量子近似优化算法收敛性的初等证明

Lennart Binkowski,[1], * Gereon Koßmann,[1], † Timo Ziegler,[1,2,] ‡ 以及 Ren´e Schwonnek[1], §
[1] *Institut für Theoretische Physik, Leibniz Universität Hannover* [2] *Volkswagen AG, Berliner Ring 2, 38440 Wolfsburg*(日期：2023年2月13日)

量子交替算子拟设（量子近似优化算法的前身为量子近似优化算法）是求解组合优化问题最广泛使用的量子算法之一。然而，由于目前尚无关于量子近似优化算法收敛性的严格证明，本文给出了这样的证明。证明通过重新追溯量子绝热算法与量子近似优化算法之间的联系展开，并由此自然地提出了对“相位分离器”和“混合器”这两个术语更精确的定义。

arXiv:2302.04968v1 [quant-ph] 9 Feb 2023

I. 引言

在当前基于门的含噪量子计算机时代，变分量子算法（VQA）是研究的核心。首先，量子近似优化算法[1] 受到了学术界和工业界的广泛关注。与许多其他 VQA 一样，它旨在借助经典优化器，在量子计算机上求解组合优化问题（COP）（最大化 $f\colon\ \{0,1\}^n \to \mathbb{R}$ 并满足一些约束）。在某种程度上，它是量子绝热算法（QAA，[2]）的离散化、基于门的版本，而后者是一种连续时间算法。QAA 及与之密切相关的量子退火[3] 依靠缓慢演化量子系统（或某些外部参数），将一个已知的初始态转变为代表某个最优解的状态。由于其模拟结构，它们无法在基于门的架构上运行，而是运行在量子退火机（概述见[4] ）上；量子退火机构成了量子计算机架构的第二大类别。

在其最初的形式中，量子近似优化算法仅适用于无约束问题。将其应用范围扩展到约束问题的一种常见技术，是对约束进行软编码。也就是说，将约束作为附加项纳入目标函数，以惩罚不可行输入。然而，在若干实例中，人们观察到这种方法会产生不理想的输出分布，导致优化质量较差或出现可行性违背（例如参见 [5–7]）。为改进对约束问题的处理，Hadfield等人将量子近似优化算法扩展为量子交替算子拟设（QAOA，[8]），该拟设也支持对约束进行硬编码。也就是说，目标函数保持不变，转而严格确保可行性保持。

简而言之，QAOA电路由参数化相位分离门 U_{P} 和混合器门 U_{M}组成。这两类门都应保持可行性，从而在理想情况下使可行态经过映射后仍为可行态。随后，通过经典方法迭代优化电路参数，理应得到最优解的良好近似。只有在每个可行态都可达时，这一启发式论证才成立：只要参数值合适，QAOA电路就能够（近似地）生成每个可行态。通常，可行态的可达性仅取决于混合器 U_{M}的性质。

关于量子近似优化算法为何应当对每个仅有一个最优解的（无约束）组合优化问题收敛，[1]中已经给出了一份或多或少严谨的证明。这份证明概要则建立在量子近似优化算法与量子绝热算法之间的紧密联系，以及绝热演化/量子退火所依据的基本原理之上（数学处理见 [9] ）。然而，该证明既未以充分的数学细节展开，也未试图做到尽可能一般化。此外，由于量子近似优化算法包含与量子近似优化算法相似的原理，因此有理由认为，一旦将这一结果进行适当形式化，就可以将其推广到量子近似优化算法；令人意外的是，这项工作迄今尚未有人着手完成。本文将探讨这一问题，并提出相位分离器和混合器门的改进定义，使其与量子近似优化算法之间的联系更加清晰。

首先，我们证明在 第三节中，采用合适的初始哈密顿量和初始态时，量子绝热算法的收敛性。该证明基于 [1] 中上述证明思路。我们提炼出其基本原理，并由此得到了对 混合器哈密顿量的精确定义。然而，通过调用一个不要求 能隙

* lennart.binkowski@itp.uni-hannover.de
† kossmann@stud.uni-hannover.de
‡ timo.ziegler@volkswagen.de§
rene.schwonnek@itp.uni-hannover.de

条件的绝热定理版本，我们得到了一个更一般的结果，该结果不要求所考虑的优化问题只有一个最优解。

其次，我们在第四节中证明具有适当初始态的量子近似优化算法的收敛性。为此，我们推广量子近似优化算法中所提出的原始混合器的所有性质。我们定义了同时和顺序混合器的相应版本，它们直接利用了刚刚推广的性质。随后，收敛性证明建立在量子绝热算法实例的收敛性之上，该实例将相应的混合器哈密顿量作为初始哈密顿量。其基本思想同样源于 Farhi 等人，但针对约束问题和顺序混合器进行了适当推广。

II. 预备知识

1. 组合优化问题

下文中，我们仅讨论最大化问题，因为最小化任务可以用类似方式处理。选择这种优化方向只是为了使我们能够更简洁地陈述收敛性证明。一个规模为 N 的通用组合优化问题具有如下形式

$$\max_{\boldsymbol{z}\in S} f(\boldsymbol{z}), \quad S \subseteq Z(N), \tag{1}$$

其中，$Z(N)$表示长度为 N的比特串集合，$f\colon Z(N) \to \mathbb{R}$是目标函数，而 $S \subseteq Z(N)$是可行比特串集合或解集。当 $S = Z(N)$ 时，该问题称为无约束问题。此外，我们将使 f 达到最大值的所有解的集合记为 $S_{\max}$。

