

# 计算材料科学中的优化问题、理论与算法

## 博士学位论文答辩

答辩人：胡雨宽

指导教师：刘 故 研究员



中国科学院数学与系统科学研究院

2024 年 5 月 10 日

实验试错  
理论模型  
数值模拟



材料性能与行为



材料研发与应用

- 实验试错成本高昂、理论模型大多难以显式求解.

“描写大部分物理学和整个化学的基本物理定律所必需的数学理论已经完全清楚了，而困难之处仅在于方程过于复杂无法求解。” [Dirac 1929]



Paul Dirac  
(1902–1984)

非线性、高维数、多尺度  $\Rightarrow$  数值求解困难.

- 计算材料科学中的能量极小化问题  $\Leftrightarrow$  优化理论与算法.

关注 电子 与 原子 层面的优化问题.

# 一、强关联电子体系计算中的 优化问题、理论与算法

# 电子结构计算的密度泛函理论 (DFT)<sup>(1)</sup>

[Pauli 1925; Schrödinger 1926; Born-Oppenheimer 1927; Hohenberg-Kohn 1964; Levy 1982; Lieb 1983; ...]

$$\begin{aligned} \inf_{\rho} \quad & F_{\text{LL}}[\rho] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) \, d\mathbf{r}, \\ \text{s. t.} \quad & \rho \geq 0, \sqrt{\rho} \in \mathcal{H}^1(\mathbb{R}^d; \mathbb{R}), \int \rho(\mathbf{r}) \, d\mathbf{r} = N. \end{aligned}$$

- $d \in \{1, 2, 3\}$ : 体系维数.  $N \in \mathbb{N}$ : 电子个数.
- $F_{\text{LL}}$ : Levy-Lieb 泛函.  $v_{\text{ext}} : \mathbb{R}^d \rightarrow \mathbb{R}$ : 外势.
- $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ : 单电子密度.

$F_{\text{LL}}$  缺乏显式表达式!

<sup>(1)</sup> 我们仅考虑孤立的、非相对论的、定态的、无自旋的多电子体系，并采用原子单位.

# DFT 的两个极限

弱相互作用极限  $\Rightarrow$  Kohn-Sham DFT (KSDFT) [Kohn-Sham 1965]

$$\min_{\Psi} \left\{ \frac{1}{2} \sum_{i=1}^N \int \|\nabla_{\mathbf{r}_i} \Psi\|^2 : \Psi \text{ 反对称}, \int |\Psi|^2 = 1, \right. \\ \left. \int |\Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 d\mathbf{r}_2 \cdots d\mathbf{r}_N = \rho(\mathbf{r})/N, \forall \mathbf{r} \in \mathbb{R}^d \right\}.$$

$\Psi : (\mathbb{R}^d)^N \rightarrow \mathbb{C}$ : 波函数.  $\mathbf{r}_i \in \mathbb{R}^d$ : 第  $i$  个电子的位置 ( $i = 1, \dots, N$ ).

**最常用的数学模型.**

- 非线性特征值问题.

[Rooshaan 1951; Pulay 1980/1982; Cancès-Le Bris 2000; Cancès 2001; Marks-Luke 2008; Gao et al. 2009;  
Yang et al. 2009; Liu et al. 2014; Liu et al. 2015; Bai et al. 2022; ...]

- 正交约束优化问题.

[Schneider et al. 2009; Wen-Yin 2013; Zhang et al. 2014; Jiang-Dai 2015; Dai et al. 2017; Hu et al. 2018; Dai et al. 2019; Gao et al. 2019; Hu et al. 2019; Wang et al. 2021; Gao et al. 2022; Liu et al. 2024; Xiao et al. 2024; ...]

**强关联电子体系的应用 (KSDFT 显著系统误差)**

磁存储 [Salamon-Jaime 2001]、高温超导 [Capone et al. 2002]、热电材料 [Costi-Zlatić 2010] 等.

# DFT 的两个极限 (续)

## 强相互作用极限 $\Rightarrow$ 严格关联电子 DFT (SCEDFT)

$\Updownarrow$  [Seidl 1999; Seidl et al. 1999; Seidl et al. 2000], [Buttazzo et al. 2012; Cotar et al. 2013]

### 带 Coulomb 费用的多边际最优运输问题 (MMOT)

$$V_{ee, SCE}[\rho] := \inf_{\Psi} \left\{ \sum_{1 \leq i < j}^N \int \frac{|\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)|^2}{\|\mathbf{r}_i - \mathbf{r}_j\|} d\mathbf{r}_1 \cdots d\mathbf{r}_N : \dots \right\}$$

$$= \min_{\Pi} \sum_{1 \leq i < j}^N \int \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N),$$

s. t.  $\Pi$  是  $(\mathbb{R}^d)^N$  上的一个概率测度,

$$\mathcal{M}_i(\Pi) = \rho/N, \quad i = 1, \dots, N.$$

- $\mathcal{M}_i$ : 第  $i$  个边际算子 ( $i = 1, \dots, N$ ).

$$[\mathcal{M}_i(\Pi)](\mathbf{r}) := \int d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_{i-1}, \mathbf{r}, \mathbf{r}_{i+1}, \dots, \mathbf{r}_N), \quad \forall \mathbf{r} \in \mathbb{R}^d.$$

- $V_{ee, SCE}$  及其泛函导数 (SCE 势)  $\Leftrightarrow$  最优值与最优对偶势. [Chen et al. 2014]

维数灾难!

# MMOT 的已有转化模型及其数值方法

表 1: MMOT 的已有转化模型及其数值方法.

转化模型	相关优化问题	现有数值方法	现有模型/方法主要缺陷	代表性参考文献
Kantorovich 对偶	非光滑优化问题	无导数优化算法	模型具有维数灾难	[Mendi-Lin 2013]
熵正则化	凸优化问题	块坐标下降算法		[Benamou et al. 2016]
$N$ 可表示	半定规划 (松弛)	半定规划算法	离散规模较小	[Khoo-Ying 2019]
矩约束松弛	非线性约束优化问题	投影梯度下降算法	数值表现对测试函数敏感	[Alfonsi et al. 2022]
Monge 拟设	带 PDE 约束的优化问题	—	—	[Seidl 1999]
类 Monge 拟设	带广义互补约束的优化问题	线性规划算法	无法处理 $N \geq 3$ 情形	[Chen et al. 2014]
拟 Monge 拟设	混合整数规划	—	—	[Friesecke-Vögler 2018]

类 Monge 拟设:  $d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{\rho(\mathbf{r}_1)}{N} \gamma_2(\mathbf{r}_1, \mathbf{r}_2) \cdots \gamma_N(\mathbf{r}_1, \mathbf{r}_N) d\mathbf{r}_1 \cdots d\mathbf{r}_N,$

其中  $\gamma_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  为第 1 和第  $i$  个电子位置的耦合函数 ( $i = 2, \dots, N$ ).

# 带广义互补约束的优化问题 (MPGCC)

MMOT + 类 Monge 拟设 + 有限元离散  $\Rightarrow$  MPGCC

$$\begin{aligned} \min_{Y_2, \dots, Y_N} \quad & \frac{1}{N} \sum_{i=2}^N \langle Y_i, \text{Diag}(\boldsymbol{\varrho}) C \rangle + \frac{1}{N} \sum_{2 \leq i < j}^N \langle Y_i, \text{Diag}(\boldsymbol{\varrho}) Y_j C \rangle, \\ \text{s. t.} \quad & Y_i \mathbf{1}_K = \mathbf{1}_K, \quad Y_i^\top \boldsymbol{\varrho} = \boldsymbol{\varrho}, \quad \text{Tr}(Y_i) = 0, \quad \underline{Y_i \geq 0}, \quad i = 2, \dots, N, \\ & \langle Y_i, Y_j \rangle = 0, \quad \forall i, j \in \{2, \dots, N\} : i \neq j. \end{aligned} \tag{P}$$

- $K \in \mathbb{N}$ : 离散规模.  $\mathbf{1}_K \in \mathbb{R}^K$ : 全一向量.
- $\boldsymbol{\varrho} := [\varrho_1, \dots, \varrho_K]^\top \in \mathbb{R}_+^K$ : 单电子密度  $\rho$  的离散.
- $C := (c_{ij}) \in \mathbb{R}^{K \times K}$ : 两体 Coulomb 势  $1/\|\mathbf{r} - \mathbf{r}'\|$  的离散 (对角线置零).
- $Y_i := (y_{i,jk})_{jk} \in \mathbb{R}^{K \times K}$ : 耦合函数  $\gamma_i$  的离散 ( $i = 2, \dots, N$ ).
- 非负约束 + 零内积约束  $\Rightarrow$  广义互补约束: 防止电子间的碰撞.

已有工作:  $N = 2 \Rightarrow$  线性规划. [Chen et al. 2014]

约束规范条件不满足  $\Rightarrow$  KKT 条件的必要性不成立.

