迈向求解基数约束二元优化的指数级量子改进

Haomu Yuan,[1,2,*] Hanqing Wu,[3,*] Kuan-Cheng Chen,[4] Bin Cheng,[2] 以及 Crispin H. W. Barnes[1†]英国剑桥 *CB3 0US*，剑桥大学物理系卡文迪许实验室[2]新加坡，新加坡国立大学量子技术中心[†3]瑞典隆德 *220 07*，隆德大学统计系[4]英国伦敦 *SW7 2AZ*，伦敦帝国理工学院帝国理工量子工程、科学与技术中心（*QuEST*）（日期：2026年3月17日）

arXiv:2603.14744v1 [quant-ph] 16 Mar 2026

基数约束二元优化是一种基础计算原语，广泛应用于机器学习、金融和科学计算。本文提出一种基于Grover的量子算法，在关于解存在性的自然承诺条件下，利用固定基数可行子空间的结构。对于二次目标函数，无论固定基数 k 和最优解简并度 M 为何，我们的方法都只需 $\mathcal{O}\left(\sqrt{\binom{n}{k}_M}\right)$ 次Grover旋转；与在 $\{0,1\}^n$ 上进行无结构搜索相比，这使 Grover 迭代次数呈指数级减少。在此结果基础上，我们进一步构建了一个基于交替方向乘子法（ADMM）算法的经典—量子混合框架。该框架保证输出一个 ϵ 近似解，其一致性容差为 $\epsilon+\delta$，并且最多使用 $\mathcal{O}\left(\sqrt{\binom{n}{k}}\frac{n^6k^{3/2}}{\sqrt{M}\epsilon^2\delta}\right)$ 次二次预言机查询，同时只需 $\mathcal{O}\left(\frac{n^6k^{3/2}}{\epsilon^2\delta}\right)$ 经典开销。总体而言，我们的方法展示了量子资源的实用价值，并表明在某些参数范围内，相较于现有的基于Grover的方法可实现指数级改进，从而为在约束二元优化中实现量子优势铺平道路。

I. 引言

固定基数二元多项式规划（BPP-FC）在投资组合优化 [1–3], 生物学 [4, 5], 图与网络分析 [6–8], 以及稀疏线性回归 [9, 10]等领域有着广泛的应用。对于某些特定的 BPP-FC 问题，已有高效的近似算法。一个例子是著名的稠密 k 子图问题，Feige 等人证明，贪心算法可在多项式时间内实现近似比 $\mathcal{O}(n^\delta)$（其中$\delta<1/3$） [11]。他们还提到了一种基于与 Goemans 的私人交流、以半定规划为子程序并获得 n/k 近似比的方法。然而，一般而言，求解 BPP-FC 是NP困难的。随着量子计算蓬勃发展，许多量子算法已在无约束二次二元优化（QUBO）中展现出成功的加速，例如量子近似优化算法（QAOA） [12], 变分量子本征求解器（VQE） [13], 量子退火 [14], Grover 自适应搜索（GAS） [15],以及哈密顿量更新 [16]。由于在实际应用中展示量子优势的需求日益增长，面向固定基数问题的量子 QUBO 求解器变体近年来开始受到更多关注。这些方法可分为两种主要范式：含噪中等规模量子（NISQ）和容错量子计算（FTQC）。NISQ 算法（包括 QAOA 和 VQE）通常通过设计能够保持约束子空间结构的参数化算子来构造 [17, 18]。与此同时，在 FTQC 范式下，Gilliam 等人提出了一种用于二次规划的 Grover自适应搜索求解器，将软约束纳入目标函数 [19]。然而，关于加速效果的严格理论界限仍然鲜有研究。

本文介绍了一个针对 BPP-FC 的新型 GAS框架，并严格量化了其量子加速。该方法的关键在于为 Grover 搜索子程序中的扩散算子设计了一种新方案，使搜索动态局限于二次规划问题的固定基数子空间。作为应用，我们通过最先进的交替方向乘子法（ADMM）算法，将该方法扩展到 BPP-FC 的一种特定形式，从而结合经典优化与我们的量子求解器。整个流程都支持可追踪的复杂度分析，证实了所需计算资源能够高效扩展。具体而言，本文的第一部分研究了一般的具有固定基数约束的二元二次规划 （BQP-FC）

* 这些作者对本研究的贡献相同。
† Haomu Yuan hy374@cam.ac.uk

定义如下

$$\begin{aligned} \underset{\mathbf{x}\in\{0,1\}^n}{\arg\min} \quad & f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\intercal\boldsymbol{\Sigma}\mathbf{x} - \boldsymbol{\mu}^\intercal\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}^\intercal\mathbf{1} = k, \end{aligned} \tag{1}$$

其中， $k \in \{0,\ldots,n\}$ 是一个固定常数， $\boldsymbol{\mu} \in \mathbb{R}^n$ 和 $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ 分别表示目标函数的线性系数和二次系数。请注意，这个优化问题可以直接转化为满足 K-局部性的加权超图问题，其中 K 是多项式的最大度。此外，由于存在固定基数约束，式（1）也被视为一类图优化问题的推广，这类问题又称为最稠密 k-子集问题 [11, 20], 、最重无权子图问题 [21], 、 k-聚类问题[22], 或 k-基数子图问题[23]。一般而言，对于最大度为三的图，问题（1）是NP困难的 [24], ；对于最大度为二的图，则可在多项式时间内求解 [22]。

此外，我们将该框架应用于一个颇具挑战性的四次优化问题——风险平价模型，以展示其通用性。该模型在投资组合管理中对冲相对风险方面表现出显著的稳定性 [2] ，其形式如下：

$$\begin{aligned} \underset{\mathbf{x}\in\{0,1\}^n}{\arg\min} \quad & g(\mathbf{x}) = \sum_{\substack{1\le i,j\le n \\ i\ne j}} \left(x_i(\boldsymbol{\Sigma}\mathbf{x})_i - x_j(\boldsymbol{\Sigma}\mathbf{x})_j\right)^2 + \lambda\left(-\mathbf{x}^\intercal\boldsymbol{\mu} + \frac{1}{2}\mathbf{x}^\intercal\boldsymbol{\Sigma}\mathbf{x}\right) \\ \text{s.t.} \quad & \mathbf{x}^\intercal\mathbf{1} = k, \end{aligned} \tag{2}$$

其中， $\boldsymbol{\mu}$ 是预期收益向量， $\boldsymbol{\Sigma}$ 是协方差矩阵， λ 是权衡超参数。式（2）中的第一项旨在最小化资产间风险贡献的差异，以实现等风险配置。桥水基金的“全天候”策略使这一方法广为人知，该策略凭借其内在的稳定性和多元化实现了稳健的表现 [25, 26]。然而，风险平价问题属于四次优化问题，通常是NP困难的 [11], ，这给大规模投资组合带来了巨大的计算挑战。

本文其余部分安排如下。第二节回顾用于求解二次无约束二元优化问题的 Grover 自适应搜索及其 Grover 搜索子程序。第三节介绍一种用于求解固定基数二次二元规划的新型 Grover 自适应搜索算法。第四节详细介绍用于求解风险平价模型的经典-量子混合ADMM框架。第五节对我们的方法进行完整的资源分析。最后，我们在第六节总结全文并讨论未来方向。

II. Grover 搜索与 Grover 自适应搜索

本节简要回顾 Grover 搜索算法 [27] 和 Grover 自适应搜索（GAS）算法 [15, 28, 29], ，后者将在本文中作为求解固定基数二元优化问题的量子子程序。

a. Grover 搜索。考虑只有一个目标的无结构搜索问题。设 $\boldsymbol{\omega} \in \{0,1\}^n$ 为目标比特串，且 f： $\{0,1\}^n \to \{0,1\}$ 满足：当且仅当$\mathbf{x} = \boldsymbol{\omega}$时， f（$\mathbf{x}$） $= 1$。标准 Grover 搜索 [27] 从 $|s\rangle = \frac{1}{\sqrt{N}}\sum_{\mathbf{x}\in\{0,1\}^n}|\mathbf{x}\rangle$开始，即计算基态的均匀叠加态。在 $|s\rangle$ 上观测到目标的概率为 $1/N$。通过交替迭代应用预言机和扩散算子，可以放大这一概率。具体而言，预言机算子定义为

$$\hat{O}\,|\mathbf{x}\rangle = (-1)^{f(\mathbf{x})}\,|\mathbf{x}\rangle\,, \tag{3}$$

扩散算子定义为

$$\hat{A} = 2\,|s\rangle\langle s| - \hat{I} = \hat{H}^{\otimes n}(2\,|0^n\rangle\langle 0^n| - \hat{I})\hat{H}^{\otimes n}\ , \tag{4}$$

ˆ其中 H 是哈达玛门， I 是恒等算子。标准格罗弗算法可写为对 AO重复执行 r ˆ ˆ 次：

$$|\psi_r\rangle = (\hat{A}\hat{O})^r\,|s\rangle\,, \tag{5}$$

其中，在 $|\psi_r\rangle$ 上测得 $|\boldsymbol{\omega}\rangle$ 的概率为 $\sin^2((r+\frac{1}{2})\theta)$，且 $\theta = 2\arcsin\frac{1}{\sqrt{N}}$。假设 N 是一个大数， $\theta/2$ 位于区间 $[0,\pi/2]$ 内，与 $r\theta$ 相比可忽略不计。因此，当 $r\theta \approx \frac{\pi}{2}$ 时，我们 –

有 $r = \frac{\pi}{2\theta} - \frac{1}{2} \approx \frac{\pi}{4}\sqrt{N}$，并且测得 $|\omega\rangle$ 的概率接近 1。在具有M 个目标解的一般情况下，量子计数算法以 $\mathcal{O}(\sqrt{N/M})$ [30]的复杂度估计 M。

注 1. 我们将 AO 的一次应用称为 Grover 旋转，因为 AO 在由 $|\omega\rangle$ 和 $|\omega^{\perp}\rangle$ 张成的二维子空间中表现为旋转：$= \frac{1}{\sqrt{N-1}} \sum_{\mathbf{x}\neq\omega} |\mathbf{x}\rangle$。

b. Grover 自适应搜索。 Grover 自适应搜索算法 [15, 28, 29] 可通过自适应更新 Grover 搜索中使用的预言机，来求得二次无约束二元优化问题的最小值。例如，给定有界二次无约束二元优化目标函数 f：$\{0,1\}^n \to \mathbb{R}$，其中 $f(\mathbf{x}) < y'$，我们可以定义预言机算子 O_y 如下：

$$\hat{O}_y |\mathbf{x}\rangle = \mathrm{sgn}(f(\mathbf{x}) - y) |\mathbf{x}\rangle \tag{6}$$

其中 $y \leq y'$。为实现这一点，我们可以先将 $(f(\mathbf{x}) - y) \in [-2^{m-1}, 2^{m-1})$ 的二进制补码表示编码到 m 个辅助量子比特中。由于最后一位编码了数值的符号，因此可以对该符号位应用 Pauli-Z （由编码所需的任何额外工作量子比特控制），从而实现符号函数。 [19, 31]中给出了使用结构化旋转序列和逆量子傅里叶变换技术实现这种算术编码的高效构造——详情请参阅 B 节和表 I。此外，还可以使用量子加法电路和量子信号处理技术来构造符号函数预言机 [32, 33]。

GAS 随后根据 Grover 搜索子程序的结果迭代降低阈值 y，最终收敛到全局最小值（可能存在简并）。本文考虑 [15] 中提出的随机 Grover 自适应搜索，详情见算法 1。对于 $\xi = 1.34$，预期迭代次数的上界为 $1.32(N/t)^{1/2}$，预期预言机查询次数的上界为 $1.32\sqrt{N}\sum_{r=t+1}^{N} \frac{1}{\sqrt{r-1}} \approx 2.46(N/t)^{1/2}$ [29]，其中 t 是全局最优解的简并度。