2. 量子计算机上的问题编码

为了借助量子计算机处理组合优化问题，首先必须将问题转换为量子力学语言。标准编码过程将每个比特串 $\boldsymbol{z}$ 对应到 N 量子比特空间 $\mathcal{H}$ 中的一个计算基态$|\boldsymbol{z}\rangle :=\mathbb{C}^{2^N}$。此外，还将经典目标函数 f视为目标哈密顿量 C，定义如下:

$$C \coloneqq \sum_{\boldsymbol{z}\in Z(N)} f(\boldsymbol{z})\, |\boldsymbol{z}\rangle\langle\boldsymbol{z}|\,. \tag{2}$$

在这种设定下，最优解比特串张成最优解空间

$$\mathcal{S}_{\max} \coloneqq \operatorname{span}\{|\boldsymbol{z}\rangle \,:\, \boldsymbol{z} \in S_{\max}\} \subseteq \mathcal{S} \coloneqq \operatorname{span}\{|\boldsymbol{z}\rangle \,:\, \boldsymbol{z} \in S\} \subseteq \mathcal{H}. \tag{3}$$

现在，最大化任务等价于在 $\mathcal{S}_{\max}$ 中寻找一个计算基态。根据构造，$\mathcal{S}_{\max}$ 是 $C|_{\mathcal{S}}$ 对应于其最大特征值的特征子空间。下文中，我们将稍微放宽量子优化任务的要求，因为我们会将 $C|_{\mathcal{S}}$ 的任意最高能态视为最优解。

3. 量子绝热算法

简而言之，连续时间量子绝热算法(QAA，[2]) 通过相对于含时哈密顿量的准绝热演化，对初始态 $|\iota\rangle$ 进行本征态搜索。含时哈密顿量 $H(t)$ 在初始哈密顿量 H_{I} 与目标哈密顿量 C之间插值。对于最大化任务，$|\iota\rangle$ 应为 H_{I}的最高能态。插值哈密顿量通常由以下凸组合给出：[10]

$$H(t) = H_{\mathrm{lin}(H_{\mathrm{I}},C)}(t) \coloneqq (1-t)H_{\mathrm{I}} + tC, \quad t \in [0,1]. \tag{4}$$

演化速度由参数 $T > 0$控制：实际的时间演化是相对于 $H(s/T)$ 进行的。$s \in [0,T]$量子绝热算法背后的直觉是，如果能级始终保持分离，那么将 $H(0)$ 中的最高能态充分缓慢地演化（即 $T \gg 1$），就会得到 C 的最高能态。绝热定理为此提供了严格的数学依据（见第三节）。

4. 量子近似优化算法

量子近似优化算法 [1] 在某种意义上可以看作是量子绝热算法的离散版本，且初始态固定

$$|+\rangle \coloneqq \frac{1}{\sqrt{2^N}} \sum_{\boldsymbol{z}\in Z(N)} |\boldsymbol{z}\rangle \tag{5}$$

和初始哈密顿量

$$B \coloneqq \sum_{n=1}^{N} \sigma_x^{(n)}. \tag{6}$$

注意， $|+\rangle$ 是 B的非简并最高能态。 B 和 C 被纳入参数化门中：

$$U_B(\beta) \coloneqq e^{-i\beta B} = \prod_{n=1}^{N} e^{-i\beta\sigma_x^{(n)}} \quad \text{and} \quad U_C(\gamma) \coloneqq e^{-i\gamma C}. \tag{7}$$

指定深度 $p \in \mathbb{N}$后，通过以下方式构造参数化试探态：

$$|\vec{\beta},\vec{\gamma}\rangle \coloneqq V(\vec{\beta},\vec{\gamma})\,|+\rangle \coloneqq \left(\prod_{q=1}^{p} U_B(\beta_q)U_C(\gamma_q)\right)|+\rangle \tag{8}$$

在迭代过程中，通过经典优化规则更新参数，以最大化期望值

$$F_p(\vec{\beta},\vec{\gamma}) = \langle\vec{\beta},\vec{\gamma}|C|\vec{\beta},\vec{\gamma}\rangle\,. \tag{9}$$

在计算基中测量最终结果 $|\beta_{\text{opt}},\ \vec{\gamma}_{\text{opt}}\rangle$ ，即可得到最优解近似的分布。

5. 量子交替算子拟设

基于量子近似优化算法的思想，量子交替算子拟设（QAOA， [8]）将其设计扩展到一般的约束问题。给定一个目标哈密顿量为 C 、解空间为 $\mathcal{S}$ 的组合优化问题，参数化门 $U_B(\beta)$ 将被针对具体问题的“混合器”门取代。为简便起见，我们将聚焦于每次迭代都使用相同混合器的情况。因此，我们可以将它们重新合并为一个混合器门 $U_{\text{M}}(\beta)$。它需要满足两个重要性质：

- 可行性保持：对于所有参数值 $\beta \in \mathbb{R}$ $U_{\text{M}}(\beta)(\mathcal{S}) \subseteq \mathcal{S}$ 都应成立。
- 解的完全混合：对于所有可行计算基态 $|\boldsymbol{z}\rangle, |\boldsymbol{z}'\rangle \in \mathcal{S}$，都应存在一个幂次$r \in \mathbb{N}$ 和一个参数值 $\beta \in \mathbb{R}$ ，使得 $\langle\boldsymbol{z}|U_{\text{M}}^r(\beta)|\boldsymbol{z}'\rangle \neq 0$。

此外，参数化门 $U_C(\gamma)$ 可以替换为更一般的“相位分离器”门 $U_{\text{P}}(\gamma)$，该门的行为类似于经典目标函数。为了使讨论更加具体，我们将进一步聚焦于 $U_{\text{M}}(\beta)$ 和 $U_{\text{P}}(\gamma)$ 由哈密顿量的指数算子（的乘积）给出的情况。

