de um sistema físico por meio de vetores pertencentes a um espaço de Hilbert complexo. Na notação de Dirac, um vetor de estado é representado por um *ket* $|\psi\rangle$, enquanto seu dual é representado pelo correspondente *bra* $\langle\psi|$. O produto interno entre dois estados $|\phi\rangle$ e $|\psi\rangle$ é denotado por

$$\langle\phi|\psi\rangle, \tag{10}$$

e fornece uma quantidade complexa cuja magnitude está diretamente associada a amplitudes de probabilidade (1, 2).

Um estado físico puro deve ser normalizado, isto é,

$$\langle \psi | \psi \rangle = 1. \quad (11)$$

Essa condição garante que a soma das probabilidades de todos os resultados possíveis de uma medição seja unitária. Além disso, dois vetores que diferem apenas por uma fase global complexa representam o mesmo estado físico. Em outras palavras, os estados $|\psi\rangle$ e $e^{i\gamma} |\psi\rangle$, com $\gamma \in \mathbb{R}$, são fisicamente indistinguíveis (1).

A unidade básica de informação quântica é o *qubit*. Diferentemente de um bit clássico, que só pode assumir os valores 0 ou 1, um qubit pode estar em uma superposição linear desses dois estados básicos. Escolhendo-se a base computacional $\{|0\rangle, |1\rangle\}$, um estado arbitrário de um único qubit pode ser escrito como

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad (12)$$

onde $\alpha, \beta \in \mathbb{C}$ e

$$|\alpha|^2 + |\beta|^2 = 1. \quad (13)$$

As quantidades $|\alpha|^2$ e $|\beta|^2$ correspondem, respectivamente, às probabilidades de se obter os resultados 0 e 1 em uma medição na base computacional (1, 2).

Desprezando-se a fase global, o estado de um qubit também pode ser parametrizado por dois ângulos reais, usualmente associados à esfera de Bloch, na forma

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle, \quad (14)$$

com $\theta \in [0, \pi]$ e $\phi \in [0, 2\pi)$. Essa representação evidencia que, embora o espaço vetorial associado ao qubit seja complexo bidimensional, os estados físicos puros podem ser representados geometricamente sobre a superfície da esfera de Bloch (1).

Uma característica central dos sistemas quânticos é o princípio da superposição: se $|\psi_1\rangle$ e $|\psi_2\rangle$ são estados válidos, então qualquer combinação linear normalizada

$$|\chi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle \quad (15)$$

também representa um estado físico admissível. Essa propriedade está na base da expressividade dos circuitos quânticos parametrizados, pois permite que a evolução do sistema explore simultaneamente múltiplas configurações da base computacional (2, 25).

Para sistemas compostos, o espaço de Hilbert total é obtido pelo produto tensorial dos espaços individuais. Assim, se $\mathcal{H}_A$ e $\mathcal{H}_B$ descrevem dois subsistemas, então o sistema conjunto é descrito por

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (16)$$

No caso de n qubits, o espaço total possui dimensão 2^n , e a base computacional é formada pelos vetores

$$|x\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_n\rangle, \quad x_i \in \{0, 1\}. \quad (17)$$

Equivalentemente, costuma-se escrever $|x\rangle = |x_1 x_2 \cdots x_n\rangle$, onde $x_1 x_2 \cdots x_n$ é uma cadeia binária. Um estado geral de n qubits assume então a forma

$$|\psi\rangle = \sum_{x \in \{0,1\}^n} c_x |x\rangle, \quad \sum_{x \in \{0,1\}^n} |c_x|^2 = 1. \quad (18)$$

Um estado puro de um sistema composto é dito separável quando pode ser escrito como um produto tensorial de estados individuais. Caso contrário, o estado é emaranhado. O emaranhamento constitui um recurso essencial da computação quântica, pois introduz correlações não clássicas entre os subsistemas (1, 2).

Embora muitos algoritmos sejam descritos em termos de estados puros, é conveniente introduzir também o operador densidade, que fornece uma formulação mais geral para estados puros e mistos. Para um estado puro $|\psi\rangle$, define-se

$$\rho = |\psi\rangle \langle \psi|. \quad (19)$$

Nesse caso, ρ satisfaz

$$\rho^\dagger = \rho, \quad \rho \geq 0, \quad \text{Tr}(\rho) = 1, \quad \rho^2 = \rho. \quad (20)$$

Já um estado misto é descrito por uma combinação convexa de estados puros,

$$\rho = \sum_k p_k |\psi_k\rangle \langle \psi_k|, \quad p_k \geq 0, \quad \sum_k p_k = 1, \quad (21)$$

na qual a condição $\rho^2 = \rho$ deixa, em geral, de ser válida (26, 1).

A formulação em termos de ρ será útil mais adiante porque grandezas físicas mensuráveis, como valores esperados e probabilidades de resultados, podem ser escritas de forma unificada nesse formalismo.

3.2 Operadores unitários e evolução temporal

Em mecânica quântica, a evolução temporal de um sistema isolado é descrita por operadores unitários. Um operador U é dito unitário quando satisfaz

$$U^\dagger U = U U^\dagger = \mathbb{1}, \quad (22)$$

onde $U^\dagger$ denota o adjunto de U . A unitariedade assegura a preservação da norma do vetor de estado e, conseqüentemente, a conservação da probabilidade total (1, 2).

Grandezas físicas mensuráveis são representadas por operadores hermitianos. Um operador O é hermitiano se

$$O^\dagger = O. \quad (23)$$

Pelo teorema espectral, um operador hermitiano admite decomposição em autovalores reais e projetores ortogonais, isto é,

$$O = \sum_m o_m \Pi_m, \quad (24)$$

onde $o_m \in \mathbb{R}$ são os autovalores e Π_m são projetores sobre os subespaços próprios correspondentes (27, 1). No contexto desta dissertação, um caso especialmente importante é o Hamiltoniano H , que representa a energia do sistema e, simultaneamente, o gerador da evolução temporal.

Para um sistema fechado, a dinâmica do vetor de estado é governada pela equação de Schrödinger

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle. \quad (25)$$

Em grande parte da literatura de computação quântica, adota-se a convenção $\hbar = 1$, o que simplifica a expressão para

$$i \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle. \quad (26)$$

Se o Hamiltoniano é independente do tempo, a solução formal é dada por

$$|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle, \quad (27)$$

onde o operador

$$U(t) = e^{-iHt} \quad (28)$$

é unitário. Quando $H(t)$ depende explicitamente do tempo, a solução formal envolve o operador de ordenação temporal $\mathcal{T}$,

$$|\psi(t)\rangle = \mathcal{T} \exp \left(-i \int_0^t H(s) ds \right) |\psi(0)\rangle. \quad (29)$$

O operador $\mathcal{T}$ organiza os Hamiltonianos na expansão da exponencial de acordo com sua ordem temporal, sendo necessário porque, em geral, Hamiltonianos avaliados em instantes distintos não comutam.

Como H é hermitiano, pode-se escrever

$$H = \sum_j E_j |E_j\rangle \langle E_j|, \quad (30)$$

onde E_j são autovalores reais e $|E_j\rangle$ são autovetores ortonormais. Se o estado inicial for expandido nessa base,

$$|\psi(0)\rangle = \sum_j c_j |E_j\rangle, \quad (31)$$

então a evolução temporal assume a forma

$$|\psi(t)\rangle = \sum_j c_j e^{-iE_j t} |E_j\rangle. \quad (32)$$

Essa expressão mostra que a dinâmica unitária preserva os módulos $|c_j|$, alterando apenas as fases relativas entre os componentes espectrais do estado (1, 2).

No contexto de problemas de otimização, o menor autovalor de H corresponde à energia fundamental. Como discutido no Capítulo 2, quando o problema é codificado em um Hamiltoniano de custo, o estado fundamental passa a representar uma solução ótima ou uma superposição de soluções ótimas degeneradas.

3.3 Medição quântica e valores esperados

A medição quântica conecta a descrição matemática do estado às quantidades observáveis experimentalmente. De modo geral, toda medida de uma grandeza física associada a um observável tem como resultado um de seus autovalores e, após a medição, o estado do sistema colapsa para o autovetor associado ao autovalor observado (1, 2). Em sua forma projetiva, a medição de um observável hermitiano O , decomposto conforme a Equação (24), produz o resultado o_m com probabilidade dada pela regra de Born,

$$p_m = \langle \psi | \Pi_m | \psi \rangle. \quad (33)$$

Caso o resultado o_m seja obtido, o estado pós-medição é

$$|\psi_m\rangle = \frac{\Pi_m |\psi\rangle}{\sqrt{p_m}}, \quad p_m \neq 0. \quad (34)$$

Essa transformação explícita que a medição não apenas revela informação sobre o sistema, mas também altera seu estado (1, 2, 27).

A quantidade central para algoritmos variacionais e de feedback não é, em geral, o resultado de uma única medição, mas o valor esperado de um observável. Para um estado puro $|\psi\rangle$, o valor esperado de O é dado por

$$\langle O \rangle_\psi = \langle \psi | O | \psi \rangle. \quad (35)$$

Utilizando a decomposição espectral de O , tem-se também

$$\langle O \rangle_\psi = \sum_m o_m p_m. \quad (36)$$

Portanto, o valor esperado é uma média ponderada dos possíveis resultados de medição.

No formalismo de operador densidade, a expressão correspondente é

$$\langle O \rangle_\rho = \text{Tr}(\rho O). \quad (37)$$

Essa forma é particularmente útil em cenários experimentais ruidosos ou quando se trabalha com estados mistos.

Na prática, computadores quânticos atuais retornam amostras de resultados de medição, e não valores esperados exatos. Assim, para estimar $\langle O \rangle$, o circuito é repetido diversas vezes, produzindo uma média empírica. Se N_s repetições independentes forem realizadas, uma estimativa amostral pode ser escrita como

$$\widehat{\langle O \rangle} = \frac{1}{N_s} \sum_{r=1}^{N_s} o^{(r)}, \quad (38)$$

onde $o^{(r)}$ é o resultado da r -ésima execução. Sob hipóteses usuais, o erro estatístico decai tipicamente como $N_s^{-1/2}$, o que explica o papel central do número de *shots* na qualidade das estimativas experimentais (9, 25).

Esse ponto é especialmente relevante para algoritmos de otimização quântica, pois tanto a função custo quanto quantidades de feedback são, em geral, acessadas via médias de medições repetidas. Em particular, no FALQON, o termo de realimentação depende da estimação de um observável associado a um comutador entre Hamiltonianos (10, 12).

3.4 Qubit, portas e circuitos quânticos

O modelo de circuitos quânticos descreve a computação como uma sequência discreta de operações unitárias aplicadas a um registrador de qubits, seguida por medições. Tal modelo é particularmente adequado à implementação de algoritmos variacionais e de feedback em dispositivos quânticos digitais (1, 2).

Na representação vetorial, os estados da base computacional de um único qubit são definidos por

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (39)$$

Entre as operações elementares mais importantes estão os operadores de Pauli,

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (40)$$

que atuam sobre a base computacional como

$$X |0\rangle = |1\rangle, \quad X |1\rangle = |0\rangle, \quad (41)$$

$$Z|0\rangle = |0\rangle, \quad Z|1\rangle = -|1\rangle. \quad (42)$$

A porta de Hadamard, denotada por H , é dada por

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (43)$$

e satisfaz

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} = |+\rangle, \quad H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |-\rangle. \quad (44)$$

A ação de H ilustra como superposições podem ser geradas a partir de estados da base computacional.

Também são amplamente utilizadas as rotações em torno dos eixos x , y e z , definidas por

$$R_\alpha(\theta) = e^{-i\theta\sigma_\alpha/2}, \quad \alpha \in \{x, y, z\}, \quad (45)$$

onde $\sigma_x = X$, $\sigma_y = Y$ e $\sigma_z = Z$. Tais operações dependem de parâmetros contínuos e, por isso, são particularmente adequadas para a construção de circuitos parametrizados (28, 1).

A computação quântica torna-se expressivamente mais poderosa quando se permite a atuação de portas multiqubit. Um exemplo fundamental é a porta CNOT, que atua como

$$\text{CNOT}|c, t\rangle = |c, t \oplus c\rangle, \quad (46)$$

onde c é o qubit de controle, t é o qubit alvo e $\oplus$ denota soma módulo 2. Em particular,

$$\text{CNOT}|00\rangle = |00\rangle, \quad \text{CNOT}|01\rangle = |01\rangle, \quad \text{CNOT}|10\rangle = |11\rangle, \quad \text{CNOT}|11\rangle = |10\rangle. \quad (47)$$

A combinação de uma porta de Hadamard em um qubit seguida de uma CNOT permite produzir estados emaranhados. Por exemplo,

$$|00\rangle \xrightarrow{H \otimes \mathbb{1}} \frac{|00\rangle + |10\rangle}{\sqrt{2}} \xrightarrow{\text{CNOT}} \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad (48)$$

Esse é um dos estados de Bell e constitui um exemplo de estado emaranhado, pois não pode ser escrito na forma

$$|\psi_A\rangle \otimes |\psi_B\rangle,$$

em que $|\psi_A\rangle$ e $|\psi_B\rangle$ representam estados individuais dos dois qubits. Embora o estado seja uma combinação linear dos estados $|00\rangle$ e $|11\rangle$, não é possível fatorá-lo como um produto tensorial de estados de um único qubit. Essa propriedade evidencia a presença de correlações quânticas entre os subsistemas (1, 2).

Um circuito quântico pode ser entendido como a composição ordenada de portas unitárias,

$$U = U_L U_{L-1} \cdots U_2 U_1, \quad (49)$$

de modo que, a partir de um estado inicial $|\psi_0\rangle$, o estado final é

$$|\psi_f\rangle = U |\psi_0\rangle. \quad (50)$$

Em algoritmos variacionais, diversas dessas portas dependem de parâmetros ajustáveis, levando à forma

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta}) |\psi_0\rangle, \quad (51)$$

onde $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$ é o vetor de parâmetros. O problema algorítmico consiste então em escolher $\boldsymbol{\theta}$ de modo a minimizar, maximizar ou controlar determinada função objetivo, frequentemente definida como o valor esperado de um Hamiltoniano de custo (5, 9, 8, 25).

Essa estrutura será retomada nas seções finais do presente capítulo, quando forem discutidas as conexões entre circuitos parametrizados, algoritmos variacionais e estratégias de feedback.

3.5 Formalismo de Hamiltonianos

O Hamiltoniano ocupa papel central tanto na mecânica quântica quanto nos algoritmos de otimização quântica. Do ponto de vista físico, ele representa a energia do sistema; do ponto de vista dinâmico, ele gera a evolução temporal; e, do ponto de vista algorítmico, ele define a função custo a ser minimizada (1, 2).

Qualquer operador hermitiano atuando sobre n qubits pode ser expandido na base ortogonal formada pelas cadeias de Pauli (1, 9),

$$H = \sum_{\mu} h_{\mu} P_{\mu}, \quad (52)$$

onde $h_\mu \in \mathbb{R}$ e cada P_μ pertence ao conjunto

$$P_\mu \in \{\mathbb{1}, X, Y, Z\}^{\otimes n}. \quad (53)$$

Essa decomposição é importante porque valores esperados de H podem ser obtidos a partir da medição de valores esperados das cadeias de Pauli que o compõem. Em algoritmos variacionais, isso permite escrever a função custo como uma soma de termos localmente mensuráveis (9, 8).

Conforme derivado na Seção 2.3, uma função custo QUBO pode ser reescrita como um Hamiltoniano de Ising na representação de Pauli- Z . Nesse caso, o Hamiltoniano de problema assume a forma

$$H_p = \sum_i h_i Z_i + \sum_{i < j} J_{ij} Z_i Z_j + c \mathbb{1}, \quad (54)$$

onde h_i , J_{ij} e c são determinados pelos coeficientes da formulação QUBO. Seus autovalores codificam os valores da função custo clássica associada ao problema.

Entretanto, se o sistema fosse preparado em um autovetor de H_p , a dinâmica gerada apenas por esse Hamiltoniano produziria somente fases, sem alterar as probabilidades na base computacional. Para induzir transições entre configurações e explorar o espaço de soluções, introduz-se um Hamiltoniano auxiliar, frequentemente denominado *driver* ou *mixer*, denotado por H_d .

Uma escolha comum é o Hamiltoniano de campo transversal,

$$H_d = \sum_{i=1}^n X_i, \quad (55)$$

composto por operadores de um qubit que promovem transições entre os estados $|0\rangle$ e $|1\rangle$. Essa, contudo, não é a única possibilidade. Dependendo da estrutura e das restrições do problema, H_d pode incluir também interações entre dois ou mais qubits, como termos do tipo $X_i X_j + Y_i Y_j$, utilizados para preservar determinados subespaços de soluções viáveis (29, 10).

A condição de não comutatividade

$$[H_p, H_d] \neq 0 \quad (56)$$

é essencial para que a ação conjunta de ambos os Hamiltonianos produza uma dinâmica não trivial. Se H_p e H_d comutassem, existiria uma base comum de autovetores, e a evolução tenderia a perder a capacidade de explorar eficientemente o espaço de soluções.

Se E_0 é o menor autovalor de um Hamiltoniano hermitiano H , então, para qualquer estado normalizado $|\psi\rangle$, vale

$$E_0 \leq \langle \psi | H | \psi \rangle. \quad (57)$$

A igualdade ocorre quando $|\psi\rangle$ pertence ao subespaço fundamental de H . Esse resultado, conhecido como princípio variacional, fundamenta diretamente os algoritmos de otimização quântica: ao procurar estados que minimizam $\langle \psi | H_p | \psi \rangle$, busca-se aproximar o estado fundamental do Hamiltoniano de problema (5, 9, 25).