## 例 1 (全局最优解不满足 KKT 条件的 MPCC)

考虑如下 MPCC:

$$\begin{aligned} \min_{x_1, x_2, x_3} \quad & x_1^2 + x_2^2 - x_3, \\ \text{s. t.} \quad & -4x_1 + x_3 \leq 0, \quad -4x_2 + x_3 \leq 0, \\ & x_1, x_2 \geq 0, \quad x_1 x_2 = 0. \end{aligned}$$

可以验证, KKT 条件在其最优解  $[0, 0, 0]^\top$  处不成立.

# MPCC 的精确罚理论

$$\begin{array}{ll} \min_{\mathbf{x}_1, \mathbf{x}_2} & \tilde{f}(\mathbf{x}_1, \mathbf{x}_2), \\ \text{s. t.} & \tilde{\mathbf{g}}(\mathbf{x}_1, \mathbf{x}_2) \geq 0, \\ & \mathbf{x}_1, \mathbf{x}_2 \geq 0, \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0. \end{array} \quad \stackrel{?}{\iff} \quad \begin{array}{ll} \min_{\mathbf{x}_1, \mathbf{x}_2} & \tilde{f}(\mathbf{x}_1, \mathbf{x}_2) + p(\mathbf{x}_1, \mathbf{x}_2), \\ \text{s. t.} & \tilde{\mathbf{g}}(\mathbf{x}_1, \mathbf{x}_2) \geq 0, \mathbf{x}_1, \mathbf{x}_2 \geq 0. \end{array}$$

表 2: MPCC 的精确罚结果.

罚项	主要假设条件	参考文献
$\ \min\{\mathbf{x}_1, \mathbf{x}_2\}\ $	可行域紧多面体	[Luo et al. 1996]
$\sqrt{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle}$	可行域紧多面体	[Luo et al. 1996]
$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle$	<b>严格互补、可行域误差界</b>	[Luo et al. 1996]
$\ \min\{\mathbf{x}_1, \mathbf{x}_2\}\  + r(\mathbf{x}_1, \mathbf{x}_2)^{(2)}$	可行域误差界	[Scholtes-Stöhr 1999]
$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle$	$f$ 仿射、可行域紧多面体	[Campêlo-Scheinberg 2001]
$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle$	可行域有内点、正乘子非退化	[Liu et al. 2001]

<sup>(2)</sup> 这里, “ $r(\mathbf{x}_1, \mathbf{x}_2)$ ” 表示其他约束的罚项.

模型选择 DFT  $\xrightarrow[\text{作用极限}]{\text{强相互}} \text{MMOT} \xrightarrow[\text{有限元离散}]{\text{类 Monge 拟设}} \text{MPGCC.}$

1. 一类特殊 MPGCC 的  $\ell_1$  精确罚**理论**.
2. 求解具有块状结构优化问题的**算法**.
3. 求解运输多胞体上分块矩阵优化问题的**算法**.

# 工作一: 一类特殊 MPGCC 的 $\ell_1$ 精确罚理论

$$\begin{aligned} & \min_{\mathbf{x}_1, \dots, \mathbf{x}_s} f(\mathbf{x}_1, \dots, \mathbf{x}_s), \\ \text{s. t. } & \mathbf{x}_i \in \mathcal{F}_i, \quad \mathbf{x}_i \geq 0, \quad i = 1, \dots, s, \\ & \langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0, \quad \forall i, j \in \{1, \dots, s\} : i \neq j. \end{aligned} \tag{MA-MPGCC}$$

- $\mathbf{x}_i \in \mathbb{R}^m$ : 第  $i$  个变量块 ( $i = 1, \dots, s$ ).
- $f: (\mathbb{R}^m)^s \rightarrow \mathbb{R}$ : 多仿射 目标函数.
- $\mathcal{F}_i \subseteq \mathbb{R}^m$ : 第  $i$  个变量块的 多面体 可行域 ( $i = 1, \dots, s$ ).
- 应用: 强关联电子体系计算、生产运输中的序列决策过程等.

[Labbé et al. 1998; Bard et al. 2000; Côté et al. 2003; Bouhtou et al. 2006; ...]

命题 2 (已有  $\ell_1$  精确罚理论不适用)

对任给  $2 \leq s \in \mathbb{N}$ , 存在 MA-MPGCC 的实例, 对其已有 MPCC 的  $\ell_1$  精确罚理论 **均不适用**.

主要贡献: 较弱假设条件下的  $\ell_1$  精确罚.

# MA-MPGCC 的 $\ell_1$ 精确罚理论

$$\begin{aligned} \min_{\mathbf{x}_1, \dots, \mathbf{x}_s} \quad & f(\mathbf{x}_1, \dots, \mathbf{x}_s) + \beta \sum_{1 \leq i < j}^s \langle \mathbf{x}_i, \mathbf{x}_j \rangle, \\ \text{s. t.} \quad & \mathbf{x}_i \in \mathcal{F}_i, \quad \mathbf{x}_i \geq 0, \quad i = 1, \dots, s, \end{aligned} \quad (\text{MA-MPGCC}_\beta)$$

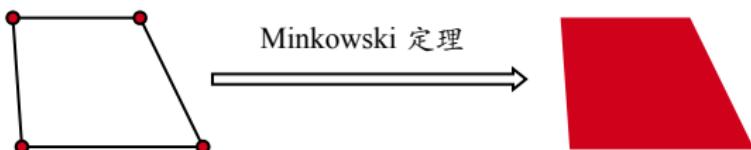
其中  $\beta \geq 0$  是罚参数.

## 定理 3 ( $\ell_1$ 罚函数的精确性)

假设  $\text{MA-MPGCC}_\beta$  可行域非空且紧. 则存在  $\bar{\beta} > 0$ , 使得对任意  $\beta \geq \bar{\beta}$ ,  $\text{MA-MPGCC}$  和  $\text{MA-MPGCC}_\beta$  具有相同的最优解集.

### 主要创新点

证明关键:



## 工作二: 求解具有块状结构优化问题的算法

### 问题 (P) 的 $\ell_1$ 罚问题

$$\begin{aligned} \min_{Y_2, \dots, Y_N} \quad & \frac{1}{N} \sum_{i=2}^N \langle Y_i, \text{Diag}(\boldsymbol{\varrho}) C + \beta \mathbf{I} \rangle + \frac{1}{N} \sum_{2 \leq i < j}^N \langle Y_i, \text{Diag}(\boldsymbol{\varrho}) Y_j C + \beta Y_j \rangle, \quad (\text{P}_\beta) \\ \text{s. t.} \quad & Y_i \mathbf{1}_K = \mathbf{1}_K, \quad Y_i^\top \boldsymbol{\varrho} = \boldsymbol{\varrho}, \quad Y_i \geq 0, \quad i = 2, \dots, N. \end{aligned}$$

### 一般具有块状结构的优化问题

$$\begin{aligned} \min_{\mathbf{x}_1, \dots, \mathbf{x}_s} \quad & f(\mathbf{x}_1, \dots, \mathbf{x}_s), \\ \text{s. t.} \quad & \mathbf{x}_i \in \mathcal{F}_i := \{\mathbf{y}_i \in \mathbb{R}^{m_i} : \mathbf{g}_i(\mathbf{y}_i) \leq 0\}, \quad i = 1, \dots, s. \end{aligned}$$

- $f: \times_{i=1}^s \mathbb{R}^{m_i} \rightarrow \mathbb{R}$ : 目标函数, 可微未必凸.
- $\mathbf{g}_i: \mathbb{R}^{m_i} \rightarrow \mathbb{R}^{p_i}$ : 不等式约束函数, 凸可微 ( $i = 1, \dots, s$ ).
- 应用: **强关联电子体系计算**、弹性力学 [Kučera 2008]、信号处理 [He et al. 2010]、图像去噪 [Bonettini et al. 2018]、拓扑优化 [Liu et al. 2020] 等.

# 邻近交替线性化极小化 (PALM) 算法

步 1. 初始化  $\mathbf{x}^{(0)} := (\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_s^{(0)}) \in \bigtimes_{i=1}^s \mathbb{R}^{m_i}$ . 置  $k := 0$ .

步 2. 对  $i = 1, \dots, s$ :

步 2.1 选取邻近参数  $\sigma_{i,k} > 0$ .

步 2.2 求解邻近线性化子问题

$$\min_{\mathbf{x}_i \in \mathcal{F}_i} \quad \left\langle \nabla_{\mathbf{x}_i} f\left(\mathbf{x}_{<i}^{(k+1)}, \mathbf{x}_{\geq i}^{(k)}\right), \mathbf{x}_i - \mathbf{x}_i^{(k)} \right\rangle + \frac{\sigma_{i,k}}{2} \left\| \mathbf{x}_i - \mathbf{x}_i^{(k)} \right\|^2$$

直至满足特定条件, 得到  $\mathbf{x}_i^{(k+1)} \in \mathbb{R}^{m_i}$ .