下文的收敛性考量自然会给出 $U_{\text{M}}(\beta)$ 的正确定义，具体见第四节。我们现在先定义相位分离器：

定义 1. 给定一个解空间为 $\mathcal{S}$ 、最优解空间为 $\mathcal{S}_{\max}$ 的组合优化问题，若哈密顿量 H 满足以下两个条件，则称其为相位分离哈密顿量：

(i) H 在计算基下是对角的。

(ii) 与 $H|_{\mathcal{S}}$ 的最大特征值对应的特征子空间是 $\mathcal{S}_{\max}$。

相应的（参数化）相位分离器为

$$U_{\text{P}}(H,\gamma) \coloneqq e^{-i\gamma H}. \tag{10}$$

III. 量子绝热算法的收敛性证明

我们首先考察量子绝热算法（QAA）的收敛性。尽管量子绝热算法最初是针对无约束问题提出的，但我们可以轻松地将这一思想扩展到具有非平凡解空间 $\mathcal{S}$ 的组合优化问题：初始哈密顿量 H_{I} 应保持可行性，即 $H(\mathcal{S}) \subseteq \mathcal{S}$，且初始态 $|\iota\rangle$ 应位于 $\mathcal{S}$ 内。此外，我们将目标哈密顿量 C 替换为更一般的相位分离哈密顿量 H_{P}，后者显然能够保持可行性。于是，对于每个 $t \in [0,1]$，相对于 $H_{\mathrm{lin}(H_{\mathrm{I}},H_{\mathrm{P}})}(t)$ 的时间演化作用于 $|\iota\rangle$ 后，仍将得到一个可行态。因此，我们实际上将自身限制在子空间 $\mathcal{S} \subseteq \mathcal{H}$ 中。

量子绝热算法的基本概念由绝热定理所描述。在我们的分析中，我们采用了比 Farhi 等人在 [2]中使用的更一般的版本。

定理 2（绝热定理，[11]）。设 $\{H(t): 0 \le t \le 1\} \subseteq \mathcal{L}(\mathcal{H})$ 是一族自伴算子，满足 $H(\cdot) \in C^2([0,1], \mathcal{L}(\mathcal{H}))$。对于 $T > 0$，令 $\tilde{U}_T$为下式的解

$$\frac{\mathrm{d}}{\mathrm{d}\,s}\tilde{U}_T(s) = -iH(s/T)\tilde{U}_T(s), \quad 0 \le s \le T; \quad \tilde{U}_T(0) = \mathbb{1} \tag{11}$$

并令 $U_T(t) := \tilde{U}_T(tT)$, $0 \le t \le 1$。设 $\lambda(t)$ 分别是 $H(t)$ 的一个特征值，其对应的谱投影为 $\mathrm{P}(t)$。此外，设 $P \in C^2([0,1], \mathcal{L}(\mathcal{H}))$，使得对于每个 $0 \le t \le 1$, $P(t)$ 都是一个投影，且满足$H(t)P(t) = \lambda(t)P(t)$。另外，对于几乎所有的 $t \in [0,1]$, $P(t) = \mathrm{P}(t)$ 应成立。于是

$$\lim_{T\to\infty} (\mathbb{1} - P(t))U_T(t)P(0) = 0 \tag{12}$$

在 t 中一致。

定理 2 基本上指出，在绝热极限下，若从 $H(0)$ 中对应于特征值 $\lambda(0)$ 的特征子空间（或其子空间）内出发，只要遵循由 H生成的时间演化，并且谱投影曲线 P 可以 C^2- 连续地延拓穿过所有可能的能级交叉，就会始终处于 $H(t)$ 中对应于特征值 $\lambda(t)$ 的特征子空间内。$0 \le t \le 1$相比之下，Farhi 等人使用的绝热定理版本禁止出现任何能级交叉（见[12]）。

文献 [1] 给出了量子绝热算法收敛性证明的一个概要，作为论证量子近似优化算法收敛性的中间步骤。除了绝热定理之外，该证明主要基于佩龙-弗罗贝尼乌斯定理。首先回顾矩阵语境下不可约性的定义。

定义 3. 矩阵 $A \in \mathbb{C}^{n\times n}$ 称为不可约，当且仅当不存在$\mathbb{C}^n$的真 A-不变坐标子空间。也就是说，被 A 保持不变的坐标子空间只有 $\{0\}$ 和 $\mathbb{C}^n$。

定理 4（佩龙-弗罗贝尼乌斯）。设 $A \in \mathbb{C}^{n\times n}$ 逐分量非负且不可约。那么 A 具有一个非简并的最大特征值。

关键的观察是，初始哈密顿量的矩阵表示（6）在计算基下满足佩龙-弗罗贝尼乌斯定理的两个要求。由于这也将在整个收敛性证明中发挥重要作用，我们利用这些性质首先给出混合器的定义。

定义 5. 哈密顿量 $B \in \mathcal{L}(\mathcal{H})$ 称为具有解空间 $\mathcal{S}$ 的组合优化问题的混合器，当且仅当 $B(\mathcal{S}) \subseteq \mathcal{S}$ 且$B|_{\mathcal{S}} \in \mathcal{L}(\mathcal{S})$ 在计算基下逐分量非负且不可约。

现在的思路是将佩龙-弗罗贝尼乌斯定理应用于线性插值 $H_{\mathrm{lin}(B,C)}(t)$，在每个时刻$0 \le t < 1$ 由此得出存在一条特征值曲线 $\lambda_{\max}$，连接$H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(0) = B|_{\mathcal{S}}$和 $H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(1) = C|_{\mathcal{S}}$的最大特征值。为此，我们需要下面这个显然的结果，其证明相当简单。

推论 6. 设 $A \in \mathbb{C}^{n\times n}$ 为对角的，且 $B \in \mathbb{C}^{n\times n}$ 为不可约的。那么 $A + B$ 也不可约。

定理 7 (量子绝热算法的收敛性)。考虑一个解空间为 $\mathcal{S} \subseteq \mathcal{H}$、最优解空间为 $\mathcal{S}_{\mathrm{opt}} \subseteq \mathcal{S}$且相位分离哈密顿量为 C的 $B \in \mathcal{L}(\mathcal{H})$。若 $B \in \mathcal{L}(\mathcal{H})$ 是定义 5 意义下的混合器哈密顿量，且 $|\iota\rangle \in \mathcal{S}$ 是 $B|_{\mathcal{S}}$的最高能态，那么

$$\lim_{T\to\infty} U_T(1)\,|\iota\rangle \in \mathcal{S}_{\mathrm{opt}}, \tag{13}$$

where U_T 是相对于 B 和 C之间线性插值的准绝热演化 .

