



量化压缩感知中最优阈值算法的数据时间权衡

献给韩继业教授 90 寿辰

徐加樑^{1*}, 赵云彬²

1. 北京科技大学数理学院, 北京 100083;

2. 深圳市大数据研究院深圳国际工业与应用数学中心, 深圳 518172

E-mail: jialiangu@ustb.edu.cn, yunbinzhao@cuhk.edu.cn

收稿日期: 2024-04-30; 接受日期: 2024-09-22; 网络出版日期: 2025-01-03; * 通信作者

国家自然科学基金 (批准号: 12301393, 12071307 和 12471295)、基本科研业务费 (批准号: FRF-TP-22-103A1) 和河套深港科技创新合作区深圳园区项目 (批准号: HZQSWS-KCCYB-2024016) 资助项目

摘要 松弛最优阈值算法是压缩感知中的一种新的阈值技术, 用于从少量带噪线性测量中高效恢复稀疏信号, 克服了传统硬阈值算法由于残差函数振荡导致的缺点. 本文考虑一种新颖的数据时间权衡分析, 以说明在任何固定恢复精度下, 松弛最优阈值算法在压缩感知、量化压缩感知还原信号过程中测量数量和迭代次数之间的权衡. 理论结果还给出在压缩感知、量化压缩感知中实现信号恢复所需要的有效测量阶数. 最后, 通过数字实验也展示了松弛最优阈值算法在稀疏重建问题中的数据时间权衡.

关键词 压缩感知 量化压缩感知 稀疏重建问题 松弛最优阈值算法 数据时间权衡

MSC (2020) 主题分类 94A12, 15A29, 90C25, 90C20

1 引言

近年来, 稀疏表示理论在信号和图像处理领域备受关注 (参见文献 [2, 8, 9, 19, 35]). 为了有效解决科学和工程领域的信号处理问题, 研究者们已经提出了许多用于恢复稀疏信号的算法 [10, 21, 22, 24, 43]. 具体而言, 在特定情境下, 假设信号是稀疏的, 若测量次数远远小于信号维数, 则有可能从测量数据中恢复出信号的关键特征. 稀疏信号的恢复等同于求解稀疏优化模型或稀疏重建问题. 用 $\|\mathbf{x}\|_0$ 表示 “ ℓ_0 -范数”, 即 $\mathbf{x} \in \mathbb{R}^n$ 中非零分量的个数. 从欠定线性方程组中寻找原始信号的模型可以被表述为 ℓ_0 最小化问题:

$$\min_{\mathbf{x}} \{\|\mathbf{x}\|_0 : \mathbf{y} = \mathbf{A}\mathbf{x}\},$$

其中 $\mathbf{y} := \mathbf{A}\mathbf{x}^*$ 表示目标信号 $\mathbf{x}^* \in \mathbb{R}^n$ 的测量, $\mathbf{A} \in \mathbb{R}^{m \times n}$ 是 $m < n$ 的测量 (感知) 矩阵. 上述模型的目的是最好地适应测量结果 $\mathbf{y}$, 找到一个最稀疏的信号 $\hat{\mathbf{x}}$, 从而实现恢复 $\hat{\mathbf{x}} = \mathbf{x}^*$. 若按照实际问题出

英文引用格式: Xu J L, Zhao Y B. Data-time tradeoffs for optimal k -thresholding algorithms in quantized compressed sensing (in Chinese). Sci Sin Math, 2025, 55: 343–360, doi: 10.1360/SSM-2024-0142

发, 考虑重建目标信号的最显著信息, 这些信息可以表示为信号在其冗余基上的几个绝对值最大的系数. 基于这一考虑, 给定目标信号的稀疏度水平 k , 稀疏重建模型可以表述为以下带有稀疏性约束的优化模型:

$$\min_{\mathbf{x}} \{\|\mathbf{y} - \mathbf{Ax}\|_2^2 : \|\mathbf{x}\|_0 \leq k\}. \quad (1.1)$$

在这个模型中, $\mathbf{y} := \mathbf{Ax}^* + \mathbf{n}$ 表示测量结果, $\mathbf{n} \in \mathbb{R}^m$ 代表噪声. 模型 (1.1) 是压缩感知 (compressed sensing, CS) 理论和算法发展中不可或缺模型之一, 与其他一些稀疏重建问题密切相关, 如量化压缩感知 (quantized compressed sensing, QCS) [6, 26, 42] 和损坏感知 [41]. 具体地, 为了有限精度的数据表示和数字化融合, QCS 的目标是从少量量化测量中估计出稀疏信号 $\mathbf{x}^* \in \mathbb{R}^n$,

$$\mathbf{y} = \mathcal{Q}(\mathbf{Ax}^* + \mathbf{n} + \boldsymbol{\tau}), \quad (1.2)$$

$\mathcal{Q}(\cdot)$ 是量化方案 (详见第 3.1 小节), $\boldsymbol{\tau} \in \mathbb{R}^m$ 是随机均匀抖动.

尽管具有稀疏目标函数或约束的优化模型是 NP (non-deterministic polynomial) 难问题 [30], 但它们实际上可以通过一些高效的算法以很高的概率实现 (参见文献 [4, 10, 15–17, 29, 44, 45]). 解决这类问题的可行算法通常被分为三类: 凸优化方法 [10, 15, 45]、启发式方法 [16, 29] 和阈值方法 [4, 17, 44]. 阈值类算法是这些高效算法中的一种重要方法, 它最初由 Donoho 和 Johnstone [20] 提出, 用于解决信号去噪问题. 根据阈值算子的性质, 阈值类的算法可以分为两类: 软阈值 [17, 18, 22, 39] 和硬阈值 [4, 5, 17, 23]. 软阈值算法是通过考虑某些凸问题的最优条件来建立的. 使用硬阈值算子的基本算法是迭代硬阈值 (iterative hard thresholding, IHT) 算法 [4, 24] 和硬阈值追踪 (hard thresholding pursuit, HTP) 算法 [23, 24], 其中还有几种衍生算法, 如分级 IHT [3, 25]、带自适应步长的 IHT [34] 和压缩采样匹配追踪 (compressive sampling matching pursuit, CoSaMP) [31]. 从数值性能的角度来看, 基于硬阈值算子的算法在迭代过程中远未达到稳健和稳定, 并且数值结果表明它们对步长和稀疏度的选择非常敏感. 为了克服基于硬阈值算子算法的固有缺陷, Zhao [44] 提出了一种新的阈值算子, 称为最优 k 阈值算子, 基于此提出了一类新颖的阈值算法.

为了描述传统的硬阈值方法和最优 k 阈值方法, 我们引入一些符号. 用 $\mathbb{R}^n$ 表示 n 维 Euclid 空间, 用 $\mathbb{O}^{n-1}$ 表示 n 维 Euclid 空间中的单位球. 用 $\mathbf{e}$ 表示元素全为 1 的向量. 给定一个向量 $\mathbf{z} \in \mathbb{R}^n$, 硬阈值算子 $\mathcal{H}_k(\mathbf{z})$ 保留 $\mathbf{z}$ 中绝对值最大的 k 个分量, 并将其余分量设为 0. 给定一个集合 $\mathcal{S} \subseteq \{1, 2, \dots, n\}$, 用 $|\mathcal{S}|$ 表示集合 $\mathcal{S}$ 的基数, 用 $\bar{\mathcal{S}} = \{1, 2, \dots, n\} \setminus \mathcal{S}$ 表示 $\mathcal{S}$ 的补集. 给定一个向量 $\mathbf{x}$ 和一个矩阵 $\mathbf{A}$, $\mathbf{x}^T$ 和 $\mathbf{A}^T$ 分别表示 $\mathbf{x}$ 和 $\mathbf{A}$ 的转置, 用 $\text{supp}(\mathbf{x})$ 表示 $\mathbf{x}$ 的支撑集. 若 $\|\mathbf{x}\|_0 \leq k$, 则称向量 $\mathbf{x} \in \mathbb{R}^n$ 是 k -稀疏的.

Blumensath 和 Davies [4] 将硬阈值算子 $\mathcal{H}_k(\mathbf{z})$ 与最小化残差函数 $\|\mathbf{y} - \mathbf{Ax}\|_2^2$ 的梯度方法结合起来, 提出了问题 (1.1) 的一种硬阈值算法. 它被称作 IHT 算法, 有如下的迭代步骤:

$$\mathbf{x}^{p+1} = \mathcal{H}_k(\mathbf{x}^p + \mathbf{A}^T(\mathbf{y} - \mathbf{Ax}^p)).$$

通过添加追踪步骤, 例如在 IHT 确定的支撑集上最小化残差函数, Foucart [23] 提出了一种 IHT 的加速版本, 称为硬阈值追踪 (HTP). 然而, 在大部分数值分析中, 它们表现出残差减少与阈值步骤的无关性, 这导致了 IHT 甚至 HTP 的发散. Zhao [44] 提出了一种迭代方案: 在阈值步骤之前寻找使得残差最小化的 k 个位置,

$$\mathbf{w}(\mathbf{u}^p) = \min_{\mathbf{w}} \{\|\mathbf{y} - \mathbf{A}(\mathbf{u}^p \odot \mathbf{w})\|_2^2, \mathbf{e}^T \mathbf{w} = k, \mathbf{w} \in \{0, 1\}^n\},$$

其中 $\mathbf{u}^p = \mathbf{x}^p + \mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^p)$, $\odot$ 表示 Hadamard 乘积, 并且使用 $\mathbf{x}^{p+1} = \mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p)$ 进行下一次迭代. 这个算法被称为最优阈值 (optimal k -thresholding, OT) 算法, 其在 $\mathbf{w}(\mathbf{u}^p)$ 的支撑集上带有追踪步骤的算法被称为最优阈值追踪 (optimal k -thresholding pursuit, OTP) 算法. 为了使算法易于处理, Zhao^[44] 用 $0 \leq \mathbf{w} \leq \mathbf{e}$ 替换 $\mathbf{w} \in \{0, 1\}^n$, 并研究了它们的凸松弛版本, 即松弛最优阈值 (relaxed optimal k -thresholding, ROT) 算法和松弛最优阈值追踪 (relaxed optimal k -thresholding pursuit, ROTP) 算法. 它们分别在算法 1 和 2 中给出.