Tal princípio conecta de forma natural a formulação Hamiltoniana apresentada no capítulo anterior aos métodos quânticos discutidos nos capítulos seguintes.

3.6 Conceitos básicos de controle quântico

Controle quântico consiste em manipular a dinâmica de um sistema quântico por meio de campos externos ou parâmetros ajustáveis, de forma a conduzir o sistema a um objetivo desejado. Esse objetivo pode ser, por exemplo, preparar um estado-alvo, implementar uma operação unitária específica, maximizar a fidelidade com determinado subespaço ou minimizar o valor esperado de um Hamiltoniano (30, 31).

Em sua forma mais comum, a dinâmica de um sistema quântico fechado sob controle pode ser escrita como (30, 31)

$$H(t) = H_0 + \sum_{k=1}^m u_k(t) H_k, \quad (58)$$

onde H_0 é o Hamiltoniano livre ou de deriva, H_k são Hamiltonianos de controle e $u_k(t)$ são funções escalares do tempo que representam os sinais de controle aplicados ao sistema. A dinâmica resultante satisfaz

$$i \frac{d}{dt} |\psi(t)\rangle = \left(H_0 + \sum_{k=1}^m u_k(t) H_k \right) |\psi(t)\rangle. \quad (59)$$

No formalismo de operador densidade, a equação equivalente é

$$\frac{d\rho}{dt} = -i[H(t), \rho(t)]. \quad (60)$$

Essa formulação é particularmente conveniente quando se deseja descrever ensembles estatísticos ou incorporar, em extensões futuras, efeitos de ruído e dissipação.

Do ponto de vista conceitual, o controle quântico busca explorar a estrutura dinâmica da mecânica quântica para tornar acessíveis determinadas trajetórias de evolução ou regiões do espaço de estados. Em vez de se trabalhar apenas com a evolução livre imposta pelo Hamiltoniano natural do sistema, introduzem-se termos adicionais capazes de redirecionar essa dinâmica segundo critérios previamente definidos. Em muitos casos, o problema central não é simplesmente evoluir o sistema no tempo, mas fazê-lo evoluir de modo orientado, aproximando-o de uma configuração de interesse físico ou computacional (30).

Entre os objetivos mais comuns em controle quântico, podem ser citados a preparação de estado, a implementação de portas quânticas, a minimização de energia e a maximização de fidelidade. Na presente dissertação, o caso de maior interesse é a minimização do valor esperado de um Hamiltoniano de problema,

$$J(t) = \langle \psi(t) | H_p | \psi(t) \rangle, \quad (61)$$

pois essa quantidade está diretamente associada à qualidade da solução aproximada produzida pelo algoritmo de otimização (10, 12).

Essa perspectiva é particularmente relevante em problemas de otimização quântica. Nesses casos, não se deseja apenas controlar o sistema para atingir um estado arbitrário, mas sim guiá-lo em direção a estados de menor energia associados a um Hamiltoniano de custo. Dessa forma, o controle quântico passa a oferecer uma linguagem natural para interpretar algoritmos em que parâmetros são ajustados de modo a favorecer trajetórias que reduzam progressivamente o valor esperado da função objetivo. Essa conexão será essencial nas seções seguintes, especialmente quando o problema for reinterpretado em termos de algoritmos de feedback inspirados em controle por Lyapunov.

3.7 Introdução ao controle por Lyapunov

O método de Lyapunov é uma ferramenta clássica da teoria de estabilidade de sistemas dinâmicos. Sua ideia central consiste em construir uma função escalar V , chamada função de Lyapunov, que seja não crescente ao longo da trajetória do sistema. Em termos intuitivos, essa função atua como uma medida de progresso da dinâmica em direção ao objetivo desejado: se $V(t)$ é monotonicamente não crescente ao longo do tempo, a evolução do sistema passa a ser guiada por uma direção preferencial (30, 31). No contexto quântico, a adaptação dessa ideia é especialmente natural quando o objetivo é reduzir a energia esperada de um Hamiltoniano de problema, pois isso permite associar diretamente a dinâmica controlada ao processo de busca por estados de menor custo.

Em controle quântico, o método de Lyapunov não consiste em resolver explicitamente um problema global de otimização a cada instante de tempo. Em vez disso, escolhe-se uma função escalar adequada e projeta-se a lei de controle de modo que sua derivada temporal seja não positiva. Com isso, garante-se que a trajetória do sistema evolua de forma monotônica em relação ao funcional escolhido. Essa perspectiva é particularmente relevante para algoritmos de feedback quântico, nos quais a informação extraída de medições é utilizada para atualizar os parâmetros do circuito de forma iterativa, preservando a tendência de redução da função custo (10, 12).

3.7.1 Escolha da função de Lyapunov

Se o objetivo é minimizar a energia associada a um Hamiltoniano H_p , uma escolha natural de função de Lyapunov é

$$V(t) = \langle \psi(t) | H_p | \psi(t) \rangle. \quad (62)$$

Essa escolha é particularmente apropriada em problemas de otimização, pois a diminuição de $V(t)$ corresponde diretamente à redução do valor esperado da função custo.

Considere agora um Hamiltoniano controlado da forma

$$H(t) = H_p + \beta(t)H_d, \quad (63)$$

onde H_p é o Hamiltoniano de problema, H_d é o Hamiltoniano de controle ou driver, e $\beta(t)$ é o sinal escalar de controle. Assumindo H_p e H_d independentes do tempo, a

derivada temporal de $V(t)$ é

$$\frac{dV}{dt} = \frac{d}{dt} \langle \psi(t) | H_p | \psi(t) \rangle. \quad (64)$$

Usando a equação de Schrödinger (26), obtém-se

$$\frac{dV}{dt} = i \langle \psi(t) | [H(t), H_p] | \psi(t) \rangle. \quad (65)$$

Substituindo (63) em (65), segue que

$$\frac{dV}{dt} = i \langle \psi(t) | [H_p + \beta(t)H_d, H_p] | \psi(t) \rangle. \quad (66)$$

Como $[H_p, H_p] = 0$, resta

$$\frac{dV}{dt} = \beta(t) \langle \psi(t) | i[H_d, H_p] | \psi(t) \rangle. \quad (67)$$

Definindo

$$A(t) = \langle \psi(t) | i[H_d, H_p] | \psi(t) \rangle, \quad (68)$$

pode-se escrever

$$\frac{dV}{dt} = \beta(t)A(t). \quad (69)$$

A flexibilidade do método de Lyapunov reside no fato de que o controle $\beta(t)$ pode ser escolhido de modo a tornar dV/dt não positivo. Uma escolha simples e muito utilizada é

$$\beta(t) = -\kappa A(t), \quad \kappa > 0, \quad (70)$$

o que implica

$$\frac{dV}{dt} = -\kappa A(t)^2 \leq 0. \quad (71)$$

Logo, o valor esperado de H_p é monotonicamente não crescente ao longo da dinâmica.

De forma mais geral, pode-se adotar

$$\beta(t) = -\omega f(t, A(t)), \quad \omega > 0, \quad (72)$$

desde que $f(t, 0) = 0$ e $A(t)f(t, A(t)) > 0$ para $A(t) \neq 0$. Nessas condições,

$$\frac{dV}{dt} = -\omega A(t)f(t, A(t)) \leq 0. \quad (73)$$

Essa forma geral mostra que o aspecto essencial não é a linearidade da lei de controle, mas a garantia de que o funcional de Lyapunov decresça ao longo do tempo (30, 12).

No contexto de otimização combinatória, a interpretação de (71) é direta: se $V(t) = \langle H_p \rangle$ decresce monotonicamente, então a dinâmica conduz o sistema a estados de energia cada vez menor. Isso não implica, em geral, garantia automática de alcance do estado fundamental global em tempo finito, mas fornece um princípio construtivo para o desenho de algoritmos que melhoram sistematicamente a estimativa da solução à medida que a dinâmica evolui (10, 12).

Além disso, a condição estacionária

$$A(t) = \langle \psi(t) | i[H_d, H_p] | \psi(t) \rangle = 0 \quad (74)$$

indica que a dinâmica alcançou um ponto em que, sob a lei de controle escolhida, a redução instantânea de $V(t)$ se anula. Em linguagem de controle, a trajetória aproxima-se de um conjunto invariante; em linguagem algorítmica, esse comportamento está relacionado à saturação da melhoria obtida pelo protocolo.

3.8 Relação entre controle quântico e algoritmos

Embora o controle quântico seja frequentemente formulado em tempo contínuo, a implementação em computadores quânticos digitais exige uma discretização temporal. Essa passagem do contínuo para o discreto é precisamente o ponto em que uma estratégia de controle passa a adquirir a forma de um algoritmo quântico executável em camadas de circuito (4, 10).

Suponha-se que o Hamiltoniano controlado seja dado por

$$H(t) = H_p + \beta(t)H_d, \quad (75)$$

onde H_p é o Hamiltoniano de problema, H_d é o Hamiltoniano *driver* e $\beta(t)$ é um parâmetro de controle dependente do tempo. Dividindo-se a evolução total em intervalos discretos de duração Δt e assumindo que o controle permaneça constante, com valor β_k , durante o k -ésimo intervalo, a evolução correspondente é dada por

$$U_k = e^{-i(H_p + \beta_k H_d)\Delta t}. \quad (76)$$

A fórmula de Lie–Trotter, correspondente à decomposição de Trotter–Suzuki de primeira ordem, permite aproximar essa evolução por (32, 33)

$$e^{-i(H_p+\beta_k H_d)\Delta t} = \left(e^{-i\beta_k H_d \Delta t/r} e^{-iH_p \Delta t/r}\right)^r + \mathcal{O}\left(\frac{\Delta t^2}{r}\right), \quad (77)$$

onde r representa o número de subdivisões do intervalo. O erro dominante está associado à não comutatividade entre H_p e H_d e depende do comutador $\beta_k[H_d, H_p]$ (34). Em cada subdivisão, o erro local é da ordem de $\mathcal{O}((\Delta t/r)^2)$ e, após as r subdivisões, o erro acumulado no intervalo é da ordem de $\mathcal{O}(\Delta t^2/r)$.

No caso utilizado na construção do circuito, considera-se uma única aplicação por intervalo, isto é, $r = 1$, resultando em

$$e^{-i(H_p+\beta_k H_d)\Delta t} = e^{-i\beta_k H_d \Delta t} e^{-iH_p \Delta t} + \mathcal{O}(\Delta t^2). \quad (78)$$

Definindo

$$U_p = e^{-iH_p \Delta t}, \quad U_d(\beta_k) = e^{-i\beta_k H_d \Delta t}, \quad (79)$$

a evolução após ℓ etapas pode ser escrita como

$$U_\ell = U_d(\beta_\ell)U_p \cdots U_d(\beta_2)U_p U_d(\beta_1)U_p. \quad (80)$$

Aplicada a um estado inicial $|\psi_0\rangle$, essa sequência produz

$$|\psi_\ell\rangle = U_\ell |\psi_0\rangle. \quad (81)$$

Essa estrutura já possui a forma típica de um circuito quântico em camadas: em cada passo, aplica-se uma evolução associada ao Hamiltoniano de problema e uma evolução associada ao Hamiltoniano driver, parametrizada por um coeficiente que pode ser ajustado conforme a estratégia de controle adotada. Em outras palavras, uma dinâmica contínua controlada pode ser reinterpretada como uma sequência discreta de operações unitárias implementáveis em hardware quântico digital.

No regime discreto, a quantidade associada ao feedback passa a ser avaliada camada a camada. Define-se

$$A_k = \langle \psi_k | i[H_d, H_p] | \psi_k \rangle, \quad (82)$$

de modo que uma regra de atualização inspirada no controle por Lyapunov pode assumir a forma

$$\beta_{k+1} = -\kappa A_k, \quad (83)$$

com $\kappa > 0$. A Equação (82) fornece a definição matemática de A_k . Em uma simulação por vetor de estado, essa quantidade pode ser calculada diretamente a partir de $|\psi_k\rangle$. Em um dispositivo quântico, entretanto, o estado não é conhecido explicitamente, e o valor esperado do observável $i[H_d, H_p]$ deve ser estimado a partir de medições repetidas do circuito correspondente à camada k . Assim, na implementação experimental, utiliza-se uma estimativa $\hat{A}_k$, sujeita a erro estatístico devido ao número finito de *shots*, e a regra de atualização passa a ser implementada como

$$\beta_{k+1} = -\kappa \hat{A}_k. \quad (84)$$

Em consequência, o algoritmo realiza uma realimentação baseada em valores esperados estimados estatisticamente (10, 12).

A importância conceitual dessa construção está no fato de que os parâmetros do circuito deixam de ser determinados por um processo de otimização clássica global. Em cada camada, o valor de A_k , definido pela Equação (82), é avaliado numericamente: de forma direta em simulações por vetor de estado ou por meio de estimativas obtidas a partir de medições em um dispositivo quântico. Esse valor é então utilizado pela regra de feedback para determinar o parâmetro da camada seguinte. Assim, um problema originalmente formulado na linguagem de controle é convertido em um protocolo algorítmico composto por etapas sucessivas de preparação do estado, aplicação de operações unitárias, avaliação de observáveis e atualização dos parâmetros.

Desse modo, a discretização da dinâmica controlada transforma um problema contínuo de controle em um algoritmo digital executável em hardware quântico. Essa é a ponte conceitual entre controle quântico e algoritmos de feedback para otimização, e é justamente essa ponte que fundamenta a formulação do FALQON e de suas variantes como estratégias de otimização quântica guiadas por realimentação.

3.9 Conexão com algoritmos variacionais e de feedback

Os algoritmos quânticos de otimização em arquitetura NISQ frequentemente utilizam circuitos parametrizados rasos e uma função custo definida por um valor esperado (5, 9, 8, 25). Nesse sentido, há uma relação estrutural clara entre algoritmos variacionais, como VQE e QAOA, e algoritmos baseados em feedback, como FALQON. A diferença central reside no modo como os parâmetros são escolhidos.

No paradigma variacional, considera-se um sistema quântico descrito por um Hamiltoniano H e um estado de teste parametrizado,

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta}) |\psi_0\rangle, \quad (85)$$

onde $U(\boldsymbol{\theta})$ é um operador unitário dependente de um conjunto de parâmetros reais $\boldsymbol{\theta}$ e $|\psi_0\rangle$ é um estado de referência inicial, frequentemente escolhido na base computacional.

A energia associada ao estado variacional é dada por

$$E(\boldsymbol{\theta}) = \langle \psi(\boldsymbol{\theta}) | H | \psi(\boldsymbol{\theta}) \rangle. \quad (86)$$

Os otimizadores clássicos empregados nesse laço híbrido não são específicos da computação quântica. Eles correspondem a métodos gerais de otimização, baseados ou não em gradientes, que podem ser utilizados em uma ampla classe de problemas, desde que sejam respeitadas as propriedades da função objetivo, as restrições existentes e as hipóteses de aplicação de cada método. No contexto dos algoritmos variacionais quânticos, a particularidade está no fato de que os valores da função objetivo são estimados a partir de medições realizadas em um circuito quântico, enquanto a atualização dos parâmetros é executada por um otimizador clássico (9, 8, 25).

De acordo com o princípio variacional, esse valor esperado é limitado inferiormente pela energia E_g do estado fundamental de H ,

$$E(\boldsymbol{\theta}) \geq E_g. \quad (87)$$

Dessa forma, ao ajustar os parâmetros $\boldsymbol{\theta}$ para minimizar $E(\boldsymbol{\theta})$, obtém-se uma aproximação para a energia fundamental,

$$\min_{\boldsymbol{\theta}} E(\boldsymbol{\theta}) \approx E_g. \quad (88)$$

No contexto desta dissertação, toma-se $H = H_p$, de modo que a minimização do valor esperado do Hamiltoniano de problema corresponde à busca por uma solução de baixa energia. Esse procedimento constitui o princípio subjacente ao VQE (5, 9, 25).

No caso do QAOA, o *ansatz* assume uma forma alternante particularmente próxima ao contexto desta dissertação:

$$|\gamma, \beta\rangle = \left(\prod_{k=1}^p e^{-i\beta_k H_d} e^{-i\gamma_k H_p} \right) |\psi_0\rangle, \quad (89)$$

onde p representa o número de camadas alternadas do circuito. Na motivação original do QAOA, essas camadas estão relacionadas à discretização de uma evolução adiabática, produzindo uma sequência alternada de evoluções geradas pelos Hamiltonianos de problema e *driver* (4, 29). Na formulação variacional, entretanto, p é tratado como um hiperparâmetro que controla a profundidade do *ansatz*. Os vetores

$$\gamma = (\gamma_1, \dots, \gamma_p) \quad \text{e} \quad \beta = (\beta_1, \dots, \beta_p)$$

contêm os parâmetros variacionais associados, respectivamente, às evoluções geradas por H_p e H_d . A semelhança estrutural com a Equação (80) decorre da alternância entre essas duas evoluções. A principal diferença é que, no QAOA, os parâmetros são determinados por um processo de otimização clássica externa (4, 29, 8).