在以下证明中，我们直接将所有出现的算子 $\mathcal{L}(\mathcal{S})$视为其在计算基下的矩阵表示。

证明。分别用 $\lambda_{\max}(t)$ 表示 $H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(t)$ 的最大特征值，其中 $0 \le t \le 1$。令 $0 \le t_0 < 1$。由于 $B|_{\mathcal{S}}$ 不可约，$(1-t_0)B|_{\mathcal{S}}$ 也不可约。由于 C 是对角的，根据推论 6，$H_{\mathrm{lin}(B,C)}(t_0)|_{\mathcal{S}} = (1-t_0)B|_{\mathcal{S}} + t_0 C|_{\mathcal{S}}$ 也不可约。不失一般性，假设 C 的谱 [13] 非负。那么，$B|_{\mathcal{S}}$ 和 $C|_{\mathcal{S}}$ 的各分量也都非负。综上，$H_{\mathrm{lin}(B,C)}(t_0)|_{\mathcal{S}}$ 各分量非负且不可约。根据佩龙-弗罗贝尼乌斯定理，$\lambda_{\max}(t_0)$ 是非简并的。此外，映射

$$H_{\mathrm{lin}(B,C)}|_{\mathcal{S}} : \mathbb{R} \to \mathcal{L}(\mathcal{S}), \quad t \mapsto H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(t) = (1-t)B|_{\mathcal{S}} + tC|_{\mathcal{S}} \tag{14}$$

是解析的，并且对于所有 $t \in \mathbb{R}$，$H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(t)$ 都是对称的。令 L 表示 $H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}$ 的能级交叉/特征值分裂所构成的离散集合。根据 [14, 定理 6.1], , $H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(t)$ 的瞬时特征值 $t \in \mathbb{R}$ 可排序为 $\{\lambda_m(t) : 1 \le m \le M\}, M \le 2^N\}$，使得 $[t \mapsto \lambda_m(t)] \in C^\omega(\mathbb{R}, \mathbb{R})$；对于相应的谱投影 $\mathrm{P}_m(t)$，有 $[t \mapsto \mathrm{P}_m(t)] \in C^\omega\big(\mathbb{R} \setminus L, \mathcal{L}(\mathcal{S})\big)$，对每个 $1 \le m \le M$ 均成立。此外，这些谱投影在 L 上具有可去奇点；也就是说，存在定义于整个 $\mathbb{R}$ 上的解析延拓 P_m，使得对于所有 $1 \le m \le M$，$P_m(t) = \mathrm{P}_m(t)$ 对 $t \in \mathbb{R} \setminus L$ 成立。由连续性可知，这些解析延拓本身是具有恒定秩的正交投影，并满足

$$H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(t) P_m(t) = \lambda_m(t) P_m(t)$$

对所有 $t \in \mathbb{R}$。不失一般性，假设 $\lambda_1(0) = \lambda_{\max}(0)$。由于 $\lambda_{\max}(t_0)$ 在 $0 \le t_0 < 1$上保持非简并，因此可得 $\lambda_1 \equiv \lambda_{\max}$ 在 $[0,1)$ 上成立；并且由 λ_1 的连续性可知，$\lambda_1 \equiv \lambda_{\max}$ 在 $[0,1]$上成立。此外，相应的谱投影$\mathbb{P}_1$ 在 $[0,1)$ 上有定义。因此，其延拓 P_1 满足应用 定理 2所需的所有性质，即（12）成立。由于 $P_1(0) = \mathbb{P}_1(0)$，特别可以得到

$$\begin{aligned} & 0 = \lim_{T\to\infty} (\mathbb{1} - P_1(1)) U_T(1) P_1(0) \, |\iota\rangle = \lim_{T\to\infty} (\mathbb{1} - P_1(1)) U_T(1) \, |\iota\rangle \\ \Leftrightarrow \quad & \lim_{T\to\infty} U_T(1) \, |\iota\rangle = P_1(1) \lim_{T\to\infty} U_T(1) \, |\iota\rangle \, . \end{aligned}$$

由于 $P_1(1)$ 是一个投影，且 $H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}(1) P_1(1) = \lambda_{\max}(1) P_1(1)$，因此可得（13）。 □

根据上述证明可以看出，特征值曲线 $\lambda_{\max}$ 不会与$H_{\mathrm{lin}(B,C)}|_{\mathcal{S}}$ 的其他任何特征值曲线相交，可能在 $t = 1$处相交的情况除外。在 [1], 中，通过假设组合优化问题只有一个最优解，也避免了在 $t = 1$ 处发生能级交叉，这意味着 $\lambda_{\max}(1)$ 是非简并的。不过，借助更一般形式的绝热定理，我们得以摆脱这一假设。

IV. 量子近似优化算法的收敛性证明

接下来，我们考察量子近似优化算法的收敛行为，该算法包含量子近似优化算法这一特例。其组成部分与我们推广后的量子绝热算法基本相同。不过，从面向应用的角度来看，将混合器哈密顿量分解为局部哈密顿量极具价值，而这一做法也是由量子近似优化算法引入的。依照定义 5的思路，我们提出对Hadfield等人的定义作如下调整。