算法 1 松弛最优阈值算法 (ROT)

输入: 测量矩阵 $\mathbf{A}$, 测量 $\mathbf{y}$, 稀疏度 k , 最大迭代步数 P .

输出: 估计结果 $\mathbf{x}^*$.

- 1: 初始化 $p = 0$, ε , η , and $\mathbf{x}^0$.
 - 2: **repeat**
 - 3: $\mathbf{u}^p = \mathbf{x}^p + \eta \mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^p)$,
 - 4: $\mathbf{w}(\mathbf{u}^p) = \arg \min_{\mathbf{w}} \{\|\mathbf{y} - \mathbf{A}(\mathbf{u}^p \odot \mathbf{w})\|_2^2, \mathbf{e}^T \mathbf{w} = k, 0 \leq \mathbf{w} \leq \mathbf{e}\}$,
 - 5: $\mathbf{x}^{p+1} = \mathcal{H}_k(\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p))$,
 - 6: $p = p + 1$,
 - 7: **until** $p > P$ or $\|\mathbf{x}^p - \mathbf{x}^{p-1}\|_2 < \varepsilon$.
-

算法 2 松弛最优阈值追踪算法 (ROTP)

输入: 测量矩阵 $\mathbf{A}$, 测量 $\mathbf{y}$, 稀疏度 k , 最大迭代步数 P .

输出: 估计结果 $\mathbf{x}^*$.

- 1: 初始化 $p = 0$, ε , η , and $\mathbf{x}^0$.
 - 2: **repeat**
 - 3: $\mathbf{u}^p = \mathbf{x}^p + \eta \mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^p)$,
 - 4: $\mathbf{w}(\mathbf{u}^p) = \arg \min_{\mathbf{w}} \{\|\mathbf{y} - \mathbf{A}(\mathbf{u}^p \odot \mathbf{w})\|_2^2, \mathbf{e}^T \mathbf{w} = k, 0 \leq \mathbf{w} \leq \mathbf{e}\}$,
 - 5: $\mathbf{v} = \mathcal{H}_k(\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p))$,
 - 6: $\mathbb{S}^{p+1} = \text{supp}(\mathbf{v})$,
 - 7: $\mathbf{x}^{p+1} = \arg \min_{\mathbf{x}} \{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2, \text{supp}(\mathbf{x}) \subseteq \mathbb{S}^{p+1}\}$,
 - 8: $p = p + 1$,
 - 9: **until** $p > P$ or $\|\mathbf{x}^p - \mathbf{x}^{p-1}\|_2 < \varepsilon$.
-

在感知矩阵满足受限等距性质 (restricted isometry property, RIP) 的前提下, Zhao^[44] 对 ROT 和 ROTP 的收敛性进行了分析. 但是, 基于 RIP 的收敛性的主要缺点是它无法描述还原准确性、测量次数和迭代次数之间的关系. 具体而言, 上述这些关系的理论分析被称为算法的数据 - 时间权衡分析, 这在从欠定系统中获取数据相关的领域中非常重要 (参见文献 [12, 33]). Oymak 等^[33] 分析了压缩感知中投影梯度方法的尖锐数据 - 时间权衡; Chen 和 Liu^[12] 刻画了在次 Gauss 矩阵下污染感知中投影梯度下降算法的数据 - 时间权衡. 在稀疏信号恢复领域, 还有一些相关工作也考虑了稀疏恢复问题中的数据 - 时间权衡^[1, 14, 33], 在线性反问题领域也有相关研究 (参见文献 [7, 11]).

为了克服上述缺点, 本文针对 ROT 和 ROTP 建立了一种新的收敛性分析, 并提供了在次 Gauss 测量下算法的数据 - 时间权衡. 我们分析了量化压缩感知中的线性反问题的 ROT 和 ROTP 算法的数据 - 时间权衡. 理论分析表明, 当测量次数相对较大时, 随着测量次数的增加, 成功重构所需要的迭代次数相应的减少. 并且, 当测量次数在压缩感知和量化压缩感知中的量级为 $k \log(n/k)$, 且具有相对较大的常数时, 算法可以成功地恢复目标信号. 此外, 还表明, 在无噪声的情形下, 算法可以在压缩感

知中实现线性收敛速率.

本文余下内容的结构如下. 第 2 节介绍一些有用的定义和事实, 并介绍 Gauss 复杂性和 Gauss 宽度的概念以及矩阵偏差 inequality. 第 3 节致力于提供问题 (1.1) 的 ROT 算法和 ROTP 算法的理论分析 (考虑数据 - 时间权衡), 并将分析扩展到量化压缩感知和压缩感知. 第 4 节为 ROT 算法和 ROTP 算法提供了数值实验, 展示了上述稀疏问题中测量次数和迭代次数之间的权衡. 最后, 第 5 节给出结论和未来工作.

2 预备工作

本节给出了几个定义和事实, 这些内容在接下来的内容中将要被使用. 所有的定义都可以在文献 [38] 中找到.

2.1 定义

如果随机变量 x 具有有限的次 Gauss 范数, 则称其为 Gauss 变量,

$$\|x\|_{\psi_2} = \inf \left\{ t > 0 : \mathbb{E} \exp \left(\frac{x^2}{t^2} \right) \leq 2 \right\}.$$

Gauss 变量、Bernoulli 变量和有界随机变量是典型的次 Gauss 变量.

如果对于任意的 $\mathbf{y} \in \mathbb{R}^n$, $\langle \mathbf{x}, \mathbf{y} \rangle$ 都是次 Gauss 的, 则随机向量 $\mathbf{x} \in \mathbb{R}^n$ 也被称为次 Gauss 变量. 相应的范数定义为 $\|\mathbf{x}\|_{\psi_2} = \sup_{\mathbf{y} \in \mathbb{Q}^{n-1}} \|\langle \mathbf{x}, \mathbf{y} \rangle\|_{\psi_2}$. 如果一个随机向量 $\mathbf{x}$ 满足 $\mathbb{E}[\mathbf{x}\mathbf{x}^T] = \mathbf{I}$, 则它是等向的, 有界子集 $\mathbb{T}$ 的 Gauss 复杂性被定义为

$$\gamma(\mathbb{T}) = \mathbb{E} \left[\sup_{\mathbf{x} \in \mathbb{T}} |\langle \mathbf{g}, \mathbf{x} \rangle| \right],$$

并且有界子集 $\mathbb{T}$ 的 Gauss 宽度被定义为

$$w(\mathbb{T}) = \mathbb{E} \left[\sup_{\mathbf{x} \in \mathbb{T}} \langle \mathbf{g}, \mathbf{x} \rangle \right],$$

其中 $\mathbf{g} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$. 文献 [13] 表明上述两个量具有相同的阶:

$$\frac{1}{3}(w(\mathbb{T}) + \|\mathbf{y}\|_2) \leq \gamma(\mathbb{T}) \leq 2(w(\mathbb{T}) + \|\mathbf{y}\|_2), \quad \forall \mathbf{y} \in \mathbb{T}. \quad (2.1)$$

2.2 事实

事实 2.1 (矩阵偏差 inequality ^[37]) 设 $\mathbf{A}$ 是一个 $m \times n$ 的随机矩阵, 其行 $\{\mathbf{A}_i\}_{i=1}^m$ 是独立、中心化、等向的次 Gauss 向量. 对于任意有界子集 $\mathbb{D} \subseteq \mathbb{R}^n$ 和 $t \geq 0$, 不等式

$$\sup_{\mathbf{x} \in \mathbb{D}} \|\mathbf{A}\mathbf{x}\|_2 - \sqrt{m}\|\mathbf{x}\|_2 \leq CK^2[\gamma(\mathbb{D}) + t \cdot \text{rad}(\mathbb{D})]$$

在至少 $1 - 2\exp(-t^2)$ 的概率下成立其中 $\text{rad}(\mathbb{D}) = \sup_{\mathbf{x} \in \mathbb{D}} \|\mathbf{x}\|_2$, $K = \max_i \|\mathbf{A}_i\|_{\psi_2}$, 并且 $C > 0$ 是一个绝对常数.

事实 2.2 (ℓ_0 球的 Gauss 宽度^[40]) 令 $\mathbb{D}_k = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y}\|_0 \leq k, \|\mathbf{y}\|_2 = 1\}$, 则它的 Gauss 宽度满足下述等式:

$$w(\mathbb{D}_k) \leq c \sqrt{k \log \frac{en}{k}}.$$

我们将用到以下两个事实来证明重要结果.

事实 2.3^[44] 设 $\mathbf{x} \in \mathbb{R}^n$ 是一个 k -稀疏向量, $\mathbb{S}$ 是 $\mathbf{x}$ 的支撑集. 设 $\mathbf{w}^p \in \mathbb{R}^n$ 是算法 1 或算法 2 第 4 步中问题的一个可行解, $\mathbf{v} = \mathcal{H}_k(\mathbf{u}^p \odot \mathbf{w}^p) \in \mathbb{R}^n$ 是一个具有支撑集 $\mathbb{S}^{p+1}$ 的向量. 则下面不等式成立:

$$\|\mathbf{x} - \mathbf{v}\|_2 \leq \|(\mathbf{u}^p \odot \mathbf{w}^p - \mathbf{x})_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2 + \|(\mathbf{u}^p \odot \mathbf{w}^p - \mathbf{x})_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2.$$

事实 2.4^[44] 令 k 和 $\bar{k}$ 为两个整数, 使得 $0 \leq \bar{k} \leq k$. 令 $\mathbf{w} \in \mathcal{W}^k = \{\mathbf{w} : \mathbf{e}^T \mathbf{w} = k, \mathbf{w} \in \{0, 1\}^n\}$ 为具有支撑集 $\mathbb{S}$ 的向量, $\bar{\mathbb{S}}$ 为互补集. 令 $\bar{\mathbb{S}}$ 具有不相交的分区 $\mathbb{S}_i, i = 1, \dots, \bar{n}$, 并满足以下属性:

- 对于 $i = 1, 2, \dots, \bar{n} - 1$, 有 $|\mathbb{S}_i| \leq k$ 并且 $|\mathbb{S}_{\bar{n}}| \leq \bar{k} < k$;
- $\mathbb{S}_i$ 是 $\mathbf{w}_{\bar{\mathbb{S}}}$ 中第 i 组的 k 个最大分量的索引集.