Em algoritmos de feedback, os parâmetros do circuito não são determinados por um procedimento de otimização clássica global, mas por uma regra explícita de atualização baseada em medições. No FALQON, por exemplo, considera-se uma evolução em camadas e adota-se, em sua forma básica, uma regra do tipo

$$\beta_1 = 0, \quad \beta_{k+1} = -A_k, \quad A_k = \langle \psi_k | i[H_d, H_p] | \psi_k \rangle, \quad (90)$$

com

$$|\psi_k\rangle = U_d(\beta_k)U_p \cdots U_d(\beta_1)U_p |\psi_0\rangle. \quad (91)$$

Essa construção deriva diretamente da ideia de controle por Lyapunov: a informação obtida a partir das medições é utilizada para definir a intensidade da camada seguinte, de modo a induzir redução monotônica do valor esperado de H_p com a profundidade do circuito (10, 12).

Assim, embora FALQON e QAOA utilizem *ansatz* alternantes similares, eles pertencem a paradigmas distintos. O QAOA é um algoritmo variacional híbrido dependente de otimização clássica externa; o FALQON, por sua vez, é um algoritmo de feedback cujo mecanismo de atualização é construído internamente a partir de observáveis do sistema quântico.

A formulação acima também evidencia por que o FALQON admite extensões naturais (35, 36, 37). Uma vez que o núcleo do método está na regra que associa as medições às intensidades de controle, é possível propor modificações no escalonamento temporal, aproximações de ordem superior e estratégias de inicialização mais informadas, preservando a lógica geral de feedback. Nessa direção, trabalhos recentes têm investigado aceleração por reescalonamento temporal (15) e aproximações quadráticas para redução de profundidade de circuito (13), além de extensões voltadas a cenários específicos de aplicação (11, 38).

4 Algoritmos de Feedback Quântico

No Capítulo 3, foram apresentados os elementos necessários para descrever a ideia de controle por Lyapunov em sistemas quânticos e sua reinterpretação, após discretização temporal, como um circuito em camadas. A partir desse ponto, o presente capítulo introduz o FALQON, detalhando sua formulação para problemas de otimização combinatória, sua regra de realimentação, a propriedade de monotonicidade da energia e aspectos práticos de implementação (10, 12). Em seguida, discutem-se limitações do algoritmo em sua forma padrão e duas variantes relevantes da literatura, TR-FALQON (15) e SO-FALQON (13), que buscam reduzir a profundidade de circuito ou acelerar a convergência preservando a lógica geral de feedback.

4.1 Da formulação de controle ao algoritmo FALQON

A ideia central do FALQON pode ser entendida como uma especialização algorítmica do esquema de controle por Lyapunov apresentado no Capítulo 3. Naquela formulação, escolheu-se como função de Lyapunov a energia esperada do Hamiltoniano de problema,

$$V(t) = \langle \psi(t) | H_p | \psi(t) \rangle, \quad (92)$$

e mostrou-se que, para um Hamiltoniano controlado da forma

$$H(t) = H_p + \beta(t)H_d, \quad (93)$$

a derivada temporal de $V(t)$ pode ser escrita como

$$\frac{dV}{dt} = \beta(t)A(t), \quad (94)$$

onde

$$A(t) = \langle \psi(t) | i[H_d, H_p] | \psi(t) \rangle. \quad (95)$$

Assim, uma escolha de controle do tipo $\beta(t) = -A(t)$, ou mais geralmente $\beta(t) = -\omega f(t, A(t))$, com $f(t, 0) = 0$ e $A f(t, A) > 0$ para $A \neq 0$, garante que $dV/dt \leq 0$, de modo que a energia esperada de H_p não aumenta ao longo da evolução (30, 31, 12).

No entanto, no modelo de computação quântica em circuitos, baseado na aplicação discreta de portas quânticas, não se implementa diretamente uma equação diferencial contínua. Em vez disso, executa-se uma sequência finita de operações unitárias. Por essa razão, a dinâmica controlada é discretizada em camadas por meio de uma decomposição trotterizada, como discutido na Seção 3.8. A forma discreta resultante, já introduzida na Equação (80), produz estados sucessivos $|\psi_k\rangle$ a partir da aplicação alternada dos operadores U_p e $U_d(\beta_k)$, associados, respectivamente, aos Hamiltonianos de problema e *driver*. A cada nova camada, uma estimativa do observável $i[H_d, H_p]$ é utilizada para determinar o parâmetro da camada seguinte, conforme a regra introduzida em (90).

Essa passagem do contínuo para o discreto é mais do que uma aproximação numérica. Ela transforma um problema de controle em um protocolo algorítmico iterativo. Em vez de se pensar em uma função de controle contínua $\beta(t)$, passa-se a trabalhar com uma sequência de parâmetros $\{\beta_k\}$. Em vez de acompanhar a derivada temporal de uma função de Lyapunov, acompanha-se a evolução da energia por camada. Em vez de um controle idealmente contínuo, tem-se um mecanismo de realimentação obtido por medições repetidas em um circuito quântico (10, 12).

O FALQON nasce exatamente dessa reinterpretação. Seu princípio consiste em utilizar a estrutura de controle por Lyapunov para definir, de forma construtiva, os parâmetros de um circuito quântico em camadas, de modo que a energia esperada do Hamiltoniano de problema tenda a diminuir à medida que novas camadas são adicionadas. Nesse sentido, o algoritmo pode ser visto como uma ponte entre a formulação do problema em termos de Hamiltonianos, a ideia de realimentação inspirada em controle quântico e a construção iterativa de um *ansatz* em camadas (12, 11).

4.2 Formulação do FALQON para otimização combinatória

4.2.1 Hamiltonianos H_p e H_d

As propriedades gerais dos Hamiltonianos de problema e *driver* foram apresentadas na Seção 3.5. Nesta subseção, são especificadas apenas as escolhas utilizadas na formulação do FALQON para problemas de otimização combinatória.

O Hamiltoniano de problema H_p é obtido a partir da função custo clássica e, para problemas formulados como QUBO/Ising, possui a estrutura diagonal apresentada na Equação (54). Por ser diagonal na base própria dos operadores Z_i , seus autovetores correspondem às cadeias binárias do problema, que podem ser identificadas mediante medições dos qubits na base Z . O subespaço fundamental de H_p codifica, portanto, as soluções ótimas do problema.

Para produzir transições entre essas configurações, o FALQON utiliza um Hamiltoniano *driver* H_d . Considerando um campo transversal, uma forma geral é

$$H_d = \sum_{i=1}^n k_i X_i, \quad (96)$$

onde $k_i \in \mathbb{R}$ determina a intensidade da atuação sobre o i -ésimo qubit. Na formulação padrão, adota-se frequentemente um campo uniforme, com $k_i = 1$,

$$H_d = \sum_{i=1}^n X_i. \quad (97)$$

Os coeficientes k_i também poderiam depender do tempo. Quando essa dependência pode ser fatorada como $k_i(t) = \beta(t)k_i$, ela é incorporada ao único sinal de controle $\beta(t)$ que multiplica o Hamiltoniano H_d fixo. Caso cada coeficiente varie independentemente, obtém-se uma formulação de controle com múltiplos sinais, que ultrapassa a versão padrão do FALQON considerada nesta dissertação. Em todos os casos, é necessário que H_d não comute com H_p , conforme estabelecido na Equação (56) (30, 12).

Para o campo transversal uniforme, uma escolha simples de estado inicial é

$$|\psi_0\rangle = |+\rangle^{\otimes n}, \quad (98)$$

que corresponde a uma superposição uniforme das configurações da base computacional e pode ser preparada pela aplicação de portas de Hadamard a cada qubit. A escolha do estado inicial e a convenção de sinal de H_d devem ser realizadas de forma consistente (10, 12).

Como exemplo, considere o problema MaxCut em um grafo não direcionado $G = (V, E)$. Adotando a convenção de minimização utilizada pelo FALQON, o Hamiltoniano

de problema é definido como

$$H_p^{\text{MaxCut}} = -\frac{1}{2} \sum_{(i,j) \in E} w_{ij} (\mathbb{1} - Z_i Z_j), \quad (99)$$

onde w_{ij} é o peso da aresta (i, j) , sendo $w_{ij} = 1$ no caso não ponderado. Quando os vértices i e j pertencem a partições diferentes, tem-se $Z_i Z_j = -1$, de modo que a aresta contribui com $-w_{ij}$ para a energia. Quando pertencem à mesma partição, $Z_i Z_j = +1$, e sua contribuição é nula. Consequentemente, minimizar o valor esperado de H_p^{MaxCut} equivale a maximizar o peso total das arestas cortadas (10, 12).

O observável utilizado na regra de feedback é obtido pelo comutador entre os dois Hamiltonianos:

$$i[H_d, H_p^{\text{MaxCut}}] = i \left[\sum_{\ell=1}^n X_\ell, -\frac{1}{2} \sum_{(i,j) \in E} w_{ij} (1 - Z_i Z_j) \right]. \quad (100)$$

Como o termo proporcional à identidade comuta com qualquer operador,

$$i[H_d, H_p^{\text{MaxCut}}] = \frac{i}{2} \sum_{(i,j) \in E} w_{ij} \left[\sum_{\ell=1}^n X_\ell, Z_i Z_j \right]. \quad (101)$$

Para cada aresta (i, j) , somente os operadores que atuam sobre os qubits i e j contribuem, de modo que

$$\left[\sum_{\ell=1}^n X_\ell, Z_i Z_j \right] = [X_i, Z_i Z_j] + [X_j, Z_i Z_j]. \quad (102)$$

Como operadores que atuam sobre qubits distintos comutam,

$$[X_i, Z_i Z_j] = [X_i, Z_i] Z_j, \quad [X_j, Z_i Z_j] = Z_i [X_j, Z_j]. \quad (103)$$

Utilizando as relações de comutação das matrizes de Pauli,

$$[X_i, Z_i] = -2iY_i, \quad [X_j, Z_j] = -2iY_j, \quad (104)$$

obtém-se

$$\left[\sum_{\ell=1}^n X_\ell, Z_i Z_j \right] = -2i (Y_i Z_j + Z_i Y_j). \quad (105)$$

Consequentemente,

$$i[H_d, H_p^{\text{MaxCut}}] = \sum_{(i,j) \in E} w_{ij} (Y_i Z_j + Z_i Y_j). \quad (106)$$

O observável resultante é uma soma de cadeias de Pauli de dois corpos, cujos valores esperados podem ser estimados após rotações locais de base que convertem as medições em X e Y em medições equivalentes na base Z , isto é, na base computacional. Termos mutuamente compatíveis também podem ser agrupados para medição conjunta (10, 12).

4.3 Regra de feedback e monotonicidade da energia

4.3.1 Derivação da regra de feedback

Definindo a energia esperada após a k -ésima camada como

$$E_k = \langle \psi_k | H_p | \psi_k \rangle, \quad (107)$$

considere a passagem do estado $|\psi_{k-1}\rangle$ para o estado $|\psi_k\rangle$:

$$|\psi_k\rangle = U_d(\beta_k) U_p |\psi_{k-1}\rangle, \quad (108)$$

em que $\beta_k = -A_{k-1}$ é determinado a partir da estimativa do observável de feedback no estado anterior $|\psi_{k-1}\rangle$.

Definindo, para simplificar a notação,

$$|\phi_{k-1}\rangle = U_p |\psi_{k-1}\rangle, \quad (109)$$

tem-se

$$|\psi_k\rangle = e^{-i\beta_k H_d \Delta t} |\phi_{k-1}\rangle. \quad (110)$$

Logo,

$$E_k = \langle \phi_{k-1} | e^{i\beta_k H_d \Delta t} H_p e^{-i\beta_k H_d \Delta t} | \phi_{k-1} \rangle. \quad (111)$$

Agora, utilizando a expansão de Hadamard para a conjugação unitária, que decorre da fórmula de Baker–Campbell–Hausdorff (39),

$$e^{i\lambda H_d} H_p e^{-i\lambda H_d} = H_p + i\lambda [H_d, H_p] - \frac{\lambda^2}{2} [H_d, [H_d, H_p]] + \mathcal{O}(\lambda^3), \quad (112)$$

com $\lambda = \beta_k \Delta t$, obtém-se

$$E_k = \langle \phi_{k-1} | H_p | \phi_{k-1} \rangle + \beta_k \Delta t \langle \phi_{k-1} | i[H_d, H_p] | \phi_{k-1} \rangle + \mathcal{O}(\Delta t^2). \quad (113)$$

Como $U_p = e^{-iH_p\Delta t}$ é gerado pelo próprio Hamiltoniano H_p , ele comuta com H_p . Portanto,

$$\langle \phi_{k-1} | H_p | \phi_{k-1} \rangle = \langle \psi_{k-1} | U_p^\dagger H_p U_p | \psi_{k-1} \rangle = \langle \psi_{k-1} | H_p | \psi_{k-1} \rangle = E_{k-1}. \quad (114)$$

Além disso, como $|\phi_{k-1}\rangle = U_p |\psi_{k-1}\rangle$ difere de $|\psi_{k-1}\rangle$ por uma evolução de duração Δt , segue que

$$\langle \phi_{k-1} | i[H_d, H_p] | \phi_{k-1} \rangle = A_{k-1} + \mathcal{O}(\Delta t). \quad (115)$$

Substituindo as Equações (114) e (115) na Equação (113), chega-se a

$$E_k - E_{k-1} = \beta_k \Delta t A_{k-1} + \mathcal{O}(\Delta t^2). \quad (116)$$

Com a escolha definida pela regra de feedback,

$$\beta_k = -A_{k-1}, \quad (117)$$

obtém-se

$$E_k - E_{k-1} = -\Delta t A_{k-1}^2 + \mathcal{O}(\Delta t^2). \quad (118)$$

Como o termo dominante é não positivo, conclui-se que, para Δt suficientemente pequeno, a energia esperada não aumenta de uma camada para a seguinte. Em outras palavras, o algoritmo preserva, em sua forma discretizada e até a ordem considerada, a propriedade de monotonicidade associada à dinâmica contínua inspirada em Lyapunov (10, 12).

O termo A_{k-1} desempenha papel central no FALQON. Matematicamente, ele corresponde ao valor esperado do observável hermitiano $i[H_d, H_p]$ no estado $|\psi_{k-1}\rangle$. Essa quantidade determina a sensibilidade, em primeira ordem, da energia de H_p à aplicação do Hamiltoniano *driver*. De fato, conforme a Equação (116), o produto $\beta_k A_{k-1}$ estabelece o sinal da variação dominante da energia na camada seguinte.

Quando $A_{k-1} \neq 0$, a regra $\beta_k = -A_{k-1}$ escolhe o sinal e a intensidade do controle de modo que a contribuição de primeira ordem para a variação da energia seja não positiva. Por outro lado, quando A_{k-1} se aproxima de zero, a atualização β_k também se torna pequena, indicando que o estado se aproxima de uma condição estacionária em relação ao observável de feedback e ao *driver* escolhidos. Isso não implica

necessariamente que o estado fundamental tenha sido alcançado, pois contribuições de ordem superior ou outras direções de controle ainda podem permitir reduções adicionais da energia (12).

4.4 Interpretação física e aspectos práticos de implementação

4.4.1 Interpretação física e computacional

Fisicamente, o FALQON pode ser interpretado como uma sequência de pequenos impulsos de controle aplicados a um sistema quântico com o objetivo de fazê-lo descer no relevo energético definido por H_p . A evolução associada a H_p preserva a energia esperada e, como esse Hamiltoniano é diagonal na base computacional, modifica apenas as fases relativas das componentes nessa base; a evolução associada a H_d , cuja intensidade é ajustada pela realimentação, pode redistribuir as amplitudes e, assim, alterar $\langle \psi | H_p | \psi \rangle$. O papel do feedback é decidir, após cada camada, qual deve ser o próximo impulso do *driver*.

Computacionalmente, isso significa que o FALQON explora o espaço de soluções de forma guiada. O estado inicial fornece uma superposição ampla; as camadas sucessivas vão concentrando amplitude em regiões energeticamente mais promissoras, por meio de um mecanismo interno de correção baseado em medições. O algoritmo combina, assim, elementos de circuitos parametrizados com uma filosofia típica de controle em malha fechada.

Esse caráter em malha fechada é importante porque o FALQON não é apenas um *ansatz* fixo. Cada nova camada depende do histórico da execução, mais precisamente do valor medido de A_k no estado atual. O circuito final, portanto, não está totalmente especificado antes da execução. Ele é construído durante a própria execução, à medida que o sistema fornece a informação necessária para escolher os parâmetros subsequentes (10, 11).

4.4.2 Estimativa de A_k por medições

Na prática experimental, o valor de A_k não é conhecido exatamente. Ele precisa ser estimado por medições repetidas do circuito que prepara $|\psi_k\rangle$. Para isso, explora-se a

decomposição do observável de feedback na base formada por cadeias de Pauli, também denominadas *Pauli strings*:

$$i[H_d, H_p] = \sum_{\mu} a_{\mu} P_{\mu}, \quad (119)$$

onde $a_{\mu} \in \mathbb{R}$ e $P_{\mu} \in \{\mathbb{1}, X, Y, Z\}^{\otimes n}$. Assim,

$$A_k = \sum_{\mu} a_{\mu} \langle \psi_k | P_{\mu} | \psi_k \rangle. \quad (120)$$

Cada termo $\langle \psi_k | P_{\mu} | \psi_k \rangle$ pode ser estimado experimentalmente por amostragem, e o valor de A_k é reconstruído a partir da soma ponderada correspondente (10, 12).