定义 8. 给定一个解空间为 $\mathcal{S}$的组合优化问题，若哈密顿量族 $\{B_i\}_{i\in I} \subset \mathcal{L}(\mathcal{H})$ 满足以下条件，则称其为混合族：对于任意 $i \in I$，$B_i(\mathcal{S}) \subseteq \mathcal{S}$, $B_i|_{\mathcal{S}}$ 在计算基下均逐分量非负；并且，$\mathcal{S}$ 的任何在每个 B_i 作用下都保持不变的坐标子空间都是平凡的。

该定义 8 确实是对定义 5 的一种分解形式，可以如下论证：考虑每个算子 $B_i|_{\mathcal{S}}$ 在计算基下的矩阵表示，将其视为一个图的邻接矩阵，其中顶点对应于可行计算基态。从 B_1所对应的图出发，再添加一个算子 B_i ，相当于添加由 B_i的矩阵表示中的非零元素所表示的边。实际权重（即元素的值）并不重要，但分量非负性条件意味着求和过程中不会有元素相互抵消，也就是说，邻接矩阵为 G_I 的图的边集

$$B_I|_{\mathcal{S}}, \quad B_I \coloneqq \sum_{i\in I} B_i \tag{15}$$

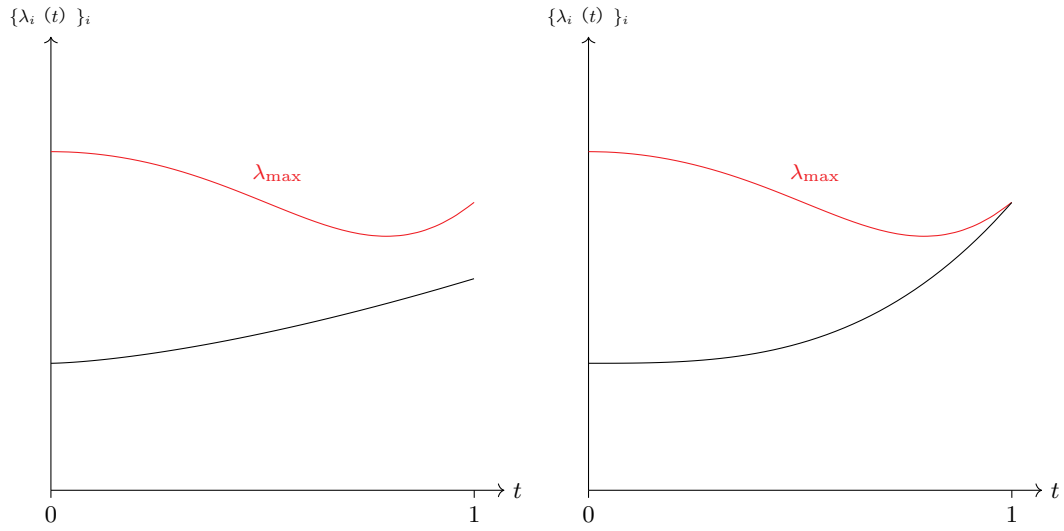


图1：特征值曲线 $\lambda_{\max}$ 在 $0 \le t < 1$ 内始终与其他所有特征值曲线保持分离。若对应的组合优化问题恰有一个最优解，则这种分离会延续至 $t = 1$ （左图）。然而，若组合优化问题有多个最优解，则 $\lambda_{\max}$ 会在 $t = 1$ 处与至少一条其他特征值曲线相交（右图）。

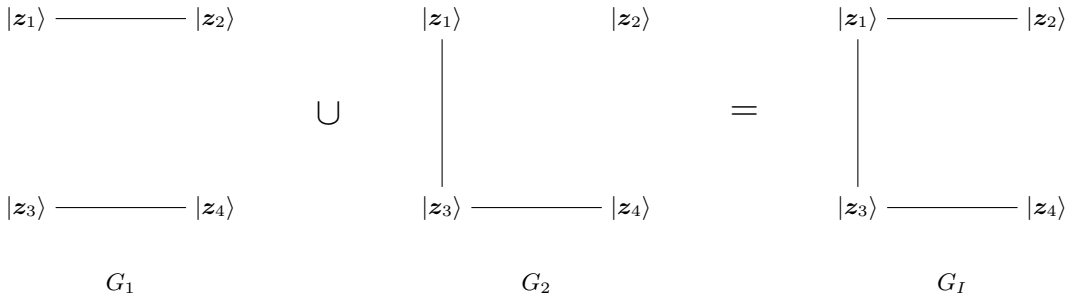


图2：命题 命题 9 的证明示意图。可行子空间 $\mathcal{S}$ 由 $|z_1\rangle$、$|z_2\rangle$、$|z_3\rangle$ 和 $|z_4\rangle$ 张成。 $B_1|_{\mathcal{S}}$ 的真不变坐标子空间为 $\mathrm{span}\{|z_1\rangle, |z_2\rangle\}$ 和 $\mathrm{span}\{|z_3\rangle, |z_4\rangle\}$，而 $B_2|_{\mathcal{S}}$ 的真不变坐标子空间为 $\mathrm{span}\{|z_1\rangle, |z_3\rangle, |z_4\rangle\}$ 和 $\mathrm{span}\{|z_2\rangle\}$。因此，它们共同的真不变坐标子空间集合为空，故 $B_I = B_1 + B_2$ 是一个混合器。

确实是各图的边集的并集，这些图分别对应邻接矩阵为 G_i 、 $B_i|_{\mathcal{S}}$、 $i \in I$ 的图。相互不变坐标子空间平凡这一条件，等价于 G_I 完全连通；而这又等价于其邻接矩阵不可约。因此，我们得出结论