设 $\alpha_{(i)}$ 是 $\mathbf{w}$ 在支撑集 $\mathbb{S}_i$ 上的最大分量. 于是有

$$\sum_{i=1}^{\bar{n}} \alpha_{(i)} \leq 2 - \frac{1}{k} < 2.$$

2.3 重要引理

为了方便收敛性分析, 首先给出一些有用的引理. 在事实 2.1 中, 令 $\mathbb{D} \in \mathbb{O}^{n-1}$ 和 $t = \gamma(\mathbb{D})$, 则得到以下引理.

引理 2.1 设 $\mathbf{A}$ 是如事实 2.1 中所定义的一个 $m \times n$ 的随机矩阵. 设 $\mathbb{D} \in \mathbb{O}^{n-1}$ 和 $t = \gamma(\mathbb{D})$, 则对于任意 $\mathbf{x} \in \mathbb{D}$, 不等式

$$\sqrt{m} - 2CK^2\gamma(\mathbb{D}) \leq \|\mathbf{Ax}\|_2 \leq \sqrt{m} + 2CK^2\gamma(\mathbb{D})$$

以至少 $1 - 2\exp(-\gamma^2(\mathbb{D}))$ 的概率成立.

下面两个引理经常会被用到主要结果的证明.

引理 2.2 设 $\mathbf{A}$ 是事实 2.1 中所定义的一个 $m \times n$ 的随机矩阵. 设 $\mathbb{D}_k = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y}\|_0 \leq k, \|\mathbf{y}\|_2 = 1\}$. 设 $\mathbb{L} \subseteq \{1, \dots, n\}$ 是一个具有基数 l 的索引集. 将 $[\mathbf{x}]_{\mathbb{L}}$ 定义为通过保留支持集 $\mathbb{L}$ 的分量并将其余分量置零而获得的向量. 设 η 是一个使得 $\eta m \leq 1$ 的步长. 我们可以得到下面结果:

(a) 对于任意的 k -稀疏向量 $\mathbf{v} \in \mathbb{R}^n$, 不等式

$$\|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}]_{\mathbb{L}}\|_2 \leq \max\{s_1, s_2\} \cdot \|\mathbf{v}\|_2$$

至少在 $1 - 4\exp(-\gamma^2(\mathbb{D}_k + \mathbb{D}_l))$ 的概率下成立, 其中

$$s_1 = 1 - \eta m \cdot \frac{\sqrt{m} - 6CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}}, \quad s_2 = \frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}} \left(\frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{2\sqrt{m}} + 1 \right);$$

(b) 对于任意 $\mathbf{x} \in \mathbb{R}^m$, 不等式

$$\|[\mathbf{A}^T \mathbf{x}]_{\mathbb{L}}\|_2 \leq \|\mathbf{x}\|_2 (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_l))$$

至少在 $1 - 2\exp(-\gamma^2(\mathbb{D}_l))$ 的概率下成立.

引理 2.3 设 $\mathbb{D}_k = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y}\|_0 \leq k, \|\mathbf{y}\|_2 = 1\}$, 则有

- (a) $\gamma(\mathbb{D}_{2k}) \leq 2\gamma(\mathbb{D}_k)$;
- (b) $\gamma(\mathbb{D}_k + \mathbb{D}_k) \leq 2\gamma(\mathbb{D}_k)$;
- (c) 对于任意 $k \leq l$, $\mathbb{D}_k \subseteq \mathbb{D}_l$, $\gamma(\mathbb{D}_k) \leq \gamma(\mathbb{D}_l)$.

3 松弛最优阈值算法的数据 - 时间权衡

ROT 算法和 ROTP 算法可以自然地用于近似解决一些与压缩感知相关的稀疏重构问题, 如量化压缩感知. 本节对在次 Gauss 测量下的量化压缩感知中 ROT 算法和 ROTP 算法进行了数据 - 时间权衡分析.

3.1 量化压缩感知

在压缩感知中, 测量被定义具有无限精度. 然而, 在实际情形下, 测量数据通常表示为有限精度和数字化的形式^[26, 42]. 这迫使我们考虑量化模型 (1.2). 量化方案 $\mathcal{Q}(\cdot)$ 有几种形式 (参见文献 [6, 26, 27, 32, 36]), 其中两种常用的量化方案^[36] 如下所示:

(a) (均匀量化) 对于量化分辨率 $\Delta > 0$, 量化算子为 $\mathcal{Q}(\mathbf{x}) = \Delta(\lfloor \frac{\mathbf{x}}{\Delta} \rfloor + \frac{1}{2})$, 随机抖动 $\boldsymbol{\tau}$ 服从 $\tau_i \in \mathcal{U}(-\Delta/2, \Delta/2]$;

(b) (一比特量化) $\mathcal{Q}(\mathbf{x}) = \text{sign}(\mathbf{x})$, 其中 $\text{sign}(\cdot)$ 表示符号函数.

本文考虑带有随机均匀抖动的均匀量化. 在目标 $\mathbf{x}^*$ 是稀疏的条件下, QCS 中的感知模型可以表示为模型 (1.1). 相应的 ROT 算法和 ROTP 算法分别与算法 1 和 2 相同. 在给出上述算法的主要收敛结果之前, 首先重新构造模型 (1.2): $\mathbf{y} = \mathcal{Q}(\mathbf{A}\mathbf{x}^* + \mathbf{n} + \boldsymbol{\tau})$. 定义两个向量

$$\bar{\mathbf{y}} = \mathbf{A}\mathbf{x}^* + \mathbf{n}, \quad \mathbf{r} = \mathcal{Q}(\bar{\mathbf{y}} + \boldsymbol{\tau}) - \bar{\mathbf{y}} - \boldsymbol{\tau}, \quad (3.1)$$

则模型 (1.2) 可以重新表述为

$$\mathbf{y} = \mathcal{Q}(\bar{\mathbf{y}} + \boldsymbol{\tau}) = \mathbf{A}\mathbf{x}^* + \mathbf{r} + \boldsymbol{\tau} + \mathbf{n}. \quad (3.2)$$

根据文献 [28], 可以得出 (3.1) 中定义的 $\mathbf{r}$ 服从均匀分布, 即 $r_i \in \mathcal{U}(-\Delta/2, \Delta/2]$, $i = 1, \dots, m$. 因此, 具有均匀量化和随机抖动的 QCS 可以被视为具有更多噪声项的标准 CS. 我们可以得出以下关于 QCS 中 ROT 算法和 ROTP 算法的主要收敛结果.

3.2 收敛性结果

定理 3.1 设 $\mathbf{A}$ 是一个如引理 2.1 中所定义的 $m \times n$ 随机矩阵, $\mathcal{Q}(\cdot)$ 是均匀量化. 令 $\mathbf{y} := \mathcal{Q}(\mathbf{A}\mathbf{x}^* + \mathbf{n} + \boldsymbol{\tau})$ 表示 k 稀疏信号 $\mathbf{x}^* \in \mathbb{R}^n$ 的测量值, 其中 $\mathbf{n} \in \mathbb{R}^m$ 是噪声向量, $\boldsymbol{\tau}$ 是随机均匀抖动. 令 $\mathbf{r}$ 如 (3.1) 中所定义, η 是满足 $\eta m = 1$ 的步长. 我们有以下主要结果: 如果测量数 m 足够大, 并满足 $\sqrt{m} = \zeta_1 CK^2 \gamma(\mathbb{D}_{2k})$, 其中 $\zeta_1 \geq 51$, 则由 ROT 生成的任何序列 $\{\mathbf{x}^p\}$ 都能以以下误差在 $1 - c \exp(-\gamma^2(\mathbb{D}_{2k}))$ 的概率下近似于 $\mathbf{x}^*$:

$$\|\mathbf{x}^* - \mathbf{x}^p\|_2 \leq \rho_1^p \|\mathbf{x}^* - \mathbf{x}^0\|_2 + \frac{\omega_1}{1 - \rho_1} \|\mathbf{n} + \mathbf{r} + \boldsymbol{\tau}\|_2, \quad (3.3)$$

其中

$$\rho_1 = \frac{48(\zeta_1 + 1)}{\zeta_1(\zeta_1 - 2)} < 1, \quad \omega_1 = \frac{6\zeta_1^2 + 12\zeta_1 + 8}{\sqrt{m}\zeta_1(\zeta_1 - 2)}, \quad (3.4)$$

$c > 0$ 是一个常数.

证明 利用 $\mathbf{u}^p = \mathbf{x}^p + \eta \mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^p)$ 和 $\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{t}$, 其中 $\mathbf{t} = \mathbf{r} + \boldsymbol{\tau} + \mathbf{n}$, 我们得到

$$\mathbf{x}^* - \mathbf{u}^p = (\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p) - \eta \mathbf{A}^T \mathbf{t}. \quad (3.5)$$

令 $\mathbf{w}(\mathbf{u}^p)$ 成为 ROT 在第 p 次迭代时如下问题的解:

$$\min_{\mathbf{w}} \{ \|\mathbf{y} - \mathbf{A}(\mathbf{u}^p \odot \mathbf{w})\|_2^2 : \mathbf{e}^T \mathbf{w} = k, \mathbf{0} \leq \mathbf{w} \leq \mathbf{e} \}. \quad (3.6)$$

下一次迭代通过 $\mathbf{x}^{p+1} \in \mathcal{H}_k(\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p))$ 进行更新. 令 $\mathbb{S}^{p+1}$ 和 $\mathbb{S}$ 分别是 $\mathbf{x}^{p+1}$ 和 $\mathbf{x}^*$ 的支撑集. 由事实 2.3, 得到

$$\|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2 \leq \|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 + \|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2. \quad (3.7)$$

首先估计 (3.7) 不等式右边第一项. 因为 $\mathbf{x}_{\mathbb{S}^{p+1} \setminus \mathbb{S}}^* = \mathbf{0}$ 和 $\mathbf{w}_i^p \in [0, 1]$, $i = 1, \dots, n$, 所以有

$$\begin{aligned} \|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 &= \|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 \\ &= \|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 \\ &= \|[(\eta \mathbf{A}^T \mathbf{t} - (\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p)) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 \\ &\leq \eta \|[\mathbf{A}^T \mathbf{t}]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 + \|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p)]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2, \end{aligned} \quad (3.8)$$