No caso do MaxCut com driver transverso padrão, o observável de feedback é composto por cadeias de Pauli de dois corpos, como na Equação (106). Consequentemente, a estimação de A_k requer a medição de valores esperados de operadores do tipo $Y_i Z_j$ e $Z_i Y_j$, ou equivalentes segundo a convenção escolhida. O número exato de termos depende da estrutura do grafo, mas cresce com o número de arestas do problema, o que já antecipa um custo experimental não desprezível (10, 12).

Seja $\hat{A}_k$ a estimativa obtida a partir de um número finito de *shots*. Então a atualização implementada experimentalmente é, na realidade,

$$\beta_{k+1} = -\hat{A}_k, \quad (121)$$

e não $-A_k$ exato. Portanto, a monotonicidade ideal do algoritmo passa a ser apenas aproximada em cenários ruidosos ou amostrais. Em dispositivos reais, erros estatísticos e ruído experimental podem introduzir flutuações que perturbam o comportamento monotônico previsto pelo modelo ideal.

4.5 Limitações práticas do FALQON

O FALQON é executado iterativamente até que um critério de parada seja satisfeito. Uma possibilidade consiste em interromper o algoritmo quando a energia esperada deixa de apresentar uma redução relevante dentro da precisão estatística disponível.

Uma das principais limitações do FALQON está relacionada à condição de monotonicidade da energia. Como mostrado na Equação (118), a regra de feedback

garante uma contribuição não positiva em primeira ordem para valores suficientemente pequenos de Δt . Quando Δt ultrapassa o regime em que essa contribuição domina a variação da energia, os termos de ordem superior tornam-se relevantes e podem comprometer a diminuição monotônica prevista pela aproximação (12, 13).

Por outro lado, a escolha de passos temporais pequenos tende a produzir variações modestas da energia em cada camada. Como consequência, um número elevado de camadas pode ser necessário para que o estado se aproxime do subespaço fundamental de H_p . A profundidade total do circuito cresce, portanto, com o número de iterações, criando uma tensão entre a preservação da monotonicidade e a necessidade de manter circuitos rasos (10, 12).

Além da profundidade L , o custo físico do circuito depende da decomposição das evoluções em portas elementares. Em dispositivos NISQ, as portas de dois qubits constituem, em geral, um dos principais gargalos de fidelidade e profundidade. Esse custo pode aumentar ainda mais quando a conectividade limitada do processador exige a inserção de portas adicionais de roteamento, como operações *SWAP* (40).

O custo total do algoritmo também inclui as medições necessárias para determinar os parâmetros de realimentação. Em cada etapa, a obtenção de β_k requer uma estimativa de A_{k-1} , cuja precisão depende do número de *shots*. Além disso, o observável $i[H_d, H_p]$ pode decompor-se em diversas cadeias de Pauli, de modo que o custo de medição tende a aumentar com o tamanho e a estrutura da instância. Assim, embora o FALQON dispense uma otimização clássica global dos parâmetros, ele permanece dependente de ciclos sucessivos de preparação do estado, execução do circuito e estimação de observáveis.

Há ainda uma limitação de natureza dinâmica. Em determinadas instâncias, o valor de A_{k-1} pode aproximar-se de zero antes que a distribuição de probabilidade esteja suficientemente concentrada nas soluções ótimas. Nesse regime, a regra $\beta_k = -A_{k-1}$ produz controles progressivamente menores, fazendo com que a evolução se aproxime de uma condição estacionária mesmo quando a probabilidade de sucesso ainda não é elevada (10).

4.6 Variantes da literatura para redução de profundidade

As limitações do FALQON padrão tornam-se particularmente relevantes em dispositivos NISQ, nos quais o aumento da profundidade do circuito intensifica os efeitos de erros de porta, decoerência e ruído de leitura. Nesse contexto, a eficiência prática do método depende simultaneamente do número de camadas, do custo de estimação dos observáveis e da robustez da realimentação frente a erros estatísticos e experimentais (41, 12). A necessidade de obter reduções mais expressivas da energia com circuitos menos profundos motivou o desenvolvimento de variantes como o TR-FALQON (15) e o SO-FALQON (13), apresentadas nas subseções seguintes.

4.6.1 TR-FALQON: reescalonamento temporal

Uma das formas de modificar a dinâmica do FALQON sem abandonar sua lógica de realimentação consiste em reparametrizar o tempo de evolução. Essa é a ideia central do TR-FALQON (*Time-Rescaled FALQON*), proposto como uma adaptação do esquema de controle por Lyapunov a uma escala temporal reparametrizada (15).

Considere a mudança de variável temporal

$$t = f(\tau), \quad (122)$$

onde f é uma função monótona crescente e suficientemente regular. Sua derivada em relação ao tempo reescalado é denotada por

$$f'(\tau) = \frac{df(\tau)}{d\tau} > 0. \quad (123)$$

O estado na nova parametrização temporal é definido por

$$|\tilde{\psi}(\tau)\rangle := |\psi(f(\tau))\rangle. \quad (124)$$

Assim, a variável τ não descreve uma trajetória distinta no espaço de estados, mas uma reparametrização da trajetória original.

Se a dinâmica original é governada por

$$i \frac{d}{dt} |\psi(t)\rangle = (H_p + \beta(t)H_d) |\psi(t)\rangle, \quad (125)$$

então,

$$\frac{d}{d\tau} |\tilde{\psi}(\tau)\rangle = f'(\tau) \left. \frac{d}{dt} |\psi(t)\rangle \right|_{t=f(\tau)}. \quad (126)$$

Definindo a função de controle no tempo reescalado como

$$\tilde{\beta}(\tau) = \beta(f(\tau)), \quad (127)$$

obtém-se

$$i \frac{d}{d\tau} |\tilde{\psi}(\tau)\rangle = f'(\tau) (H_p + \tilde{\beta}(\tau) H_d) |\tilde{\psi}(\tau)\rangle. \quad (128)$$

O Hamiltoniano efetivo associado à evolução na variável τ é, portanto,

$$\tilde{H}(\tau) = f'(\tau) (H_p + \tilde{\beta}(\tau) H_d). \quad (129)$$

Definindo a função custo reescalada como a energia esperada de H_p , sua derivada é dada por

$$\frac{d}{d\tau} \langle \tilde{\psi}(\tau) | H_p | \tilde{\psi}(\tau) \rangle = \tilde{\beta}(\tau) f'(\tau) \tilde{A}(\tau), \quad (130)$$

onde

$$\tilde{A}(\tau) = \langle \tilde{\psi}(\tau) | i[H_d, H_p] | \tilde{\psi}(\tau) \rangle. \quad (131)$$

Como $f'(\tau) > 0$, uma escolha de controle que preserva a condição de monotonicidade é

$$\tilde{\beta}(\tau) = -\frac{\tilde{A}(\tau)}{f'(\tau)}. \quad (132)$$

Substituindo essa regra na Equação (130), obtém-se

$$\frac{d}{d\tau} \langle \tilde{\psi}(\tau) | H_p | \tilde{\psi}(\tau) \rangle = -\tilde{A}(\tau)^2 \leq 0. \quad (133)$$

Desse modo, a condição de Lyapunov é preservada, enquanto a evolução passa a ser distribuída de acordo com a escala temporal determinada por $f(\tau)$ (15).

No modelo de circuitos, considera-se o intervalo reescalado discretizado em passos de duração $\Delta\tau$, com

$$\tau_k = k\Delta\tau. \quad (134)$$

Aplicando uma decomposição de Trotter de primeira ordem à evolução gerada por $\tilde{H}(\tau_k)$, definem-se os operadores da k -ésima camada como

$$U_{p,k}^{\text{TR}} = e^{-iH_p f'(\tau_k) \Delta\tau}, \quad U_{d,k}^{\text{TR}}(\tilde{\beta}_k) = e^{-i\tilde{\beta}_k H_d f'(\tau_k) \Delta\tau}. \quad (135)$$

O estado após a aplicação da k -ésima camada é, então,

$$|\tilde{\psi}_k\rangle = U_{d,k}^{\text{TR}}(\tilde{\beta}_k)U_{p,k}^{\text{TR}}|\tilde{\psi}_{k-1}\rangle. \quad (136)$$

Como o parâmetro da k -ésima camada deve ser definido antes da preparação de $|\tilde{\psi}_k\rangle$, utiliza-se o observável estimado no estado produzido pela camada anterior. A regra discreta de realimentação é, portanto,

$$\tilde{\beta}_1 = 0, \quad \tilde{\beta}_k = -\frac{\tilde{A}_{k-1}}{f'(\tau_k)}, \quad k \geq 2, \quad (137)$$

onde

$$\tilde{A}_{k-1} = \langle \tilde{\psi}_{k-1} | i[H_d, H_p] | \tilde{\psi}_{k-1} \rangle. \quad (138)$$

O TR-FALQON preserva o observável utilizado na realimentação e a estrutura alternada entre os Hamiltonianos H_p e H_d . A principal modificação encontra-se nos intervalos temporais efetivos de cada camada, que passam a ser ponderados por $f'(\tau_k)$. Dessa forma, a função de reescalonamento determina quanto da evolução efetiva é realizado em cada etapa do circuito.

Na evolução contínua ideal, essa transformação corresponde a uma reparametrização temporal da mesma trajetória. Entretanto, na implementação discretizada e para um número fixo de camadas, a redistribuição da evolução efetiva pode acelerar a aproximação ao estado fundamental nas etapas iniciais. Por essa razão, o TR-FALQON pode apresentar maior probabilidade de sucesso do que o FALQON padrão em regimes de circuitos rasos, desde que a função $f(\tau)$ e seus parâmetros sejam escolhidos adequadamente (15).

4.6.2 SO-FALQON: correção de segunda ordem

Outra limitação do FALQON padrão decorre do fato de que sua lei de atualização é obtida a partir de uma aproximação essencialmente linear em Δt . A variante SO-FALQON (*Second-Order FALQON*) procura contornar esse problema incorporando termos de segunda ordem na aproximação da variação da energia por camada (13).

A ideia central é a seguinte. Em vez de considerar apenas a aproximação de primeira ordem

$$\Delta E_k \equiv E_k - E_{k-1} \approx \Delta t A_{k-1} \beta_k, \quad (139)$$

retêm-se também os termos quadráticos relevantes em Δt . O resultado é uma aproximação da forma

$$\Delta E_k \approx \Delta t A_{k-1} \beta_k + (\Delta t)^2 (B_{k-1} \beta_k^2 + C_{k-1} \beta_k), \quad (140)$$

onde

$$B_{k-1} = \langle \psi_{k-1} | \frac{1}{2} [[H_d, H_p], H_d] | \psi_{k-1} \rangle, \quad (141)$$

e

$$C_{k-1} = \langle \psi_{k-1} | [[H_d, H_p], H_p] | \psi_{k-1} \rangle. \quad (142)$$

Aqui, B_{k-1} captura a curvatura associada à ação repetida do driver sobre o Hamiltoniano de problema, enquanto C_{k-1} incorpora a correção que surge da interação entre a evolução por H_p e o observável de feedback (13).

A expressão (140) é quadrática em β_k . Completando o quadrado, obtém-se

$$\Delta E_k \approx (\Delta t)^2 B_{k-1} \left(\beta_k + \frac{A_{k-1} + \Delta t C_{k-1}}{2\Delta t B_{k-1}} \right)^2 - \frac{(A_{k-1} + \Delta t C_{k-1})^2}{4B_{k-1}}. \quad (143)$$

Quando $B_{k-1} > 0$, o valor de β_k que minimiza a aproximação quadrática acima é

$$\beta_k^{\text{SO}} = -\frac{A_{k-1} + \Delta t C_{k-1}}{2\Delta t B_{k-1}}. \quad (144)$$

Se $B_{k-1} < 0$, a análise mostra que é necessário ajustar o sinal da atualização para manter $\Delta E_k < 0$ no nível da aproximação de segunda ordem (13).

Essa lei de atualização é a essência do SO-FALQON. Em comparação com a regra de primeira ordem $\beta_k = -A_{k-1}$, a nova expressão usa não apenas o valor local do observável de feedback, mas também informação de segunda ordem sobre a resposta da energia à aplicação do driver. Isso permite trabalhar com valores de Δt maiores do que aqueles tipicamente compatíveis com o FALQON padrão, preservando por mais tempo o comportamento monotônico da energia (13).

A interpretação física do SO-FALQON também é clara. Enquanto o FALQON padrão escolhe a intensidade do driver com base em uma previsão local linear da queda de energia, o SO-FALQON corrige essa previsão incluindo a curvatura da paisagem efetiva vista pelo controle em torno do estado atual. Em outras palavras, em vez de

dar um passo baseado apenas na inclinação local, o algoritmo tenta levar em conta também a forma local da função custo ao longo da direção induzida pelo driver.

Naturalmente, essa melhora vem acompanhada de novos cuidados práticos. Em primeiro lugar, a implementação passa a exigir a estimação não apenas de A_k , mas também de B_k e C_k , o que aumenta o custo de medição. Em segundo lugar, a expressão (144) pode gerar valores muito grandes para β_k quando B_{k-1} é pequeno em módulo. Por essa razão, implementações práticas do SO-FALQON costumam combinar a regra de segunda ordem com limitações no valor de β_k ou com esquemas híbridos que recorrem à atualização de primeira ordem em regimes problemáticos (13).

Ainda assim, o SO-FALQON mostra que a profundidade elevada do FALQON padrão não é uma consequência necessária da utilização de feedback, mas está relacionada, em parte, à aproximação de primeira ordem empregada na escolha dos parâmetros. Ao incorporar os termos B_k e C_k , a regra de controle pode produzir reduções mais expressivas da energia por camada e, conseqüentemente, diminuir o número de camadas necessário. Essas quantidades também podem ser acompanhadas como indicadores auxiliares da dinâmica, embora seus valores isolados não constituam critérios suficientes de convergência. Para esse propósito, é mais apropriado considerar a variação de energia prevista pela aproximação de segunda ordem em conjunto com a variação efetivamente observada (13).

4.7 Comparação conceitual

As três formulações discutidas neste capítulo compartilham a mesma estrutura conceitual. Em todas elas, o problema clássico é representado por um Hamiltoniano H_p , enquanto um Hamiltoniano *driver* H_d é incorporado para induzir transições entre configurações do espaço de estados. Os parâmetros do circuito são atualizados iterativamente a partir de valores esperados de observáveis estimados no próprio sistema quântico. Permanece, portanto, a ideia fundamental de substituir a otimização clássica global por uma lei explícita de realimentação inspirada no controle por Lyapunov (10, 12).

A diferença entre as variantes aparece no modo como cada uma escolhe essa lei de atualização. No FALQON padrão, a atualização é governada pela regra mais simples

possível,

$$\beta_{k+1} = -A_k, \quad (145)$$

o que torna o método conceitualmente direto e relativamente econômico em termos do número de observáveis necessários. Sua principal vantagem é a simplicidade: medir A_k basta para determinar a próxima camada. Sua principal limitação é o caráter local e linear dessa regra, que tende a exigir Δt pequeno e, por consequência, muitos passos até a convergência.

No TR-FALQON, a realimentação continua baseada no mesmo observável A_k , mas a evolução deixa de ser distribuída uniformemente no tempo. O reescalonamento temporal modifica o peso efetivo de cada camada na dinâmica total, procurando tornar a evolução mais eficiente nas regiões em que isso é mais vantajoso. Pode-se dizer, portanto, que o TR-FALQON não altera o tipo de informação usada no feedback, mas altera a forma como essa informação é inserida na dinâmica (15).

Já o SO-FALQON modifica o próprio conteúdo da regra de atualização. Em vez de usar somente o termo linear A_k , ele incorpora as quantidades B_k e C_k , associadas à aproximação quadrática da variação da energia. Em termos práticos, isso aumenta o custo por camada, mas tende a reduzir o número total de camadas ao permitir passos mais agressivos e ainda assim controlados (13).

Desse modo, o FALQON padrão, o TR-FALQON e o SO-FALQON correspondem a implementações distintas de uma mesma abordagem de feedback quântico. Essas modificações não são, em princípio, mutuamente exclusivas, de modo que seria possível investigar uma formulação que combine o reescalonamento temporal com a aproximação de segunda ordem. Entretanto, essa integração exigiria uma nova derivação da regra de feedback para intervalos temporais efetivos dependentes de $f'(\tau_k)$, além da estimação conjunta dos observáveis A_k , B_k e C_k . Como essa combinação não é analisada nos trabalhos considerados, sua investigação permanece como uma possível extensão futura (15, 13).

Apesar dessas melhorias na dinâmica e na regra de realimentação, nenhuma das variantes discutidas modifica diretamente a inicialização do algoritmo. Em circuitos rasos, entretanto, a escolha do estado inicial pode influenciar significativamente a rapidez com que a distribuição de probabilidade se concentra em configurações promissoras.

Essa observação motiva naturalmente a discussão do capítulo seguinte. Em vez de agir apenas sobre a dinâmica camada a camada, pode-se introduzir um mecanismo *externo* de aprimoramento, no qual informações obtidas por medições em execuções anteriores são utilizadas para construir estados iniciais mais informados para execuções subsequentes. Essa é a ideia do MGI-FALQON, que preserva a lógica de feedback do FALQON, mas acrescenta uma camada adicional de adaptação orientada por dados de medição. Assim, se o presente capítulo concentrou-se na dinâmica do algoritmo e em suas variantes de atualização, o próximo capítulo voltará a atenção para a questão complementar da inicialização guiada por medições.