Alg法 1： Grover 自适应搜索 (GAS) 输入：f：$\{0,1\}^n \to \mathbb{R}$, $\xi > 1$, $r_{\max} = 1$, $i = 1$结果：y均匀选择 $\mathbf{u}_1 \in \{0,1\}^n$，设定 $y_1 = f(\mathbf{u}_1)$，$\mathbf{x} = \mathbf{u}_1$，，且 $y = y_1$；重复执行：均匀选择旋转次数 r_i，取自 $\{0, \ldots, \lceil r_{\max} - 1 \rceil\}$；使用预言机$O_{y_i}$执行 r_i 次旋转的 Grover 搜索，并输出 $\mathbf{x}$ 和 y；若 $y < y_i$，则设定 $\mathbf{u}_{i+1} = \mathbf{x}$，$y_{i+1} = y$，且 $r_{\max} = 1$；否则设定 $\mathbf{u}_{i+1} = \mathbf{u}_i$，$y_{i+1} = y_i$，且 $r_{\max} = \xi r_{\max}$；结束$i = i + 1$；直至满足终止条件；

III. 用于固定基数二次规划的 Grover 自适应搜索

本节介绍用于硬约束规划的 Grover 搜索算法，并使用 Grover 自适应搜索框架求解式（1）的最小值。与用于软约束规划的 Grover 搜索算法相比，我们的方法可将迭代次数呈指数级减少。此外，我们还分析了该算法的门数量和电路深度。

在 [19], Gilliam等人 提出了一种格罗弗算法，通过在目标函数中添加惩罚项，求解软约束的式 (1) 二元二次规划。也就是说，针对软约束问题的格罗弗算法会基于加入约束惩罚项的新目标函数构造预言机。具体而言，可将格罗弗算法应用于以下优化问题

$$\arg\min_{\mathbf{x}\in\{0,1\}^n} \mathbf{x}^{\intercal}\boldsymbol{\Sigma}\mathbf{x} - \boldsymbol{\mu}^{\intercal}\mathbf{x} + \lambda(\sum_i x_i - k)^2, \tag{7}$$

[1]请注意，单个解的最优旋转次数为 $\lceil \pi\sqrt{N/4} \rceil$，因此无需进一步增加 $\lceil k \rceil$。在用于硬约束二次规划的 Grover 搜索框架中，我们将给出一个由问题约束定义的更一般的界——详见第 III 节。

其中， λ 是一个较大的数，用于确保 Grover 搜索满足约束。然而，惩罚变量 λ 只能对约束施加宽松的界限，因为取值过大可能导致原目标函数出现病态条件，而取值过小则可能导致违反约束。在这一假设下，最坏情况下的时间复杂度仍为 $\mathcal{O}(\sqrt{N/M})$。若按二项式定理考虑从 n 中选取 k 的所有组合，则存在一种经典约束蛮力算法，其时间复杂度为 $\binom{n}{k}$，可穷举满足固定基数约束的二次二元优化问题的所有可行解——详情见 [18] 中的附录 B。因此，尚不清楚软约束 Grover 算法是否比经典蛮力算法更高效。在本工作中，我们为硬约束问题构造了一种新颖的扩散算子，拓展了 [34]提出的框架。值得注意的是， [35] 开展的一项独立并行研究也采用 Dicke 态制备来施加固定基数约束；不过，他们的方法是在标准 Grover 搜索的背景下提出的，而非本文所探讨的 Grover 自适应搜索（GAS）范式。

在定理 1 中，我们定义了一种用于硬约束问题的 Grover 搜索算法，该算法保证相较于所有其他算法都能获得非平凡的加速。

定理 1（硬约束规划的 Grover 搜索）。 考虑一个受固定基数约束的二元二次规划，其可行集为 $\mathcal{C}$。令 $\mathcal{F} \subseteq \mathcal{C}$ 表示目标值严格优于 $\mathcal{C}\backslash\mathcal{F}$ 中目标值的解集，并令 $|\mathcal{F}| = M$。假设存在一个高效的量子电路 C，用于制备 $|s_c\rangle = \frac{1}{\sqrt{|\mathcal{C}|}}\sum_{\mathbf{x}\in\mathcal{C}}|\mathbf{x}\rangle$。那么，该硬约束问题的 *Grover* 搜索会在以下集合上生成均匀分布：

$\mathcal{F}$，其中每个解在约经过 $\frac{\pi}{4}\sqrt{\frac{|\mathcal{C}|}{|\mathcal{F}|}}$ 次迭代后被测得的概率为 $\frac{1}{M}$。特别地，当 $\mathcal{C}$ 由固定基数 k 定义时，迭代次数为 $\frac{\pi}{4}\sqrt{\binom{n}{k}/M}$。

证明。 硬约束问题的 Grover 扩散算子定义如下：

$$\hat{A}_c := 2\,|s_c\rangle\langle s_c| - \hat{I} = \hat{C}(2\,|0^n\rangle\langle 0^n| - \hat{I})\hat{C}^\dagger. \tag{8}$$

与原始格罗弗算法类似，初始态可写为：

$$|s_c\rangle = \cos(a)|\bar{0}\rangle + \sin(a)\,|\bar{1}\rangle\,, \tag{9}$$

其中

$$\begin{aligned} \sin(a) &= \sqrt{\frac{|\mathcal{F}|}{|\mathcal{C}|}}, \qquad & \cos(a) &= \sqrt{1-\frac{|\mathcal{F}|}{|\mathcal{C}|}} \\ |\bar{1}\rangle &= \frac{1}{\sqrt{|\mathcal{F}|}}\sum_{\mathbf{x}\in\mathcal{F}}|\mathbf{x}\rangle\,, \qquad & |\bar{0}\rangle &= \frac{1}{\sqrt{|\mathcal{C}\setminus\mathcal{F}|}}\sum_{\mathbf{x}\in\mathcal{C}\setminus\mathcal{F}}|\mathbf{x}\rangle\,. \end{aligned} \tag{10}$$

因此，在实现 GAS 式（6）中的预言机算子时，我们有

$$\hat{O}\,|s_c\rangle = \cos(a)|\bar{0}\rangle - \sin(a)\,|\bar{1}\rangle\,. \tag{11}$$

那么，根据 A_c 作用于每个基态的结果，即

$$\begin{aligned} \hat{A}_c|\bar{0}\rangle &= (2\cos(a)^2-1)|\bar{0}\rangle + 2\sin(a)\cos(a)\,|\bar{1}\rangle \\ &= \cos(2a)|\bar{0}\rangle + \sin(2a)\,|\bar{1}\rangle\,, \\ \hat{A}_c\,|\bar{1}\rangle &= (2\sin(a)^2-1)\,|\bar{1}\rangle + 2\sin(a)\cos(a)|\bar{0}\rangle \\ &= -\cos(2a)\,|\bar{1}\rangle + \sin(2a)|\bar{0}\rangle. \end{aligned} \tag{12}$$

因此，经过一次迭代后，我们有

$$\begin{aligned} \hat{A}_c\hat{O}\,|s_c\rangle &= [\cos(a)\cos(2a) - \sin(a)\sin(2a)]\,|\bar{0}\rangle + [\sin(a)\cos(2a) + \cos(a)\sin(2a)]\,|\bar{1}\rangle \\ &= \cos(3a)\,|\bar{0}\rangle + \sin(3a)\,|\bar{1}\rangle\,. \end{aligned} \tag{13}$$

根据归纳法，将 Grover 迭代算子 $\hat{A}_c\hat{O}$ 应用 r 次后， $|\bar{1}\rangle$ 的振幅变为 $\sin(a+2ar)$。因此，我们可以设置 $(2r+1)$ $a=\pi/2$，使 $|\bar{1}\rangle$ 态的振幅达到最大值 1。结合式 (10)，这意味着

$$r = \frac{\pi}{4\arcsin\left(\sqrt{\frac{|\mathcal{F}|}{|\mathcal{C}|}}\right)} \approx \frac{\pi}{4}\sqrt{\frac{|\mathcal{C}|}{|\mathcal{F}|}}, \tag{14}$$

并推导出结论。 □

在接下来的小节中，我们将继续介绍定理 1 框架下符合固定基数约束的扩散算子和预言机的设计与分析。

A. 扩散算子

如定理 1 所示，针对硬约束问题的 Grover 搜索的核心思想是构造一个扩散算子，将状态空间限制在可行解范围内。在本小节中，我们给出用于生成以下状态的量子电路 C 的显式构造：当 $\mathcal{C} = \{\mathbf{x} \in \{0,1\}^n : \ |\mathbf{x}| = k\}$ 是基数约束子集时

$$|s_c\rangle = \hat{C}\,|0^n\rangle = \frac{1}{\sqrt{|\mathcal{C}|}} \sum_{\mathbf{x}\in\mathcal{C}} |\mathbf{x}\rangle \ , \tag{15}$$

这将给出硬约束 Grover 扩散算子的构造。我们将利用用于制备 Dicke 态的量子电路来制备状态 $|s_c\rangle$。为区别于其他约束，我们将 $|h_k\rangle$ 专门记作表示固定基数约束的状态，其中 k 表示和。

Dicke 态制备是一种映射：从汉明重量为 $|\mathbf{x}| = k$ 的一个基态出发，映射到所有汉明重量为 $|\mathbf{x}| = k$ 的状态的均匀叠加态：

$$\hat{U}_k^n : \left|0^{n-k}\right\rangle \left|1^k\right\rangle \to |h_k\rangle := \binom{n}{k}^{-\frac{1}{2}} \sum_{\substack{\mathbf{x}\in\{0,1\}^n \\ |\mathbf{x}|=k}} |\mathbf{x}\rangle \ . \tag{16}$$

在量子计算机中，通过归纳构造可以确定性地生成该映射对应的酉算子 [36], ，其形式如下：

$$\hat{U}_k^n = \prod_{\ell=2}^{k} \left(\widehat{SCS}_{\ell,\ell-1} \otimes \mathrm{Id}^{\otimes n-\ell}\right) \cdot \prod_{\ell=k+1}^{n} \left(\mathrm{Id}^{\otimes \ell-k-1} \otimes \widehat{SCS}_{\ell,k} \otimes \mathrm{Id}^{\otimes n-\ell}\right) \tag{17}$$

其中 $\widehat{SCS}_{i,\ j}$ 称为分裂与循环移位酉算子，定义如下。对于所有 $\ell \in 1, \ldots, \ j$，其中 $j \le i -$ 1

$$\widehat{SCS}_{i,j} \left|0^{j+1}\right\rangle = \left|0^{j+1}\right\rangle \tag{18}$$

$$\widehat{SCS}_{i,j} \left|0^{j+1-\ell}\right\rangle \left|1^{\ell}\right\rangle = \sqrt{\frac{\ell}{i}} \left|0^{j+1-\ell}\right\rangle \left|1^{\ell}\right\rangle + \sqrt{\frac{i-\ell}{i}} \left|0^{j-\ell}\right\rangle \left|1^{\ell}\right\rangle |0\rangle \tag{19}$$

$$\widehat{SCS}_{i,j} \left|1^{j+1}\right\rangle = \left|1^{j+1}\right\rangle . \tag{20}$$

因此，固定基数约束的扩散算子定义为

$$\hat{A}_n^k = \hat{U}_k^n (2\left|0^{n-k}1^k\right\rangle\!\left\langle 0^{n-k}1^k\right| - \hat{I})\hat{U}_k^{n\dagger} = 2\,|h_k\rangle\,\langle h_k| - \hat{I}. \tag{21}$$

对于算符 $2\left|0^{n-k}1^k\right\rangle\!\left\langle 0^{n-k}1^k\right| - \hat{I}$，可按如下方式构造。首先注意， $2\,|1^n\rangle\langle 1^n| - \hat{I}$ 是多重受控的 Z 量子门 $\widehat{CZ}_n$，但相差一个负号。然后对其进行 $\hat{X}^{\otimes(n-k)}\hat{I}^{\otimes k}$ 共轭变换，即可得到 $2\left|0^{n-k}1^k\right\rangle\!\left\langle 0^{n-k}1^k\right| - \hat{I}$ 的量子电路。具体而言，我们有

$$\hat{A}_n^k = \hat{U}_k^n (\hat{X}^{\otimes(n-k)}\hat{I}^{\otimes k})\widehat{CZ}_n (\hat{X}^{\otimes(n-k)}\hat{I}^{\otimes k})\hat{U}_k^{n\dagger} \ . \tag{22}$$