其中第三个等式源于 (3.5). 令 $\mathbf{v} = \mathbf{x}^* - \mathbf{x}^p \in \mathbb{D}_{2k}$ 和 $\mathbb{L} = \mathbb{S}^{p+1} \setminus \mathbb{S}$, 使用引理 2.2, 可以估计上述不等式中的两个项. 注意到 $\mathbb{D}_l = \mathbb{D}_k$, 可得

$$\|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p)]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 \leq \max\{r_1, r_2\} \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 \quad (3.9)$$

和

$$\|[\mathbf{A}^T \mathbf{t}]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 \leq \|\mathbf{t}\|_2 (\sqrt{m} + 2CK^2 \gamma(\mathbb{D}_k)), \quad (3.10)$$

其中

$$r_1 = 1 - \eta m \cdot \frac{\sqrt{m} - 6CK^2 \gamma(\mathbb{D}_{2k} + \mathbb{D}_k)}{\sqrt{m}}, \quad r_2 = \frac{3CK^2 \gamma(\mathbb{D}_{2k} + \mathbb{D}_k)}{\sqrt{m}} \left(\frac{3CK^2 \gamma(\mathbb{D}_{2k} + \mathbb{D}_k)}{2\sqrt{m}} + 1 \right).$$

由引理 2.3(b) 和 2.3(c), 有 $\gamma(\mathbb{D}_{2k} + \mathbb{D}_k) \leq \gamma(\mathbb{D}_{2k} + \mathbb{D}_{2k}) \leq 2\gamma(\mathbb{D}_{2k})$. 在符号略有滥用的情形下, 常数 r_1 和 r_2 更新为

$$r_1 = 1 - \eta m \cdot \frac{\sqrt{m} - 12CK^2 \gamma(\mathbb{D}_{2k})}{\sqrt{m}}, \quad r_2 = \frac{6CK^2 \gamma(\mathbb{D}_{2k})}{\sqrt{m}} \left(\frac{3CK^2 \gamma(\mathbb{D}_{2k})}{\sqrt{m}} + 1 \right). \quad (3.11)$$

结合 (3.8)–(3.10), 可得

$$\|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \setminus \mathbb{S}}\|_2 \leq \max\{r_1, r_2\} \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 + \eta \|\mathbf{t}\|_2 (\sqrt{m} + 2CK^2 \gamma(\mathbb{D}_k)). \quad (3.12)$$

下一步估计 (3.7) 不等式右边的第二项. 令 $\mathbb{D} = \mathbb{D}_{2k}$, 使用引理 2.1, 可得不等式

$$\left\| \mathbf{A} \frac{[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}}{\|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2} \right\|_2 \geq \sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k}) \quad (3.13)$$

在至少 $1 - 2\exp(-\gamma^2(\mathbb{D}_{2k}))$ 的概率下成立. 于是得到

$$\begin{aligned} \|\mathbf{y} - \mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p)]\|_2 &= \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*] - \mathbf{t}\|_2 \\ &= \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}} + \mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}} - \mathbf{t}\|_2 \\ &\geq \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2 - \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}\|_2 - \|\mathbf{t}\|_2 \\ &\geq \|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2 (\sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k})) \\ &\quad - \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}\|_2 - \|\mathbf{t}\|_2, \end{aligned} \quad (3.14)$$

其中第一个等式使用了 $\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{t}$, 第一个不等式使用了三角不等式, 最后一个不等式源于 (3.13). 因此

$$\|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2 \leq \frac{1}{\sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k})} \{\|\mathbf{y} - \mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p)]\|_2 + T + \|\mathbf{t}\|_2\}, \quad (3.15)$$

其中 $T = \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}\|_2$. 令 $\hat{\mathbf{w}} \in \mathcal{W}^{(k)} = \{\mathbf{w} : \mathbf{e}^T \mathbf{w} = k, \mathbf{w} \in \{0, 1\}^n\}$ 并满足 $\mathbf{x}^* = \mathbf{x}^* \odot \hat{\mathbf{w}}$, 于是

$$\begin{aligned} \|\mathbf{y} - \mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p)]\|_2 &\leq \|\mathbf{y} - \mathbf{A}(\mathbf{u}^p \odot \hat{\mathbf{w}})\|_2 \\ &= \|\mathbf{A}\mathbf{x}^* + \mathbf{t} - \mathbf{A}(\mathbf{u}^p \odot \hat{\mathbf{w}})\|_2 \\ &= \|[\mathbf{A}(\mathbf{x}^* - \mathbf{u}^p) \odot \hat{\mathbf{w}}] + \mathbf{t}\|_2 \\ &\leq \|\mathbf{A}[(\mathbf{x}^* - \mathbf{u}^p) \odot \hat{\mathbf{w}}]\|_2 + \|\mathbf{t}\|_2 \\ &\leq (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|(\mathbf{x}^* - \mathbf{u}^p) \odot \hat{\mathbf{w}}\|_2 + \|\mathbf{t}\|_2, \end{aligned} \quad (3.16)$$

其中第一个不等式源于 $\mathbf{w}(\mathbf{u}^p)$ 的最优性, 第二个不等式使用了三角不等式, 并且最后一个不等式使用了引理 2.1 (令 $\mathbb{D} = \mathbb{D}_k$). 基于事实 $\mathbf{x}^* - \mathbf{u}^p = (\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p) - \eta \mathbf{A}^T \mathbf{t}$ 和引理 2.2, 可得

$$\begin{aligned} \|(\mathbf{x}^* - \mathbf{u}^p) \odot \hat{\mathbf{w}}\|_2 &\leq \|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p)] \odot \hat{\mathbf{w}}\|_2 + \eta \|[\mathbf{A}^T \mathbf{t}] \odot \hat{\mathbf{w}}\|_2 \\ &\leq \max\{r_1, r_2\} \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 + \eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2, \end{aligned} \quad (3.17)$$

其中 r_1 和 r_2 在 (3.11) 给定. 结合 (3.16) 和 (3.17) 可得

$$\begin{aligned} \|\mathbf{y} - \mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p)]\|_2 &\leq (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \max\{r_1, r_2\} \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 \\ &\quad + [\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 + 1] \|\mathbf{t}\|_2. \end{aligned} \quad (3.18)$$

在证明的最后部分, 我们估计 (3.15) 中的 T . 由于 $\mathbf{x}^*_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}} = 0$, 所以 T 具有以下形式:

$$T = \|\mathbf{A}[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}\|_2 = \|\mathbf{A}[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}\|_2.$$

令 $\mathbb{T}_i, i = 1, 2, \dots, \bar{n}$ 成为 $\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}$ 的不相交子集, 并满足以下性质:

- 对于所有 $i = 1, 2, \dots, \bar{n}$, 有 $|\mathbb{T}_i| = k$ 且 $|\mathbb{T}_{\bar{n}}| = \bar{k} < k$;

• $\mathbb{T}_i$ 是 $\mathbf{w}(\mathbf{u}^p)_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}$ 中第 i 组的 k 个最大分量的索引集.

显然, $|\mathbb{S}^{p+1} \cup \mathbb{S}| = (\bar{n} - 1)k + \bar{k}$. 根据这样的划分, $[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}}$ 可以写成

$$[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\overline{\mathbb{S}^{p+1} \cup \mathbb{S}}} = \sum_{i=1}^{\bar{n}} [(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}.$$

因此 T 可以简化为

$$\begin{aligned} T &= \left\| \sum_i^{\bar{n}} \mathbf{A}[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i} \right\|_2 \\ &\leq \sum_i^{\bar{n}} \|\mathbf{A}[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2 \\ &\leq (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \sum_i^{\bar{n}} \|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2. \end{aligned} \quad (3.19)$$

利用 (3.5) 和三角不等式, 对于每一项 $\|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2$, 有

$$\begin{aligned} \|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2 &= \|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p) \odot \mathbf{w}(\mathbf{u}^p) - \eta \mathbf{A}^T \mathbf{t} \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2 \\ &\leq \|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2 + \eta \|\mathbf{A}^T \mathbf{t} \odot \mathbf{w}(\mathbf{u}^p)\|_{\mathbb{T}_i}. \end{aligned}$$

定义 $\alpha_i \triangleq \max\{(\mathbf{w}(\mathbf{u}^p))_j : j \in \mathbb{T}_i\}$, 于是

$$\begin{aligned} \|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2 &\leq \alpha_i \|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^p)]_{\mathbb{T}_i}\|_2 + \eta \alpha_i \|\mathbf{A}^T \mathbf{t}\|_{\mathbb{T}_i}, \\ &\leq \alpha_i \max\{r_1, r_2\} \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 + \eta \alpha_i (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2, \end{aligned}$$

其中 r_1 和 r_2 在 (3.11) 给定, 因此

$$\sum_i^{\bar{n}} \|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{T}_i}\|_2 \leq \sum_i^{\bar{n}} \alpha_i \max\{r_1, r_2\} \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 + \eta \sum_i^{\bar{n}} \alpha_i (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2.$$

利用事实 2.4 ($\mathbf{w} = \mathbf{w}(\mathbf{u}^p)$, $\mathbb{S} = \mathbb{S}^{p+1} \cup \mathbb{S}$) 和 $\mathbb{S}_i = \mathbb{T}_i$ 可以得到

$$T \leq 2 \max\{r_1, r_2\} (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \cdot \|\mathbf{x}^* - \mathbf{x}^p\|_2 + 2\eta (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 \|\mathbf{t}\|_2. \quad (3.20)$$

结合 (3.20), (3.15) 和 (3.18), 可得

$$\begin{aligned} \|[\mathbf{u}^p \odot \mathbf{w}(\mathbf{u}^p) - \mathbf{x}^*]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2 &\leq \|[(\mathbf{u}^p - \mathbf{x}^*) \odot \mathbf{w}(\mathbf{u}^p)]_{\mathbb{S}^{p+1} \cup \mathbb{S}}\|_2 \\ &\leq \frac{3 \max\{r_1, r_2\} (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))}{\sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k})} \|\mathbf{x}^* - \mathbf{x}^p\|_2 \\ &\quad + \frac{3\eta (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 + 2}{\sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k})} \|\mathbf{t}\|_2. \end{aligned}$$