5 Inicialização Guiada por Medições

Ao final do Capítulo 4, discutiu-se que variantes como TR-FALQON e SO-FALQON atuam sobre a dinâmica interna do algoritmo, buscando tornar mais eficiente a atualização camada a camada. Ainda assim, a profundidade L necessária para que o FALQON alcance soluções de alta qualidade pode permanecer elevada. Como discutido na Introdução, em análises do MaxCut sobre grafos 3-regulares, a profundidade requerida pelo FALQON pode ser quase duas ordens de grandeza superior à do QAOA para atingir níveis comparáveis de precisão (10, 13). Como referência experimental, implementações de QAOA em processadores quânticos supercondutores com até 127 qubits foram avaliadas em profundidades reduzidas, com resultados favoráveis principalmente entre $L = 1$ e $L = 3$, antes que o aumento da profundidade passasse a ser limitado pelos efeitos do ruído (42, 43).

Nesta dissertação, adota-se como regime de circuitos rasos uma profundidade interna de até $L = 5$ camadas. Essa convenção é motivada pela demonstração experimental original do FALQON, na qual a redução monotônica da energia e o aumento da probabilidade de sucesso foram observados até a quinta camada. Para $L > 5$, naquela instância e no dispositivo utilizado, os efeitos do ruído de hardware tornaram-se suficientemente relevantes para impedir a continuidade da redução da energia, em contraste com a evolução ideal simulada (10). Esse valor não constitui um limite universal de profundidade para dispositivos NISQ, mas fornece uma referência operacional para os experimentos realizados neste trabalho. Embora camadas de FALQON e QAOA não sejam diretamente equivalentes em termos de número de portas e profundidade física após a compilação, esses resultados ilustram a dificuldade de implementar métodos que dependem de valores elevados de L em hardware quântico ruidoso.

Essa observação motiva a introdução de um mecanismo complementar, que não modifica a regra de feedback interna do FALQON, mas sim a forma de funcionamento do algoritmo. A ideia central consiste em reaproveitar informação extraída de medições realizadas em execuções rasas anteriores para construir estados iniciais progressivamente mais informados (44). Em vez de depender apenas do aprofundamento do circuito para

concentrar amplitude em boas soluções, busca-se transferir parte desse trabalho para um processo iterativo externo de refinamento do estado inicial. Essa é a proposta do MGI aplicada ao FALQON, denominada MGI-FALQON (16).

Do ponto de vista conceitual, o método preserva a estrutura fundamental do FALQON discutida no Capítulo 4. O Hamiltoniano de problema continua sendo H_p , o Hamiltoniano *driver* continua sendo H_d , e os parâmetros β_k continuam sendo gerados por realimentação a partir do observável $i[H_d, H_p]$. A diferença está em que, após uma execução de profundidade fixa (L_{fixo}), as *bitstrings* medidas no estado final deixam de ser utilizadas apenas para estimar observáveis ou avaliar a qualidade da solução: elas passam também a alimentar a construção do estado inicial da próxima execução.

Assim, o MGI-FALQON pode ser entendido como um esquema de adaptação em dois níveis. No nível interno, permanece a dinâmica em camadas do FALQON, já apresentada no capítulo anterior. No nível externo, introduz-se uma iteração adicional, orientada pelos dados de medição, que atualiza a distribuição inicial de probabilidade do circuito. O efeito desejado é que, mesmo mantendo-se fixa a profundidade L de cada execução, o algoritmo passe a iniciar suas evoluções em regiões do espaço de Hilbert com maior sobreposição com configurações promissoras, reduzindo a profundidade necessária para obter boas soluções (16).

5.1 Formulação geral do loop externo do MGI-FALQON

Considere um problema de otimização codificado em um Hamiltoniano de custo H_p e um Hamiltoniano *driver* H_d , ambos já definidos anteriormente. No FALQON padrão, para um estado inicial $|\psi_0\rangle$, uma execução com profundidade L produz a sequência de estados

$$|\psi_\ell\rangle = U_d(\beta_\ell)U_p \cdots U_d(\beta_2)U_p U_d(\beta_1)U_p |\psi_0\rangle, \quad \ell = 1, \dots, L, \quad (146)$$

onde

$$U_p = e^{-iH_p\Delta t}, \quad U_d(\beta_\ell) = e^{-i\beta_\ell H_d\Delta t}, \quad (147)$$

e os coeficientes de feedback obedecem à regra já estabelecida no Capítulo 4,

$$\beta_1 = 0, \quad \beta_{\ell+1} = -\langle \psi_\ell | i[H_d, H_p] | \psi_\ell \rangle. \quad (148)$$

No MGI-FALQON, essa dinâmica interna é mantida, mas deixa de ser executada apenas uma vez. Em vez disso, considera-se um laço externo com R iterações. Em cada iteração externa r , com $r = 0, 1, \dots, R - 1$, escolhe-se um estado inicial $|\psi_0^{(r)}\rangle$, executa-se um FALQON de profundidade fixa L e obtém-se um estado final

$$|\psi_L^{(r)}\rangle = U_d(\beta_L^{(r)}) U_p \cdots U_d(\beta_2^{(r)}) U_p U_d(\beta_1^{(r)}) U_p |\psi_0^{(r)}\rangle. \quad (149)$$

Os coeficientes $\beta_\ell^{(r)}$ continuam sendo definidos pela mesma lógica de realimentação:

$$\beta_1^{(r)} = 0, \quad \beta_{\ell+1}^{(r)} = -\langle \psi_\ell^{(r)} | i[H_d, H_p] | \psi_\ell^{(r)} \rangle, \quad \ell = 1, \dots, L - 1. \quad (150)$$

A diferença essencial é que, após a obtenção de $|\psi_L^{(r)}\rangle$, realiza-se uma etapa adicional de medição e processamento estatístico. O estado final é medido N_{shots} vezes na base computacional, produzindo um multiconjunto de *bitstrings*. O parâmetro N_{shots} corresponde ao número de repetições independentes do circuito utilizadas para estimar as frequências dos resultados de medição.

$$\mathcal{M}^{(r)} = \{x^{(r,1)}, x^{(r,2)}, \dots, x^{(r,N_{\text{shots}})}\}, \quad x^{(r,s)} \in \{0, 1\}^{n_q}, \quad (151)$$

onde n_q denota o número total de qubits efetivamente utilizados na instância do problema.

A partir desse multiconjunto, constrói-se uma nova inicialização $|\psi_0^{(r+1)}\rangle$ para a iteração seguinte. Em outras palavras, o MGI-FALQON pode ser esquematicamente resumido como

$$|\psi_0^{(r)}\rangle \xrightarrow{\text{FALQON}} |\psi_L^{(r)}\rangle \xrightarrow{\mathcal{M}} |\psi_0^{(r+1)}\rangle$$

É importante notar que o parâmetro controlado no laço externo não é a profundidade do circuito, que permanece fixa em L , mas a própria distribuição inicial do estado quântico. Assim, enquanto no FALQON padrão a melhoria da solução é obtida principalmente pelo aumento do número de camadas, no MGI-FALQON parte desse processo é substituída por refinamentos sucessivos da inicialização. Em termos práticos, isso significa que o método troca parte da profundidade unitária por um procedimento iterativo baseado em medições.

A escolha do estado inicial da primeira iteração, $|\psi_0^{(0)}\rangle$, pode seguir a convenção usual do FALQON padrão.

A partir da segunda iteração, entretanto, o estado inicial passa a ser determinado pelas estatísticas coletadas na iteração anterior.

Uma vez obtido o multiconjunto $\mathcal{M}^{(r)}$ de resultados de medição da iteração r , pode-se associar a ele uma distribuição empírica sobre a base computacional. Para cada bitstring $x \in \{0, 1\}^{n_q}$, define-se sua contagem

$$N_x^{(r)} = \sum_{s=1}^{N_{\text{shots}}} \mathbf{1}[x^{(r,s)} = x], \quad (152)$$

onde $\mathbf{1}[\cdot]$ denota a função indicadora. A distribuição empírica correspondente é então

$$\hat{p}_r(x) = \frac{N_x^{(r)}}{N_{\text{shots}}} \quad (153)$$

onde,

$$\sum_{x \in \{0,1\}^{n_q}} \hat{p}_r(x) = 1. \quad (154)$$

Em princípio, seria possível utilizar toda a distribuição empírica $\hat{p}_r$ para definir a próxima inicialização. Entretanto, em regime raso, essa distribuição ainda pode ser bastante dispersa, contendo peso estatístico não desprezível em configurações pouco relevantes, oriundas tanto da própria dinâmica limitada quanto de flutuações de amostragem. Por essa razão, o MGI introduz uma etapa de filtragem, cujo objetivo é concentrar a atualização apenas nos desfechos mais representativos.

5.1.1 Seleção das *bitstrings* dominantes

Sejam $b_1^{(r)}, b_2^{(r)}, \dots, b_M^{(r)}$ as *bitstrings* distintas observadas na iteração r , ordenadas de modo que

$$N_{b_1}^{(r)} \geq N_{b_2}^{(r)} \geq \dots \geq N_{b_M}^{(r)}. \quad (155)$$

A partir dessa ordenação, seleciona-se um subconjunto formado pelas *bitstrings* mais frequentes. Na forma mais simples do método, escolhe-se um inteiro fixo $n \leq M$,

constante ao longo de todas as iterações externas. Nesse caso, o subconjunto filtrado é dado por

$$\mathcal{S}_n^{(r)} = \left\{ b_k^{(r)} \right\}_{k=1}^n, \quad (156)$$

isto é, o conjunto formado pelas n *bitstrings* mais frequentes observadas na iteração r .

Entretanto, essa escolha não precisa ser fixa. O número de *bitstrings* selecionadas depende do problema estudado, da instância, da profundidade considerada e do comportamento da distribuição de saída. Até o presente momento, não há uma regra teórica geral para determinar o valor ótimo de n . Por essa razão, n é tratado como um hiperparâmetro empírico, escolhido por meio de testes numéricos para cada configuração experimental. Em distribuições mais espalhadas, pode ser conveniente manter um número maior de *bitstrings*, preservando uma parcela mais ampla da informação estatística. Em distribuições mais concentradas, um subconjunto menor pode ser suficiente, reduzindo a influência de configurações menos representativas.

Uma estratégia variável consiste em permitir que o número de *bitstrings* selecionadas se altere ao longo das iterações externas. Denotando por n_r a quantidade de *bitstrings* retidas na iteração r , define-se o subconjunto filtrado de forma mais geral como

$$\mathcal{S}_{n_r}^{(r)} = \left\{ b_k^{(r)} \right\}_{k=1}^{n_r}. \quad (157)$$

Nesse caso, a versão com n fixo é recuperada como caso particular, bastando tomar $n_r = n$ para todo r .

Uma estratégia particularmente útil consiste em fazer n_r decrescer ao longo das iterações, começando com um filtro mais amplo e tornando-o progressivamente mais seletivo à medida que a inicialização vai sendo refinada. Mantendo a indexação externa $r = 0, 1, \dots, R - 1$, essa estratégia pode ser escrita como

$$n_r = \left\lfloor n_{\max} - (n_{\max} - n_{\min}) \frac{r}{R - 1} \right\rfloor, \quad (158)$$

onde $n_{\max}$ e $n_{\min}$ são, respectivamente, os números máximo e mínimo de *bitstrings* retidas ao longo do procedimento, e $\lfloor \cdot \rfloor$ denota a função piso. Pela própria construção, quando $r = 0$, obtém-se $n_0 = n_{\max}$, enquanto na última iteração, $r = R - 1$, tem-se $n_{R-1} = n_{\min}$. Para valores intermediários de r , o número de *bitstrings* decresce aproximadamente de forma linear.

Essa formulação possui uma interpretação clara. Nas primeiras iterações, quando a distribuição de saída ainda tende a ser mais difusa, convém utilizar um valor maior de n_r , de modo a capturar uma fração mais ampla dos desfechos dominantes. À medida que o processo iterativo avança e a distribuição tende a se concentrar em configurações mais promissoras, torna-se natural reduzir n_r , tornando a atualização mais seletiva. Desse modo, o filtro variável permite ajustar a agressividade da pós-seleção ao estágio da evolução externa.

Para cada bitstring pertencente ao conjunto filtrado, seja

$$m_k^{(r)} = N_{b_k}^{(r)} \quad (159)$$

a sua multiplicidade no multiconjunto de medições.

Essa etapa de pós-seleção possui caráter heurístico. As *bitstrings* mais frequentes apresentam maior probabilidade empírica na distribuição de saída e, como a dinâmica do FALQON busca reduzir a energia esperada, espera-se que estejam tipicamente associadas a regiões de menor energia do espaço de busca. Entretanto, essa correspondência não é garantida para cada configuração individual. O filtro seleciona a parcela da distribuição concentrada nos resultados mais frequentes, sem avaliar diretamente a energia de cada *bitstring*. Desse modo, busca-se preservar as estruturas dominantes da distribuição e reduzir a influência de resultados pouco frequentes e de flutuações estatísticas (16).

5.1.2 Derivação explícita das marginais de um qubit

A partir do subconjunto filtrado, define-se uma distribuição empírica renormalizada,

$$\tilde{p}_r(x) = \frac{\hat{p}_r(x) \mathbf{1}[x \in \mathcal{S}_{n_r}^{(r)}]}{\sum_{y \in \mathcal{S}_{n_r}^{(r)}} \hat{p}_r(y)}. \quad (160)$$

Por construção, $\tilde{p}_r(x) \geq 0$ e

$$\sum_{x \in \{0,1\}^{n_q}} \tilde{p}_r(x) = 1. \quad (161)$$

Se $x_i \in \{0, 1\}$ denota o valor do i -ésimo qubit em uma bitstring x , a marginal empírica de um qubit correspondente à probabilidade de observar $x_i = 1$ sob a

distribuição filtrada é

$$c_i^{(r)} = \sum_{x \in \{0,1\}^{n_q}} \tilde{p}_r(x) x_i, \quad i = 1, \dots, n_q. \quad (162)$$

Substituindo (160) em (162), obtém-se

$$c_i^{(r)} = \frac{\sum_{x \in \mathcal{S}_{n_r}^{(r)}} \hat{p}_r(x) x_i}{\sum_{y \in \mathcal{S}_{n_r}^{(r)}} \hat{p}_r(y)}. \quad (163)$$

Agora, escrevendo $\hat{p}_r(x) = N_x^{(r)} / N_{\text{shots}}$, segue que

$$c_i^{(r)} = \frac{\sum_{x \in \mathcal{S}_{n_r}^{(r)}} N_x^{(r)} x_i}{\sum_{y \in \mathcal{S}_{n_r}^{(r)}} N_y^{(r)}}. \quad (164)$$

Como os elementos de $\mathcal{S}_{n_r}^{(r)}$ são precisamente as *bitstrings* $b_k^{(r)}$, com multiplicidades $m_k^{(r)}$, e se $b_{k,i}^{(r)} \in \{0, 1\}$ denota o valor do i -ésimo bit na *bitstring* $b_k^{(r)}$, chega-se à expressão operacional

$$c_i^{(r)} = \frac{\sum_{k=1}^{n_r} m_k^{(r)} b_{k,i}^{(r)}}{\sum_{k=1}^{n_r} m_k^{(r)}}, \quad i = 1, \dots, n_q. \quad (165)$$

A Equação (165) merece uma interpretação cuidadosa. O numerador soma, para o qubit i , quantas vezes o valor 1 aparece entre as *bitstrings* retidas, ponderando cada ocorrência pela frequência com que aquela *bitstring* apareceu. O denominador, por sua vez, corresponde ao número total de amostras efetivamente utilizadas após a filtragem. Assim, $c_i^{(r)}$ nada mais é do que a frequência relativa, dentro do conjunto filtrado, com que o qubit i é encontrado no estado $|1\rangle$.

Do ponto de vista estatístico, trata-se de uma média ponderada de variáveis binárias. Se, para simplificar, todas as *bitstrings* filtradas tivessem a mesma multiplicidade, a Equação (165) se reduziria à média aritmética simples dos bits $b_{k,i}^{(r)}$. A presença dos pesos $m_k^{(r)}$ apenas reflete que *bitstrings* mais frequentes devem contribuir mais para a atualização da inicialização.

Do ponto de vista físico, essa etapa comprime em um conjunto de probabilidades locais uma parte da informação contida na distribuição de saída do circuito. Em vez de tentar reconstruir toda a distribuição conjunta sobre n_q qubits, o método retém apenas

estatísticas de um corpo. Essa simplificação é precisamente o que torna a atualização leve e compatível com circuitos rasos, embora, como será discutido adiante, também imponha limitações importantes.

5.2 Construção do estado inicial enviesado

5.2.1 Estado de um qubit e correspondência com R_y

Uma vez determinadas as marginais $c_i^{(r)}$, o objetivo é construir um novo estado inicial em que a probabilidade de medir o qubit i no estado $|1\rangle$ coincida exatamente com esse valor. O modo mais simples de realizar isso é preparar um estado produto em que cada qubit é ajustado independentemente.