$\widehat{SCS}_{l,\ k}$ 将涉及三个量子比特上的量子比特激发交换：第 l 个量子比特、第（$l-k$）个量子比特以及第（$l-k+1$）个量子比特。因此，量子电路 $SCS_{l,\ k}$ 可通过一个双量子比特门和 $k-1$ 个三量子比特门构建——详情见 A 节。根据式（17）中 $\hat{U}_k^n$ 的乘积结构，其电路深度为 $\mathcal{O}$（n），使用 $\mathcal{O}$（kn）个量子门。我们还注意到，对 $\hat{U}_k^n$ 内部的电路块进行并行化优化后，若采用全连接电路 [37]，电路深度可降至 $\mathcal{O}$（$k \log \frac{n}{k}$）。

B. 预言机

本文在 BQP-FC 量子求解器中使用量子字典（QD）预言机 [31]——设计详情及量子门使用量汇总见 B 节和表I。给定一个二次函数（如式（7）所示）以及 m 位精度，QD 预言机需要 mn 1 个控制旋转门，以及 mn（$n-1$）/2 个双控制旋转门，以将 f（$\mathbf{x}$） $-y$ 编码到制备态的相位中。此外，我们还需要在辅助量子比特上实现逆量子傅里叶变换，将相位编码转换为基编码；根据 [38] 的建议，这需要 $\mathcal{O}$（$m$log m）个托福利门。

对象	Gate	QD-GAS[19]	ADMM-GAS-hard
二次式 (1)	$\widehat{C^1R(\theta)}$	mn	mn
	$\widehat{C^2R(\theta)}$	$m\binom{n}{2}$	$m\binom{n}{2}$
四次公式 (2)	$\widehat{C^1R(\theta)}$	mn	mn
	$\widehat{C^2R(\theta)}$	$m\binom{n}{2}$	$m\binom{n}{2}$
	$\widehat{C^3R(\theta)}$	$m\binom{n}{3}$	-
	$\widehat{C^4R\ (\theta)}$	$m\binom{n}{4}$	-
K 次非稀疏多项式 $K>4$	$\widehat{C^1R\ (\theta)}$	mn	mn
	$\widehat{\ C^2R\ }$ (θ)	$m\binom{n}{2}$	$m\binom{n}{2}$
	$\widehat{C^kR(\theta)}$ $k=3,\ldots K$	$m\binom{n}{k}$	-

表I. 在额外使用 $\mathcal{O}$($m$log m) 个用于逆量子傅里叶变换的托福利门的情况下，GAS 应用于具有固定基数约束的二次、四次及高次二元规划时，每次 Grover 搜索中各预言机的门数量。QD-GAS 是基于量子字典预言机的 GAS 的简称；ADMM-GAS-hard 是用于硬约束规划的基于 ADMM 的混合 GAS 方法。

然而，对于更高次的多项式函数，即式 (2) 中的四次目标函数，式 (B3) 中的受控旋转 $\widehat{CU}_{\mathbf{x}}(2\pi/2^m)$ 将需要 $m\binom{n}{4}$ 4 控制旋转、 $m\binom{n}{3}$ 3 控制旋转、 $m\binom{n}{2}$ 2 控制旋转和 mn 1 控制旋转。事实上，多项式次数的增加会导致量子门数量和电路深度呈指数增长。因此，直接将 QD 预言机用于高次多项式规划效率低下。由此，减少高次多项式二元优化中量子门使用量的方法，对于将 GAS 扩展到更广泛的应用至关重要。在下一节中，我们将介绍一种基于 ADMM 的方法，将风险平价模型的优化分解为若干二次子优化问题，这些问题可由我们的 GAS 算法和经典求解器高效求解。一般而言，所提出的基于 ADMM 的方法也适用于其他 BPP-FC 实例，以减少量子资源和电路深度。

IV. 风险平价问题的交替方向乘子法

本节介绍用于求解公式（2）中风险平价模型的交替方向乘子法算法。近期研究 [39]表明，将量子计算算法集成到用于求解混合二元规划问题的交替方向乘子法（ADMM） [40] 方案中具有潜力。由于当前量子电路难以求解四次优化问题，我们转而提出以下目标函数，将四次优化转换为双二次优化。新的目标函数与交替方向乘子法方案无缝衔接，并且如 [41]所确立的，具有理论上的收敛性保证。

在介绍交替方向乘子法算法之前，我们对均值向量 $\boldsymbol{\mu}$ 和协方差矩阵 $\boldsymbol{\Sigma}$作出以下假设:

（**数据有界性**） 存在常数 $c_1, c_2 > 0$，使得 $\|\boldsymbol{\mu}\|_\infty \le c_1$ 且 $\Sigma_{ii} \le c_2$ 对所有 $i = 1, \ldots, n$均成立。

(非简并性) 存在一个常数 $c_3 > 0$，使得 $\gamma_{\min}(\mathbf{\Sigma}) \geq c_3$。

在金融场景中，这两个假设都较为温和且合理。数据有界性假设反映了这样一种预期：资产收益率和方差都是有限且有界的。特别是，该假设还意味着$\gamma_{\max}(\mathbf{\Sigma}) \leq c_2 n$，因为对于半正定 $\mathbf{\Sigma}$，$\gamma_{\max} \leq \mathrm{tr}(\mathbf{\Sigma})$。同时，采用 Ledoit-Wolf 收缩技术 [42, 43], 很容易满足非退化假设，该技术可确保样本协方差矩阵的估计具有良好的条件数。

在上述两个假设下，我们可以对式 (2) 中原优化问题的等价重构应用 ADMM：

$$\begin{aligned} \min_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}\in\mathbb{R}^n} \quad & \phi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}) := g(\mathbf{x}_1,\mathbf{x}_2) + \iota_{\mathcal{S}}(\mathbf{x}_1) + f(\mathbf{x}_2) + h(\mathbf{y}) \\ s.t. \quad & \mathbf{x}_1 - \mathbf{x}_2 - \mathbf{y} = \mathbf{0}, \end{aligned} \tag{23}$$

其中

$$\begin{aligned} g(\mathbf{x}_1,\mathbf{x}_2) &= \sum_{\substack{1\leq i,j\leq n \\ i\neq j}} \left(x_{1,i}(\mathbf{\Sigma}\mathbf{x}_2)_i - x_{1,j}(\mathbf{\Sigma}\mathbf{x}_2)_j\right)^2, \\ f(\mathbf{x}_2) &= \lambda\left(-\mathbf{x}_2^\intercal\boldsymbol{\mu} + \frac{1}{2}\mathbf{x}_2^\intercal\mathbf{\Sigma}\mathbf{x}_2\right), \\ h(\mathbf{y}) &= \frac{\zeta}{2}\|\mathbf{y}\|^2, \end{aligned}$$

且 $\iota_{\mathcal{S}}(\cdot)$ 是 $\mathcal{S} = \{\mathbf{x} \in \{0,1\}^n$ 的示性函数：$\mathbf{x}^\intercal\mathbf{1} = k\}$，使得

$$\iota_{\mathcal{S}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{S} \\ +\infty, & \text{otherwise.} \end{cases}$$

这里，$\lambda > 0$ 是权衡超参数，$\zeta > 0$ 是一个足够大的常数 .

混合 ADMM 算法使用由下式给出的增广拉格朗日函数求解优化问题式 (23)

$$\mathcal{L}_\beta(\mathbf{x}_1,\mathbf{x}_2,\mathbf{y},\mathbf{w}) = \phi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}) + \mathbf{w}^\intercal(\mathbf{x}_1 - \mathbf{x}_2 - \mathbf{y}) + \frac{\beta}{2}\|\mathbf{x}_1 - \mathbf{x}_2 - \mathbf{y}\|^2. \tag{24}$$

为方便记号，我们将第 $(t+1)$ 次迭代中的更新记为 $(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^+,\mathbf{w}^+) = (\mathbf{x}_1^{t+1},\mathbf{x}_2^{t+1},\mathbf{y}^{t+1},\mathbf{w}^{t+1})$。所提出的方法总结于算法 2 中。

两个子问题 Eq. (26) 和 Eq. (27) 均可显式求解。具体而言，对于 Eq. (26)，注意到

$$\begin{aligned} \sum_{\substack{1\leq i,j\leq n \\ i\neq j}} \left(x_{1,i}^+(\mathbf{\Sigma}\mathbf{x}_2)_i - x_{1,j}^+(\mathbf{\Sigma}\mathbf{x}_2)_j\right)^2 &= \sum_{i=1}^n 2(n-1){x_{1,i}^+}^2(\mathbf{\Sigma}\mathbf{x}_2)_i^2 - 2\sum_{\substack{1\leq i,j\leq n \\ i\neq j}} x_{1,i}^+x_{1,j}^+(\mathbf{\Sigma}\mathbf{x}_2)_i(\mathbf{\Sigma}\mathbf{x}_2)_j \\ &= \mathbf{x}_2^\intercal\mathbf{Q}\mathbf{x}_2, \end{aligned} \tag{28}$$

其中 $\mathbf{Q} = \mathbf{\Sigma}\mathbf{H}\mathbf{\Sigma}$ 是一个半正定矩阵，而 $\mathbf{H}$ 是一个 $n \times n$ 对称矩阵，满足

$$H_{ij} = \begin{cases} 2(n-1){x_{1,i}^+}^2 & i = j \\ -2x_{1,i}^+x_{1,j}^+ & i \neq j. \end{cases}$$

因此,

$$\begin{aligned} \mathbf{x}_2^+ &= \arg\min_{\mathbf{x}_2\in\mathbb{R}^n} \mathbf{x}_2^\intercal\mathbf{Q}\mathbf{x}_2 + \lambda\left(-\mathbf{x}_2^\intercal\boldsymbol{\mu} + \frac{1}{2}\mathbf{x}_2^\intercal\mathbf{\Sigma}\mathbf{x}_2\right) - \mathbf{w}^{t\intercal}\mathbf{x}_2 + \frac{\beta}{2}\|\mathbf{x}_1^+ - \mathbf{x}_2 - \mathbf{y}^t\|^2 \\ &= (2\mathbf{Q} + \lambda\mathbf{\Sigma} + \beta\mathbf{I})^{-1}(\lambda\mu + \mathbf{w}^t + \beta(\mathbf{x}_1^+ - \mathbf{y}^t)). \end{aligned}$$

对于公式（27），

$$\mathbf{y}^+ = \frac{1}{\zeta+\beta}\left(\mathbf{w}^t + \beta(\mathbf{x}_1^+ - \mathbf{x}_2^+)\right). \tag{29}$$

算法 2： 用于求解风险平价问题（式（23））的混合 ADMM 算法

输入： $t=0$，$\mathbf{x}_1^0 \in \mathcal{S}$，$\mathbf{x}_2^0 = \mathbf{x}_1^0$, $\mathbf{y}^0 = \mathbf{w}^0 = \mathbf{0}$，收益向量 $\boldsymbol{\mu}$，协方差矩阵 $\boldsymbol{\Sigma}$，权衡超参数 λ，容差超参数 $\epsilon > 0$，正则化参数 $\zeta > 0$，增广拉格朗日罚参数 $\beta > \sqrt{2}\zeta$（根据推论 4 的要求），最大迭代次数 $T_{\max}$，并初始化误差项 $\Delta = +\infty$。

结果： $\mathbf{x}_1^t$

当 $t < T_{\max}$ **且** $\Delta \geq \dfrac{\epsilon}{\beta + 1}$ **时**

1. 使用高效量子优化算法求解二元优化子问题：

$$\begin{aligned}\mathbf{x}_1^+ = \arg\min_{\mathbf{x}_1 \in \mathcal{S}} & \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} \left(x_{1,i}(\boldsymbol{\Sigma}\mathbf{x}_2^t)_i - x_{1,j}(\boldsymbol{\Sigma}\mathbf{x}_2^t)_j\right)^2 \\ & + \mathbf{w}^{t\intercal}\mathbf{x}_1 + \frac{\beta}{2}\|\mathbf{x}_1 - \mathbf{x}_2^t - \mathbf{y}^t\|^2\end{aligned} \tag{25}$$