结合上述结果、(3.7) 和 (3.12), 可得

$$\|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2 \leq c_1 \max\{r_1, r_2\} \|\mathbf{x}^* - \mathbf{x}^p\|_2 + c_2 \|\mathbf{t}\|_2, \quad (3.21)$$

其中

$$c_1 = \frac{4\sqrt{m} + 4CK^2\gamma(\mathbb{D}_{2k})}{\sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k})}, \quad c_2 = \frac{3\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 + 2}{\sqrt{m} - 2CK^2\gamma(\mathbb{D}_{2k})} + \eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)).$$

令 $\rho_1 = c_1 \max\{r_1, r_2\}$. 当测量的数量满足

$$\sqrt{m} = \zeta_1 CK^2\gamma(\mathbb{D}_{2k})$$

时, 其中 $\zeta_1 \geq 51$, 可得

$$\rho_1 = \frac{48(\zeta_1 + 1)}{\zeta_1(\zeta_1 - 2)} < 1.$$

结合 $\gamma(\mathbb{D}_k) \leq \gamma(\mathbb{D}_{2k})$ 可得

$$c_2 < \frac{6\zeta_1^2 + 12\zeta_1 + 8}{\sqrt{m}\zeta_1(\zeta_1 - 2)} \triangleq \omega_1.$$

利用 $\sum_{j=1}^{\infty} s^j = \frac{1}{1-s}$, 结果 (3.21) 可以简化为

$$\|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2 \leq \rho_1^{p+1} \|\mathbf{x}^* - \mathbf{x}^0\|_2 + \frac{\omega_1}{1 - \rho_1} \|\mathbf{t}\|_2. \quad (3.22)$$

证毕. □

定理 3.2 假设信号模型定义如定理 3.1 中所述. 令 η 为 ROTP 算法中的步长, 并满足 $\eta m = 1$. 如果测量数量满足

$$\sqrt{m} = \zeta_2 CK^2\gamma(\mathbb{D}_{2k}),$$

并且 $\zeta_2 \geq 53$, 则由 ROTP 生成的任何序列 $\{\mathbf{x}^p\}$ 都能以概率 $1 - c' \exp(-\gamma^2(\mathbb{D}_{2k}))$ 满足下列不等式:

$$\|\mathbf{x}^* - \mathbf{x}^p\|_2 \leq \rho_2^p \|\mathbf{x}^* - \mathbf{x}^0\|_2 + \frac{\omega_2}{1 - \rho_2} \|\mathbf{n} + \mathbf{r} + \boldsymbol{\tau}\|_2, \quad (3.23)$$

其中 $c' > 0$ 是一个常数,

$$\rho_2 = \frac{48(\zeta_2 + 1)}{\sqrt{\zeta_2^2 - 144}(\zeta_2 - 2)} < 1, \quad \omega_2 = \frac{6\zeta_2^2 + 12\zeta_2 + 8}{\sqrt{m}(\zeta_2^2 - 144)(\zeta_2 - 2)} + \frac{\zeta_2 + 2}{\sqrt{m}(\zeta_2 - 12)}. \quad (3.24)$$

证明 由于 ROTP 算法和 ROT 算法有相同的步骤 3-5, 因此可以简化证明, 利用 (3.21) 来估计 $\|\mathbf{x}^* - \mathbf{v}\|_2$:

$$\|\mathbf{x}^* - \mathbf{v}\|_2 \leq \rho_1 \|\mathbf{x}^* - \mathbf{x}^p\|_2 + c_2 \|\mathbf{t}\|_2. \quad (3.25)$$

接下来, 将 $\|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2$ 与 $\|\mathbf{x}^* - \mathbf{v}\|_2$ 联系起来. 由于 $\mathbf{x}^{p+1}$ 是下面问题的解:

$$\mathbf{x}^{p+1} = \arg \min_{\mathbf{x}} \{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2, \text{supp}(\mathbf{x}) \subseteq \mathbb{S}^{p+1}\},$$

可得 $[\mathbf{A}^T(\mathbf{y} - \mathbf{A}\mathbf{x}^{p+1})]_{\mathbb{S}^{p+1}} = 0$. 结合 $\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{t}$, 可得

$$[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^{p+1})]_{\mathbb{S}^{p+1}} = (\mathbf{x}^* - \mathbf{x}^{p+1})_{\mathbb{S}^{p+1}} + \eta(\mathbf{A}^T \mathbf{t})_{\mathbb{S}^{p+1}}.$$

利用三角不等式, 得到

$$\|(\mathbf{x}^* - \mathbf{x}^{p+1})_{\mathbb{S}^{p+1}}\|_2 \leq \|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^{p+1})]_{\mathbb{S}^{p+1}}\|_2 + \eta \|(\mathbf{A}^T \mathbf{t})_{\mathbb{S}^{p+1}}\|_2.$$

由于 $\mathbb{S}^{p+1}$ 的基数为 k , 且 $(\mathbf{x}^* - \mathbf{x}^{p+1})$ 是一个 $2k$ -稀疏向量, 所以通过引理 2.2 可推导出

$$\|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})(\mathbf{x}^* - \mathbf{x}^{p+1})]_{\mathbb{S}^{p+1}}\|_2 \leq \max\{r_1, r_2\} \|(\mathbf{x}^* - \mathbf{x}^{p+1})\|_2$$

和

$$\|(\mathbf{A}^T \mathbf{t})_{\mathbb{S}^{p+1}}\|_2 \leq (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2.$$

所以

$$\|(\mathbf{x}^* - \mathbf{x}^{p+1})_{\mathbb{S}^{p+1}}\|_2 \leq \max\{r_1, r_2\} \|(\mathbf{x}^* - \mathbf{x}^{p+1})\|_2 + \eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2.$$

因为 $(\mathbf{x}^{p+1})_{\overline{\mathbb{S}^{p+1}}} = 0$ 和 $\mathbf{v}_{\overline{\mathbb{S}^{p+1}}} = 0$, 所以

$$\begin{aligned} \|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2^2 &= \|(\mathbf{x}^* - \mathbf{x}^{p+1})_{\mathbb{S}^{p+1}}\|_2^2 + \|(\mathbf{x}^* - \mathbf{x}^{p+1})_{\overline{\mathbb{S}^{p+1}}}\|_2^2 \\ &= \|(\mathbf{x}^* - \mathbf{x}^{p+1})_{\mathbb{S}^{p+1}}\|_2^2 + \|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2^2 \\ &\leq (\max\{r_1, r_2\})^2 \|(\mathbf{x}^* - \mathbf{x}^{p+1})\|_2^2 + \eta^2(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 \|\mathbf{t}\|_2^2 \\ &\quad + 2\eta \max\{r_1, r_2\}(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \cdot \|(\mathbf{x}^* - \mathbf{x}^{p+1})\|_2 \|\mathbf{t}\|_2 \\ &\quad + \|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2^2, \end{aligned}$$

等同于

$$\begin{aligned} [1 - (\max\{r_1, r_2\})^2] \|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2^2 &\leq 2\eta \max\{r_1, r_2\}(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|(\mathbf{x}^* - \mathbf{x}^{p+1})\|_2 \|\mathbf{t}\|_2 \\ &\quad + \eta^2(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 \|\mathbf{t}\|_2^2 + \|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2^2. \end{aligned} \quad (3.26)$$

利用二次方程根的公式, 可以得到 $\|(\mathbf{x}^* - \mathbf{x}^{p+1})\|_2$ 的上界如下:

$$\begin{aligned} \|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2 &\leq \frac{2\eta \max\{r_1, r_2\}(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2}{2[1 - (\max\{r_1, r_2\})^2]} \\ &\quad + \frac{\sqrt{4\eta^2(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))^2 \|\mathbf{t}\|_2^2 + 4[1 - (\max\{r_1, r_2\})^2] \|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2^2}}{2[1 - (\max\{r_1, r_2\})^2]} \\ &\leq \frac{2\eta \max\{r_1, r_2\}(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2}{2[1 - (\max\{r_1, r_2\})^2]} \\ &\quad + \frac{2\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2 + 2\sqrt{1 - (\max\{r_1, r_2\})^2} \|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2}{2[1 - (\max\{r_1, r_2\})^2]} \\ &= \frac{\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2}{1 - \max\{r_1, r_2\}} + \frac{\|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2}{\sqrt{1 - (\max\{r_1, r_2\})^2}} \\ &\leq \frac{\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2}{1 - \max\{r_1, r_2\}} + \frac{\|\mathbf{x}^* - \mathbf{v}\|_2}{\sqrt{1 - (\max\{r_1, r_2\})^2}}, \end{aligned} \quad (3.27)$$

其中第二个不等式使用了如下事实: 对于任意 $a, b \geq 0$, $a^2 + b^2 \leq (a + b)^2$ 成立, 并且最后一个不等式源于 $\|(\mathbf{x}^* - \mathbf{v})_{\overline{\mathbb{S}^{p+1}}}\|_2 \leq \|\mathbf{x}^* - \mathbf{v}\|_2$. 结合 (3.25), 可以完成证明

$$\|\mathbf{x}^* - \mathbf{x}^{p+1}\|_2 \leq \frac{\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k)) \|\mathbf{t}\|_2}{1 - \max\{r_1, r_2\}} + \frac{\rho_1 \|\mathbf{x}^* - \mathbf{x}^p\|_2 + c_2 \|\mathbf{t}\|_2}{\sqrt{1 - (\max\{r_1, r_2\})^2}}$$

$$= \rho_2 \|\mathbf{x}^* - \mathbf{x}^p\|_2 + c_3 \|\mathbf{t}\|_2,$$

其中

$$\rho_2 = \frac{\rho_1}{\sqrt{1 - (\max\{r_1, r_2\})^2}}, \quad c_3 = \frac{c_2}{\sqrt{1 - (\max\{r_1, r_2\})^2}} + \frac{\eta(\sqrt{m} + 2CK^2\gamma(\mathbb{D}_k))}{1 - \max\{r_1, r_2\}}.$$

令 $\rho_2 = \rho_1 / \sqrt{1 - (\max\{r_1, r_2\})^2}$. 当测量的数量满足 $\sqrt{m} = \zeta_2 CK^2 \gamma(\mathbb{D}_{2k})$, $\zeta_2 \geq 53$ 时, 有

$$\rho_2 = \frac{48(\zeta_2 + 1)}{\sqrt{\zeta_2^2 - 144}(\zeta_2 - 2)} < 1.$$

结合事实 $\gamma(\mathbb{D}_k) \leq \gamma(\mathbb{D}_{2k})$, 以下不等式成立:

$$c_3 < \frac{6\zeta_2^2 + 12\zeta_2 + 8}{\sqrt{m(\zeta_2^2 - 144)}(\zeta_2 - 2)} + \frac{\zeta_2 + 2}{\sqrt{m(\zeta_2 - 12)}} \triangleq \omega_2.$$

与定理 3.1 的证明类似, 我们得到了所需要的结果 (3.23). $\square$

我们在注 3.1 中简要讨论了上述主要结果的数据 - 时间权衡, 并对成功恢复所需的测量数量的阶数在注 3.2 中进行了讨论. 另外, 在压缩感知中, 算法的数据 - 时间权衡结果在注 3.4 解释.