Considere, inicialmente, um único qubit. Seja

$$|\varphi_i^{(r)}\rangle = \sqrt{1 - c_i^{(r)}} |0\rangle + \sqrt{c_i^{(r)}} |1\rangle. \quad (166)$$

Como

$$\left| \sqrt{1 - c_i^{(r)}} \right|^2 + \left| \sqrt{c_i^{(r)}} \right|^2 = 1, \quad (167)$$

o estado dado pela Equação (166) já se encontra normalizado. Além disso, quando medido na base computacional, produz

$$\Pr(0) = 1 - c_i^{(r)}, \quad \Pr(1) = c_i^{(r)}. \quad (168)$$

Para implementar esse estado por meio de uma porta elementar, utiliza-se uma rotação em torno do eixo y , introduzida no Capítulo 3. Relembrando sua definição,

$$R_y(\theta) = e^{-i\theta Y/2} = \cos\left(\frac{\theta}{2}\right) \mathbb{1} - i \sin\left(\frac{\theta}{2}\right) Y. \quad (169)$$

Aplicando $R_y(\theta)$ ao estado $|0\rangle$, obtém-se

$$R_y(\theta) |0\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) |1\rangle. \quad (170)$$

A última igualdade segue da ação da matriz de Pauli Y sobre $|0\rangle$, ou, equivalentemente, da forma matricial explícita de $R_y(\theta)$.

Comparando (166) com (170), vê-se que basta impor

$$\cos\left(\frac{\theta_i^{(r)}}{2}\right) = \sqrt{1 - c_i^{(r)}}, \quad \sin\left(\frac{\theta_i^{(r)}}{2}\right) = \sqrt{c_i^{(r)}}. \quad (171)$$

Elevando ao quadrado a segunda relação, tem-se

$$\sin^2\left(\frac{\theta_i^{(r)}}{2}\right) = c_i^{(r)}. \quad (172)$$

Como $c_i^{(r)} \in [0, 1]$, segue que

$$\theta_i^{(r)} = 2 \arcsin\left(\sqrt{c_i^{(r)}}\right). \quad (173)$$

Portanto, a marginal estimada a partir das medições pode ser convertida diretamente no parâmetro de uma rotação R_y , de forma que o estado de um qubit preparado na iteração seguinte reproduza exatamente a probabilidade local desejada.

5.2.2 Construção do estado produto global

Aplicando o procedimento anterior a todos os qubits, define-se o estado inicial da iteração seguinte como

$$|\psi_0^{(r+1)}\rangle = \bigotimes_{i=1}^{n_q} R_y(\theta_i^{(r)}) |0\rangle, \quad (174)$$

ou, de forma equivalente,

$$|\psi_0^{(r+1)}\rangle = \bigotimes_{i=1}^{n_q} \left(\sqrt{1 - c_i^{(r)}} |0\rangle + \sqrt{c_i^{(r)}} |1\rangle \right). \quad (175)$$

A expressão (175) mostra explicitamente que o novo estado inicial é um estado produto. Portanto, ele é inteiramente determinado pelas marginais individuais $c_i^{(r)}$ e não codifica correlações entre os qubits no momento da reinicialização. Essas correlações podem ser geradas posteriormente pela evolução do circuito FALQON. Em contrapartida, a preparação do estado inicial permanece simples do ponto de vista experimental, pois requer apenas rotações locais de um qubit, sem a necessidade de portas emaranhadoras adicionais.

5.3 Interpretação conceitual e considerações sobre o método

A lógica do MGI-FALQON pode agora ser sintetizada de forma mais clara. Após cada execução rasa do FALQON, os resultados de medição definem uma distribuição empírica sobre as *bitstrings* da base computacional. Em vez de utilizar essa informação apenas para avaliar o desempenho da execução corrente, o método a reaproveita para redefinir a condição inicial da execução seguinte. Mais especificamente, as *bitstrings* dominantes são selecionadas, suas marginais de um qubit são estimadas e essas marginais são então convertidas em um novo estado produto preparado por rotações locais. Desse modo, a informação estatística acumulada ao longo das medições passa a atuar diretamente sobre a preparação do estado inicial. O fluxo geral desse procedimento em uma única iteração é ilustrado na Fig. 2.

Sob esse ponto de vista, o MGI pode ser interpretado como um mecanismo que substitui parte da concentração de probabilidade normalmente obtida pelo aumento da profundidade do circuito por um processo iterativo externo de refinamento orientado por medições. A repetição desse procedimento busca concentrar progressivamente a probabilidade em regiões promissoras do espaço de *bitstrings*, preservando as probabilidades marginais individuais extraídas da distribuição filtrada. Essa interpretação ajuda a explicar por que o método permanece simples, leve do ponto de vista computacional e compatível com circuitos rasos.

No entanto, a efetividade do método depende diretamente da combinação entre a profundidade fixa L do circuito FALQON executado em cada iteração interna e o número R de iterações externas do MGI. Para determinadas configurações de L e R , o ganho obtido ao longo do processo pode ser expressivo, enquanto, em outras, a melhoria tende a saturar mais cedo. Isso indica que a relação entre a profundidade interna do FALQON e o refinamento externo promovido pelo MGI deve ser analisada cuidadosamente para cada problema e instância. Além disso, a escolha da quantidade de *bitstrings* retidas pelo filtro, seja por meio de um valor fixo n ou de uma estratégia variável n_r , permanece empírica, não havendo ainda uma regra geral que determine a seleção mais adequada. Ainda assim, o MGI-FALQON preserva características centrais do FALQON: mantém o caráter não variacional do método, pois não introduz uma rotina de otimização clássica global sobre parâmetros livres, e permanece compatível com

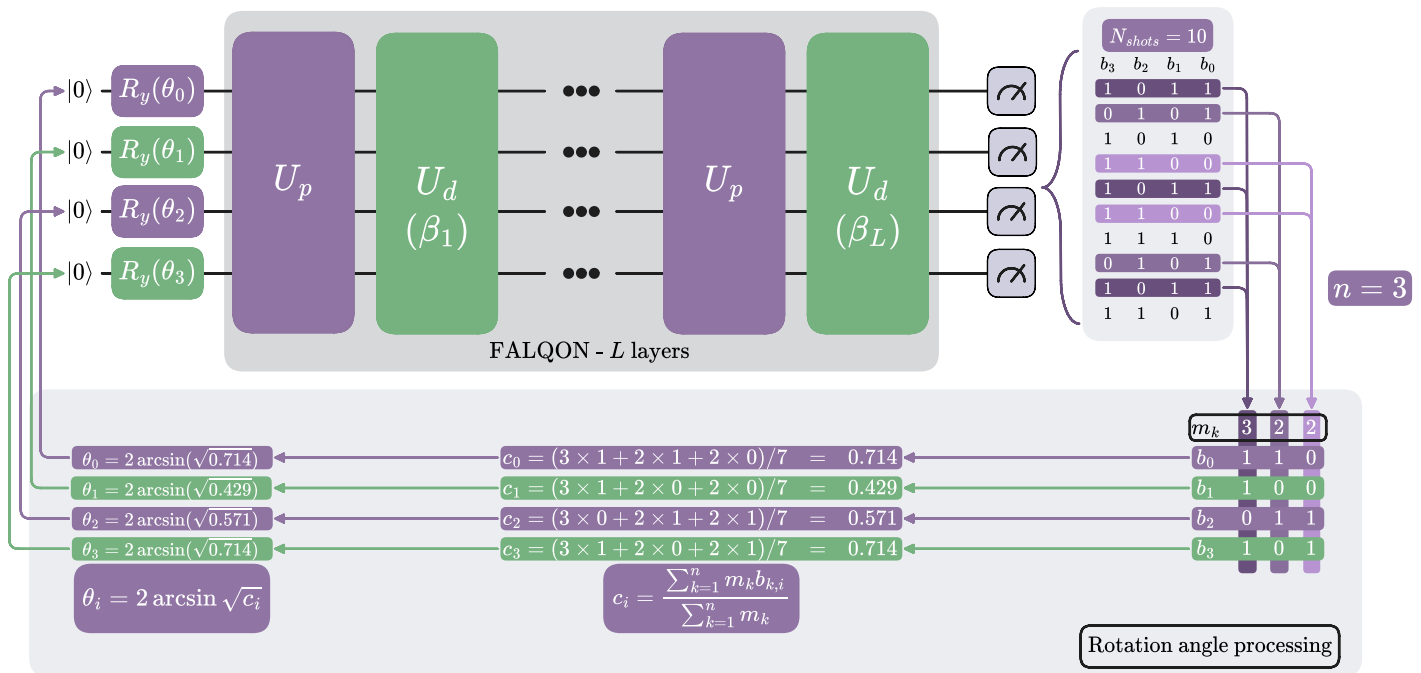


Figura 2 – Exemplo de uma iteração do MGI-FALQON para um sistema de 4 qubits. Em cada iteração, o estado inicial é preparado por rotações $R_y(\theta_i)$, o circuito FALQON com profundidade fixa L é executado, o estado final é medido N_{shots} vezes, selecionam-se as $n = 3$ *bitstrings* mais frequentes e, a partir desse conjunto reduzido, estimam-se as marginais c_i usadas para definir os ângulos da próxima iteração.

Fonte: (16).

circuitos rasos, uma vez que a atualização do estado inicial exige apenas processamento clássico simples e rotações locais de um qubit. Por isso, constitui uma extensão natural do esquema de *feedback* estudado nesta dissertação.

Por fim, a formulação apresentada neste capítulo prepara o terreno para a análise experimental do capítulo seguinte. Tendo sido estabelecidos o loop externo do MGI-FALQON, as estratégias de filtragem das *bitstrings*, a derivação das marginais e a construção do estado inicial enviesado, torna-se possível investigar de forma sistemática em que medida essa estratégia consegue, na prática, substituir profundidade de circuito

por refinamento iterativo guiado por medições.

6 Aplicações e Resultados

Este capítulo apresenta os resultados obtidos para o problema MaxCut a partir da aplicação do FALQON, do TR-FALQON, do SO-FALQON e do MGI-FALQON. O objetivo é analisar o comportamento desses métodos em termos de energia, probabilidade de sucesso e profundidade de circuito, evidenciando que, embora as variantes do FALQON possam acelerar a convergência, a profundidade ainda pode permanecer elevada para o contexto NISQ. Além de retomar os resultados iniciais que culminaram no artigo em pré-publicação *Measurement-Guided State Refinement for Shallow Feedback-Based Quantum Optimization Algorithm* (16), este capítulo também apresenta comparações adicionais entre filtros e variantes, bem como uma discussão crítica sobre as limitações atuais do método e suas perspectivas de implementação em *hardware* real.

6.1 Protocolo experimental

Todos os experimentos apresentados neste capítulo foram realizados no problema MaxCut, já discutido no Capítulo 2. Considerou-se a instância de 8 vértices apresentada na Figura 3. Como o MaxCut possui uma simetria global associada à troca entre as duas partições, cada solução x possui uma solução complementar equivalente $\bar{x}$, dada por

$$\bar{x}_i = 1 - x_i, \quad C(x) = C(\bar{x}), \quad (176)$$

o que permite fixar o valor de um dos vértices sem perda de generalidade. Neste trabalho, o vértice 0 foi fixado na partição representada pelo valor binário 0. Essa fixação elimina a redundância entre soluções complementares e reduz a formulação efetiva de 8 para 7 variáveis binárias. Consequentemente, a implementação quântica requer apenas 7 qubits, preservando o valor ótimo e a estrutura relevante do problema.

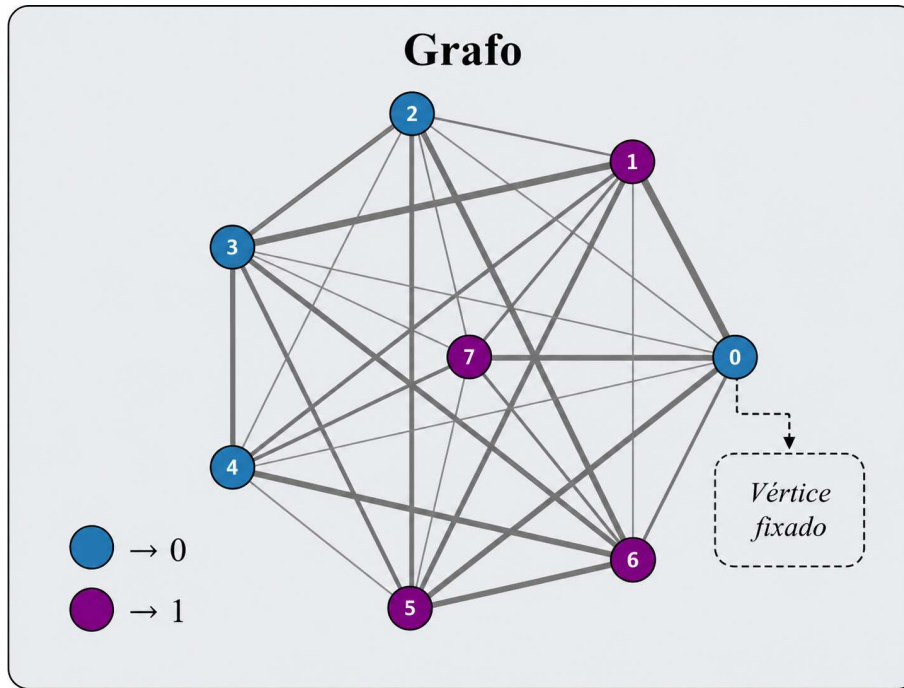


Figura 3 – Instância do problema MaxCut com 8 vértices utilizada nos experimentos. As cores azul e roxa representam, respectivamente, as atribuições binárias 0 e 1 das duas partições. O vértice 0 é fixado para eliminar a redundância decorrente da simetria global entre soluções complementares, reduzindo a implementação para 7 qubits.

Fonte: Elaborado pelo autor.

A implementação computacional foi realizada em *Python* com a biblioteca *qtsuru* (45), desenvolvida por Lucas Rattighieri, integrante do mesmo grupo de pesquisa, e utilizada como base para a construção dos Hamiltonianos, dos operadores e das rotinas de evolução quântica consideradas neste trabalho. Em todos os casos, empregou-se o Hamiltoniano de problema associado ao MaxCut e o Hamiltoniano *driver* transversal, conforme a formulação apresentada no Capítulo 4. No caso do MGI-FALQON, a evolução interna foi combinada ao laço externo descrito no Capítulo 5, no qual o estado inicial é atualizado iterativamente a partir dos resultados de medição.

As métricas analisadas foram a energia média final,

$$E_{\text{final}} = \langle \psi_{\text{final}} | H_p | \psi_{\text{final}} \rangle, \quad (177)$$

e a probabilidade de sucesso associada ao conjunto de soluções ótimas,

$$P_{\text{opt}} = \sum_{x \in \mathcal{S}_*} p(x), \quad (178)$$

onde $\mathcal{S}_*$ representa o conjunto de *bitstrings* ótimas e $p(x)$ é a probabilidade de observar a *bitstring* x na base computacional. Nos experimentos com múltiplas sementes, também foi considerada a dispersão entre diferentes realizações, indicada nas figuras pelas regiões sombreadas.

6.2 Resultados do FALQON e de suas variantes

A primeira etapa experimental consistiu em comparar o FALQON padrão, o TR-FALQON e o SO-FALQON em um regime de até $L = 200$. O objetivo dessa análise foi verificar não apenas se as variantes aceleram a evolução, mas também se essa aceleração é suficiente para tornar o método compatível com restrições realistas de profundidade.

Para o TR-FALQON, adotou-se a função trigonométrica de reescalonamento

$$f(\tau) = a\tau - \frac{t_f}{2\pi a}(a-1)\sin\left(\frac{2\pi a}{t_f}\tau\right), \quad (179)$$

cuja derivada é

$$f'(\tau) = a - (a-1)\cos\left(\frac{2\pi a}{t_f}\tau\right). \quad (180)$$

O parâmetro a controla a intensidade do reescalonamento temporal: para $a = 1$, recupera-se $f(\tau) = \tau$ e, portanto, a evolução sem reescalonamento; para $a > 1$, o intervalo total é contraído segundo $\tau_f = t_f/a$, enquanto $f'(\tau)$ redistribui de forma não uniforme o tempo efetivo ao longo de L . O tempo final foi definido como $t_f = L(a\Delta t)$, garantindo que o intervalo reescalonado contenha os L passos utilizados na simulação (15).