步 3. 检查终止准则是否满足. 若是, 停机; 若否, 置  $k := k + 1$ , 转步 2.

- 记号:  $\mathbf{x} := (\mathbf{x}_1, \dots, \mathbf{x}_s)$ ,  $\mathbf{x}_{<i} := (\mathbf{x}_1, \dots, \mathbf{x}_{i-1})$ ,  $\mathbf{x}_{\geq i} := (\mathbf{x}_i, \dots, \mathbf{x}_s)$ .
- 求解子问题  $\Leftrightarrow$  将  $\tilde{\mathbf{x}}_i^{(k)} := \mathbf{x}_i^{(k)} - \nabla_{\mathbf{x}_i} f(\mathbf{x}_{<i}^{(k+1)}, \mathbf{x}_{\geq i}^{(k)}) / \sigma_{i,k}$  正交投影至  $\mathcal{F}_i$ .

# 不同类型的 PALM 算法

## 精确的 PALM (PALM-E) 算法

[Razaviyayn et al. 2013; Xu-Yin 2013; Bolte et al. 2014; Wang et al. 2018; ...]

- 子问题被精确求解.
- 目标函数值序列充分下降性:

$$\exists a_1 > 0, \text{ s. t. } \textcolor{violet}{a}_1 \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|^2 \leq f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k+1)}).$$

- 迭代点稳定性违反度上界:

$$\exists a_2 > 0, \text{ s. t. } \text{dist}\left(0, \partial F(\mathbf{x}^{(k+1)})\right) \leq a_2 \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|,$$

其中  $F := f + \sum_{i=1}^s \delta_{\mathcal{F}_i}$ ,  $\delta_{\mathcal{F}_i}$  是集合  $\mathcal{F}_i$  的指示函数 ( $i = 1, \dots, s$ ).

## 可行非精确的 PALM (PALM-F) 算法

[Hua-Yamashita 2016; Bonettini et al. 2018; Ochs 2019; Gur et al. 2022; ...]

- 子问题被非精确求解且恒有  $\mathbf{x}_i^{(k)} \in \mathcal{F}_i$ .
- (近似) 目标函数值序列充分下降性 + (近似) 迭代点稳定性违反度上界.

## 不可行非精确的 PALM (PALM-I) 算法

- 子问题被非精确求解但可能  $\mathbf{x}_i^{(k)} \notin \mathcal{F}_i$ .
- 分析难点: 目标函数值序列**非单调性无法避免**.
- 场景举例: 刻画可行域的代数系统**严重欠定**  $\Rightarrow$  不可行算法.
- 唯一已有工作**不具有实际应用价值**. [Frankel et al. 2015]
  - 存在  $\beta_1, \beta_2 > 0$ , 使得对  $i = 1, \dots, s$  和  $k \geq 0$ ,

$$\left\{ \begin{array}{l} \sum_{j=1}^{i-1} \|\mathbf{x}_j^{(k+1)} - \bar{\mathbf{x}}_j^{(k+1)}\| + \sum_{j=i}^s \|\mathbf{x}_j^{(k)} - \bar{\mathbf{x}}_j^{(k)}\| \leq \beta_1 \|\bar{\mathbf{x}}_i^{(k+1)} - \bar{\mathbf{x}}_i^{(k)}\|; \\ \left\langle \mathbf{x}_i^{(k)} - \bar{\mathbf{x}}_i^{(k)}, \bar{\mathbf{x}}_i^{(k+1)} - \bar{\mathbf{x}}_i^{(k)} \right\rangle \leq \beta_2 \|\bar{\mathbf{x}}_i^{(k+1)} - \bar{\mathbf{x}}_i^{(k)}\|^2, \end{array} \right.$$

其中  $\bar{\mathbf{x}}_i^{(k+1)}$  是子问题的唯一最优解.

主要贡献: 可实现的 PALM-I 算法及其理论性质.

# PALM-I 算法

步 1. 初始化  $\mathbf{x}^{(0)} := (\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_s^{(0)}) \in \bigtimes_{i=1}^s \mathbb{R}^{m_i}$ ,  $\bar{\varepsilon} \in (0, 1)$ . 置  $k := 0$ .

步 2. 对  $i = 1, \dots, s$ :

步 2.1 选取邻近参数  $\sigma_{i,k} > 0$  与误差控制参数  $\varepsilon_k \in [0, \bar{\varepsilon}]$ .

步 2.2 求解邻近线性化子问题

$$\min_{\mathbf{x}_i \in \mathcal{F}_i} \quad \left\langle \nabla_{\mathbf{x}_i} f\left(\mathbf{x}_{<i}^{(k+1)}, \mathbf{x}_{\geq i}^{(k)}\right), \mathbf{x}_i - \mathbf{x}_i^{(k)} \right\rangle + \frac{\sigma_{i,k}}{2} \left\| \mathbf{x}_i - \mathbf{x}_i^{(k)} \right\|^2$$

得到  $\mathbf{x}_i^{(k+1)} \in \mathbb{R}^{m_i}$ , 使得存在  $\boldsymbol{\lambda}_i^{(k+1)} \in \mathbb{R}_+^{p_i}$  满足

$$\sqrt{r_i(\mathbf{x}_i^{(k+1)}, \boldsymbol{\lambda}_i^{(k+1)}, \tilde{\mathbf{x}}_i^{(k)})} \leq \varepsilon_k. \quad (1)$$

步 3. 检查终止准则是否满足. 若是, 停机; 若否, 置  $k := k + 1$ , 转步 2.

$$r_i(\mathbf{x}_i, \boldsymbol{\lambda}_i, \tilde{\mathbf{x}}_i) := \underbrace{\max \{ \langle \mathbf{x}_i, \mathbf{x}_i - \tilde{\mathbf{x}}_i + \nabla g_i(\mathbf{x}_i) \boldsymbol{\lambda}_i \rangle, 0 \} + \| \mathbf{x}_i - \tilde{\mathbf{x}}_i + \nabla g_i(\mathbf{x}_i) \boldsymbol{\lambda}_i \|_\infty}_{\text{次稳定性违反度}} \\ + \underbrace{\| \max \{ g_i(\mathbf{x}_i), 0 \} \|_\infty}_{\text{可行性违反度}} + \underbrace{\max \{ -\langle \boldsymbol{\lambda}_i, g_i(\mathbf{x}_i) \rangle, 0 \}}_{\text{互补松弛违反度}}. \quad \text{可满足性、可计算性.}$$

# PALM-I 算法的理论性质

## 条件 4 (目标函数的光滑性)

目标函数  $f$  对每个变量块都局部 Lipschitz 连续可微, 即对  $i = 1, \dots, s$ , 对任一有界  $\bar{\mathcal{F}}_i \supseteq \mathcal{F}_i$ , 存在常数  $L_i > 0$  使得如下不等式成立: 对任意  $\mathbf{x}, \mathbf{x}' \in \bigtimes_{i=1}^s \bar{\mathcal{F}}_i$ ,

$$\|\nabla_{\mathbf{x}_i} f(\mathbf{x}) - \nabla_{\mathbf{x}_i} f(\mathbf{x}')\| \leq L_i \|\mathbf{x} - \mathbf{x}'\|.$$

## 条件 5 (可行域的正则性)

对  $i = 1, \dots, s$ , 可行域  $\mathcal{F}_i$  是  $\mathbb{R}^{m_i}$  中的紧凸集, 且  $\mathbf{g}_i$  满足如下两个条件中的一个:

- (1)  $\mathbf{g}_i$  是线性函数.
- (2)  $\mathbf{g}_i$  满足 Slater 约束规范条件且在某个  $\tilde{\mathcal{F}}_i \supseteq \mathcal{F}_i$  有如下 Hoffman 型不等式成立:

$$\exists \tilde{c}_i > 0, \text{ s. t. } \text{dist}(\mathbf{x}_i, \mathcal{F}_i) \leq \tilde{c}_i \|\max\{\mathbf{g}_i(\mathbf{x}_i), 0\}\|, \quad \forall \mathbf{x}_i \in \tilde{\mathcal{F}}_i.$$

## 条件 6 (误差控制序列的可和性)

- (1) 序列  $\{\varepsilon_k\}$  非负且平方可和.
- (2) 序列  $\{\varepsilon_k\}$  非负、可和, 且存在  $\bar{\theta} \in (0, 1)$  使得  $\{e_k^{\bar{\theta}}\}$  可和, 其中

$$e_k := \sum_{t=k}^{\infty} \varepsilon_t^2, \quad \forall k \geq 0.$$

**主要创新点:** 构造**单调代理序列**

$$v_k := F(\bar{\mathbf{x}}^{(k)}) + u_k + u_{k+1}, \quad u_k := \sum_{t=k}^{\infty} C_t \|\mathbf{x}^{(t)} - \bar{\mathbf{x}}^{(t)}\|^2,$$

其中  $\{C_t\}$  是一个正序列.