注 3.1 如果 m 大于某个值, 则收敛速度 ρ_1 或 ρ_2 小于 1; 否则, 算法可能会发散. 此外, 显然, 值 c_1, r_1 和 r_2 随着 m 的增加而减小. 因此, ρ_1 或 ρ_2 减小, 从而算法有快速的收敛速度. 因此可以得出结论, 当测量数量足够大时, 测量数量越大, 收敛速度越快, 在固定精度下所需的迭代次数越少. 例如, 令步长 η 等于 $1/m$. 在这种情形下, 对于 ROT 算法, 满足 $\rho_1 < 1$ 的收敛速度由 $\sqrt{m} > C'CK^2\gamma(\mathbb{D}_{2k})$ 确保, 其中 $C' \approx 50.943$ 是方程 $t^2 - 50t - 48 = 0$ 的正实根, 当 $\sqrt{m} \leq C'CK^2\gamma(\mathbb{D}_{2k})$ 时, 收敛速度 $\rho_1 \geq 1$, 导致 ROT 算法的发散. 类似地, 在 ROTP 算法中, 满足 $\rho_2 < 1$ 的收敛速度由 $\sqrt{m} > C''CK^2\gamma(\mathbb{D}_{2k})$ 确保, 其中 $C'' \approx 52.2614$ 是方程 $t^4 - 4t^3 - 2444t^2 - 4032t - 2880 = 0$ 的正实根.

注 3.2 需要指出的是, 根据方程 (2.1), Gauss 复杂度 $\gamma(\mathbb{D}_k)$ 和 Gauss 宽度 $w(\mathbb{D}_k)$ 具有相同的阶. 结合事实 2.2, 可以得知 $\gamma(\mathbb{D}_k)$ 的阶数为 $\sqrt{k \log(n/k)}$, 其中 n 和 k 分别是目标信号的维数和稀疏度. 根据定理 3.1、3.2 和注 3.1 中的主要结果, 不难看出 ROT 或 ROTP 中成功恢复所需的测量数量与 $k \log(n/k)$ 具有相同的阶.

注 3.3 定理 3.1 保证了序列 $\mathbf{x}^p$ 收敛到以 $\mathbf{x}^*$ 为中心、半径为 $c\|\mathbf{n} + \mathbf{r} + \boldsymbol{\tau}\|_2$ 的邻域里, 其中 ROT 中 $c = \frac{\omega_1}{1 - \rho_1}$, ROTP 中 $c = \frac{\omega_2}{1 - \rho_2}$. 特别地, 当没有噪声时 ($\mathbf{n} = \mathbf{0}$), 半径变为 $c\sqrt{m}\Delta$.

注 3.4 利用定理 3.1 和 3.2, 将 (3.3) 和 (3.23) 中的 $\|\mathbf{n} + \mathbf{r} + \boldsymbol{\tau}\|_2$ 替代为 $\|\mathbf{n}\|_2$, 得到

$$\|\mathbf{x}^* - \mathbf{x}^p\|_2 \leq \rho_1^p \|\mathbf{x}^* - \mathbf{x}^0\|_2 + \frac{\omega_1}{1 - \rho_1} \|\mathbf{n}\|_2, \quad (3.28)$$

$$\|\mathbf{x}^* - \mathbf{x}^p\|_2 \leq \rho_2^p \|\mathbf{x}^* - \mathbf{x}^0\|_2 + \frac{\omega_2}{1 - \rho_2} \|\mathbf{t}\|_2. \quad (3.29)$$

我们也可以得到在压缩感知中 ROT 算法和 ROTP 算法的主要收敛结果. 注意, 在无噪声的情形下, ROT 算法可以实现线性收敛率. 类似地, 当满足一定条件时, ROTP 算法也可以实现线性收敛率.

4 数值分析

本节进行数值实验, 展示了 ROT 算法和 ROTP 算法的数据 - 时间权衡. 我们绘制了压缩感知问题 (1.1)、量化压缩感知问题 (1.2) 中 ROT 算法和 ROTP 算法的相变图.

对于所有问题, 目标信号 $\mathbf{x}^*$ 是一个 k 稀疏的 n 维向量, 其非零元素服从标准 Gauss 分布. 感知矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 是一个 Bernoulli 矩阵, 其元素是独立对称 Bernoulli 变量. 设 ε 为停止准则的容差值. 对于一组 m 和 p , 我们进行了 50 次 Monte Carlo 实验, 并统计了满足以下条件的成功实验次数: 对于问题 (1.1), 满足 $\|\mathbf{x}^* - \mathbf{x}^p\|_2 / \|\mathbf{x}^*\|_2 \leq \varepsilon$. 然后, 我们计算了每对的成功概率, 并得到了相变结果.

我们将维数设置为 $n = 64$. 初始值 $\mathbf{x}^0$ 被设置为 $\mathbf{0}$, $\mathbf{x}^*$ 的稀疏度被设为 $k = 3$, 量化分辨率被设为 $\Delta = 0.05$. 在压缩感知问题中, 用于停止准则的容差值被设为 $\varepsilon = 0.01$; 而在量化压缩感知中, 为了减轻量化的影响, 容差值被设为 $\varepsilon = 0.1$. 如定理 3.1 和 3.2 中所述, 将步长设为 $\eta = 1/m$.

松弛最优阈值算法 (ROT 和 ROTP) 在压缩感知、量化压缩感知问题中在设定环境下的相变曲线如图 1 所示. 值得注意的是, 当样本数量大约超过 20 时, ROTP 算法用于恢复目标信号所需的迭代次数比 ROT 算法少, 这意味着 ROTP 算法在恢复稀疏信号方面比 ROT 算法性能更好. 值得注意的是, ROTP 算法对于足够恢复目标信号所需的迭代次数非常少 (在所呈现的模拟中, 一两次迭代就足够了). 两个稀疏重建问题中的两种算法都说明了我们在以上备注中讨论的“数据 - 时间权衡”. 曲线还表明, 当测量数量小于一定数量时, 算法无法恢复目标信号. 此外, 当测量数量大于一定数量时, 信

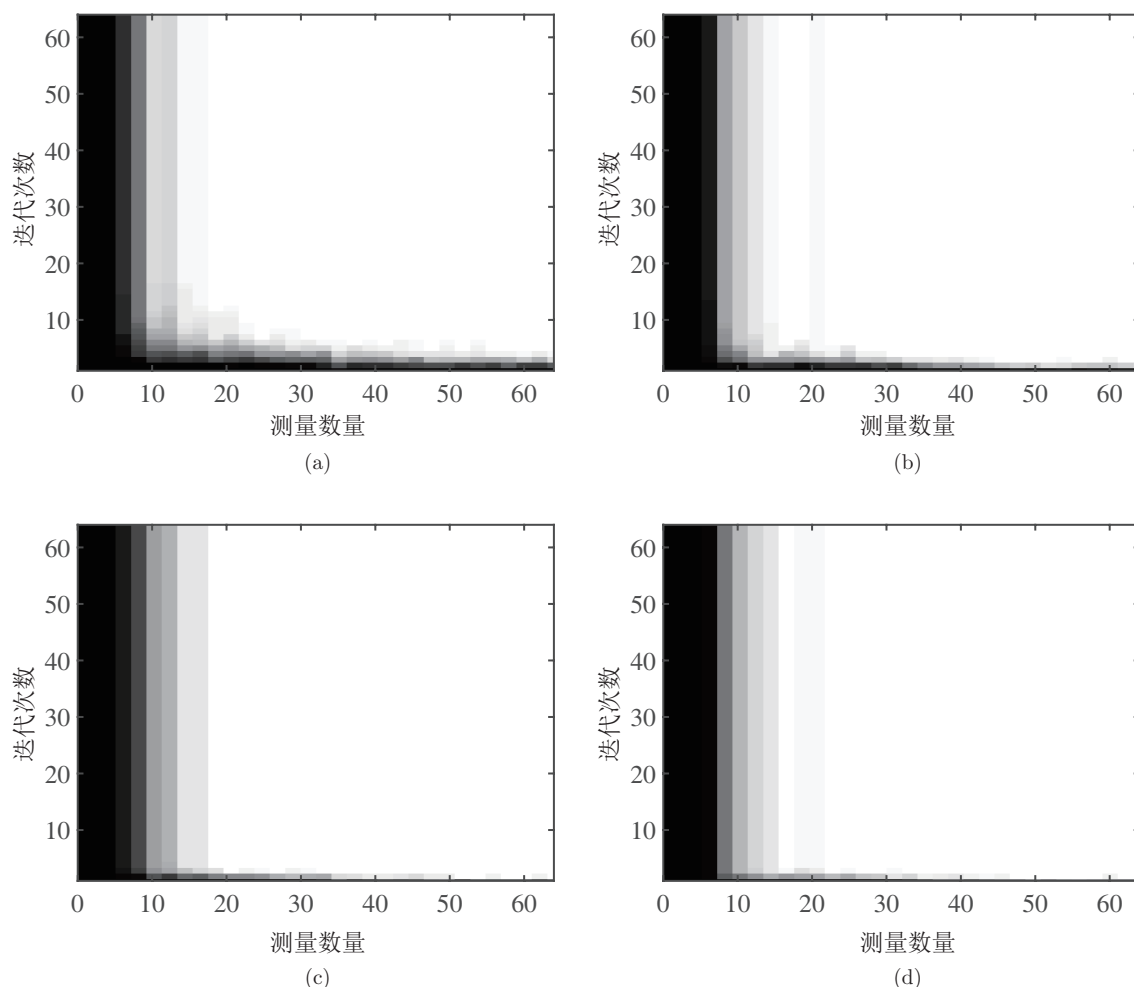


图 1 在压缩感知和量化压缩感知中松弛最优阈值算法的相变曲线. (a) CS: ROT; (b) QCS: ROT; (c) CS: ROTP; (d) QCS: ROTP. 每个点的亮度表示成功重构的概率, 范围从黑色 (0%) 到白色 (100%)

号成功恢复所需的迭代次数随着测量数量的增加而减少. 例如, 图 1(c) 显示, 当 m 大于 18 时, ROTP 算法中成功恢复所需的迭代次数与测量数量呈负相关. 数字模拟中观察到的现象与我们在定理中的理论结果相符.