命题 9. 给定一个解空间为 $\mathcal{S}$ 的组合优化问题，一族哈密顿算符 $\{B_i\}_{i\in I} \subset \mathcal{L}(\mathcal{H})$ 是混合族，当且仅当 B_I 是混合器哈密顿量。

利用混合族的定义，我们现在介绍我们所定义的“同时”混合器和“顺序”混合器。

定义10. 设 $\mathsf{H} = \{H_i\}_{i\in I} \subset \mathcal{L}(\mathcal{H})$ 是给定组合优化问题的一个混合族。相应的（参数化）同时混合器定义为

$$U_{\mathrm{M},0}(\mathsf{H}, \beta) \coloneqq e^{-i\beta \sum_{i\in I} H_i}. \tag{16}$$

指定一个 排列 $\sigma \in S(I)$ 后，相应的（参数化）顺序混合器定义为

$$U_{\mathrm{M},\sigma}(\mathsf{H}, \beta) \coloneqq \prod_{i\in I} e^{-i\beta H_{\sigma(i)}}. \tag{17}$$

根据定义可以立即得出，16和17均满足 QAOA 的原始要求：可行性保持和解的完全混合。然而，得益于我们改进后的定义，现在可以将[1] 中的收敛性证明概要扩展到一般的 QAOA 设置。具体步骤如下：

1. 将准绝热时间演化离散化 U_T

2. 使用（多元）李乘积公式分解 $H_{\mathrm{lin}(B,C)}$

3. 利用相应量子绝热算法实例的收敛性

我们先给出一个简单结论，讨论因子彼此接近的算子乘积之间的距离。

引理 11. 对于 $\varepsilon > 0$ 和 $m \in \mathbb{N}$，设 $\{V_j\}_{j=1}^m$、$\{W_j\}_{j=1}^m$、$\subset \mathcal{L}$（$\mathcal{H}$）是酉算子族，并满足

$$\|V_j - W_j\| < \varepsilon \tag{18}$$

对所有 $j \in [m]$ 均成立。则以下估计成立：

$$\left\| \prod_{j=1}^m V_j - \prod_{j=1}^m W_j \right\| < (1+\varepsilon)^m - 1. \tag{19}$$

证明。由于（18）成立，因此可以找到线性算子 $R_j \in \mathcal{L}(\mathcal{H})$，使得 $\|R_j\| \leq 1$ 且 $V_j = W_j + \varepsilon R_j$ 对每个 $j \in [m]$，分别成立。（19）显然对 $m = 1$ 成立。因此，只需证明：若（19）对某个 $m \in \mathbb{N}$ 成立，则它对 $m+1$ 也成立：

$$\begin{aligned}
\left\| \prod_{j=1}^{m+1} V_j - \prod_{j=1}^{m+1} W_j \right\| &= \left\| \left(\prod_{j=1}^m V_j \right) V_{m+1} - \left(\prod_{j=1}^m W_j \right) W_{m+1} \right\| \\
&= \left\| \left(\prod_{j=1}^m V_j \right) (W_{m+1} + \varepsilon R_{m+1}) - \left(\prod_{j=1}^m W_j \right) W_{m+1} \right\| \\
&= \left\| \left(\prod_{j=1}^m V_j \right) (\mathbb{1} - \varepsilon R_{m+1} W_{m+1}^*) - \left(\prod_{j=1}^m W_j \right) \right\| \\
&\leq \left\| \prod_{j=1}^m V_j - \prod_{j=1}^m W_j \right\| + \left\| \left(\prod_{j=1}^m V_j \right) \varepsilon R_{m+1} W_{m+1}^* \right\| \\
&< (1+\varepsilon)^m - 1 + \varepsilon \\
&\leq (1+\varepsilon)^m - 1 + \varepsilon(1+\varepsilon)^m = (1+\varepsilon)^{m+1} - 1.
\end{aligned}$$

□

定理 12（量子近似优化算法的收敛性）。考虑一个解空间为 $\mathcal{S} \subseteq \mathcal{H}$、最优解空间为 $\mathcal{S}_{\mathrm{opt}} \subseteq \mathcal{S}$、相位分离哈密顿量为 C、混合族为 $\{B_i\}_{i \in I}$ 的组合优化问题。令 U_{P} 和 U_{M} 分别为相应的相位分离器和（同时或顺序）混合器。此外，令 $|\iota\rangle \in \mathcal{S}$ 为 $B_I|_{\mathcal{S}}$ 的最高能态。那么，对于每个 $\varepsilon > 0$，都可以选择有限个参数 β 和 $\vec{\gamma}$，使得

$$\mathrm{dist}(|\vec{\beta}, \vec{\gamma}\rangle, \mathcal{S}_{\mathrm{opt}}) < \varepsilon, \tag{20}$$

其中

$$|\vec{\beta}, \vec{\gamma}\rangle := V(\vec{\beta}, \vec{\gamma}) |\iota\rangle := \left(\prod_q U_{\mathrm{M}}(\beta_q) U_{\mathrm{P}}(\gamma_q) \right) |\iota\rangle . \tag{21}$$

证明。令 U_T， $T > 0$表示关于 $H_{\mathrm{lin}(B_I,C)}$ 的准绝热演化。根据 命题 9， B_I 是一个符合 定义 5所述含义的混合器哈密顿量。因此，对于任意 $\varepsilon > 0$，定理 7 意味着存在一个 $T > 0$ 使得

$$\|(\mathbb{1} - P_1(1)) U_T(1) |\iota\rangle\| < \frac{\varepsilon}{2},$$

其中， P_1 是投影到$H_{\mathrm{lin}(B_I,C)}|_{\mathcal{S}}$的最高能特征子空间上的谱投影曲线的 C^2-延拓。不失一般性，假设 $\dim(\mathcal{S}) > 1$，否则结论显然成立。于是， $\alpha := \|\mathbb{1} - P_1(1)\|_{\mathcal{L}(\mathcal{S})} > 0$，因为