2. 求解凸优化子问题：

$$\begin{aligned}\mathbf{x}_2^+ = \arg\min_{\mathbf{x}_2 \in \mathbb{R}^n} & \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} \left(x_{1,i}^+(\boldsymbol{\Sigma}\mathbf{x}_2)_i - x_{1,j}^+(\boldsymbol{\Sigma}\mathbf{x}_2)_j\right)^2 \\ & + \lambda\left(-\mathbf{x}_2^\intercal\boldsymbol{\mu} + \frac{1}{2}\mathbf{x}_2^\intercal\boldsymbol{\Sigma}\mathbf{x}_2\right) - \mathbf{w}^{t\intercal}\mathbf{x}_2 + \frac{\beta}{2}\|\mathbf{x}_1^+ - \mathbf{x}_2 - \mathbf{y}^t\|^2\end{aligned} \tag{26}$$

3. 求解凸优化子问题：

$$\mathbf{y}^+ = \arg\min_{\mathbf{y} \in \mathbb{R}^n} \ \frac{\zeta}{2}\|\mathbf{y}\|^2 - \mathbf{w}^{t\intercal}\mathbf{y} + \frac{\beta}{2}\|\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}\|^2 \tag{27}$$

4. 更新对偶变量：$\mathbf{w}^+ = \mathbf{w}^t + \beta(\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}^+)$。5. 更新误差项：$\Delta = \|\mathbf{y}^+ - \mathbf{y}^t\|$6. 设置 $t = t+1$

end

返回 $\mathbf{x}_1^t$

[41] 中的工作通过一系列正则性假设，利用交替方向乘子法分析了非凸非光滑优化的一个相当一般的情形。在我们具体的风险平价背景下，我们通过确立以下四个关键性质，直接验证了所提出的算法 2 的收敛性（事实上，我们的优化问题也满足这些正则性假设）：

P1（**序列有界性**）序列 $\{\mathbf{x}_1^t,\ \mathbf{x}_2^t,\ \mathbf{y}^t,\ \mathbf{w}^t\}$ 有界，并且 $\mathcal{L}_\beta(\mathbf{x}_1^t,\ \mathbf{x}_2^t,\ \mathbf{y}^t,\ \mathbf{w}^t)$ 有下界。

P2（**充分下降**）存在一个常数 $C_1(\beta) > 0$，使得对所有 t，我们都有

$$\mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) - \mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+) \geq C_1(\beta)\left(\|\mathbf{x}_2^+ - \mathbf{x}_2^t\|^2 + \|\mathbf{y}^+ - \mathbf{y}^t\|^2\right).$$

P3（**充分梯度界**）存在 $C_2(\beta) > 0$ 和 $\mathbf{d}^+ \in \partial\mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+)$，使得

$$\|\mathbf{d}^+\| \leq C_2(\beta)\left(\|\mathbf{x}_2^+ - \mathbf{x}_2^t\| + \|\mathbf{y}^+ - \mathbf{y}^t\|\right).$$

P4（极限连续性）若 $(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{y}^*, \mathbf{w}^*)$ 是子序列 $(\mathbf{x}_1^{t_s}, \mathbf{x}_2^{t_s}, \mathbf{y}^{t_s}, \mathbf{w}^{t_s})$ 对于 $s \in \mathbb{N}$ 的极限点，则 $\mathcal{L}_\beta(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{y}^*, \mathbf{w}^*) = \lim_{s\to\infty}\mathcal{L}_\beta(\mathbf{x}_1^{t_s}, \mathbf{x}_2^{t_s}, \mathbf{y}^{t_s}, \mathbf{w}^{t_s})$。

我们按 P2 $\to$ P1 $\to$ P3 $\to$ P4 的顺序验证这四个性质。为证明充分下降性质 P2，我们先给出以下两个命题：

命题 2 (参见 [41] 中的引理 4**)。** *算法 2 中的迭代序列满足*

$$\mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) - \mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^t, \mathbf{w}^t) \geq \frac{\beta}{2}\|\mathbf{x}_2^+ - \mathbf{x}_2^t\|^2. \tag{30}$$

证明。 算法 2 的更新意味着

$$\begin{aligned}\mathcal{L}_\beta(\mathbf{x}_1^t,\mathbf{x}_2^t,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)&\geq\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^t,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)\\&=f(\mathbf{x}_2^t)-f(\mathbf{x}_2^+)+g(\mathbf{x}_1^+,\mathbf{x}_2^t)-g(\mathbf{x}_1^+,\mathbf{x}_2^+)-\mathbf{w}^{t\intercal}(\mathbf{x}_2^t-\mathbf{x}_2^+)+\\&\quad+\frac{\beta}{2}\|\mathbf{x}_1^+-\mathbf{x}_2^t-\mathbf{y}^t\|^2-\frac{\beta}{2}\|\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^t\|^2\end{aligned}\tag{31}$$

而子问题$\mathbf{x}_2$, $\nabla_{\mathbf{x}_2}\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2,\mathbf{y}^t,\mathbf{w}^t)$ 的最优性条件 $=\mathbf{0}$ 意味着

$$\nabla f(\mathbf{x}_2^+)+\nabla_2 g(\mathbf{x}_1^+,\mathbf{x}_2^+)=\mathbf{w}^t+\beta(\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^t)\tag{32}$$

结合公式（31）和公式（32），可得

$$\begin{aligned}\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^t,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)&=f(\mathbf{x}_2^t)-f(\mathbf{x}_2^+)-[\nabla f(\mathbf{x}_2^+)]^\intercal(\mathbf{x}_2^t-\mathbf{x}_2^+)+g(\mathbf{x}_1^+,\mathbf{x}_2^t)-g(\mathbf{x}_1^+,\mathbf{x}_2^+)\\&\quad-[\nabla_2\ g(\mathbf{x}_1^+,\mathbf{x}_2^+)]^\intercal(\mathbf{x}_2^t-\mathbf{x}_2^+)+\frac{\beta}{2}\|\mathbf{x}_2^t-\mathbf{x}_2^+\|^2\end{aligned}\tag{33}$$

具体而言，由于在我们的模型中， $f(\mathbf{x}_2)$ 是凸且可微的，因此可得

$$f(\mathbf{x}_2^t)-f(\mathbf{x}_2^+)-[\nabla f(\mathbf{x}_2^+)]^\intercal(\mathbf{x}_2^t-\mathbf{x}_2^+)\geq 0.\tag{34}$$

此外，令 $\hat{}g(\mathbf{x}_2)=g(\mathbf{x}_1^+,\mathbf{x}_2)$，则 $\nabla_2\ g(\mathbf{x}_1^+,\mathbf{x}_2)=\nabla\hat{}g(\mathbf{x}_2)$。根据公式（28），$\hat{}g(\mathbf{x}_2)=\mathbf{x}_2^\intercal\mathbf{Q}\mathbf{x}_2,$，其中 $\mathbf{Q}$ 是一个半正定矩阵。这意味着 $\hat{}g(\mathbf{x}_2)$ 也是一个凸可微函数，并且

$$\hat{g}(\mathbf{x}_2^t)-\hat{g}(\mathbf{x}_2^+)-[\nabla g(\mathbf{x}_2^+)]^\intercal(\mathbf{x}_2^t-\mathbf{x}_2^+)\geq 0\tag{35}$$

结合公式（33）、公式（34）和公式（35），显然可得

$$\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^t,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)\geq\frac{\beta}{2}\|\mathbf{x}_2^+-\mathbf{x}_2^t\|^2.$$

至此，命题得证。 □

命题 3 (参见 [41] 中的引理 5)。 算法 *2* 中的迭代序列满足

$$\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^+,\mathbf{w}^+)\geq\left(\frac{\beta}{2}-\frac{\zeta^2}{\beta}\right)\|\mathbf{y}^+-\mathbf{y}^t\|^2\tag{36}$$

证明。 在我们的情形下,

$$\begin{aligned}\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^+,\mathbf{w}^+)&=h(\mathbf{y}^t)-h(\mathbf{y}^+)+(\mathbf{w}^t)^\intercal(\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^t)-(\mathbf{w}^+)^\intercal(\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^+)\\&\quad+\frac{\beta}{2}\|\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^t\|^2-\frac{\beta}{2}\|\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^+\|^2\end{aligned}\tag{37}$$

由于 $\mathbf{w}^+-\mathbf{w}^t=\beta(\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^+)$ 且 $\zeta\mathbf{y}^+=\mathbf{w}^+$ （由公式（29）），公式（37）可改写为

$$\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^+,\mathbf{w}^+)=h(\mathbf{y}^t)-h(\mathbf{y}^+)-(\mathbf{w}^+)^\intercal(\mathbf{y}^t-\mathbf{y}^+)+\left(\frac{\beta}{2}-\frac{\zeta^2}{\beta}\right)\|\mathbf{y}^t-\mathbf{y}^+\|^2.\tag{38}$$

$\mathbf{y}$ 的子问题的最优性条件， $\nabla_{\mathbf{y}}\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2,\mathbf{y}^t,\mathbf{w}^t)=\mathbf{0}$ 意味着

$$\nabla h(\mathbf{y}^+)-\mathbf{w}^t-\beta(\mathbf{x}_1^+-\mathbf{x}_2^+-\mathbf{y}^+)=\nabla h(\mathbf{y}^+)-\mathbf{w}^+=0.\tag{39}$$

将公式（39）代入公式（38），得到

$$\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^t,\mathbf{w}^t)-\mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^+,\mathbf{w}^+)=h(\mathbf{y}^t)-h(\mathbf{y}^+)-[\nabla h(\mathbf{y}^+)]^\intercal(\mathbf{y}^t-\mathbf{y}^+)+\left(\frac{\beta}{2}-\frac{\zeta^2}{\beta}\right)\|\mathbf{y}^t-\mathbf{y}^+\|^2.$$

然而，由于 $h(\mathbf{y})=\frac{\zeta}{2}\|\mathbf{y}\|^2$ 是一个凸函数，我们必有 $h(\mathbf{y}^t)-h(\mathbf{y}^+)-[\nabla h(\mathbf{y}^+)]^\intercal(\mathbf{y}^t-\mathbf{y}^+)\geq 0$。因此，公式（36）成立。 □

结合命题 2 和 由 3，我们立即得到关于充分下降性质 P2 的以下推论 :

推论 4. 设 $\beta > \sqrt{2}\zeta$，充分下降性质 *P2* 在 $C_1(\beta) = \frac{\beta}{2} - \frac{\zeta^2}{\beta}$ 下成立。

现在我们来讨论序列有界性 P1。为了验证这一性质，我们首先证明目标函数的强制性（或按照 [44] 的术语，水平有界性），具体如下命题所述:

命题 *5*（强制性）。定义可行集 $\mathcal{F} := \{(x_1, x_2, y) \in \mathbb{R}^{3n} : x_1 - x_2 - y = 0\}$。目标函数 $\phi(x_1, x_2, y)$ 在该集合上具有强制性，即 $\phi(x_1, x_2, y) \to \infty$ *if* $(x_1, x_2, y) \in \mathcal{F}$ *and* $\|(x_1, x_2, y)\| \to \infty$。

证明。如果 $\|\mathbf{x}_1\| \to +\infty$，那么根据 $\iota_{\mathcal{S}}(\mathbf{x}_1)$ 的定义，我们立即得到 $\phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \to +\infty$。因此，我们只需考虑 $\|(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})\| \to +\infty$ 且 $\|\mathbf{x}_1\|$ 有界的情形。根据约束，必有 $\|\mathbf{x}_2\| \to +\infty$ 且 $\|\mathbf{y}\| \to +\infty$。

$g(\mathbf{x}_1, \mathbf{x}_2)$ 和 $\iota_{\mathcal{S}}(\mathbf{x}_1)$ 均非负。由于协方差矩阵 $\mathbf{\Sigma}$ 正定，我们有 e

$$f(\mathbf{x}_2) \geq \lambda \left(\frac{\gamma_{\min}(\mathbf{\Sigma})}{2} \|\mathbf{x}_2\|^2 - \|\boldsymbol{\mu}\| \|\mathbf{x}_2\| \right) \to +\infty \quad \text{as } \|\mathbf{x}_2\| \to +\infty,$$