5 结论

本文提供了针对松弛最优阈值算法在典型稀疏重建问题中 (包括压缩感知、量化压缩感知) 的数据 - 时间权衡分析. 理论结果表明, 为了保证松弛最优阈值算法的收敛性, 测量数量应适当地增加. 此外, 当测量数量大于一定数量时, 成功重建信号所需的迭代次数与测量数量呈负相关. 我们提供了数字模拟实验以验证我们理论的正确性. 对于未来的工作, 我们将松弛最优阈值算法的数据 - 时间权衡分析扩展到其他复杂的稀疏优化问题, 如损坏压缩感知问题, 或者将该分析从稀疏结构扩展到其他类型的结构, 如低秩结构或低秩加稀疏结构.

参考文献

- 1 Agarwal A, Negahban S, Wainwright M. Fast global convergence rates of gradient methods for high-dimensional statistical recovery. *Ann Statist*, 2012, 40: 2452–2482
- 2 Aubry A, Carotenuto V, De M A, et al. Multi-snapshot spectrum sensing for cognitive radar via block-sparsity exploitation. *IEEE Trans Signal Process*, 2019, 67: 1396–1406
- 3 Blanchard J D, Tanner J, Wei K. CGIHT: Conjugate gradient iterative hard thresholding for compressed sensing and matrix completion. *Inf Inference*, 2015, 4: 289–327
- 4 Blumensath T, Davies M E. Iterative hard thresholding for compressed sensing. *Appl Comput Harmon Anal*, 2009, 27: 265–274
- 5 Bouchot J L, Foucart S, Hitczenko P. Hard thresholding pursuit algorithms: Number of iterations. *Appl Comput Harmon Anal*, 2016, 41: 412–435
- 6 Boufounos P T. Universal rate-efficient scalar quantization. *IEEE Trans Inform Theory*, 2012, 58: 1861–1872
- 7 Bruer J J, Tropp J A, Cevher V, et al. Designing statistical estimators that balance sample size, risk, and computational cost. *IEEE J Sel Top Signal Process*, 2015, 9: 612–624
- 8 Candès E J. Compressive sampling. *Proc Int Congr Math*, 2006, 3: 1433–1452
- 9 Candès E J, Tao T. Decoding by linear programming. *IEEE Trans Inform Theory*, 2005, 51: 4203–4215
- 10 Candès E J, Wakin M B, Boyd S P. Enhancing sparsity by reweighted ℓ_1 minimization. *J Fourier Anal Appl*, 2008, 14: 877–905
- 11 Chandrasekaran V, Jordan M. Computational and statistical tradeoffs via convex relaxation. *Proc Natl Acad Sci USA*, 2013, 110: 1181–1190
- 12 Chen J, Liu Y. Data-time tradeoffs for corrupted sensing. *IEEE Signal Process Lett*, 2018, 25: 941–945
- 13 Chen J, Liu Y. Stable recovery of structured signals from corrupted sub-gaussian measurements. *IEEE Trans Inform Theory*, 2019, 65: 2976–2994
- 14 Chen J, Xu P, Wang L, et al. Covariate adjusted precision matrix estimation via nonconvex optimization. In: *Proceedings of the 35th International Conference on Machine Learning, ICML 2018*. Stockholm: International Machine Learning Society (IMLS), 2018, 922–931
- 15 Chen S S, Donoho D L, Saunders M A. Atomic decomposition by basis pursuit. *SIAM Rev*, 2001, 43: 129–159
- 16 Dai W, Milenkovic O. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans Inform Theory*, 2009, 55: 2230–2249
- 17 Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm Pure Appl Math*, 2004, 57: 1413–1457
- 18 Donoho D L. De-noising by soft-thresholding. *IEEE Trans Inform Theory*, 1995, 41: 613–627
- 19 Donoho D L. Compressed sensing. *IEEE Trans Inform Theory*, 2006, 52: 1289–1306
- 20 Donoho D L, Johnstone I M. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 1994, 81: 425–455
- 21 Elhamifar E, Vidal R. Sparse subspace clustering: Algorithm, theory, and applications. *IEEE Trans Pattern Anal Mach Intell*, 2013, 35: 2765–2781
- 22 Fornasier M, Rauhut H. Iterative thresholding algorithms. *Appl Comput Harmon Anal*, 2008, 25: 187–208

- 23 Foucart S. Hard thresholding pursuit: An algorithm for compressive sensing. *SIAM J Numer Anal*, 2011, 49: 2543–2563
- 24 Foucart S, Rauhut H. An invitation to compressive sensing. In: *Mathematics and Statistics*. New York: Birkhäuser, 2013, 1–39
- 25 Garg R, Khandekar R. Gradient descent with sparsification: An iterative algorithm for sparse recovery with restricted isometry property. In: *Proceedings of the 26th Annual International Conference on Machine Learning*. New York: ACM, 2009, 337–344
- 26 Güntürk C, Lammers M, Powell A, et al. Sobolev duals for random frames and $\Sigma\Delta$ quantization of compressed sensing measurements. *Found Comput Math*, 2013, 13: 1–36
- 27 Kamilov U S, Goyal V K, Rangan S. Message-passing de-quantization with applications to compressed sensing. *IEEE Trans Signal Process*, 2012, 60: 6270–6281
- 28 Lipshitz P, Wannamaker A, Vanderkooy J. Quantization and dither: A theoretical survey. *J Audio Eng Soc*, 1992, 40: 355–375
- 29 Mallat S G, Zhang Z F. Matching pursuits with time-frequency dictionaries. *IEEE Trans Signal Process*, 1993, 41: 3397–3415
- 30 Natarajan B. Sparse approximate solutions to linear systems. *SIAM J Comput*, 1995, 24: 227–234
- 31 Needell D, Tropp J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl Comput Harmon Anal*, 2009, 26: 301–321
- 32 Nguyen H Q, Goyal V K, Varshney L R. Frame permutation quantization. *Appl Comput Harmon Anal*, 2011, 31: 74–97
- 33 Oymak S, Recht B, Soltanolkotabi M. Sharp time-data tradeoffs for linear inverse problems. *IEEE Trans Inform Theory*, 2017, 64: 4129–4158
- 34 Suggala A, Bhatia K, Ravikumar P, et al. Adaptive hard thresholding for near-optimal consistent robust regression. In: *Proceedings of the Thirty-Second Conference on Learning Theory*. San Diego: JMLR, 2019, 2892–2897
- 35 Tan X, Roberts W, Li J, et al. Sparse learning via iterative minimization with application to MIMO radar imaging. *IEEE Trans Signal Process*, 2011, 59: 1088–1101
- 36 Thrampoulidis C, Rawat A S. The generalized Lasso for sub-Gaussian measurements with dithered quantization. *IEEE Trans Inform Theory*, 2020, 66: 2487–2500
- 37 Vershynin R. Four lectures on probabilistic methods for data science. arXiv:1612.06661, 2016
- 38 Vershynin R. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge: Cambridge Univ Press, 2018
- 39 Voronin S, Woerdeman H J. A new iterative firm-thresholding algorithm for inverse problems with sparsity constraints. *Appl Comput Harmon Anal*, 2013, 35: 151–164
- 40 Wainwright M. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge: Cambridge Univ Press, 2019
- 41 Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell*, 2009, 31: 210–227
- 42 Xu C, Jacques L. Quantized compressive sensing with RIP matrices: The benefit of dithering. *Inf Inference*, 2020, 9: 543–586
- 43 Zhao Y B. *Sparse Optimization Theory and Methods*. Boca Raton: CRC Press, 2018
- 44 Zhao Y B. Optimal k -thresholding algorithms for sparse optimization problems. *SIAM J Optim*, 2020, 30: 31–55
- 45 Zhao Y B, Li D. Reweighted ℓ_1 -minimization for sparse solutions to underdetermined linear systems. *SIAM J Optim*, 2012, 22: 1065–1088

附录 A 引理的证明

附录 A.1 引理 2.2 的证明

首先证明引理 2.2(a). 设

$$\mu = \frac{[(I - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}]_{\mathbb{L}}}{\|[(I - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}]_{\mathbb{L}}\|_2} \in \mathbb{D}_l, \quad \nu = \frac{\mathbf{v}}{\|\mathbf{v}\|_2} \in \mathbb{D}_k,$$

则可得

$$\|[(I - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}]_{\mathbb{L}}\|_2 = \langle [(I - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}]_{\mathbb{L}}, \mu \rangle = \langle (I - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}, \mu \rangle = \|\mathbf{v}\|_2 \langle (I - \eta \mathbf{A}^T \mathbf{A})\nu, \mu \rangle.$$

因为 $\langle (I - \eta A^T A)\nu, \mu \rangle$, 所以有

$$\begin{aligned}\langle (I - \eta A^T A)\nu, \mu \rangle &= \langle \nu, \mu \rangle - \eta \langle A\nu, A\mu \rangle \\ &= \frac{1}{4}(\|\nu + \mu\|_2^2 - \eta\|A(\nu + \mu)\|_2^2 - \|\nu - \mu\|_2^2 + \eta\|A(\nu - \mu)\|_2^2)\end{aligned}$$

利用 $\mathbb{D}_l = -\mathbb{D}_l$, 得到 $\nu + \mu \in \mathbb{D}_k + \mathbb{D}_l$, $\nu - \mu \in \mathbb{D}_k - \mathbb{D}_l = \mathbb{D}_k + \mathbb{D}_l$. 令 $\mathbb{D} = \mathbb{D}_k + \mathbb{D}_l$, $t = \gamma(\mathbb{D}_k + \mathbb{D}_l)$ 和 $\text{rad}(\mathbb{D}_k + \mathbb{D}_l) \leq 2$, 代入事实 2.1, 可得到下面两个不等式:

$$\begin{aligned}\|A(\nu + \mu)\|_2^2 &\geq (\max\{\sqrt{m}\|\nu + \mu\|_2 - 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l), 0\})^2, \\ \|A(\nu - \mu)\|_2^2 &\leq [\sqrt{m}\|\nu - \mu\|_2 + 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)]^2\end{aligned}$$

都在至少在 $1 - 2\exp(-\gamma^2(\mathbb{D}_k + \mathbb{D}_l))$ 的概率下成立.