A Figura 4 mostra, no painel a), que os três métodos promovem uma queda monotônica da energia à medida que L aumenta, mas com taxas distintas. O SO-FALQON apresenta a redução mais rápida para os menores valores de L , seguido pelo TR-FALQON, enquanto o FALQON padrão evolui de forma mais lenta e gradual. No

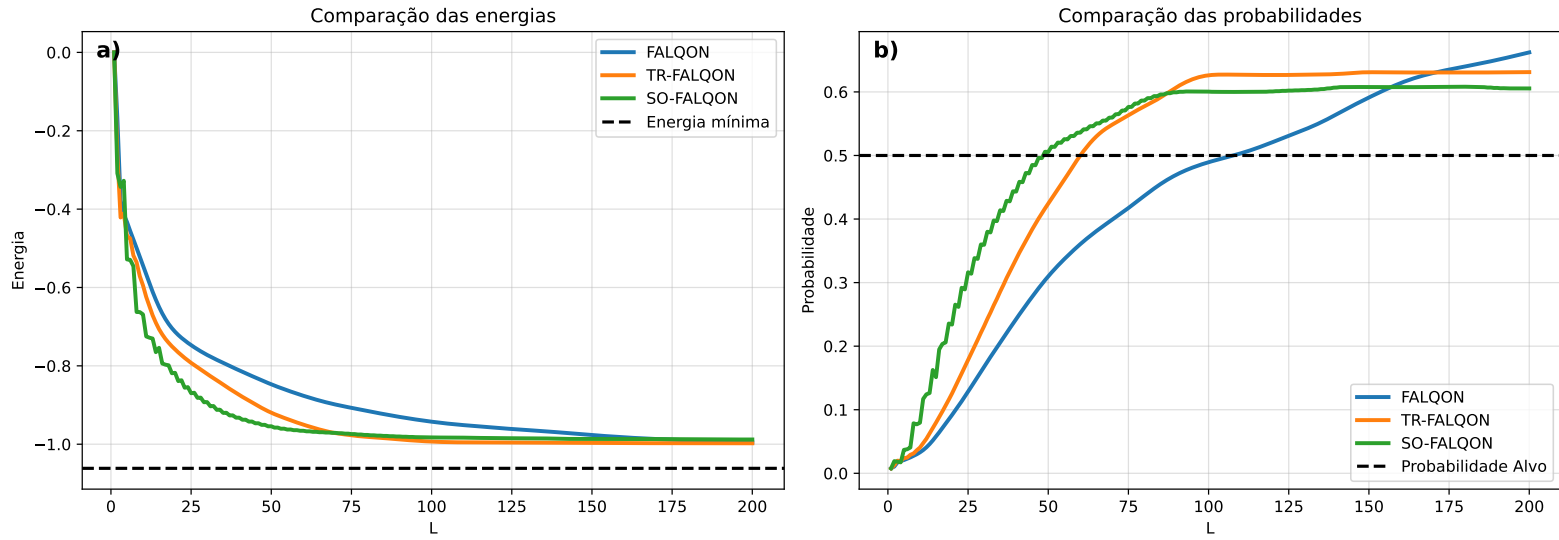


Figura 4 – Comparação entre o FALQON e suas variantes no problema MaxCut com 7 qubits. Foram usados $L = 200$ em todos os casos; para o FALQON, $\Delta t = 0.2$ e $\beta_1 = 0$; para o TR-FALQON, $\Delta t = 0.24$, $a = 1.5$, $\beta_1 = 0$, $t_f = L(a\Delta t)$ e $f'(\tau) = a - (a - 1) \cos\left(\frac{2\pi a}{t_f}\tau\right)$; para o SO-FALQON, $\Delta t = 0.4$ na versão híbrida. No painel a), energia em função de L ; no painel b), probabilidade da solução ótima.

entanto, à medida que L cresce, as curvas passam a se aproximar e, em $L = 200$, os três métodos permanecem acima da energia mínima exata, embora TR-FALQON e SO-FALQON terminem ligeiramente mais próximos desse limite do que o FALQON padrão. No painel b), observa-se novamente um comportamento distinto entre as variantes. O SO-FALQON apresenta o crescimento inicial mais rápido e ultrapassa cedo a marca de 0.5, o que é relevante porque indica que, a partir desse ponto, a solução ótima passa a ser o resultado mais provável de uma medição individual. Em seguida, o TR-FALQON cresce de forma mais consistente e supera o SO-FALQON para valores intermediários de L , estabilizando-se em um patamar mais elevado. Já o FALQON padrão, embora seja mais lento para valores de L nas primeiras dezenas, continua melhorando por mais tempo e, em $L = 200$, alcança a maior probabilidade média entre os três métodos.

Essa diferença entre energia e probabilidade é relevante. As variantes TR-FALQON e SO-FALQON de fato aceleram a dinâmica para os menores valores de L e melhoram o desempenho em regimes intermediários, mas, para a instância estudada, o FALQON padrão ainda consegue apresentar a maior probabilidade final quando se permite uma profundidade suficientemente grande. Assim, o principal ganho das variantes não é necessariamente produzir o melhor resultado final absoluto, mas sim melhorar a eficiência da convergência em profundidades menores. Ainda assim, mesmo com essa aceleração, os valores de L necessários permanecem da ordem de centenas, o que é pouco atrativo para *hardware* quântico de curto prazo.

6.3 Resultados do MGI-FALQON com diferentes filtros

A segunda etapa experimental concentrou-se no MGI-FALQON em regime raso. Em cada iteração externa, o circuito FALQON foi executado com profundidade fixa $L = 2$; em seguida, o estado final foi medido, as *bitstrings* foram filtradas e as marginais resultantes foram utilizadas para preparar o estado inicial da execução seguinte. Esse procedimento foi repetido durante $R = 30$ iterações externas. Foram empregadas 100 sementes pseudoaleatórias, cada uma associada a uma realização independente da amostragem das medições, com o objetivo de avaliar a variabilidade estatística do método e verificar se as tendências observadas não dependem de uma única sequência

de resultados. As curvas apresentam as médias entre essas realizações, enquanto as regiões sombreadas representam a dispersão correspondente. O objetivo principal foi comparar o impacto de diferentes estratégias de filtragem das *bitstrings* observadas.

A Figura 5 mostra que o comportamento do MGI-FALQON depende de forma decisiva da estratégia de filtragem adotada. No painel a), observa-se que o filtro fixo com $n = 5$ produz a queda mais rápida da energia nas primeiras iterações, mas essa melhora inicial logo se esgota. A curva praticamente entra em platô após poucas iterações e termina significativamente afastada da energia mínima. Em contraste, o filtro com $n = 30$ e, sobretudo, o filtro variável $30 \rightarrow 5$ continuam promovendo redução progressiva da energia ao longo de todo o processo. Entre eles, o filtro variável apresenta o melhor resultado final, aproximando-se mais do mínimo energético.

No painel b), esse comportamento também se manifesta na probabilidade de sucesso. O filtro com $n = 5$ cresce rapidamente no início, mas satura cedo em um patamar inferior e com maior dispersão entre sementes. Já o filtro com $n = 30$ mantém crescimento mais estável ao longo das iterações, alcançando desempenho final superior. O melhor resultado, entretanto, é obtido pela estratégia variável, cuja curva continua crescendo de forma consistente ao longo do processo e termina com a maior probabilidade média de sucesso entre os filtros considerados.

Os resultados indicam, portanto, que o desempenho do MGI-FALQON não depende apenas da profundidade interna do circuito, mas também do modo como a informação estatística extraída das medições é incorporada à etapa seguinte. No caso do filtro variável, a utilização de um conjunto mais amplo de *bitstrings* nas primeiras iterações, seguida de uma seleção progressivamente mais restritiva, parece favorecer um equilíbrio mais adequado entre diversidade estatística e concentração em estados promissores. Para a instância estudada, essa estratégia conduziu simultaneamente aos melhores resultados em energia e em probabilidade de sucesso.

Outro ponto importante é que, mesmo com profundidade interna $L = 2$ por execução, o MGI-FALQON com filtro variável atinge desempenho final superior ao observado anteriormente para os métodos profundos com $L = 200$. Isso indica que, a princípio, o refinamento iterativo da inicialização consegue compensar de maneira muito eficiente a limitação severa de profundidade.

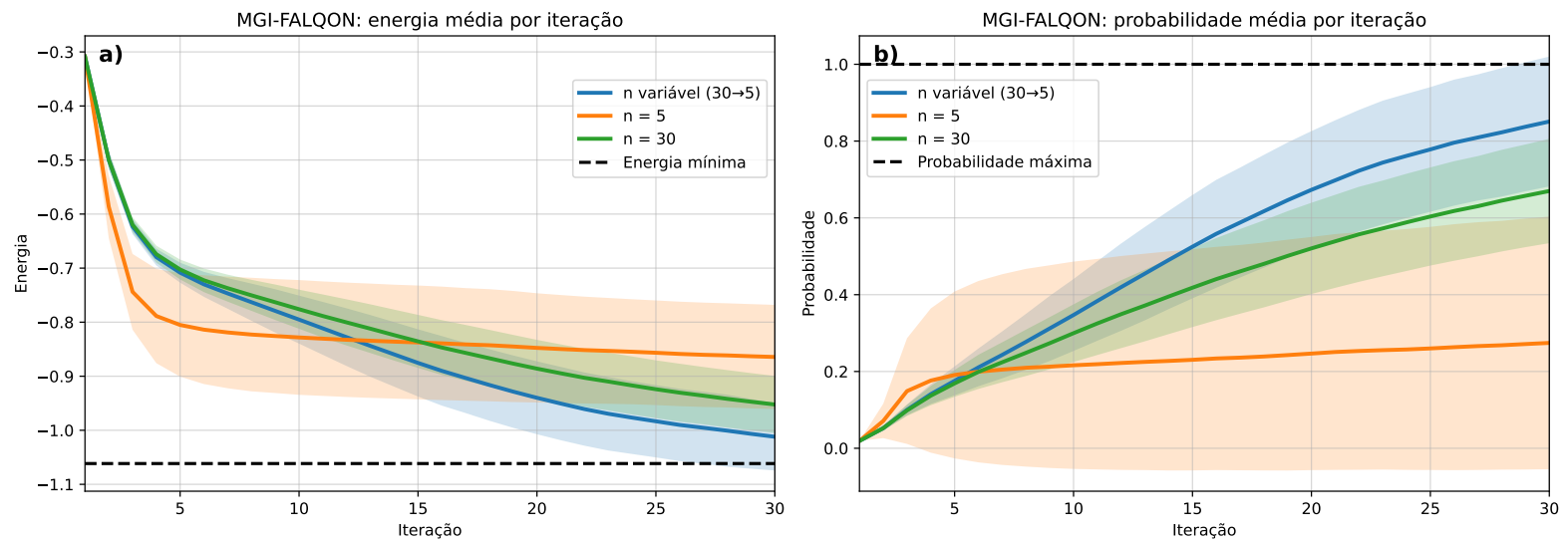


Figura 5 – Resultados do MGI-FALQON com profundidade interna $L = 2$, $R = 30$ iterações externas e 100 sementes. Foram comparados os filtros $n = 5$, $n = 30$ e n variável de $30 \rightarrow 5$. No painel a), energia média por iteração; no painel b), probabilidade média da solução ótima. As regiões sombreadas indicam a dispersão entre sementes.

6.4 Resultados do MGI aplicado às variantes do FALQON

Uma vez estabelecida a vantagem do MGI-FALQON com filtro variável, a etapa seguinte consistiu em aplicar o mesmo mecanismo às variantes internas do algoritmo, com o objetivo de avaliar se o ganho promovido pelo laço externo é compatível com diferentes regras de evolução.

A Figura 6 mostra que o MGI pode ser combinado com diferentes variantes do FALQON. No painel a), observa-se que o MGI-FALQON e o MGI-SO-FALQON apresentam uma queda inicial de energia mais rápida do que o MGI-TR-FALQON. Entretanto, após aproximadamente as primeiras 10 iterações, a curva do MGI-TR-FALQON cruza as demais e passa a apresentar a menor energia média, mantendo essa vantagem até o final do experimento. O MGI-SO-FALQON termina ligeiramente à frente do MGI-FALQON, mas ambos permanecem acima do desempenho final do MGI-TR-FALQON.

O painel b) mostra um comportamento análogo na probabilidade de sucesso. Embora as três curvas cresçam de forma semelhante nas primeiras iterações, o MGI-TR-FALQON passa a se destacar de forma nítida a partir da região intermediária do processo, alcançando a maior probabilidade média final entre os três métodos. O MGI-SO-FALQON aparece em segundo lugar, enquanto o MGI-FALQON padrão termina com o menor valor médio entre as variantes testadas, ainda que com desempenho bastante competitivo.

Esses resultados indicam que o MGI atua como um mecanismo complementar e não como um substituto da dinâmica interna do algoritmo. A realimentação por medições melhora o ponto de partida de cada nova execução, mas o modo como essa informação é explorada ainda depende da variante interna utilizada. No caso da instância analisada, a combinação entre o MGI e o TR-FALQON foi a mais eficiente, sugerindo que o reescalonamento temporal beneficia-se de forma particularmente forte do refinamento iterativo da inicialização.

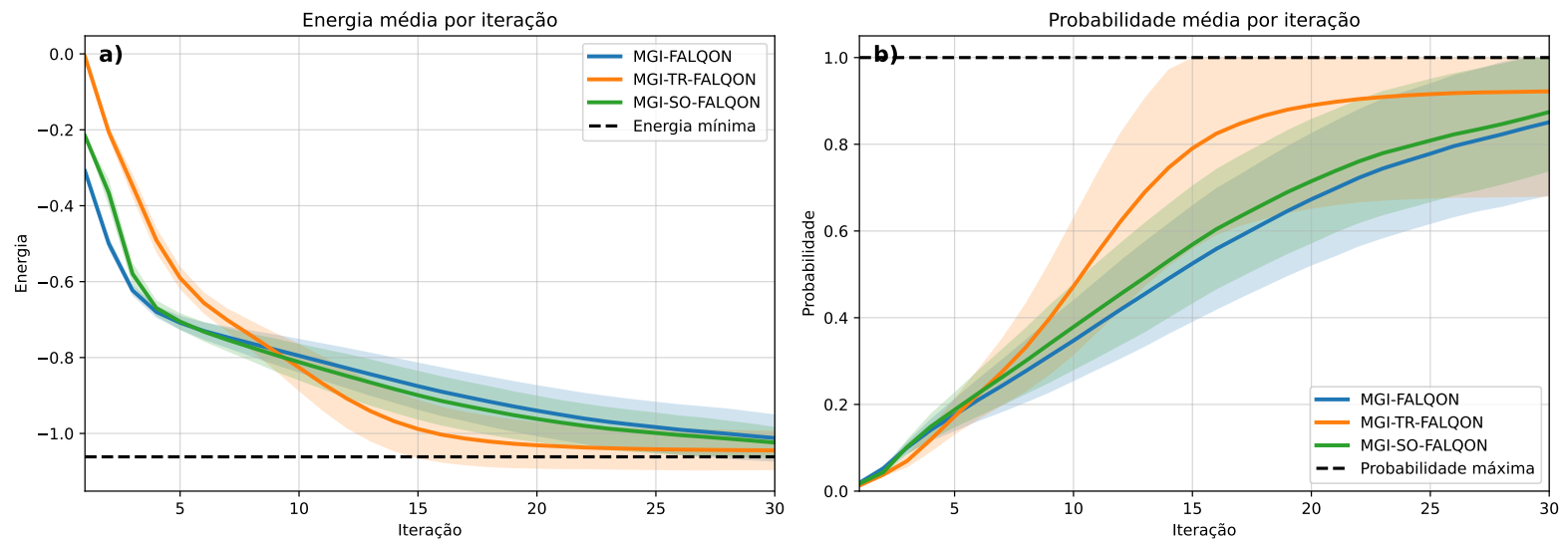


Figura 6 – Comparação entre as variantes acopladas ao MGI com filtro variável, profundidade interna $L = 2$, $R = 30$ iterações externas e 100 sementes. Foram usados $\Delta t = 0.4$ para o FALQON, $\Delta t = 0.2$ e $a = 2$ para o TR-FALQON, e $\Delta t = 0.42$ para o SO-FALQON. No painel a), energia média por iteração; no painel b), probabilidade média da solução ótima.

6.5 Comparação quantitativa entre métodos

Os resultados das Figuras 4, 5 e 6 permitem uma comparação direta entre dois paradigmas distintos. No primeiro, representado pelo FALQON e por suas variantes profundas, a melhora da solução é obtida majoritariamente por meio do aumento de L . No segundo, representado pelo MGI-FALQON, parte desse aprofundamento é substituída por uma sequência de execuções rasas intercaladas com atualização do estado inicial.

Do ponto de vista da profundidade interna por circuito, a diferença é imediata. Nos experimentos com FALQON, TR-FALQON e SO-FALQON, utilizou-se $L = 200$, enquanto nos experimentos com MGI a profundidade interna foi fixada em $L = 2$. Isso corresponde, por execução individual, a uma redução de profundidade de duas ordens de grandeza. Mais importante do que esse fator isolado é observar que, na instância estudada, essa redução drástica não comprometeu o desempenho; ao contrário, o MGI com filtro variável superou os métodos profundos tanto em energia quanto em probabilidade média de sucesso. Além disso, quando combinado com variantes internas, o melhor resultado global foi obtido justamente pelo MGI-TR-FALQON.

Naturalmente, essa comparação não deve ser interpretada como uma redução uniforme do custo total em todos os recursos. O MGI troca profundidade interna por iterações externas, amostragem repetida e pós-processamento clássico. A cada iteração, é necessário executar novamente o circuito raso, coletar *bitstrings*, aplicar o filtro escolhido, estimar marginais e preparar um novo estado inicial. Ainda assim, no contexto NISQ, essa troca tende a ser vantajosa, pois o aumento da profundidade costuma ser um dos principais fatores de degradação experimental devido à decoerência e ao acúmulo de erros de porta.

Os resultados obtidos permitem destacar alguns pontos importantes. O primeiro é que TR-FALQON e SO-FALQON cumprem bem o papel de acelerar a dinâmica do FALQON em regimes de profundidade intermediária. Isso fica particularmente evidente pela rapidez com que aumentam a probabilidade de sucesso para valores de L nas primeiras dezenas. No entanto, essa aceleração não elimina o gargalo da profundidade: mesmo em versões mais eficientes, ainda se trabalha com circuitos relativamente longos.

O segundo ponto é que a análise conjunta de energia e probabilidade revela

nuances que seriam perdidas caso apenas uma métrica fosse considerada. Na Figura 4, por exemplo, TR-FALQON e SO-FALQON terminam ligeiramente melhores em energia, mas o FALQON padrão alcança a maior probabilidade final. Isso mostra que a aproximação da energia mínima não se traduz necessariamente, de forma direta, no maior peso probabilístico sobre a solução ótima individual.