**定理 7 (PALM-I 算法的全局依子列收敛性)**

假设条件 4 与 5 成立. 令  $\{\mathbf{x}^{(k)}\}$  为 PALM-I 算法产生的迭代点序列, 其中  $\{\varepsilon_k\}$  满足条件 6 (1). 则  $\{\mathbf{x}^{(k)}\}$  至少有一个聚点且其每个聚点都是  $F$  的稳定点.

**定理 8 (PALM-I 算法的全局依点列收敛性)**

假设条件 4 与 5 成立,  $F$  在任一稳定点  $\bar{\mathbf{x}}$  处满足 Łojasiewicz 性质: 存在  $c > 0, \theta \in [0, 1)$  以及  $\eta > 0$ , 使得

$$|F(\mathbf{x}) - F(\bar{\mathbf{x}})|^\theta \leq c \cdot \text{dist}(0, \partial F(\mathbf{x})), \quad \forall \mathbf{x} \in \bigtimes_{i=1}^s \mathbb{R}^{m_i} : \|\mathbf{x} - \bar{\mathbf{x}}\| \leq \eta,$$

其中  $\theta$  称为  $F$  在  $\bar{\mathbf{x}}$  处的 Łojasiewicz 指数. 令  $\{\mathbf{x}^{(k)}\}$  为 PALM-I 算法产生的迭代点序列, 其中  $\{\varepsilon_k\}$  满足条件 6 (2). 则  $\{\mathbf{x}^{(k)}\}$  收敛于  $F$  的一个一阶稳定点.

# PALM-I 算法的理论性质 (续)

表 3: PALM 算法在不同情形下的收敛速度.  
其中  $\theta$  为  $F$  在极限点处的 Łojasiewicz 指数.

$\theta$	$\varepsilon_k$	假设	渐进收敛速度	参考文献
0	0	-	有限终止	[Xu-Yin 2013] [Bolte et al. 2014]
0	$\tilde{\rho}^k$	$\tilde{\rho} \in (0, 1)$	$\mathcal{O}(\rho_1^k)$ , 其中 $\rho_1 \in (0, 1)$	本文
	$\frac{1}{(k+1)^l}$	$l \in (1, \infty)$	$\mathcal{O}(k^{-(l-1)})$	
0	0	-	$\mathcal{O}(\hat{\rho}^k)$ , 其中 $\hat{\rho} \in (0, 1)$	[Xu-Yin 2013] [Bolte et al. 2014]
$(0, \frac{1}{2}]$	$\tilde{\rho}^k$	$\tilde{\rho} \in (0, 1)$	$\mathcal{O}(\rho_2^k)$ , 其中 $\rho_2 \in (0, 1)$	本文
	$\frac{1}{(k+1)^l}$	$l \in (1, \infty)$	$\mathcal{O}(k^{-(l-1)})$	
0	0	-	$\mathcal{O}\left(k^{-\frac{1-\theta}{2\theta-1}}\right)$	[Xu-Yin 2013] [Bolte et al. 2014]
$(\frac{1}{2}, 1)$	$\tilde{\rho}^k$	$\tilde{\rho} \in (0, 1)$	$\mathcal{O}\left(k^{-\frac{1-\theta}{2\theta-1}}\right)$	
	$\frac{1}{(k+1)^l}$	$l \in (1, \infty)$	$\mathcal{O}\left(k^{-\frac{1-\theta}{2\theta-1}}\right)$ , 若 $l \geq \frac{\theta}{2\theta-1}$ $\mathcal{O}(k^{-(l-1)})$ , 若 $l < \frac{\theta}{2\theta-1}$	本文

## 测试问题

$$\min_{\mathbf{x}} \frac{1}{2} \langle \mathbf{x}, A\mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{x} \rangle, \quad \text{s. t. } \frac{1}{2} \langle \mathbf{x}_i, B_i \mathbf{x}_i \rangle \leq 1, \quad i = 1, \dots, s.$$

- $\{\mathbf{x}_i\}_{i=1}^s \subseteq \mathbb{R}^m$ .  $\mathbf{b} \in \mathbb{R}^{ms}$ .
- $A \in \mathbb{R}^{ms \times ms}$  对称.  $\{B_i = (b_{i,jk})_{jk}\}_{i=1}^s \subseteq \mathbb{R}^{m \times m}$  对称正定.

## 问题设置

- $s = 5, m = 500 \Rightarrow 2500$  个变量.
- $\mathbf{b}$  使用 “`randn`” 随机生成.  $A$  使用 “`randn`” + 对称化随机生成.
- $\{B_i = (b_{i,jk})_{jk}\}_{i=1}^5$  为对角矩阵, 定义为 [Dai 2006; Jia et al. 2017]

$$b_{i,jk} := \begin{cases} 10^{\frac{j-1}{m-1} ncond_i}, & \text{若 } j = k; \\ 0, & \text{否则,} \end{cases} \quad j, k = 1, \dots, m,$$

其中  $\{ncond_i\}_{i=1}^5 = \{3.00, 3.25, 3.50, 3.75, 4.00\}$ .

## 比较算法及其设置

- PALM-E 算法: 交替方向乘子法 [Jia et al. 2017] + 子问题非精确求解准则 (1),  
 $\varepsilon_k \equiv 10^{-6}$ .
- PALM-F 算法: 混合投影算法 [Dai 2006] + 子问题非精确求解准则

$$\left\| \mathbf{x}_i^{(k+1)} - \tilde{\mathbf{x}}_i^{(k)} + \lambda_i^{(k+1)} B_i \mathbf{x}_i^{(k+1)} \right\| \leq \frac{0.99}{2} \left\| \mathbf{x}_i^{(k+1)} - \mathbf{x}_i^{(k)} \right\|_{\infty},$$

其中  $\lambda_i^{(k+1)} \geq 0$  是对应于椭球约束的 Lagrange 乘子 [Hua-Yamashita 2016].

- PALM-I 算法: 交替方向乘子法求解子问题, 子问题非精确求解准则 (1),

$$\varepsilon_k = \max \left\{ \frac{10^{-1}}{(k+1)^{0.75}}, 10^{-6} \right\}, \quad \forall k \geq 0.$$

- 公共参数: 邻近参数  $\sigma_{i,k} \equiv 1$ .
- 终止准则: 相对 KKT 违反度小于  $10^{-5}$ .

## 数值实验—算法比较 (续)

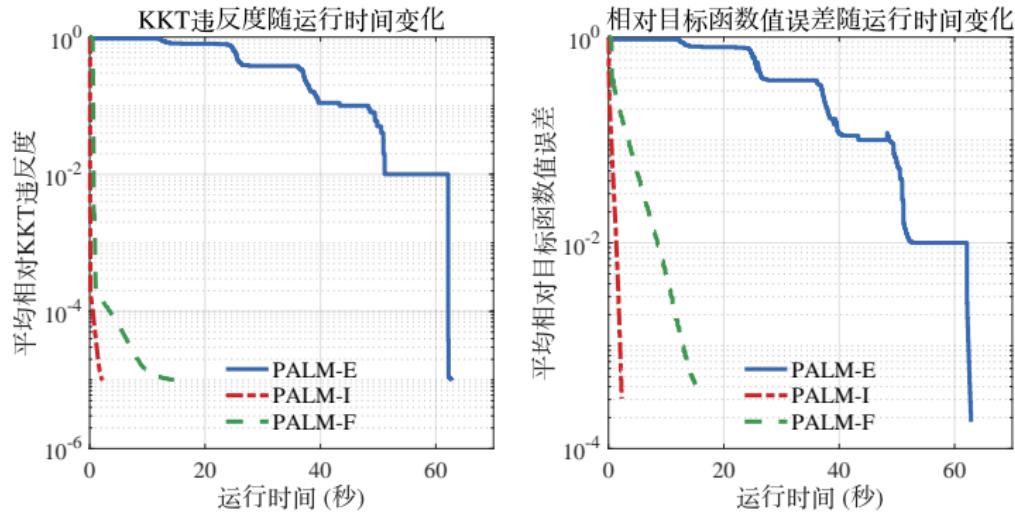


图 1: 从 100 个随机初始点出发, 不同算法的平均相对 KKT 违反度与相对目标函数值误差随运行时间的变化.

运行时间: PALM-E: 30.15 秒, PALM-F: 9.26 秒, **PALM-I: 1.74 秒.**

# 数值实验—强关联电子体系计算

通过求解问题  $(P_\beta)$  模拟强关联电子体系.