接下来, 将问题分成两种情形.

情形 I 当 $\sqrt{m}\|\nu + \mu\|_2 - 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l) \geq 0$ 时, 得到

$$\begin{aligned}\langle (I - \eta A^T A)\nu, \mu \rangle &= \langle \nu, \mu \rangle - \eta \langle A\nu, A\mu \rangle \\ &= \frac{1}{4}[\|\nu + \mu\|_2^2 - \eta\|A(\nu + \mu)\|_2^2 - \|\nu - \mu\|_2^2 + \eta\|A(\nu - \mu)\|_2^2] \\ &\leq \frac{1}{4}[\|\nu + \mu\|_2^2 - \eta(\sqrt{m}\|\nu + \mu\|_2 - 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l))^2 - \|\nu - \mu\|_2^2 \\ &\quad + \eta(\sqrt{m}\|\nu - \mu\|_2 + 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l))^2] \\ &= \frac{1}{4}(1 - \eta m)(\|\nu + \mu\|_2^2 - \|\nu - \mu\|_2^2) \\ &\quad + \frac{3}{2}\eta\sqrt{m}CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)(\|\nu + \mu\|_2 + \|\nu - \mu\|_2) \\ &= (1 - \eta m)\langle \nu, \mu \rangle + \frac{3}{2}\eta\sqrt{m}CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)(\|\nu + \mu\|_2 + \|\nu - \mu\|_2) \\ &\leq 1 - \eta m + 6\eta\sqrt{m}CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l) \\ &= 1 - \eta m \cdot \frac{\sqrt{m} - 6CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}},\end{aligned}\tag{A.1}$$

其中最后一个不等式使用了 $\|\nu + \mu\|_2 \leq 2$ 和 $\|\nu - \mu\|_2 \leq 2$.

情形 II 当 $\sqrt{m}\|\nu + \mu\|_2 - 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l) < 0$ 时, 有

$$\begin{aligned}\langle (I - \eta A^T A)\nu, \mu \rangle &= \langle \nu, \mu \rangle - \eta \langle A\nu, A\mu \rangle \\ &= \frac{1}{4}[\|\nu + \mu\|_2^2 - \eta\|A(\nu + \mu)\|_2^2 - \|\nu - \mu\|_2^2 + \eta\|A(\nu - \mu)\|_2^2] \\ &\leq \frac{1}{4}[\|\nu + \mu\|_2^2 - \|\nu - \mu\|_2^2 + \eta(\sqrt{m}\|\nu - \mu\|_2 + 3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l))^2] \\ &= \frac{1}{4}\|\nu + \mu\|_2^2 - \frac{1}{4}(1 - \eta m)\|\nu - \mu\|_2^2 \\ &\quad + \frac{3}{4}\eta CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)(3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l) + 2\sqrt{m}\|\nu - \mu\|_2) \\ &\leq \frac{1}{4}\|\nu + \mu\|_2^2 + \frac{3}{4}\eta CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)(3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l) + 2\sqrt{m}\|\nu - \mu\|_2) \\ &\leq \frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}} \left[\frac{1}{4} \frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}} + \frac{1}{4}\eta m \left(\frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}} + 2\|\nu - \mu\|_2 \right) \right]\end{aligned}$$

$$\leq \frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}} \left(\frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{2\sqrt{m}} + 1 \right). \quad (\text{A.2})$$

其中第二个不等式源于 $\eta m \leq 1$, 最有一个不等式源于 $\eta m \leq 1$, $\|\boldsymbol{\nu} - \boldsymbol{\mu}\|_2 \leq 2$. 最后, 得到

$$\|[(\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})\mathbf{v}]_{\mathbb{L}}\|_2 = \|\mathbf{v}\|_2 \langle (\mathbf{I} - \eta \mathbf{A}^T \mathbf{A})\boldsymbol{\nu}, \boldsymbol{\mu} \rangle \leq \max\{s_1, s_2\} \cdot \|\mathbf{v}\|_2,$$

其中

$$s_1 = 1 - \eta m \cdot \frac{\sqrt{m} - 6CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}}, \quad s_2 = \frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{\sqrt{m}} \left(\frac{3CK^2\gamma(\mathbb{D}_k + \mathbb{D}_l)}{2\sqrt{m}} + 1 \right).$$

其次, 证明引理 2.2(b). 设

$$\mathbf{z} = \frac{[\mathbf{A}^T \mathbf{x}]_{\mathbb{L}}}{\|[\mathbf{A}^T \mathbf{x}]_{\mathbb{L}}\|_2} \in \mathbb{D}_l, \quad (\text{A.3})$$

则可得

$$\begin{aligned} \|[\mathbf{A}^T \mathbf{x}]_{\mathbb{L}}\|_2 &= \langle [\mathbf{A}^T \mathbf{x}]_{\mathbb{L}}, \mathbf{z} \rangle = \langle \mathbf{A}^T \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{A} \mathbf{z} \rangle \\ &\leq \|\mathbf{x}\|_2 \|\mathbf{A} \mathbf{z}\|_2 \leq \|\mathbf{x}\|_2 (\sqrt{m} + 2CK^2\gamma(\mathbb{D}_l)) \end{aligned}$$

以至少 $1 - 2\exp(-\gamma^2(\mathbb{D}_l))$ 的概率成立, 其中分别使用 Hölder 不等式和引理 2.1 得到第一个不等式和第二个不等式.

附录 A.2 引理 2.3 的证明

首先, 证明引理 2.3(a). 对于任意 $\mathbf{x} \in \mathbb{D}_{2k}$, 存在两个向量 $\mathbf{y}, \mathbf{z} \in \mathbb{D}_k$, 使得 $\mathbf{x} = (\mathbf{y} + \mathbf{z})/\|\mathbf{y} + \mathbf{z}\|_2$ 和 $y_i z_i = 0, i = 1, \dots, n$, 这意味着两个向量 $\mathbf{y}$ 和 $\mathbf{z}$ 的非零元位置完全不一样. 于是得到

$$\begin{aligned} \gamma(\mathbb{D}_{2k}) &= \mathbb{E} \sup_{\mathbf{x} \in \mathbb{D}_{2k}} |\langle \mathbf{g}, \mathbf{x} \rangle| \\ &= \mathbb{E} \sup_{(\mathbf{y} + \mathbf{z})/\|\mathbf{y} + \mathbf{z}\|_2 \in \mathbb{D}_{2k}} \left| \left\langle \mathbf{g}, \frac{\mathbf{y} + \mathbf{z}}{\|\mathbf{y} + \mathbf{z}\|_2} \right\rangle \right| \\ &= \mathbb{E} \sup_{\mathbf{y} \in \mathbb{D}_k, \mathbf{z} \in \mathbb{D}_k} \left| \left\langle \mathbf{g}, \frac{\mathbf{y} + \mathbf{z}}{\|\mathbf{y} + \mathbf{z}\|_2} \right\rangle \right| \\ &\leq \mathbb{E} \sup_{\mathbf{y} \in \mathbb{D}_k} \left| \left\langle \mathbf{g}, \frac{\mathbf{y}}{\|\mathbf{y} + \mathbf{z}\|_2} \right\rangle \right| + \mathbb{E} \sup_{\mathbf{z} \in \mathbb{D}_k} \left| \left\langle \mathbf{g}, \frac{\mathbf{z}}{\|\mathbf{y} + \mathbf{z}\|_2} \right\rangle \right| \\ &\leq \mathbb{E} \sup_{\mathbf{y} \in \mathbb{D}_k} \left| \left\langle \mathbf{g}, \frac{\mathbf{y}}{\|\mathbf{y}\|_2} \right\rangle \right| + \mathbb{E} \sup_{\mathbf{z} \in \mathbb{D}_k} \left| \left\langle \mathbf{g}, \frac{\mathbf{z}}{\|\mathbf{z}\|_2} \right\rangle \right| \\ &\leq \mathbb{E} \sup_{\mathbf{y} \in \mathbb{D}_k} \|\langle \mathbf{g}, \mathbf{y} \rangle\| + \mathbb{E} \sup_{\mathbf{z} \in \mathbb{D}_k} \|\langle \mathbf{g}, \mathbf{z} \rangle\| \\ &= 2\gamma(\mathbb{D}_k), \end{aligned}$$

其次, 引理 2.3(b) 和 2.3(c) 的证明是显然的.

Data-time tradeoffs for optimal k -thresholding algorithms in quantized compressed sensing

Jialiang Xu & Yunbin Zhao

Abstract Relaxed optimal k -thresholding (ROT) algorithms are designed as a new thresholding technique in compressed sensing to efficiently recover a k -sparse signal from a small number of noisy linear measurements, which overcome the shortcomings of traditional hard thresholding algorithms caused by the oscillation of the residual function. In this paper, we carry out a novel data-time tradeoff analysis to characterize the tradeoff between the number of measurements and the number of iterations under any fixed recovery accuracy for the relaxed optimal k -thresholding algorithms in quantized compressed sensing. Both the analysis and numerical results from synthetic signal recovery demonstrate the data-time tradeoffs of ROT algorithms in sparse reconstruction problems. Furthermore, the theory presents the order of the number of measurements required for successful recovery in compressed sensing, quantized compressed sensing, and corrupted sensing.

Keywords compressed sensing, quantized compressed sensing, sparse reconstruction problems, relaxed optimal k -thresholding, data-time tradeoffs

MSC(2020) 94A12, 15A29, 90C25, 90C20

doi: 10.1360/SSM-2024-0142