其中 $\gamma_{\min}(\cdot)$ 表示相应矩阵的最小特征值。显然，当 $\|\mathbf{y}\| \to +\infty$ 时， $\frac{\zeta}{2}\|\mathbf{y}\|^2 \to +\infty$。总之，强制性条件成立。

□

现在我们证明，我们的优化问题满足性质 P1:

命题 6 (序列有界性，参见 [41] 中的引理 6)。 当 $\beta > \zeta$ 时，由算法 *2* 生成的序列 $\{\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t\}$ 满足性质 *P1*。

证明。我们首先证明 $\inf_{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \in \mathcal{F}} \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) > -\infty$。根据 [44], 中的定理 1.9，如果扩展实值函数 $f: \mathbb{R}^m \to \mathbb{R}$ 是下半连续、水平有界且适当的，那么 $\inf f$ 是有限的。在本例中，我们考虑扩展目标函数 $\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) + \iota_{\mathcal{F}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$。由命题 5 可知其水平有界性成立，并且显然 $\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ 是适当函数（即，它不恒等于 $+\infty$，且从不取值 $-\infty$）。接下来需要证明 $\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ 是下半连续的，这归结为证明 $\iota_{\mathcal{S}}(\mathbf{x}_1)$ 和 $\iota_{\mathcal{F}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ 的下半连续性，因为 $g(\mathbf{x}_1, \mathbf{x}_2)$ $f(\mathbf{x}_2)$ 和 $h(\mathbf{y})$ 都是连续函数。回顾一下，若函数对任意开区间 $(u, \infty]$ 的原像都是开集，则该函数是下半连续的。对于任意 $u \geq 0$，我们知道 $\iota_{\mathcal{S}}^{-1}((u, \infty]) = \mathbb{R}^n \setminus \mathcal{S}$，且 $\iota_{\mathcal{F}}^{-1}((u, \infty]) = \mathbb{R}^{3n} \setminus \mathcal{F}$，由于 $\mathcal{S}$ 和 $\mathcal{F}$ 都是闭集，因此这两个集合都是开集。而对于任意 $u < 0$， $\iota_{\mathcal{S}}^{-1}((u, \infty]) = \mathbb{R}^n$，且 $\iota_{\mathcal{F}}^{-1}((u, \infty]) = \mathbb{R}^{3n}$，这两个集合也都是开集。因此， $\iota_{\mathcal{S}}(\mathbf{x}_1)$ 和 $\iota_{\mathcal{F}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ 都是下半连续的。由此可知， $\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ 是下半连续的，因此 $\inf_{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \in \mathcal{F}} \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) > -\infty$。

现在我们证明 $\mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t)$ 的下有界性。

$$\begin{aligned} \mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) &= \phi(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t) + \mathbf{w}^{t\intercal}(\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t) + \frac{\beta}{2}\|\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t\|^2 \\ &\geq \phi(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{x}_1^t - \mathbf{x}_2^t) + h(\mathbf{y}^t) - h(\mathbf{x}_1^t - \mathbf{x}_2^t) + [\nabla h(\mathbf{y}^t)]^\intercal(\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t) + \frac{\beta}{2}\|\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t\|^2. \end{aligned} \tag{40}$$

显然,

$$\phi(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{x}_1^t - \mathbf{x}_2^t) \geq \inf_{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \in \mathcal{F}} \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) > -\infty. \tag{41}$$

另一方面，由于 $\nabla h(y) = \zeta \mathbf{y}$ 的利普希茨常数为 ζ，因此我们有（例如，参见 [45] 中的定理 2.1.5）

$$h(\mathbf{y}^t) - h(\mathbf{x}_1^t - \mathbf{x}_2^t) + [\nabla h(\mathbf{y}^t)]^\intercal(\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t) \geq -\frac{\zeta}{2}\|\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t\|^2. \tag{42}$$

将式（41）和式（42）代入式（40），得到

$$\begin{aligned} \mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) &\geq \inf_{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \in \mathcal{F}} \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) + \frac{\beta - \zeta}{2}\|\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t\|^2 \\ &\geq \inf_{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \in \mathcal{F}} \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) > -\infty, \end{aligned} \tag{43}$$

这证实了 $\mathcal{L}_\beta(\mathbf{x}_1^t,\ \mathbf{x}_2^t,\ \mathbf{y}^t,\ \mathbf{w}^t)$ 的下有界性。

最后，我们验证 $\{\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t\}$ 的有界性。根据充分下降性质 P2，$\mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t)$ 由 $\mathcal{L}_\beta(\mathbf{x}_1^0, \mathbf{x}_2^0, \mathbf{y}^0, \mathbf{w}^0)$ 上界，因此 $\phi(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{x}_1^t - \mathbf{x}_2^t)$ 和 $\|\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t\|$ 也有上界。由命题 5 可知，$\{\mathbf{x}_1^t\}$ 和 $\{\mathbf{x}_2^t\}$ 均有界。结合 $\|\mathbf{x}_1^t - \mathbf{x}_2^t - \mathbf{y}^t\|$ 的有界性，可推出 $\{\mathbf{y}^t\}$ 有界。最后，由于 $\mathbf{w}^t = \nabla h(\mathbf{y}^t) = \zeta \mathbf{y}^t$，因此 $\{\mathbf{w}^t\}$ 也有界。 □

在下文的命题中，我们针对当前情境提出以下命题，以处理次梯度界性质 P3；该命题也有助于确定停止准则和进行资源估算。

命题 7 (参见 [41] 中的引理10)。 设 $\beta > \zeta > 0$。存在某个 $\mathbf{d}^+ \in \partial \mathcal{L}_\beta(\mathbf{x}_1^+,\mathbf{x}_2^+,\mathbf{y}^+,\mathbf{w}^+)$ 使得

$$\|\mathbf{d}^+\| \leq \left(\beta + \frac{\zeta}{\beta}\right) \|\mathbf{y}^+ - \mathbf{y}^t\|,$$

证明。由于 $\mathcal{S}$ 是一个有限离散集，示性函数 $\iota_{\mathcal{S}}$ 的次微分显然为 $\mathbb{R}^n$，这意味着$\partial_{\mathbf{x}_1} \mathcal{L}_\beta = \mathbb{R}^n$。因此，存在$\mathbf{d}^+ \in \partial \mathcal{L}_\beta(\mathbf{x}_1^+,$ $\mathbf{x}_2^+,\ \mathbf{y}^+,\ \mathbf{w}^+)$，使得

$$\|\mathbf{d}^+\| \leq \left\| \frac{\partial \mathcal{L}_\beta}{\partial \mathbf{x}_2}(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+) \right\| + \left\|\nabla_{\mathbf{y}} \mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+)\right\| + \left\|\nabla_{\mathbf{w}} \mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+)\right\| \tag{44}$$

现在我们分析右侧的三项。$\mathbf{x}_2$ 子问题的最优性条件意味着

$$2\mathbf{Q}\mathbf{x}_2^+ - \lambda \boldsymbol{\mu} + \lambda \boldsymbol{\Sigma} \mathbf{x}_2^+ - \mathbf{w}^t - \beta(\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}^t) = \mathbf{0}.$$

因此,

$$\begin{aligned} \left\| \frac{\partial \mathcal{L}_\beta}{\partial \mathbf{x}_2}(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+) \right\| &= \|2\mathbf{Q}\mathbf{x}_2^+ - \lambda \boldsymbol{\mu} + \lambda \boldsymbol{\Sigma} \mathbf{x}_2^+ - \mathbf{w}^+ - \beta(\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}^+)\| \\ &= \|\mathbf{w}^t - \mathbf{w}^+ + \beta(\mathbf{y}^+ - \mathbf{y}^t)\| \\ &\leq (\beta - \zeta)\|\mathbf{y}^+ - \mathbf{y}^t\|. \end{aligned} \tag{45}$$

另一方面，根据$\mathbf{y}$ 和$\mathbf{w}$子问题的最优性条件,

$$\begin{aligned} \left\|\nabla_{\mathbf{y}} \mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+)\right\| &= \|\zeta \mathbf{y}^+ - \mathbf{w}^+ - \beta(\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}^+)\| \\ &= \zeta \|\mathbf{y}^+ - \mathbf{y}^t\|, \end{aligned} \tag{46}$$

and

$$\begin{aligned} \left\|\nabla_{\mathbf{w}} \mathcal{L}_\beta(\mathbf{x}_1^+, \mathbf{x}_2^+, \mathbf{y}^+, \mathbf{w}^+)\right\| &= \|\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}^+\| \\ &= \frac{\zeta}{\beta} \|\mathbf{y}^+ - \mathbf{y}^t\|. \end{aligned} \tag{47}$$

结合式（44）、式（45）、式（46）和式（47），我们立即得知存在$\mathbf{d}^+ \in \partial \mathcal{L}_\beta(\mathbf{x}_1^+,\ \mathbf{x}_2^+,\ \mathbf{y}^+,\ \mathbf{w}^+)$，使得

$$\|\mathbf{d}^+\| \leq \left(\beta + \frac{\zeta}{\beta}\right) \|\mathbf{y}^+ - \mathbf{y}^t\|.$$

□

根据命题 7，性质P3在 $C_2(\beta) = \beta + \frac{\zeta}{\beta}$下显然成立。最后，关于极限连续性性质P4，我们在以下命题中正式证明:

命题8 （极限连续性，参见定理 1 的证明 [41]）。 由算法 *2* 生成的序列 $\{\mathbf{x}_1^t,\mathbf{x}_2^t,\mathbf{y}^t,\mathbf{w}^t\}$ 满足极限连续性性质 *P4*。

证明。我们的增广拉格朗日函数 $\mathcal{L}_\beta(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \mathbf{w})$ 按公式（24）定义，是连续函数与示性函数 $\iota_{\mathcal{S}}(\mathbf{x}_1)$之和。由于连续函数显然满足极限连续性性质，我们只需考虑 $\iota_{\mathcal{S}}(\mathbf{x}_1)$。然而，由算法 2 生成的迭代序列的任意子序列 $(\mathbf{x}_1^{t_s}, \mathbf{x}_2^{t_s}, \mathbf{y}^{t_s}, \mathbf{w}^{t_s})$ 都保证 $\mathbf{x}_1^{t_s} \in \mathcal{S}$，这意味着对于任意 $s \in \mathbb{N}$，$\iota_{\mathcal{S}}(\mathbf{x}_1^{t_s}) = 0$。另一方面，$\mathcal{S}$是有限离散集，因此是闭集。这意味着 $\mathbf{x}_1^* = \lim_{s\to\infty} \mathbf{x}_1^{t_s} \in \mathcal{S}$，因此 $\iota_{\mathcal{S}}(\mathbf{x}_1^*) = 0$。结合连续项的极限，我们得到

$$\mathcal{L}_\beta(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{y}^*, \mathbf{w}^*) = \lim_{s\to\infty} \mathcal{L}_\beta(\mathbf{x}_1^{t_s}, \mathbf{x}_2^{t_s}, \mathbf{y}^{t_s}, \mathbf{w}^{t_s}).$$

命题至此得证。 □

除了性质P1-P4之外，关于我们的增广拉格朗日函数的（KŁ）性质的下述命题还保证了算法 2生成的序列收敛到唯一极限点：

命题9. 增广拉格朗日函数式(24) 是一个*Kurdyka-Łojasiewicz*（*KŁ*）函数（例如，参见 *[46]*中的定义*2.2*）。

证明。 要证明 $\mathcal{L}_\beta$ 是一个KŁ函数，只需证明它是半代数的 [39, 47, 48]。由于多项式函数是半代数的，半代数集的示性函数是半代数的，且半代数函数的有限和也是半代数的 [48], ，因此我们只需证明集合 $\mathcal{S}$ 是半代数的，也就是说，它可以写成以下形式的集合的有限并

$$\{\mathbf{x} \in \mathbb{R}^n : p_i(\mathbf{x}) = 0,\ q_i(\mathbf{x}) < 0, i = 1, 2, \ldots, p\},$$

其中 p_i，q_i 是实多项式函数。这是成立的，因为

$$\mathcal{S} = \bigcup_{\mathbf{x}_0 \in \mathcal{S}} \left\{ \mathbf{x} : p_1(\mathbf{x}) = \sum_{i=1}^{n} (x_i - x_{0,i})^2 = 0,\ q_1(\mathbf{x}) = p_1(\mathbf{x}) - 1 < 0 \right\}$$