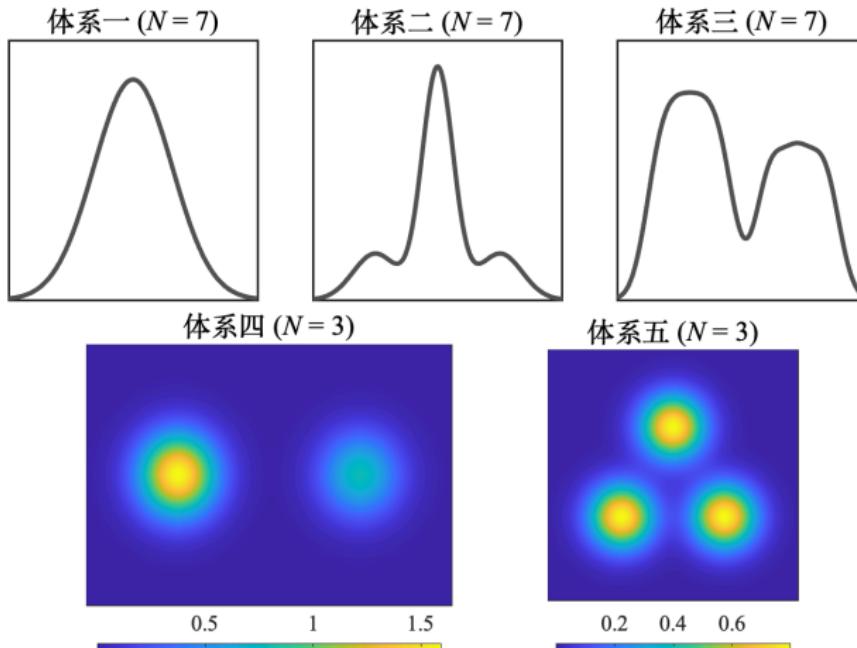


图 2: 待模拟体系单电子密度可视化.

# 数值实验—强关联电子体系计算 (续)

调用高精度优化算法求解初始网格  $\mathcal{T}^{(0)}$  上的问题, 得解  $Y_i^{(0,\star)} (i = 2, \dots, N)$

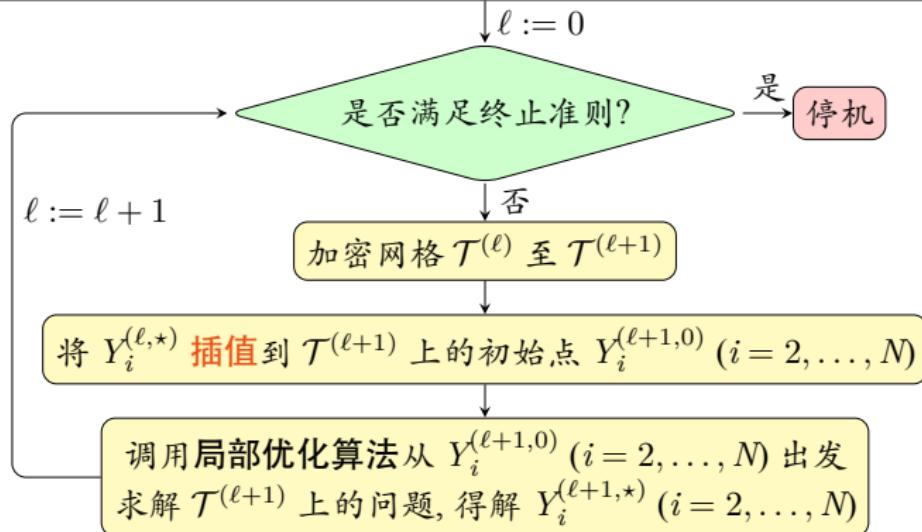


图 3: 求解问题  $(P_\beta)$  的瀑布型多重网格优化框架 CMGOPT.

**插值算子**  $\mathcal{I}_\ell^{\ell+1}(Y_i)_{j'k'} := \frac{\varrho_j^{(\ell)}}{\varrho_{j'}^{(\ell+1)}} y_{i,jk}, j: e_{j'}^{(\ell+1)} \subseteq e_j^{(\ell)}, k: e_{k'}^{(\ell+1)} \subseteq e_k^{(\ell)}.$

## 算法参数设置

- 罚参数  $\beta \sim \mathcal{O}(1/K)$ .
- CMGOPT 框架:
  - 一维: 等质量离散 + 一致加密, 随机多初始 [Hickernell-Yuan 1999] + PALM-I.
  - 二维: 自适应离散 + 一致加密, BARON [Tawarmalani-Sahinidis 2005] + PALM-I.
- PALM-I 算法:
  - 邻近参数:  $\sigma_{i,k} \equiv \sigma = 10^{-3}$ .
  - 半光滑 Newton 法求解子问题. [Li et al. 2020]
  - 子问题非精确求解准则 (1),  $\varepsilon_k \equiv 10^{-9}$ .
  - 终止准则:  $\sqrt{\sigma \sum_{i=2}^N \|Y_i^{(k+1)} - Y_i^{(k)}\|^2} \leq 10^{-4}$ .

# 数值实验—强关联电子体系计算 (续)

表 4: 一维七电子体系在每层网格上的目标函数值与映射误差.

$\ell$	体系一			体系二			体系三		
	$K$	$f_{\text{out}}$	err_map	$K$	$f_{\text{out}}$	err_map	$K$	$f_{\text{out}}$	err_map
0	14	173.951	0.052	14	151.891	0.039	14	111.964	0.030
1	28	181.474	0.019	28	158.797	0.028	28	117.223	0.010
2	56	181.929	0.025	56	158.507	0.026	56	117.050	0.008
3	112	181.989	0.012	112	158.317	0.011	112	116.914	0.008
4	224	181.954	0.013	224	158.267	0.008	224	116.876	0.006
5	448	181.942	0.002	448	158.255	0.004	448	116.864	0.003
6	896	181.939	0.001	896	158.254	0.002	896	116.861	0.001

- 电子位置间近似映射:  $\gamma_i(\mathbf{r}_1, \mathbf{r}_i) \rightarrow \hat{\mathcal{T}}_i(\mathbf{r}_1) = \mathbf{r}_i$ ,

$$\hat{\mathcal{T}}_i(\mathbf{d}_j) := \sum_{k=1}^K \mathbf{d}_k y_{i,jk}, \quad j = 1, \dots, K, \quad i = 2, \dots, N,$$

其中  $\mathbf{d}_k \in \mathbb{R}^d$  是第  $k$  个单元的重心 ( $k = 1, \dots, K$ ).

- 映射误差:

$$\text{err\_map} := \frac{1}{K|\Omega|} \sum_{k=1}^K \sum_{i=2}^N \|\mathcal{T}_i^*(\mathbf{d}_k) - \hat{\mathcal{T}}_i(\mathbf{d}_k)\|,$$

其中  $\{\mathcal{T}_i^*\}_{i=2}^N$  为最优映射 [Colombo et al. 2015],  $\Omega \subseteq \mathbb{R}^d$  为截断区域.

# 数值实验—强关联电子体系计算 (续)

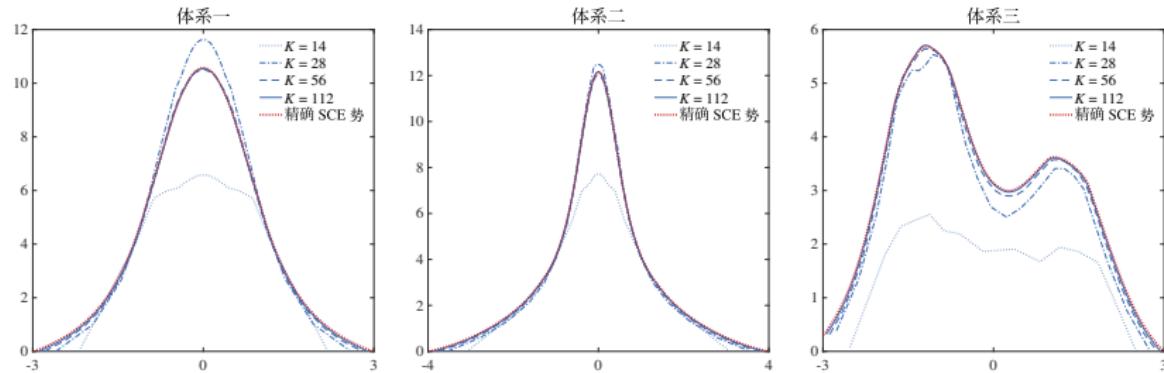


图 4: 一维七电子体系在第  $\ell = 0 \sim 3$  层网格上的  $\hat{\lambda}$  以及精确 SCE 势.

- SCE 势的近似:  $\hat{\lambda} := \bar{\lambda} - \min_{k=1}^K \{\bar{\lambda}_k\} \cdot \mathbf{1}_K$ , 其中

$$\bar{\lambda} := \frac{1}{N-1} \sum_{i=2}^N \lambda_{i,\varrho} \in \mathbb{R}^K,$$

$\lambda_{i,\varrho} \in \mathbb{R}^K$  为对应于约束  $Y_i^\top \varrho = \varrho$  的对偶变量 ( $i = 2, \dots, N$ ).

# 数值实验—强关联电子体系计算 (续)

表 5: 二维三电子体系在每层网格上的目标函数值.