由连续性可知，$P_1(1)$的秩为 1。对准绝热时间演化 $U_T(1)$进行离散化，可知存在一个$m \in \mathbb{N}$，使得

$$\left\| \prod_{j=1}^{m} e^{-iH_{\mathrm{lin}(B_I,C)}\left(j\frac{T}{m}\right)j\frac{T}{m}} - U_T(1) \right\| < \frac{\varepsilon}{4\alpha}. \tag{22}$$

下面，令

$$W_j \coloneqq e^{-iH_{\mathrm{lin}(B_I,C)}\left(j\frac{T}{m}\right)j\frac{T}{m}}$$

并区分选择混合器的两种情况。

同时混合器：李乘积公式意味着，对于所有 $j \in [m]$，存在 $n_j \in \mathbb{N}$，使得对于所有 $\tilde{n} \geq n_j$，均有

$$\left\| \left(e^{-i\frac{1-j\frac{T}{m}}{\tilde{n}}j\frac{T}{m}B_I} e^{-i\frac{\left(j\frac{T}{m}\right)^2}{\tilde{n}}C} \right)^{\tilde{n}} - W_j \right\| < \sqrt[m]{\frac{\varepsilon}{4\alpha}+1} - 1, \tag{23}$$

分别。取 取 $n \coloneqq \max\{n_j : j \in [m]\}$，则此估计对所有 $j \in [m]$ 以及（尤其是） $\tilde{n}$ =均成立 n.

顺序混合器：不失一般性，选择排列 $\sigma = \mathrm{id}_I$。多变量李乘积公式 [15, 问题 IX.8。5] 意味着对于所有 $j \in [m]$，存在 $n_j \in \mathbb{N}$ 使得对于所有 $\tilde{n} \geq n_j$ 均有

$$\left\| \left(\left(\prod_{i\in I} e^{-i\frac{\left(1-j\frac{T}{m}\right)}{\tilde{n}}j\frac{T}{m}B_i} \right) e^{-i\frac{\left(j\frac{T}{m}\right)^2}{\tilde{n}}C} \right)^{\tilde{n}} - W_j \right\| < \sqrt[m]{\frac{\varepsilon}{4\alpha}+1} - 1, \tag{24}$$

分别。$\vec{}$ $\vec{}$ $\vec{}$在这两种情况下，选择 $q = n^m$ 参数值 $\beta = (\beta_1, \ldots, \beta_m)$和 $\vec{\gamma} = (\vec{\gamma}_1, \ldots, \vec{\gamma}_m)$为

$$\left(\vec{\beta}_j\right)_k = \frac{1 - j\frac{tT}{m}}{n} j\frac{tT}{m}$$
$$\left(\vec{\gamma}_j\right)_k = \frac{\left(j\frac{tT}{m}\right)^2}{n}$$

对所有 $k \in [n]$ 以及所有 $j \in [m]$。那么，根据构造，（23）和（24）可转化为

$$\left\| V(\vec{\beta}_j, \vec{\gamma}_j) - W_j \right\| < \sqrt[m]{\frac{\varepsilon}{4\alpha}+1} - 1.$$

因此，由（22）和 引理 11，可得

$$\begin{aligned}
\left\| V(\vec{\beta}, \vec{\gamma}) - U_T(t) \right\| &= \left\| \prod_{j=1}^{m} V(\vec{\beta}_j, \vec{\gamma}_j) - U_T(t) \right\| \\
&\leq \left\| \prod_{j=1}^{m} V(\vec{\beta}_j, \vec{\gamma}_j) - \prod_{j=1}^{m} W_j \right\| + \left\| \prod_{j=1}^{m} W_j - U_T(t) \right\| \\
&< \frac{\varepsilon}{4\alpha} + \frac{\varepsilon}{4\alpha} = \frac{\varepsilon}{2\alpha}.
\end{aligned}$$

总之，可得

$$\begin{aligned}
\left\| (\mathbb{1} - P_1(1)) \left|\vec{\beta}, \vec{\gamma}\right\rangle \right\| &= \left\| (\mathbb{1} - P_1(1)) V(\vec{\beta}, \vec{\gamma}) \left|\iota\right\rangle \right\| \\
&\leq \left\| (\mathbb{1} - P_1(1))(V(\vec{\beta}, \vec{\gamma}) - U_T(1)) \left|\iota\right\rangle \right\| + \left\| (\mathbb{1} - P_1(1)) U_T(1) \left|\iota\right\rangle \right\| \\
&< \left\| \mathbb{1} - P_1(1) \right\|_{\mathcal{L}(\mathcal{S})} \frac{\varepsilon}{2\alpha} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

然后，$\mathrm{im}(P_1(1)) \subseteq \mathcal{S}_{\mathrm{opt}}$ 证明了该断言。 □

V. 结论与展望

本文给出了量子近似优化算法收敛性的一个初等证明。该证明可视为对绝热定理进行离散化并谨慎扩展后的版本，其基础是 Farhi 等人的思想。除了另一个核心定理（佩龙-弗罗贝尼乌斯定理）之外，这一扩展仅基于初等矩阵不等式。最重要的是，我们的证明所需假设更少（允许存在多个最优解），并且适用于非平凡的可行性结构（$\mathcal{S} \subsetneq \mathcal{H}$）。