且这类对象只有有限 个 数个这样的 $\mathbf{x}_0$。总之， $\mathcal{L}_\beta$ 是半代数的，因此是一个 KŁ 函数。 □

综合推论 4、命题 6 – 9，以及 [41], 中的命题 2，我们得到以下定理 ：

定理10. 从任意初始化开始，算法 *2* 生成的序列全局收敛到唯一极限点 $(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{y}^*, \mathbf{w}^*)$，该点是 $\mathcal{L}_\beta$的一个驻点。

在进行资源估算之前，我们将详细阐述 ADMM算法的停止准则以及 ζ 的选择。对于前者，关于 $\|\mathbf{y}^+ - \mathbf{y}^t\|$ 的条件用于监测对偶可行性，而这意味着原始可行性。事实上，对于原始残差，我们有

$$\begin{aligned} \|\mathbf{x}_1^+ - \mathbf{x}_2^+ - \mathbf{y}^+\| &= \frac{1}{\beta}\|\mathbf{w}^+ - \mathbf{w}^t\| \\ &= \frac{\zeta}{\beta}\|\mathbf{y}^+ - \mathbf{y}^t\|. \end{aligned}$$

到目前为止， ζ 的选择尚未明确。然而，谨慎选择这一参数（因而也包括 β）对于风险平价模型在实践中的成功至关重要。如果 ζ 过小，算法可能会在一致性误差 $\|\mathbf{x}_1^T - \mathbf{x}_2^T\|$ 仍过大、无法构成原始优化问题的有效解时终止。鉴于此，我们在以下命题中分析确保算法终止时解有效所需的 ζ 的缩放方式：

命题 11. 存在某个常数 $C_3(\lambda, \epsilon, c_1, c_2) > 0$，使得若

$$\zeta > C_3(\lambda, \epsilon, c_1, c_2)\frac{n^3\sqrt{k}}{\delta},$$

那么一致性 算法 *2* 成功终止时的误差满足 $\|\mathbf{x}_1^T - \mathbf{x}_2^T\| < \epsilon + \delta$。

证明。终止时，我们知道

$$\begin{aligned}\|\mathbf{x}_1^T-\mathbf{x}_2^T\| &\le \|\mathbf{x}_1^T-\mathbf{x}_2^T-\mathbf{y}^T\|+\|\mathbf{y}^T\| \\ &= \frac{\zeta}{\beta}\|\mathbf{y}^T-\mathbf{y}^{T-1}\|+\frac{1}{\zeta}\|\nabla h(\mathbf{y}^T)\| \\ &\le \epsilon+\frac{1}{\zeta}\|\mathbf{w}^T\|\end{aligned} \tag{48}$$

另一方面，根据算法 2 在第 T 次迭代时$\mathbf{x}_2$ 子问题的最优性条件，

$$2\mathbf{Q}\mathbf{x}_2^T-\lambda\boldsymbol{\mu}+\lambda\boldsymbol{\Sigma}\mathbf{x}_2^T-\mathbf{w}^{T-1}-\beta(\mathbf{x}_1^T-\mathbf{x}_2^T-\mathbf{y}^{T-1})=\mathbf{0},$$

这意味着

$$\begin{aligned}\|\mathbf{w}^T\| &\le \|2\mathbf{Q}\mathbf{x}_2^T\|+\lambda\|\boldsymbol{\mu}\|+\lambda\|\boldsymbol{\Sigma}\mathbf{x}_2^T\|+\beta\|\mathbf{y}^T-\mathbf{y}^{T-1}\| \\ &\le 2\gamma_{\max}(\mathbf{Q})\|\mathbf{x}_2^T\|+\lambda\sqrt{n}\|\boldsymbol{\mu}\|_\infty+\lambda\gamma_{\max}(\boldsymbol{\Sigma})\|\mathbf{x}_2^T\|+\epsilon\end{aligned} \tag{49}$$

根据数据有界性假设可得 $\|\boldsymbol{\mu}\|_\infty<c_1$, $\gamma_{\max}(\boldsymbol{\Sigma})<c_2 n$。对于 $\gamma_{\max}(\mathbf{Q})$，我们知道

$$\begin{aligned}\gamma_{\max}(\mathbf{Q}) &= \gamma_{\max}(\boldsymbol{\Sigma}^\intercal\mathbf{H}\boldsymbol{\Sigma}) \\ &= \max_{\|\mathbf{x}\|=1}\mathbf{x}\boldsymbol{\Sigma}^\intercal\mathbf{H}\boldsymbol{\Sigma}\mathbf{x} \\ &\le \gamma_{\max}(\mathbf{H})\cdot\max_{\|\mathbf{x}\|=1}\mathbf{x}^\intercal\boldsymbol{\Sigma}^2\mathbf{x} \\ &= \gamma_{\max}(\mathbf{H})\gamma_{\max}^2(\boldsymbol{\Sigma}).\end{aligned}$$

设 $\mathbf{D}$ 为一个对角矩阵，其每个对角项为 $D_{ii}=2n{x_1^+}_{,i}{}^2, i=1,2,\ldots,n$，则 $\mathbf{H}=\mathbf{D}-2\mathbf{x}_1^+\mathbf{x}_1^{+\intercal}$。由于对每个 $i=1,2,\ldots,n$, $0\le {x_1^+}_{,i}\le 1$，我们有

$$\begin{aligned}\gamma_{\max}(\mathbf{H}) &\le \gamma_{\max}(\mathbf{D}) \\ &\le \max_{\|\mathbf{x}\|=1}\mathbf{x}^\intercal\mathbf{D}\mathbf{x} \\ &= \max_{\|\mathbf{x}\|=1}2n\sum_{i=1}^n {x_{1,i}^+}^2x_i^2 \\ &\le 2n\sum_{i=1}^n x_i^2=2n.\end{aligned}$$

因此，我们有 $\gamma_{\max}(\mathbf{Q})\le 2n\gamma_{\max}^2(\boldsymbol{\Sigma})=2c_2^2n^3$。此外，我们知道 $\|\mathbf{x}_2^T\|\le\|\mathbf{x}_1^T-\mathbf{x}_2^T\|+\|\mathbf{x}_1^T\|=\|\mathbf{x}_1^T-\mathbf{x}_2^T\|+\sqrt{k}$。因此，式 (49) 导出

$$\|\mathbf{w}^T\|\le 2c_2^2n^3(\sqrt{k}+\|\mathbf{x}_1^T-\mathbf{x}_2^T\|)+c_1\lambda\sqrt{n}+c_2\lambda n(\sqrt{k}+\|\mathbf{x}_1^T-\mathbf{x}_2^T\|)+\epsilon \tag{50}$$

结合式 (48) 和式 (50)，可得

$$(\zeta-2c_2^2n^3-c_2\lambda n)\|\mathbf{x}_1^T-\mathbf{x}_2^T\|\le\epsilon\zeta+2c_2^2n^3\sqrt{k}+c_1\lambda\sqrt{n}+c_2\lambda n\sqrt{k}+\epsilon. \tag{51}$$

Let 考虑 $\zeta>2c_2^2n^3+c_2\lambda n$ 的情况。若 $\|\mathbf{x}_1^T-\mathbf{x}_2^T\|<\epsilon+\delta$，则可保证

$$\frac{\epsilon\zeta+2c_2^2n^3\sqrt{k}+c_1\lambda\sqrt{n}+c_2\lambda n\sqrt{k}+\epsilon}{\zeta-2c_2^2n^3-c_2\lambda n}<\epsilon+\delta,$$

这可简化为条件

$$\zeta>\frac{2c_2^2n^3(\sqrt{k}+\epsilon+\delta)+c_1\lambda\sqrt{n}+c_2\lambda n(\sqrt{k}+\epsilon+\delta)+\epsilon}{\delta}.$$

这意味着存在某个常数 $C_3(\lambda,\ \epsilon,\ c_1,\ c_2)>0$，使得当 $\zeta>C_3(\lambda,\ \epsilon,\ c_1,\ c_2)\frac{n^3\sqrt{k}}{\delta}$，$\|\mathbf{x}_1^T-\mathbf{x}_2^T\|<\epsilon+\delta$。

□

基于停止准则以及为我们的混合 ADMM 方案恰当选择 ζ，我们现在可以评估迭代复杂度，从而补全资源使用分析中的最后一块拼图。结果总结如下定理:

定理 12. 对于任意给定的 $0 < \epsilon, \delta < 1$，令 $\zeta = \frac{c_4 n^3 \sqrt{k}}{\delta}$，其中某个 $c_4 > C_3(\lambda, \epsilon, c_1, c_2)$，且 $\beta = c_5\zeta$，其中某个 $c_5 > \sqrt{2}$。算法 2 保证至多经过 $T = \mathcal{O}\left(\frac{n^6 k^{3/2}}{\epsilon^2 \delta}\right)$ 次迭代后收敛，并得到一个解，其中存在 $\mathbf{d}^T \in \partial\mathcal{L}_\beta(\mathbf{x}_1^T, \mathbf{x}_2^T, \mathbf{y}^T, \mathbf{w}^T)$ 使得 $\|\mathbf{d}^T\| < \epsilon$，且一致性误差为 $\|\mathbf{x}_1^T - \mathbf{x}_2^T\| < \delta + \epsilon$。

证明。我们首先分析 $\mathcal{L}_\beta$ 的下界和上界。根据命题 6，尤其是式 (43)，我们有

$$\begin{aligned}\min_{0\le t\le T} \mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) &\ge \inf_{(\mathbf{x}_1,\mathbf{x}_2,\mathbf{y})\in\mathcal{F}} \phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) \\ &\ge \min_{\mathbf{x}_2\in\mathbb{R}^n} f(\mathbf{x}_2) \\ &= -\frac{\lambda}{2}\boldsymbol{\mu}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ &\ge -\frac{\lambda\|\boldsymbol{\mu}\|^2}{2\gamma_{\min}(\boldsymbol{\Sigma})} \ge -\frac{\lambda c_1^2 n}{2c_3}\end{aligned} \tag{52}$$

另一方面,

$$\begin{aligned}\max_{0\le t\le T} \mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) &= L_\beta(\mathbf{x}_1^0, \mathbf{x}_2^0, \mathbf{y}^0, \mathbf{w}^0) \\ &\le \max_{\mathbf{x}_2\in\mathcal{S}} \left[\mathbf{x}_2^\intercal\mathbf{Q}^0\mathbf{x}_2 + \lambda\left(-\mathbf{x}_2^\intercal\boldsymbol{\mu} + \frac{1}{2}\mathbf{x}_2^\intercal\boldsymbol{\Sigma}\mathbf{x}_2\right)\right] \\ &\le \gamma_{\max}(\mathbf{Q}^0)\|\mathbf{x}_2\|^2 + \lambda\left(k\|\boldsymbol{\mu}\|_\infty + \frac{1}{2}\gamma_{\max}(\boldsymbol{\Sigma})\|\mathbf{x}_2\|^2\right) \\ &\le 2nk\gamma_{\max}^2(\boldsymbol{\Sigma}) + \frac{\lambda k}{2}\gamma_{\max}(\boldsymbol{\Sigma}) + \lambda k\|\boldsymbol{\mu}\|_\infty \\ &\le 2c_2^2 n^3 k + \frac{c_2\lambda nk}{2} + c_1\lambda k,\end{aligned} \tag{53}$$

其中，$\mathbf{Q}^0$ 是由初始化 $\mathbf{x}_1^0$ 诱导的矩阵 $\mathbf{Q}$。

假设算法 2 在经过 $T < T_{\max}$ 次迭代后成功终止，即 $\|\mathbf{y}^T - \mathbf{y}^{T-1}\| < \epsilon/(\beta+1)$。根据命题 7，$\|\mathbf{d}^T\| < \epsilon$。另一方面，根据命题 11，$\|\mathbf{x}_1^T - \mathbf{x}_2^T\| < \delta + \epsilon$ 可由 ζ 的选取保证。此外，根据推论 4、式 (52) 和式 (53),

$$\begin{aligned}\left(\frac{\beta}{2} - \frac{\zeta^2}{\beta}\right)\sum_{t=1}^T (\|\mathbf{x}_2^t - \mathbf{x}_2^{t-1}\|^2 + \|\mathbf{y}^t - \mathbf{y}^{t-1}\|^2) &\le \max_{0\le t\le T} \mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) - \min_{0\le t\le T} \mathcal{L}_\beta(\mathbf{x}_1^t, \mathbf{x}_2^t, \mathbf{y}^t, \mathbf{w}^t) \\ &\le 2c_2^2 n^3 k + \frac{c_2\lambda nk}{2} + c_1\lambda k + \frac{\lambda c_1^2 n}{2c_3},\end{aligned} \tag{54}$$