$\ell$	体系四		体系五	
	$K$	$f_{\text{out}}$	$K$	$f_{\text{out}}$
0	240	9.503	170	9.491
1	960	9.577	680	9.533
2	3840	9.598	2720	9.543
3	15360	9.604	10880	9.546
4	61440	9.606	43520	9.547

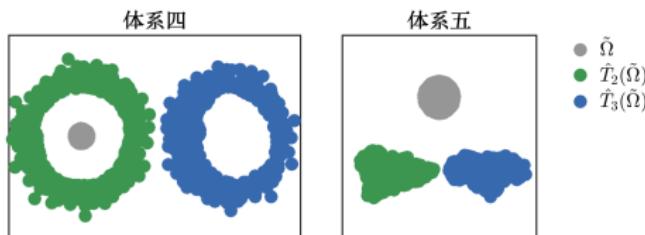
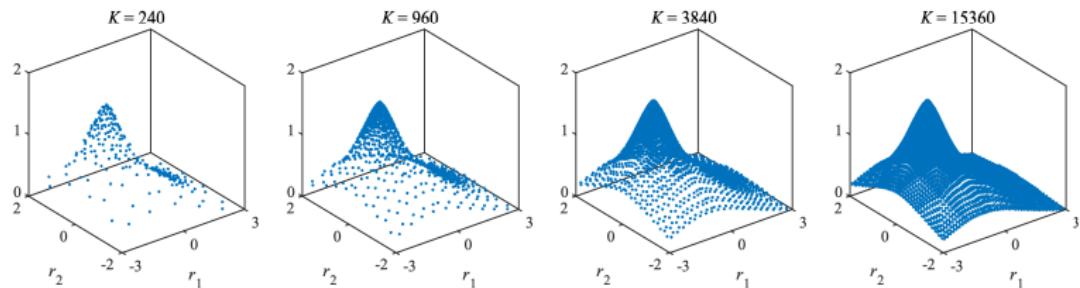


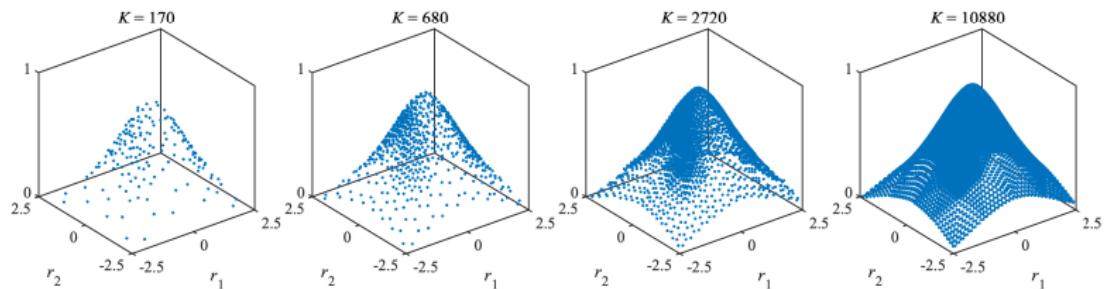
图 5: 二维三电子体系在第  $\ell = 3$  层网格上的近似映射  $\{\hat{\mathcal{T}}_i\}_{i=2}^N$  切片.

首次可视化了二维情形下电子位置之间的近似映射.

# 数值实验—强关联电子体系计算 (续)



(a) 体系四



(b) 体系五

图 6: 二维三电子体系在第  $\ell = 0 \sim 3$  层网格上的  $\hat{\lambda}$ .

# 工作三: 求解运输多胞体上分块矩阵优化问题的算法

## 问题 (P) 的 $\ell_1$ 罚问题

$$\begin{aligned} \min_{Y_2, \dots, Y_N} \quad & \frac{1}{N} \sum_{i=2}^N \langle Y_i, \text{Diag}(\boldsymbol{\varrho}) C + \beta I \rangle + \frac{1}{N} \sum_{2 \leq i < j}^N \langle Y_i, \text{Diag}(\boldsymbol{\varrho}) Y_j C + \beta Y_j \rangle, \\ \text{s. t.} \quad & Y_i \mathbf{1}_K = \mathbf{1}_K, \quad Y_i^\top \boldsymbol{\varrho} = \boldsymbol{\varrho}, \quad Y_i \geq 0, \quad i = 2, \dots, N. \end{aligned} \quad (\text{P}_\beta)$$

## 运输多胞体上的分块矩阵优化问题

$$\min_{X_1, \dots, X_s} \quad f(X_1, \dots, X_s), \quad \text{s. t. } X_i \in \mathcal{U}(\mathbf{a}_i, \mathbf{b}_i), \quad i = 1, \dots, s.$$

- $f: \bigtimes_{i=1}^s \mathbb{R}^{m_i \times n_i} \rightarrow \mathbb{R}$ : 目标函数, 可微未必凸.
- $\mathcal{U}(\mathbf{a}_i, \mathbf{b}_i) \subseteq \mathbb{R}^{m_i \times n_i}$ : 运输多胞体.

$$\mathcal{U}(\mathbf{a}_i, \mathbf{b}_i) := \left\{ T \in \mathbb{R}_+^{m_i \times n_i} : T \mathbf{1}_{n_i} = \mathbf{a}_i, \quad T^\top \mathbf{1}_{m_i} = \mathbf{b}_i \right\}.$$

- 应用: **强关联电子体系计算**、Wasserstein 重心计算 [Bigot-Klein 2012; Carlier et al. 2015]、**标签分布学习** [Geng 2016; Zhao-Zhou 2018] 等.

# 已有工作—块坐标下降 (BCD) 型算法

- 块坐标下降 (BCD) 算法.

[Wright 2015; Bertsekas 2016; Fercoq-Richtárik 2016; Hong et al. 2017; ...]

- 块条件梯度 (BCG) 算法.

[Lacoste-Julien et al. 2013; Beck et al. 2015; Braun et al. 2022; ...]

- 邻近交替线性化极小化 (PALM) 算法.

[Razaviyayn et al. 2013; Xu-Yin 2013; Bolte et al. 2014; Ochs 2019; ...]

- 随机版本 (梯度计算或块更新次序).

[Lacoste-Julien et al. 2013; Beck et al. 2015; Chen et al. 2019; Xiao et al. 2019;

Sun et al. 2020; Driggs et al. 2021; Hertrich-Steidl 2022; ...]

求解大规模问题 **存储和计算代价过高**.

# 已有工作—求解最优运输问题的算法

仿射目标 + 单块  $\Rightarrow$  最优运输问题 [Monge 1781; Kantorovich 1942; Villani 2021]

- 传统方法: 求解 Monge-Ampère 方程 (特殊费用函数)、线性规划算法 (立方计算复杂度). [Brenier 1997; Rubner et al. 1997; Pele-Werman 2009; ...]
- 多重网格型方法 (特殊费用函数、费用矩阵存储).  
[Mérigot 2011; Schmitzer 2016; Gerber-Maggioni 2017; Schmitzer 2019; Liu-Wen-Yin 2022; ...]
- 基于熵正则的方法.
  - Sinkhorn 算法及部分更新变体 (单步平方计算复杂度、全矩阵存储).  
[Cuturi 2013; Altschuler et al. 2017; Lin et al. 2019; Alaya et al. 2019; Lin et al. 2022; ...]
  - 结合随机采样的 Sinkhorn 算法变体: 复杂度更低. [Li et al. 2023]

主要贡献: 结合随机采样的 BCD 型算法及其理论性质.

# 熵正则交替线性化极小化 (ERALM) 算法

ERALM = BCG + 熵正则.

步 1. 初始化  $(X_1^{(0)}, \dots, X_s^{(0)}) \in \bigtimes_{i=1}^s \mathbb{R}^{m_i \times n_i}$ . 置  $t := 0$ .

步 2. 对  $i = 1, \dots, s$ :

步 2.1 选取正则化参数  $\lambda_{i,t} > 0$  与步长  $\alpha_{i,t} \in (0, 1]$ .

步 2.2 求解如下子问题得到  $\tilde{X}_i^{(t+1)} \in \mathbb{R}^{m_i \times n_i}$ :

$$\begin{aligned} \min_{X_i} \quad & \left\langle \nabla_{X_i} f(X_{<i}^{(t+1)}, X_{\geq i}^{(t)}), X_i - X_i^{(t)} \right\rangle + \lambda_{i,t} h(X_i), \\ \text{s. t.} \quad & X_i \mathbf{1}_{n_i} = \mathbf{a}_i, \quad X_i^\top \mathbf{1}_{m_i} = \mathbf{b}_i. \end{aligned}$$

步 2.3 更新变量块  $X_i^{(t+1)} := (1 - \alpha_{i,t}) X_i^{(t)} + \alpha_{i,t} \tilde{X}_i^{(t+1)} \in \mathbb{R}^{m_i \times n_i}$ .

步 3. 检查终止准则是否满足. 若是, 停机; 若否, 置  $t := t + 1$ , 转步 2.