O terceiro ponto, e o mais relevante para esta dissertação, é que o MGI altera qualitativamente o papel da profundidade no algoritmo. Em vez de depender apenas do aprofundamento do circuito, a melhora passa a surgir do acúmulo de informação estatística extraída das medições. A escolha do filtro torna-se então uma peça central do desempenho. Os resultados mostram com clareza que um filtro excessivamente restritivo, como $n = 5$, pode gerar ganhos rápidos no início, mas compromete a evolução posterior. Em contraste, filtros mais amplos ou adaptativos preservam uma parcela maior das frequências empíricas das *bitstrings* selecionadas. Essas frequências determinam as marginais c_i de cada qubit, que são convertidas nos ângulos das rotações R_y e, conseqüentemente, nas amplitudes do estado produto preparado para a execução seguinte. Assim, a informação estatística incorporada à física do estado corresponde às probabilidades marginais inferidas da distribuição filtrada. Ao evitar uma concentração prematura dessas marginais, filtros mais amplos ou adaptativos mantêm maior diversidade na inicialização e permitem seu refinamento ao longo das iterações externas.

Por fim, a comparação entre variantes mostra que o MGI não atua de forma idêntica sobre qualquer dinâmica interna. Embora ele beneficie todas as variantes testadas, os ganhos são mais pronunciados quando combinado ao TR-FALQON. Isso sugere que ainda há uma estrutura fina na interação entre a regra de evolução interna e o mecanismo de reinicialização, aspecto que merece investigação futura.

6.6 Limitações práticas e perspectivas

Uma questão central deste trabalho é saber se a redução de profundidade promovida pelo MGI torna o método mais próximo de uma implementação real em *hardware* quântico. O principal aspecto favorável é que o MGI preserva circuitos internos rasos, o que reduz a exposição do sistema à decoerência e ao acúmulo de

erros. Além disso, a atualização do estado inicial entre iterações pode ser implementada por rotações locais, o que constitui uma operação relativamente simples do ponto de vista experimental. Outro fator positivo é que o método não exige uma etapa global de otimização clássica contínua, como ocorre em algoritmos variacionais tradicionais. O processamento clássico envolvido consiste essencialmente em contagem de *bitstrings*, filtragem, estimação de marginais e atualização de parâmetros locais.

Por outro lado, ainda permanecem desafios relevantes. O método depende de estatísticas de medição suficientemente confiáveis, de modo que ruídos de leitura podem distorcer as marginais estimadas e prejudicar a reinicialização. Além disso, embora cada circuito seja raso, o protocolo requer muitas execuções para alimentar o laço externo, o que pode elevar o custo total de amostragem. Soma-se a isso o fato de que a estimação dos observáveis de *feedback* do próprio FALQON continua sendo uma etapa não trivial.

Desse modo, o MGI não elimina o custo experimental, mas redistribui esse custo de uma forma potencialmente mais favorável ao regime NISQ: menos profundidade por circuito e mais realimentação baseada em medições.

Apesar dos resultados promissores, ainda existem limitações importantes. Uma das limitações está associada à própria escolha do filtro. Os resultados deste capítulo mostram que o filtro tem papel decisivo, mas ainda não existe uma regra geral que determine qual política deve ser adotada em diferentes problemas, tamanhos de instância ou profundidades internas. O bom desempenho do filtro variável indica que abordagens adaptativas são promissoras, mas sua fundamentação teórica ainda está em desenvolvimento.

A segunda limitação refere-se ao compromisso entre profundidade interna, número de iterações externas e custo total de amostragem. Embora o MGI permita trabalhar com circuitos extremamente rasos, ele pode demandar muitas execuções para atingir seu melhor desempenho. Assim, permanece em aberto a questão de qual é o regime ótimo de equilíbrio entre esses recursos.

Por fim, ainda é necessário compreender de forma mais completa o comportamento do método sob ruído realista. Os resultados apresentados neste capítulo concentram-se em simulações controladas, e a próxima etapa natural é incorporar modelos de ruído e, posteriormente, realizar testes em *hardware* real.

7 Conclusão

Esta dissertação teve como objetivo investigar algoritmos quânticos baseados em *feedback*, com ênfase no FALQON e em estratégias para reduzir sua profundidade de circuito em cenários NISQ. Para isso, o trabalho reuniu três frentes principais: a consolidação dos fundamentos teóricos necessários à formulação dos problemas, a análise de variantes da literatura voltadas à melhoria da dinâmica do FALQON e a proposição da estratégia *Measurement-Guided Initialization*, denominada MGI-FALQON.

Do ponto de vista conceitual, o trabalho mostrou que o FALQON constitui uma alternativa relevante a abordagens variacionais usuais, pois substitui a otimização clássica global por uma regra explícita de realimentação inspirada em controle por Lyapunov. Essa característica torna o método mais simples em sua atualização dos parâmetros e, ao mesmo tempo, a análise também evidenciou sua principal limitação prática: a necessidade de valores elevados de L para atingir soluções de alta qualidade.

A investigação das variantes TR-FALQON e SO-FALQON mostrou que é possível melhorar a eficiência da convergência por meio de modificações na dinâmica interna do algoritmo. O TR-FALQON mostrou-se capaz de tornar a evolução mais eficaz em regimes rasos e intermediários de L , enquanto o SO-FALQON indicou que a inclusão de informação de segunda ordem pode acelerar a redução da energia para os menores valores de L . Ainda assim, embora apresentem melhora em relação ao FALQON padrão, essas variantes continuam exigindo profundidades elevadas no regime analisado.

A principal contribuição desta dissertação foi a formulação e a avaliação do MGI-FALQON. A proposta consiste em utilizar execuções rasas do algoritmo para coletar *bitstrings*, selecionar as mais frequentes e construir, a partir das marginais observadas, um novo estado inicial para a execução seguinte. Com isso, parte do esforço computacional deixa de ser concentrada apenas no aumento da profundidade interna e passa a ser distribuída em um processo iterativo externo guiado por medições.

Os resultados obtidos para o problema MaxCut indicaram que essa estratégia é promissora. Observou-se que o desempenho do método depende fortemente da política de filtragem adotada, sendo que o filtro variável apresentou o melhor comportamento

global entre os casos testados. Além disso, verificou-se que o MGI-FALQON, mesmo operando com profundidade interna $L = 2$ por execução, foi capaz de superar o desempenho dos métodos profundos com $L = 200$ na instância considerada. Esse resultado mostra que o refinamento iterativo da inicialização pode compensar de forma bastante eficiente limitações severas de profundidade.

Também foi verificado que o laço externo de reinicialização atua de maneira complementar às variantes internas do FALQON. Todas as versões analisadas foram beneficiadas pela estratégia proposta, com destaque para a combinação entre MGI e TR-FALQON, que apresentou o melhor desempenho no cenário estudado. Esse resultado sugere que existe uma interação relevante entre a reinicialização guiada por medições e a dinâmica interna do algoritmo.

De modo geral, os resultados permitem concluir que os objetivos centrais da dissertação foram alcançados. O trabalho consolidou os fundamentos dos algoritmos de *feedback* quântico, analisou variantes relevantes da literatura e mostrou que a inicialização guiada por medições constitui uma estratégia efetiva para reduzir a dependência de circuitos profundos. Assim, o MGI-FALQON se apresenta como uma alternativa consistente para ampliar a viabilidade do FALQON em dispositivos quânticos de curto prazo.

Apesar disso, a abordagem ainda apresenta limitações. O ganho em profundidade vem acompanhado de maior custo de amostragem e dependência de estatísticas confiáveis de medição. Além disso, a escolha da política de filtragem mais adequada ainda depende de análise empírica, o que indica espaço para desenvolvimento teórico adicional.

Como continuidade natural deste trabalho, uma primeira direção consiste em incorporar modelos de ruído às simulações, a fim de avaliar a robustez do método em condições mais realistas. Outra direção importante é sua implementação em *hardware* quântico real, especialmente em plataformas compatíveis com *Qiskit*, permitindo investigar o comportamento do algoritmo diante de limitações experimentais concretas.

Em síntese, esta dissertação mostrou que o MGI-FALQON representa uma estratégia promissora para tornar algoritmos quânticos baseados em *feedback* mais compatíveis com o regime NISQ. Ao explorar sistematicamente a informação estatística

obtida nas medições, a abordagem proposta amplia as possibilidades de aplicação do FALQON e estabelece uma base sólida para investigações futuras, tanto teóricas quanto experimentais.

Referências

- 1 NIELSEN, M. A.; CHUANG, I. L. *Quantum Computation and Quantum Information*. [S.l.]: Cambridge University Press, 2000.
- 2 PRESKILL, J. Quantum Computing in the NISQ era and beyond. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 2, p. 79, ago. 2018. ISSN 2521-327X. Disponível em: <<https://doi.org/10.22331/q-2018-08-06-79>>.
- 3 PEXE, G. E. L. et al. *Quantum Optimization Algorithms for Strongly Correlated Many-Body Systems*. 2026. Disponível em: <<https://doi.org/10.48550/arXiv.2606.03147>>.
- 4 FARHI, E.; GOLDSTONE, J.; GUTMANN, S. *A Quantum Approximate Optimization Algorithm*. 2014. Disponível em: <<https://arxiv.org/abs/1411.4028>>.
- 5 PERUZZO, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, Springer Science and Business Media LLC, v. 5, n. 1, jul 2014. Disponível em: <<https://doi.org/10.1038/ncomms5213>>.
- 6 LUCAS, A. Ising formulations of many np problems. *Frontiers in Physics*, v. 2, 2014. ISSN 2296-424X. Disponível em: <<https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2014.00005>>.
- 7 ALBASH, T.; LIDAR, D. A. Adiabatic quantum computation. *Reviews of Modern Physics*, v. 90, n. 1, p. 015002, 2018.
- 8 CEREZO, M. et al. Variational quantum algorithms. *Nature Reviews Physics*, Springer Science and Business Media LLC, v. 3, n. 9, p. 625–644, ago. 2021. ISSN 2522-5820. Disponível em: <<http://dx.doi.org/10.1038/s42254-021-00348-9>>.
- 9 MCCLEAN, J. R. et al. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, IOP Publishing, v. 18, n. 2, p. 023023, fev. 2016. ISSN 1367-2630. Disponível em: <<http://dx.doi.org/10.1088/1367-2630/18/2/023023>>.
- 10 MAGANN, A. B. et al. Feedback-based quantum optimization. *Physical Review Letters*, American Physical Society (APS), v. 129, n. 25, dec 2022. Disponível em: <<https://doi.org/10.1103/PhysRevLett.129.250502>>.

- 11 LARSEN, J. B. et al. Feedback-based quantum algorithms for ground state preparation. *Phys. Rev. Res.*, American Physical Society, v. 6, p. 033336, Sep 2024. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevResearch.6.033336>>.
- 12 MAGANN, A. B. et al. Lyapunov-control-inspired strategies for quantum combinatorial optimization. *Phys. Rev. A*, American Physical Society, v. 106, p. 062414, Dec 2022. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.106.062414>>.
- 13 ARAI, D. et al. Scalable circuit depth reduction in feedback-based quantum optimization with a quadratic approximation. *Physical Review Research*, American Physical Society, v. 7, n. 1, p. 013035, jan. 2025.
- 14 ZHOU, L. et al. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Physical Review X*, American Physical Society, v. 10, n. 2, p. 021067, 2020.
- 15 RATTIGHIERI, L. et al. Accelerating feedback-based quantum algorithms through time rescaling. *Physical Review A*, APS, v. 112, n. 4, p. 042607, 2025.
- 16 RATTIGHIERI, L. A. M. et al. *Measurement-Guided State Refinement for Shallow Feedback-Based Quantum Optimization Algorithm*. 2026. Disponível em: <<https://arxiv.org/abs/2602.20407>>.
- 17 KORTE, B.; VYGEN, J. *Combinatorial Optimization: Theory and Algorithms*. 6. ed. [S.I.]: Springer, 2018. v. 21. (Algorithms and Combinatorics, v. 21).
- 18 SCHRIJVER, A. *Combinatorial Optimization: Polyhedra and Efficiency*. [S.I.]: Springer, 2003. v. 24. (Algorithms and Combinatorics, v. 24).
- 19 GAREY, M. R.; JOHNSON, D. S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: W. H. Freeman, 1979. ISBN 978-0716710455.
- 20 GOEMANS, M. X.; WILLIAMSON, D. P. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, Association for Computing Machinery, New York, NY, USA, v. 42, n. 6, p. 1115–1145, nov. 1995. ISSN 0004-5411. Disponível em: <<https://doi.org/10.1145/227683.227684>>.
- 21 LEWIS, M.; GLOVER, F. Quadratic unconstrained binary optimization problem preprocessing: Theory and empirical analysis. *Networks*, v. 70, n. 2, p. 79–97, 2017. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/net.21751>>.
- 22 GLOVER, F. W.; KOCHENBERGER, G. A. A tutorial on formulating QUBO models. *CoRR*, abs/1811.11538, 2018. Disponível em: <<http://arxiv.org/abs/1811.11538>>.

- 23 DATE, P.; ARTHUR, D.; PUSEY-NAZZARO, L. Qubo formulations for training machine learning models. *Scientific Reports*, Springer Science and Business Media LLC, v. 11, n. 1, maio 2021. ISSN 2045-2322. Disponível em: <<http://dx.doi.org/10.1038/s41598-021-89461-4>>.
- 24 PRADO, P. M. et al. *Quantum Feedback Algorithms for DNA Assembly Using FALQON Variants*. 2026. Disponível em: <<https://doi.org/10.48550/arXiv.2602.21080>>.
- 25 TILLY, J. et al. The variational quantum eigensolver: A review of methods and best practices. *Physics Reports*, Elsevier BV, v. 986, p. 1–128, nov. 2022. ISSN 0370-1573. Disponível em: <<http://dx.doi.org/10.1016/j.physrep.2022.08.003>>.
- 26 KRAUS, K. *States, Effects, and Operations: Fundamental Notions of Quantum Theory*. Berlin, Heidelberg: Springer, 1983. v. 190. (Lecture Notes in Physics, v. 190).
- 27 BUSCH, P.; LAHTI, P. Observable. In: GREENBERGER, D.; HENTSCHEL, K.; WEINERT, F. (Ed.). *Compendium of Quantum Physics*. Berlin, Heidelberg: Springer, 2009. p. 425–428. ISBN 978-3-540-70626-7.
- 28 BARENCO, A. et al. Elementary gates for quantum computation. *Physical Review A*, v. 52, n. 5, p. 3457–3467, 1995.
- 29 HADFIELD, S. et al. From the quantum approximate optimization algorithm to a quantum alternating operator ansatz. *Algorithms*, v. 12, n. 2, p. 34, 2019.
- 30 CONG, S.; MENG, F. A survey of quantum lyapunov control methods. *The Scientific World Journal*, Wiley, v. 2013, n. 1, jan. 2013. ISSN 1537-744X. Disponível em: <<http://dx.doi.org/10.1155/2013/967529>>.
- 31 GRIVOPOULOS, S.; BAMIEH, B. Lyapunov-based control of quantum systems. In: *Proceedings of the 42nd IEEE Conference on Decision and Control*. Maui, HI, USA: IEEE, 2003. v. 1, p. 434–438.
- 32 TROTTER, H. F. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, v. 10, n. 4, p. 545–551, aug 1959.
- 33 SUZUKI, M. Generalized trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, v. 51, n. 2, p. 183–190, 1976.
- 34 CHILDS, A. M. et al. Theory of trotter error with commutator scaling. *Physical Review X*, v. 11, n. 1, p. 011020, feb 2021.

- 35 BRADY, L. T.; HADFIELD, S. Feedback-based optimally controlled quantum states. *Physical Review A*, v. 111, n. 6, p. 062406, 2025.
- 36 MALLA, R. K. et al. Feedback-based quantum algorithm inspired by counterdiabatic driving. *Physical Review Research*, v. 6, n. 4, p. 043068, 2024.
- 37 CHANDARANA, P. et al. *Lyapunov Controlled Counterdiabatic Quantum Optimization*. 2024. Disponível em: <<https://arxiv.org/abs/2409.12525>>.
- 38 PEXE, G. E. L. et al. Using a feedback-based quantum algorithm to analyze the critical properties of the annni model without classical optimization. *Physical Review B*, v. 110, n. 22, p. 224422, 2024.
- 39 HALL, B. C. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. 2. ed. Cham: Springer, 2015. v. 222. (Graduate Texts in Mathematics, v. 222).
- 40 WEIDENFELLER, J. et al. Scaling of the quantum approximate optimization algorithm on superconducting qubit based hardware. *Quantum*, v. 6, p. 870, 2022.
- 41 BHARTI, K. et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.*, American Physical Society, v. 94, p. 015004, Feb 2022. Disponível em: <<https://link.aps.org/doi/10.1103/RevModPhys.94.015004>>.
- 42 PELOFSKE, E.; BÄRTSCHI, A.; EIDENBENZ, S. Short-depth QAOA circuits and quantum annealing on higher-order Ising models. *npj Quantum Information*, v. 10, p. 30, mar. 2024.
- 43 PELOFSKE, E. et al. Scaling whole-chip QAOA for higher-order Ising spin glass models on heavy-hex graphs. *npj Quantum Information*, v. 10, p. 109, nov. 2024.
- 44 EGGER, D. J.; MAREČEK, J.; WOERNER, S. Warm-starting quantum optimization. *Quantum*, v. 5, p. 479, 2021.
- 45 RATTIGHIERI, L. *qtsuru: Quantum circuit and algorithm simulation library using PyTorch*. 2026. GitHub repository. Disponível em: <<https://github.com/Lucas-Rattighieri/qtsuru>>.