这意味着存在某个常数 $c_6 > 0$，使得

$$\|\mathbf{y}^T - \mathbf{y}^{T-1}\| \le \sqrt{\frac{2c_5 c_6}{(c_5^2 - 2)\zeta T}}\, n^{3/2}\sqrt{k}. \tag{55}$$

另一方面，算法 2 将在以下情况下终止:

$$\begin{aligned}\|\mathbf{y}^T - \mathbf{y}^{T-1}\| &< \frac{\epsilon}{\beta+1} \\ &= \frac{\epsilon}{c_5\zeta + 1}.\end{aligned} \tag{56}$$

结合公式 (55) 和公式 (56)，当满足以下条件时即可保证终止：

$$\sqrt{\frac{2c_5c_6}{(c_5^2-2)\zeta T}}n^{3/2}\sqrt{k}<\frac{\epsilon}{c_5\zeta+1},$$

这将在以下条件下得到满足

$$T=\left\lceil\frac{2c_5c_6n^3k(c_5\zeta+1)^2}{(c_5^2-2)\zeta\epsilon^2}\right\rceil+1=\mathcal{O}\left(\frac{n^6k^{3/2}}{\epsilon^2\delta}\right).$$

定理得证。 □

V. 资源分析

通过将 Grover 自适应搜索用于硬约束，我们提出了一种*ADMM-GAS-hard* 框架来求解风险平价问题。如算法 2 所述，我们交替使用针对硬约束问题的 GAS 求解式 (25) 中的固定基数二次二元优化，并使用经典求解器求解式 (26) 和 (27) 中的凸优化，以获得风险平价问题的解。考虑式 (2) 中带有固定基数约束 $k\leq n$ 的四次优化问题。我们的目标是在重构后的问题式 (23) 中生成高质量解，其特征由次梯度界 ϵ 和一致性容差 $\epsilon+\delta$ 表示。使用参数为 $\xi=1.34$ 的硬约束二元二次规划固定基数问题 GAS，该框架在求解式 (26) 和 (27) 中的子问题时产生 $\mathcal{O}\left(\frac{n^6k^{3/2}}{\epsilon^2\delta}\right)$ 的经典开销。二次预言机查询总数的渐近增长为：

$$\mathcal{O}\left(\sqrt{\binom{n}{k}}\frac{n^6k^{3/2}}{\sqrt{M}\epsilon^2\delta}\right)\tag{57}$$

其中， M 表示全局最优解的简并度。当 $k\ll n$ 是一个与 n无关的固定常数时，复杂度式（57）关于 n为多项式级，因此我们实现了指数级改进。更一般地，考虑 k 随 n线性增长的情形，即 $k=\alpha n$。根据 [49], 第 10 章中的引理 7，对于任意 $\alpha\in(0,1)$

$$\binom{n}{k}\leq\frac{1}{\sqrt{2\pi\alpha(1-\alpha)}}2^{nH_2(\alpha)},$$

其中 $H_2(\alpha)=-\alpha\log_2(\alpha)-(1-\alpha)\log_2(1-\alpha)$。由于 $\alpha\neq\frac{1}{2}$蕴含 $H_2(\alpha)<1$，因此我们有

$$\log\left[\binom{n}{k}\operatorname{poly}(n)\right]\leq H_2(\alpha)n+\log\operatorname{poly}(n).$$

因此，当 n 足够大时，在 $\alpha\neq\frac{1}{2}$的情况下，我们必有 $H_2(\alpha)n+\log\operatorname{poly}(n)<n$，这意味着$\binom{n}{k}\operatorname{poly}(n)\ll 2^n$。

相比之下，QD-GAS 方法需要

$$\mathcal{O}\left(\frac{2^{n/2}}{\sqrt{M}}\right)\tag{58}$$

四次预言机的查询。因此，对于任意固定的 $k\leq n$，ADMM-GAS-hard 所需的预言机总数增长速度都慢于风险平价模型中的 QD-GAS。由于 ADMM-GAS-hard 仅需二次预言机，其预言机所需的总量子门数量扩展为 $\mathcal{O}\left(\sqrt{\binom{n}{k}}\frac{n^8k^{3/2}m}{\sqrt{M}\epsilon^2\delta}\right)$，最多达到 $\widehat{C^2R}(\theta)$ 个量子门。然而，QD-GAS 方法的预言机需要 $\mathcal{O}(\frac{2^{n/2}n^4m}{\sqrt{M}})$ 至 $\widehat{C^4R}(\theta)$ 个量子门。在实现方面，Claudon 等人 [50] 声称，一个 d-受控 Pauli-X 门 $\widehat{C^dX}$ 可利用一个辅助量子比特分解为电路深度 $\Theta\left(\log(d)^3\right)$ 和门数量 $\mathcal{O}(d\log(d)^4)$，表示与 CZ_n 相同的规模。鉴于式（16）要求门数量为 $\mathcal{O}(nk)$，电路深度为$\mathcal{O}(k\log\frac{n}{k})$，所提出的框架在量子门数量和电路深度方面均实现了指数级缩减，从而带来

量子资源需求的大幅减少。为展示量子门使用量方面的总体差异，我们在表I中给出了部分预言机分析。

然而，我们注意到，交替方向乘子法并不保证在非凸问题中收敛到全局最优解。相反，该算法通常收敛到一个驻点，而该驻点可能是局部最优解，而非全局最优解。最优性差距的大小在很大程度上取决于目标函数景观。在量子-经典混合框架下对这一差距进行严格量化，是一项颇具挑战性的理论问题，我们将这项研究留待未来开展。

VI. 结论

总之，我们提出了一种 Grover 搜索子程序，专为硬约束二元二次规划问题中的 Grover 自适应搜索而设计，并证明在固定基数约束下，可在 $\frac{\pi}{4}\sqrt{\binom{n}{k}}\sqrt{M}$ 次 Grover 迭代内找到标记解，其中 k 表示规定的汉明重量，M 表示标记解的数量。与软约束二元二次规划的 Grover 搜索相比，这使 Grover 迭代次数呈指数级减少；后者的迭代复杂度上界为 $\frac{\pi}{4}\sqrt{2^n}/\sqrt{M}$。此外，我们还基于迪克态制备，为硬约束问题提供了扩散算子的显式量子线路构造，并分析了其电路深度和量子门复杂度。

此外，我们分析了应用于二元二次优化和高次多项式二元规划的 Grover 自适应搜索预言机的构造。因此，在查询复杂度层面，所提出的方法能够通过 Grover 自适应搜索，以指数级更快的速度确定固定基数二元二次规划的全局最优解。

此外，我们提出了一种经典-量子混合交替方向乘子法框架，通过在经典半正定规划求解器和我们的量子求解器之间交替迭代，来解决广泛应用于投资组合管理的风险平价模型。我们的方法总共需要对二次预言机进行 $\mathcal{O}\left(\sqrt{\binom{n}{k}}\frac{n^6k^{3/2}m}{\sqrt{M}\epsilon^2\delta}\right)$ 次查询，以获得一个 ϵ-近似解。与直接实现四次预言机相比——后者涉及诸如 $C^4R(\theta)$ 的多重受控旋转——我们的方法在整体上显著降低了量子门复杂度和电路深度，且降幅呈指数级。该框架还可在适当的正则性条件下进一步扩展至高次多项式优化问题，这将留待未来研究。

总体而言，我们的结果具有稳健性，并为量子算法的实际应用提供了见解。

VII. 致谢

作者感谢 Kelvin Koor 就复杂性问题进行讨论，并感谢 Yuemeng Sun 就风险平价模型进行讨论。

[1] 哈里·马科维茨。投资组合选择。《金融学杂志》，7:77–91，1952年3月。[2] Stéphane Maillard、Thierry Roncalli 和 Jérôme Teıletche。关于等权风险贡献投资组合的性质。《投资组合管理杂志》，36(4):60–70，2010年。[3] Federica Ricca 和 Andrea Scozzari。通过网络方法进行投资组合优化：网络同配混合与投资组合多元化。《欧洲运筹学杂志》，312(2):700–717，2024年。[4] Eugene Fratkin、Brian T. Naughton、Douglas L. Brutlag 和 Serafim Batzoglou。Motifcut：利用最大密度子图发现调控基序。《生物信息学》，22(14):e150–e157，2006年7月。[5] Barna Saha、Allison Hoch、Samir Khuller、Louiqa Raschid 和 Xiao-Ning Zhang。具有约束条件的稠密子图及其在基因注释图中的应用。载于 Bonnie Berger（编），《计算分子生物学研究》，第456–472页，柏林、海德堡，2010年。Springer 柏林·海德堡。[6] Monika R. Henzinger、Rajeev Motwani 和 Craig Silverstein。网络搜索引擎面临的挑战。《*SIGIR*论坛》，36(2):11–22，2002年9月。[7] David Gibson、Ravi Kumar 和 Andrew Tomkins。在大规模图中发现大型稠密子图。载于第*31*届超大型数据库国际会议论文集，VLDB '05，第721–732页。VLDB基金会，2005年。[8] Ulrik Brandes。网络分析：方法论基础，第3418卷。施普林格科学与商业媒体，2005年。