- $h(T) := \sum_{ij} t_{ij}(\log t_{ij} - 1)$ : 矩阵  $T = (t_{ij}) \in \mathbb{R}_+^{m \times n}$  的负熵. [Wilson 1969]

# ERALM 算法 (续)

## 求解子问题的 Sinkhorn 算法 [Sinkhorn-Knopp 1967]

$$\boxed{\begin{aligned}\check{\mathbf{u}}_i^{(t,l+1)} &:= \mathbf{a}_i \oslash \left( \Psi_i^{(t)} \check{\mathbf{v}}_i^{(t,l)} \right), \quad \check{\mathbf{v}}_i^{(t,l+1)} := \mathbf{b}_i \oslash \left( \Psi_i^{(t)\top} \check{\mathbf{u}}_i^{(t,l+1)} \right), \\ \tilde{X}_i^{(t,l+1)} &:= \text{Diag}(\check{\mathbf{u}}_i^{(t,l+1)}) \Psi_i^{(t)} \text{Diag}(\check{\mathbf{v}}_i^{(t,l+1)}).\end{aligned}}$$

- “ $\oslash$ ”: 向量逐元素相除. “ $\exp(\cdot)$ ”: 逐元素取自然指数.
- $\Psi_i^{(t)} := \exp\left(-\nabla_{X_i} f(X_{<i}^{(t+1)}, X_{\geq i}^{(t)})/\lambda_{i,t}\right) \in \mathbb{R}^{m_i \times n_i}$ : 子问题的核矩阵.
- **单步计算复杂度**  $\mathcal{O}(m_i n_i)$ , 可扩展,  $R$ —线性收敛速度 [Luo-Tseng 1993].

## 子问题的最优解

$$\boxed{\tilde{X}_i^{(t+1,\star)} := \text{Diag}\left(\exp\left(\frac{\tilde{\mathbf{u}}_i^{(t,\star)}}{\lambda_{i,t}}\right)\right) \Psi_i^{(t)} \text{Diag}\left(\exp\left(\frac{\tilde{\mathbf{v}}_i^{(t,\star)}}{\lambda_{i,t}}\right)\right) \in \mathbb{R}^{m_i \times n_i}.}$$

- $(\tilde{\mathbf{u}}_i^{(t,\star)}, \tilde{\mathbf{v}}_i^{(t,\star)}) \in \mathbb{R}^{m_i} \times \mathbb{R}^{n_i}$ : 子问题对偶问题的一个最优解.
- **零元素的继承性:**  $\psi_{i,jk}^{(t)} = 0 \Rightarrow \tilde{x}_{i,jk}^{(t+1,\star)} = 0$ .

# ERALM 算法的采样版本 (S-ERALM 算法)

## 矩阵逐元素随机近似

用 **稀疏矩阵**  $\hat{\Psi}_i^{(t)} = (\hat{\psi}_{i,jk}^{(t)})_{jk} \in \mathbb{R}^{m_i \times n_i}$  代替  $\Psi_i^{(t)} = (\psi_{i,jk}^{(t)})_{jk}$ , 其中

$$\hat{\psi}_{i,jk}^{(t)} := \begin{cases} \psi_{i,jk}^{(t)} / p_{i,jk}^{(t)\#}, & \text{以概率 } p_{i,jk}^{(t)\#} := \min\{1, \hat{n}_i \cdot p_{i,jk}^{(t)}\}; \\ 0, & \text{以概率 } 1 - p_{i,jk}^{(t)\#}. \end{cases}$$

- $\hat{n}_i \in \mathbb{N}$ : 采样参数.
- 重要性采样概率  $P_i^{(t)} = (p_{i,jk}^{(t)})_{jk} \in \mathbb{R}^{m_i \times n_i}$ :

$$p_{i,jk}^{(t)} := \gamma \frac{x_{i,jk}^{(t)}}{\sum_{j' k'} x_{i,j' k'}^{(t)}} + (1 - \gamma) \frac{\sqrt{a_{i,j} b_{i,k}}}{\sum_{j' k'} \sqrt{a_{i,j'} b_{i,k'}}}, \quad j = 1, \dots, m_i, \quad k = 1, \dots, n_i,$$

其中  $\gamma \in [0, 1]$  为插值因子. [Liu 1996/2004; Owen 2013; Ma et al. 2014; ...]

## ERALM 算法 + 随机采样 = S-ERALM 算法

采样到的指标集:  $\mathcal{I}_i^{(t)} \subseteq \{(j, k) : j = 1, \dots, m_i, k = 1, \dots, n_i\}$ .

- 仅需计算  $\nabla_{X_i} f(X_{<i}^{(t+1)}, X_{\geq i}^{(t)})$  中的  $|\mathcal{I}_i^{(t)}|$  个元素.
- Sinkhorn 算法: 单步计算复杂度  $\mathcal{O}(|\mathcal{I}_i^{(t)}|)$ .

# 基于 Kullback-Leibler (KL) 散度的交替线性化极小化 (KLALM) 算法

**KLALM = PALM + KL 散度.**

步 1. 初始化  $(X_1^{(0)}, \dots, X_s^{(0)}) \in \times_{i=1}^s \mathbb{R}^{m_i \times n_i}$ . 置  $t := 0$ .

步 2. 对  $i = 1, \dots, s$ :

步 2.1 选取邻近参数  $\mu_{i,t} > 0$ .

步 2.2 求解如下子问题得到  $X_i^{(t+1)} \in \mathbb{R}^{m_i \times n_i}$ :

$$\begin{aligned} \min_{X_i} \quad & \left\langle \nabla_{X_i} f(X_{<i}^{(t+1)}, X_{\geq i}^{(t)}), X_i - X_i^{(t)} \right\rangle + \mu_{i,t} \text{KL}(X_i; X_i^{(t)}), \\ \text{s. t.} \quad & X_i \mathbf{1}_{n_i} = \mathbf{a}_i, \quad X_i^\top \mathbf{1}_{m_i} = \mathbf{b}_i. \end{aligned}$$

步 3. 检查终止准则是否满足. 若是, 停机; 若否, 置  $t := t + 1$ , 转步 2.

- $\text{KL}(T; T') := \sum_{ij} [t_{ij}(\ln t_{ij} - \ln t'_{ij}) - (t_{ij} - t'_{ij})]$ : 矩阵  $T = (t_{ij}) \in \mathbb{R}_+^{m \times n}$  与  $T' = (t'_{ij}) \in \mathbb{R}_+^{m \times n}$  的 KL 散度. [Kullback-Leibler 1951]

## 使用 Sinkhorn 算法求解子问题

$$\boxed{\begin{aligned}\check{\mathbf{u}}_i^{(t,l+1)} &:= \mathbf{a}_i \oslash \left( \Phi_i^{(t)} \check{\mathbf{v}}_i^{(t,l)} \right), \quad \check{\mathbf{v}}_i^{(t,l+1)} := \mathbf{b}_i \oslash \left( \Phi_i^{(t)\top} \check{\mathbf{u}}_i^{(t,l+1)} \right), \\ X_i^{(t,l+1)} &:= \text{Diag} \left( \check{\mathbf{u}}_i^{(t,l+1)} \right) \Phi_i^{(t)} \text{Diag} \left( \check{\mathbf{v}}_i^{(t,l+1)} \right).\end{aligned}}$$

- “ $\odot$ ”: 矩阵 Hadamard 积.
- $\Phi_i^{(t)} := \exp(-\nabla_{X_i} f(X_{<i}^{(t+1)}, X_{\geq i}^{(t)})/\mu_{i,t}) \odot X_i^{(t)} \in \mathbb{R}^{m_i \times n_i}$ : 子问题的核矩阵.
- **单步计算复杂度  $\mathcal{O}(m_i n_i)$ .**

## 子问题的最优解

- 零元素的继承性:  $x_{i,jk}^{(t)} = 0 \Rightarrow \phi_{i,jk}^{(t)} = 0 \Rightarrow x_{i,jk}^{(t+1,\star)} = 0$ .

**KLALM 算法 + 第  $\hat{t} \in \mathbb{N}$  步采样 = S-KLALM 算法**

- 仅需计算  $\nabla_{X_i} f(X_{<i}^{(t+1)}, X_{\geq i}^{(t)})$  中的  $|\mathcal{I}_i^{(\hat{t})}|$  个元素.
- Sinkhorn 算法: 单步计算复杂度  $\mathcal{O}(|\mathcal{I}_i^{(\hat{t})}|)$ .

# S-ERALM 算法的理论性质

## 条件 9 (目标函数的光滑性)

目标函数  $f$  在  $\bigtimes_{i=1}^s \mathcal{U}(\mathbf{a}_i, \mathbf{b}_i)$  上是分块 Lipschitz 光滑的, 即存在  $L \geq 0$ , 使得对任意  $i = 1, \dots, s$  和  $X, X' \in \bigtimes_{i=1}^s \mathcal{U}(\mathbf{a}_i, \mathbf{b}_i)$ , 有  $\|\nabla_{X_i} f(X) - \nabla_{X'_i} f(X')\| \leq L \|X - X'\|$ .