此外，证明自然引出了对QAOA 混合器和QAOA 相位分离器概念的进一步精确定义。最值得注意的是，在 QAOA 框架内恰当地重现经典可行性对称性时，也会得到完全相同的概念（见[16]）。这有力地表明，本文给出的定义最恰当地抓住了 QAOA 所依据的总体原理。

我们实质上证明了，将混合器哈密顿量 B限制在可行子空间 $\mathcal{S}$后，其不可约性和分量非负性是保证量子绝热算法和量子近似优化算法收敛的充分条件。此外，可以很容易验证，不可约性在以下意义上也是必要条件：给定任意初始态$|\iota\rangle$，并且存在一个非平凡的 $B|_{\mathcal{S}}$不变坐标子空间，则总存在某个目标哈密顿量C，使得量子绝热算法和量子近似优化算法无法以任意精度逼近 $\mathcal{S}_{\max}$ 中的任何状态。另一方面，要求 $B|_{\mathcal{S}}$ 具有分量非负性并非必要条件。在我们的收敛性证明中，我们施加这一条件，是为了应用佩龙-弗罗贝尼乌斯定理。然而，该定理也存在更一般的版本（例如见[17]），这些版本以保持给定锥并分别置换其各个面的更一般性质，取代了这一条件和不可约性。遗憾的是，相关的锥仅由 $\mathcal{S}$ 中的所有正交象限构成，因为 $\mathcal{S}$ 的每个坐标子空间都必须体现在这些锥的面上。因此，除允许一些额外的矩阵符号外，最终得到的仍是相同的条件。因此，我们认为难以进一步放宽定理 7和定理 12中的假设。

一个有趣且尚未解决的问题是，对于存在多个最优解的情形，是否也能刻画量子绝热算法的收敛速率。在这种情况下，谱隙必然会在 $t \to 1$ (见图 1) 处消失，但在整个区间 $[0, 1)$ 内始终保持有限。也就是说，尽管会发生能级交叉，但它只发生一次，而且发生时间可以预测。《绝热定理》中有一些关于收敛速率的结果，适用于各种（允许的）能级交叉情形（例如，见 [11, 18]）。针对量子绝热算法的具体情形对这些结果进行精细调整，有望成为一项富有启发性的未来研究。

致谢

我们感谢 Tim Heine、Lauritz van Luijk、Tobias J. Osborne、Christoph Pohl、Antonio Rotundo、Martin Steinbach 和 Reinhard F. Werner 富有帮助的讨论。GK 感谢德国学术交流中心提供的经费支持，IIT Indore（Kapil Ahuja）感谢其接待访问期间获得的经费支持。RS 感谢下萨克森量子谷以及德国联邦教育与研究部 ATIQ 项目提供的经费支持。LB 和 TZ 感谢德国联邦教育与研究部 QuBRA 项目提供的经费支持。

[1] E. Farhi、J. Goldstone 和 S. Gutmann，《量子近似优化算法》（2014），arXiv:1411.4028 [quant-ph]。[2] E. Farhi、J. Goldstone、S. Gutmann 和 M. Sipser，《绝热演化量子计算》（2000），arXiv:0001106 [quant-ph]。[3] T. Kadowaki 和 H. Nishimori，《物理评论E》58，5355（1998）。[4] P. Hauke、H. G. Katzgraber、W. Lechner、H. Nishimori 和 W. D. Oliver，《物理学进展报告》83，054401（2020）。[5] J. S. Baker 和 S. K. Radha，《Wasserstein 解质量与量子近似优化算法：投资组合优化案例研究》（2022年），arXiv:2202.06782 [quant-ph]。[6] P. D. de la Grand'rive 和 J.-F. Hullo，《用于电池收益优化的量子近似优化算法背包问题变体》（2019），arXiv:1908.02210 [quanth-ph]。[7] W. van Dam、K. Eldefrawy、N. Genise 和 N. Parham，《量子优化启发式算法及其在背包问题中的应用》（2021），arxiv:2108.08805 [quant-ph]。[8] S. Hadfield、Z. Wang、B. O'Gorman、E. Rieffel、D. Venturelli 和 R. Biswas，《算法》12，34（2019）。[9] S. Morita 和 H. Nishimori，《数学物理学杂志》49，125210（2008）。[10] 此外，还存在更复杂的凸组合，其中 H_I 和 C 前面的系数是 t 的非线性函数（见 [19]）。采用此类问题特定系数可以显著提高量子绝热算法的收敛速率。[11] S. Teufel，《数学物理学快报》58（2001）。[12] A. Messiah，*Quantum Mechanics, Vol. II*（阿姆斯特丹：北荷兰出版社；纽约：Wiley，1976）。

[13] 否则，考虑 $\tilde{B} := B - \lambda_{\min}\mathbb{1}$ 和 $\tilde{C} = C - \lambda_{\min}\mathbb{1}$，其中 $\lambda_{\min} < 0$ 是 C 的最小特征值。于是 $H_{\mathrm{lin}(\tilde{B},\tilde{C})} = H_{\mathrm{lin}(B,C)} - \lambda_{\min}\mathbb{1}$ 生成的时间演化与其相同，但相差一个全局相位。[14] T. Kato, 《*Pertubation Theory for Linear Operators*》（施普林格，柏林、海德堡，1995年）。[15] R. Bhatia, 《 *Matrix Analysis*》（施普林格，纽约州纽约，1997年）。[16] G. Koßmann、L. Binkowski、C. Tutschku 和 R. Schwonnek, 《带硬约束的开放式车间调度》（2022年），arXiv:2211.05822 [量子物理]。[17] H. Schneider 和 M. Vidyasagar，SIAM 数值分析杂志 7，508（1970年）。[18] J. E. Avron 和 A. Elgart，数学物理通讯 203，445（1999年）。[19] J. Roland 和 N. J. Cerf, 《物理评论 A》65，042308（2002年）。