[9] Dimitris Bertsimas、Angela King 和 Rahul Mazumder。通过现代优化视角进行最佳子集选择。《统计年鉴》，44(2)，2016年。[10] Antoine Dedieu、Hussein Hazimeh 和 Rahul Mazumder。学习稀疏分类器：连续优化与混合整数优化视角。《机器学习研究杂志》，22(135)：1–47，2021年。[11] U. Feige、D. Peleg 和 G. Kortsarz。稠密 k 子图问题。Algorithmica，29(3)：410–421，2001年3月。[12] Edward Farhi、Jeffrey Goldstone 和 Sam Gutmann。一种量子近似优化算法，2014年11月。[13] M. Cerezo、Andrew Arrasmith、Ryan Babbush、Simon C. Benjamin、Suguru Endo、Keisuke Fujii、Jarrod R. McClean、Kosuke Mitarai、Xiao Yuan、Lukasz Cincio 和 Patrick J. Coles。变分量子算法。《自然评论：物理》，3(9)：625–644，2021年8月。[14] Tadashi Kadowaki 和 Hidetoshi Nishimori。横场伊辛模型中的量子退火。Phys. Rev. E，58：5355–5363，1998年11月。[15] Christoph Durr 和 Peter Hoyer。一种寻找最小值的量子算法，1999年。[16] Haomu Yuan、Daniel Stilck França、Ilia Luchnikov、Egor Tiunov、Tobias Haug 和 Leandro Aolita。通过量子吉布斯态和泡利稀疏性，对结构化 Goemans-Williamson 松弛实现指数级加速，2025年。[17] Stuart Hadfield、Zhihui Wang、Bryan O’Gorman、Eleanor G. Rieffel、Davide Venturelli 和 Rupak Biswas。从量子近似优化算法到量子交替算子拟设。《算法》，12(2)：34，2019年2月。[18] Haomu Yuan、Christopher K. Long、Hugo V. Lepage 和 Crispin H. W. Barnes。量化将量子近似算法应用于投资组合优化的优势，2024年。[19] Austin Gilliam、Stefan Woerner 和 Constantin Gonciulea。用于约束多项式二元优化的 Grover 自适应搜索。《Quantum》，5：428，2021年4月。[20] Renata Sotirov。关于大型图上稠密子图问题的求解。《优化方法与软件》，35(6)：1160–1178，2020年。[21] G. Kortsarz 和 D. Peleg。关于选择稠密子图。载于《1993年 IEEE 第34届年度计算机科学基础会议论文集》，第692–701页，1993年。[22] D.G. Corneil 和 Y. Perl。完美图中的聚类与支配。《离散应用数学》，9(1)：27–39，1984年。[23] Maurizio Bruglieri、Matthias Ehrgott、Horst W. Hamacher 和 Francesco Maffioli。具有固定基数约束的组合优化问题注释书目。《离散应用数学》，154(9)：1344–1357，2006年。第二届科隆/特温特图与组合优化研讨会（CTW 2003）。[24] U. Feige 和 M. Seltser。关于最稠密 k 子图问题。技术报告，以色列魏茨曼科学出版社，ISR，1997年。[25] Edward Qian。风险贡献的金融解释：风险预算确实可以相加。Panagora资产管理公司，2005年。[26] Clifford S. Asness、Andrea Frazzini 和 Lasse Heje Pedersen。杠杆厌恶与风险平价。《金融分析师杂志》，68(1)：47–59，2012年。[27] Lov K. Grover。一种用于数据库搜索的快速量子力学算法。载于《第二十八届 ACM 计算理论年度研讨会论文集》，STOC ’96，第212–219页，纽约，美国，1996年。ACM。[28] D. Bulger、W. P. Baritompa 和 G. R. Wood。利用 Grover 量子算法实现纯自适应搜索。《优化理论与应用杂志》，116(3)：517–529，2003年3月。[29] W. P. Baritompa、D. W. Bulger 和 G. R. Wood。将 Grover 量子算法应用于全局优化。《SIAM 优化杂志》，15(4)：1170–1184，2005年。[30] Gilles Brassard、Peter HØyer 和 Alain Tapp。量子计数。载于 Kim G. Larsen、Sven Skyum 和 Glynn Winskel 编，《自动机、语言与编程》，第820–831页，柏林、海德堡，1998年。Springer 柏林·海德堡。[31] Austin Gilliam、Charlene Venci、Sreraman Muralidharan、Vitaliy Dorum、Eric May、Rajesh Narasimhan 和 Constantin Gonciulea。高效量子计算的基础模式，2021年。[32] Steven A. Cuccaro、Thomas G. Draper、Samuel A. Kutin 和 David Petrie Moulton。一种新的量子波纹进位加法电路，2004年。[33] Robin Ollive 和 Stephane Louise。基于量子信号处理的 Grover 自适应搜索预言机，用于高阶无约束二元优化问题。载于《2024年 IEEE 量子计算与工程国际会议（QCE）》，第02卷，第256–261页，2024年。[34] Haomu Yuan。《量子优化算法及其应用》。博士论文，剑桥大学，2026年。提交于2026年1月22日。[35] Eunok Bae、Jeonghyeon Shin 和 Minjin Choi。通过约束感知初始化减少 Grover 算法中的电路资源，2026年。[36] Andreas Bärtschi 和 Stephan Eidenbenz。Dicke 态的确定性制备。载于 Leszek Antoni Gąsieniec、Jesper Jansson 和 Christos Levcopoulos 编，《计算理论基础》，第126–139页，尚，2019年。Springer 国际出版。[37] Andreas Bärtschi 和 Stephan Eidenbenz。用于 Dicke 态制备的短深度电路。载于《2022年 IEEE 量子计算与工程国际会议（QCE）》，第87–96页，2022年。[38] Yunseong Nam、Yuan Su 和 Dmitri Maslov。具有 o(n log(n)) 个 T 门的近似量子傅里叶变换。《npj 量子信息》，6(1)，2020年3月。

[39] Claudio Gambella 和 Andrea Simonetto. 经典计算机和量子计算机上混合二元优化的多块 ADMM 启发式方法。*IEEE* 量子工程汇刊，1:1–22，2020年。 [40] Stephen Boyd、Neal Parikh、Eric Chu、Borja Peleato、Jonathan Eckstein 等。通过交替方向乘子法实现分布式优化与统计学习。基础与趋势® 机器学习，3 122，2011 1 1( ): – .[41] Yu Wang、Wotao Yin 和 Ji nshan Zeng。非凸非光滑优化中 ADMM 的全局收敛性。科学计算杂志，78:29–63，2019年。 [42] Olivier Ledoit 和 Michael Wolf。亲爱的，我缩小了样本协方差矩阵。《投资组合管理杂志》，30(4):110–119，2004年7月。 [43] Olivier Ledoit 和 Michael Wolf。适用于高维协方差矩阵的良态估计量。多元分析杂志，88(2):365–411，2004年。 [44] R. Tyrrell Rockafellar 和 Roger J. B. Wets。变分分析，第317卷，属于数学科学基础丛书。Springer 柏林·海德堡，柏林，海德堡，1998年。 [45] Yurii Nesterov。凸优化讲义，第137卷，属于*Springer* 优化及其应用丛书。Springer 国际出版，尚，2018年。 [46] Guoyin Li 和 Ting Kei Pong。Kurdyka–Łojasiewicz 不等式指数的计算及其在线性收敛一阶方法中的应用。计算数学基础，18(5):1199–1232，2018年。 [47] Jérôme Bolte、Aris Daniilidis、Adrian Lewis 和 Masahiro Shiota。可分层函数的 Clarke 次梯度。《*SIAM* 优化杂志》，18(2):556–572，2007年。 [48] Hédy Attouch、Jérôme Bolte、Patrick Redont 和 Antoine Soubeyran。非凸问题的近端交替最小化与投影方法：基于 Kurdyka–Łojasiewicz 不等式的方法。运筹学数学，35(2):438–457，2010年。 [49] F.J. MacWilliams 和 N.J.A. Sloane。纠错码理论。数学文库。North-Holland 出版公司，1977年。 [50] Baptiste Claudon、Julien Zylberman、César Feniou、Fabrice Debbasch、Alberto Peruzzo 和 Jean-Philip Piquemal。无需辅助量子比特的多对数深度受控非门。自然通讯，15(1)，2024年7月。 [51] Craig Gidney。将量子加法的成本减半。《*Quantum*》

，2:74，2018年6月。

附录A：分裂与循环移位酉算子的实现

给定 $n \geq 2$，$k \geq 1$，$SCS_{n,\,k}$ 酉算子可以采用 [36], 提出的方法实现，该方法基于依次实现 $\widehat{SCS}_2^{(n)}$ 和 $\widehat{SCS}_3^{(n,t)}$ 算子：

$$\widehat{SCS}_{n,k} = \begin{cases} \widehat{SCS}_2^{(n)}, & k = 1, \\ \widehat{SCS}_2^{(n)} \prod_{t=2}^{k} \widehat{SCS}_3^{(n,t)}, & k \geq 2. \end{cases} \tag{A1}$$

其中，$\widehat{SCS}_2^{(n)}$ 和 $\widehat{SCS}_3^{(n,t)}$ 如图1所示。

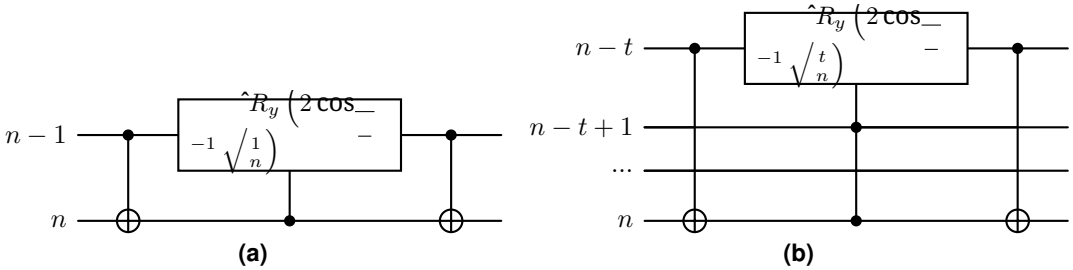

图1. $\widehat{SCS}_2^{(n)}$ 和 $\widehat{SCS}_3^{(n,t)}$ 的量子线路，用于 $\widehat{SCS}_{n,\,l}$ 酉算子：（a）$\widehat{SCS}_2^{(n)}$，（b）$\widehat{SCS}_3^{(n,t)}$。

附录 B：量子字典预言机

本节简要回顾应用于 BPP-FC [19] 的 Grover 自适应搜索（GAS）量子字典预言机实现。所需量子门和量子比特的汇总见表I。

第一步是将函数值 $f(\mathbf{x}) - y$ 编码到辅助量子比特中，即，

$$O : |\mathbf{x}\rangle_n |0\rangle_m \to |\mathbf{x}\rangle_n |f(\mathbf{x}) - y\rangle_m . \tag{B1}$$

其中 $|\mathbf{x}\rangle := |x_1 x_2, \ldots\ldots, x_i, \ldots\ldots, x_n\rangle$, $x_i \in \{0, 1\}$ 是计算基中的一种，而 $f(\mathbf{x}) - y$ 表示二进制补码表示法，即 $f(\mathbf{x}) - y \in \mathbb{Z}$, $-2^{m-1} \leq f(\mathbf{x}) < 2^{m-1}$。注意，二次函数的系数可能不是整数，因此我们需要将目标函数近似到 m位二进制小数，这会引入随 $\epsilon = \mathcal{O}(2^{-m})$ 缩放的近似误差。将量子算术预言机应用于式（1）中的二次二进制目标函数时，我们需要实现 $\mathcal{O}(n^2)$ 次加法和乘法，以在辅助量子比特上构造 $f(\mathbf{x}) - y$。在量子算术中，直接实现 m位整数的加法需要$2m+2$个量子比特和 $2m + 1$ 个托福利门 [32, 51]。注意，在乘法和加法开始时，我们可以迭代地将量子比特优化至 $m' < m$，以降低实现复杂度；因此，所需的量子门总数可能小于 [19] 中所述的数量。不过，改进这一点并非本文重点；我们对实现量子算术预言机的成本作宽松估计，认为需要 $\mathcal{O}(n^2 m)$ 个托福利门和 m 个辅助量子比特。

然而，受量子数据结构的启发，量子字典预言机可以将函数值直接映射到辅助量子比特，而无需执行算术运算。总之，量子字典预言机通过受控酉算符 $CU_{\mathbf{x}}(2\pi/2^m)$构造映射$\widehat{\ }$，该算符以基态 $|x\rangle_n$ 为控制，然后对目标辅助量子比特施加 $U_{\mathbf{x}}(2\pi/2^m)$，即

$$\hat{U}_{\mathbf{x}} (2\pi/2^m) |K\rangle_m = e^{i2\pi K[f(\mathbf{x})-y]/2^m} |K\rangle_m . \tag{B2}$$

其中，K 是基态 $|K\rangle_m$的索引。然后，受控酉算符（包括哈达玛变换）作用于初始态的结果为

$$\begin{aligned}
&\widehat{CU}_{\mathbf{x}} (2\pi/2^m) |\mathbf{x}\rangle_n \hat{H}^{\otimes m} |0\rangle_m \\
=& |\mathbf{x}\rangle_n \hat{U}_{\mathbf{x}} (2\pi/2^m) \hat{H}^{\otimes m} |0\rangle_m \\
=& |\mathbf{x}\rangle_n \frac{1}{\sqrt{N'}} \sum_{K=1}^{N'-1} \hat{U}_{\mathbf{x}} (2\pi/2^m) |K\rangle_m \\
=& |\mathbf{x}\rangle_n \frac{1}{\sqrt{N'}} \sum_{K=1}^{N'-1} e^{i2\pi K[f(\mathbf{x})-y]/2^m} |K\rangle_m
\end{aligned} \tag{B3}$$

其中 $N' = 2^m$。随后，对辅助量子比特施加逆量子傅里叶变换 $QFT^\dagger$，以将 $f(\mathbf{x})$的二进制表示 $- y$ 编码到 m 量子比特中，如式 (B1) 所示。构造 $U_{\mathbf{x}}(2\pi/2^m)$需要依次施加 m个相位门，其旋转角逐渐递增，分别为$R(2^j\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i2^j\theta} \end{pmatrix}$，$j = 0, \ldots m - 1$，并作用于每个 m 辅助量子比特。受控相位旋转以目标函数中对应于非平凡元素的量子比特为控制，并对目标辅助量子比特施加$R(2^j\theta)$。