## 条件 10 (问题数据的渐进性)

存在  $\underline{f} \in \mathbb{R}$ 、 $q > 0$ 、 $\theta \geq 0$  和  $\xi \geq 0$ , 使得以下条件对任意  $\{m_i\}_{i=1}^s, \{n_i\}_{i=1}^s \subseteq \mathbb{N}$  都成立:

- (1) 问题的最优值以  $\underline{f}$  为下界.
- (2)  $\mathbf{a}_i^\top \mathbf{1}_{m_i} = \mathbf{b}_i^\top \mathbf{1}_{n_i} = 1$ ,  $\max_j a_{i,j} \leq q \cdot \min_j a_{i,j}$ ,  $\max_k b_{i,k} \leq q \cdot \min_k b_{i,k}$ .
- (3)  $L = \mathcal{O}(\sum_{i=1}^s (m_i + n_i)^\theta)$ .
- (4)  $\max_{i=1}^s (m_i + n_i) / \min_{i=1}^s (m_i + n_i) \leq \xi$ .

## 条件 11 (采样的充分性)

- (1) 存在常数  $\nu \in (1/2, 1]$ ,  $c_1, c_2, \hat{c}_2 > 0$ , 使得对任意  $0 \leq t \leq t_{\max}$  和  $i = 1, \dots, s$ ,

$$\left\| \Psi_i^{(t)} \right\|_2 \geq \frac{(m_i + n_i)^\nu}{c_1}, \quad \kappa(\Psi_i^{(t)}) \leq c_2, \quad \kappa(\hat{\Psi}_i^{(t)}) \leq \hat{c}_2;$$

- (2) 插值因子  $\gamma$  小于 1, 且存在  $\varepsilon > 0$ , 使得对  $i = 1, \dots, s$ ,

$$\frac{1}{\max_{j,k,t} p_{i,jk}^{(t)}} \geq \hat{n}_i \geq \frac{8(m_i + n_i)^{1-2\nu} \ln^4(m_i + n_i)}{(1-\gamma) \ln^4(1+\varepsilon)} \cdot \frac{\sum_{j',k'} \sqrt{a_{i,j'} b_{i,k'}}}{\min_{j,k} \sqrt{a_{i,j} b_{i,k}}}.$$

## 主要创新点

### 定理 12 (S-ERALM 算法的渐进性质)

假设条件 9 与 10 成立. 令  $\{X^{(t)}\}$  为 S-ERALM 算法产生的迭代点序列, 其中

$$t_{\max} = \Theta\left(\sum_{i=1}^s (m_i + n_i)^\eta\right) \text{ 满足 } t_{\max} \geq \frac{f(X^{(0)}) - \underline{f}}{2\bar{d}^2 s L(2\sqrt{s} + 1)}, \quad f(X^{(0)}) \leq M,$$

$$\alpha_{i,t} \equiv \frac{1}{\bar{d}} \sqrt{\frac{f(X^{(0)}) - \underline{f}}{2sL(2\sqrt{s} + 1)t_{\max}}}, \quad \lambda_{i,t} \equiv o\left(\frac{1}{\sum_{i=1}^s \sqrt{m_i n_i} \ln(m_i + n_i)}\right),$$

$$\hat{n}_i = \Theta\left(\frac{m_i n_i}{(m_i + n_i)^{2\nu-1}} \ln^4(m_i + n_i)\right)$$

对于  $1 \leq i \leq s$  和  $0 \leq t \leq t_{\max}$  均成立, 子问题可行且被精确求解, 条件 11 成立,  $c_1$ 、 $c_2$ 、 $\hat{c}_2$ 、 $\varepsilon$ 、 $\eta (> \theta)$ 、 $\gamma$ 、 $\nu$ 、 $\xi$ 、 $M$  与  $\{m_i\}_{i=1}^s$ 、 $\{n_i\}_{i=1}^s$  无关. 则随着  $\sum_{i=1}^s (m_i + n_i)$  趋于无穷大, **以概率趋于 1**,

$$\frac{1}{t_{\max}} \sum_{t=0}^{t_{\max}-1} R(X^{(t)}) \rightarrow 0,$$

其中, 残差函数  $R$  度量了给定点处的 KKT 违反度.

# 数值实验—强关联电子体系计算

通过求解问题  $(P_\beta)$  ( $X_i := \text{Diag}(\varrho) Y_i$ ) 模拟强关联电子体系.

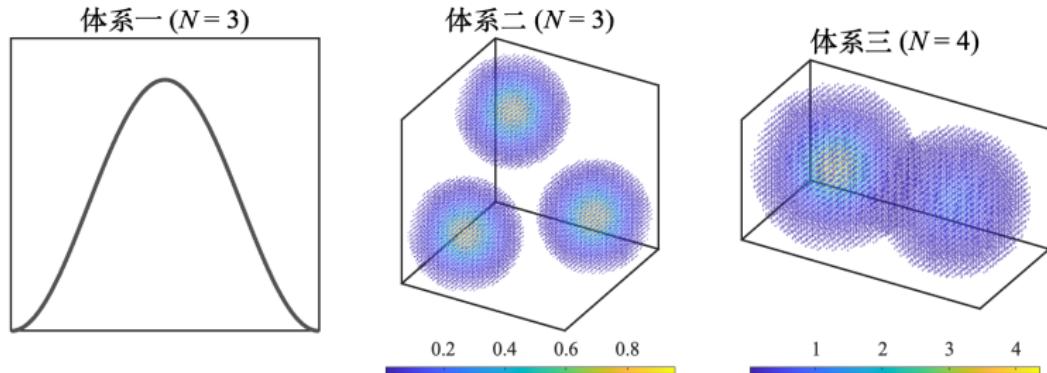


图 7: 待模拟体系单电子密度可视化. 对于三维体系, 仅展示单电子密度大于 0.01 的区域.

问题设置 体系一, 等质量剖分, 罚参数  $\beta = 1$ .

## 比较算法及其设置

- PALM-I 算法:
  - 邻近参数:  $\sigma_{i,t} \equiv 2/\ln(K)$ .
  - 半光滑 Newton 法求解子问题. [Li et al. 2020]
  - 子问题非精确求解准则 (1),  $\varepsilon_k \equiv 10^{-6}$ .
- ERALM、S-ERALM、KLALM、S-ERALM 算法:
  - 参数:  $\hat{t} = 0$ ,  $\hat{n}_i \equiv \lfloor K^{1.5} \rfloor$ ,  $\alpha_{i,t} = (t+1)^{-0.75}$ ,  $\lambda_{i,t} = \mu_{i,t} \equiv 2/\ln(K)$ .
  - Sinkhorn 算法求解子问题.
  - 子问题非精确求解准则:  $\|X_i^{(t+1)}\mathbf{1}_K - \varrho\|_\infty < 10^{-6}$ .
- 终止准则:

$$\frac{1}{N-1} \sum_{i=2}^N \|\text{Diag}(\varrho)^{-1}(X_i^{(t)} - X_i^{(t-1)})\| \leq tol.$$

## 算法比较具体设置

- 离散规模  $K = 90$ .
- 参数:  $\lambda_{i,t} = \mu_{i,t} \equiv 2\sigma / \ln(K)$  ( $\sigma \in \{2^0, 2^2, \dots, 2^8\}$ ),  $tol = 5 \times 10^{-3} / \sqrt{\sigma}$ .

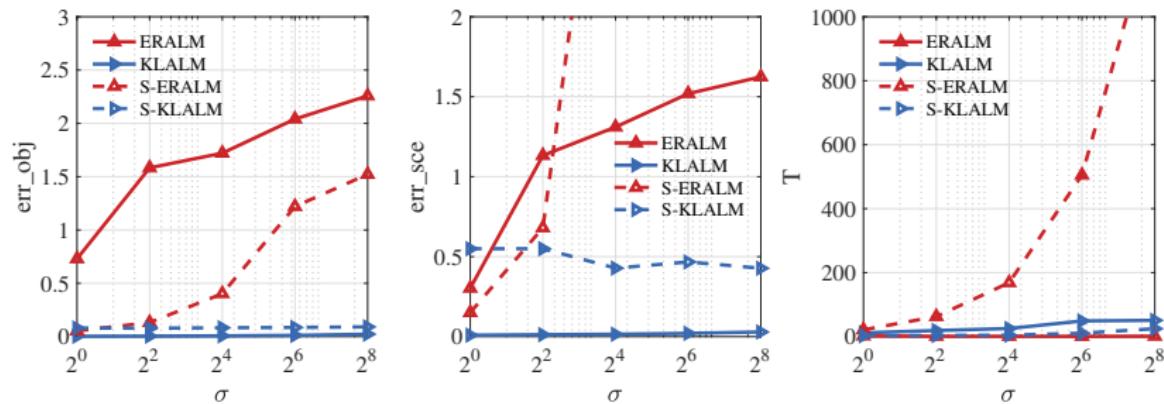


图 8: 在不同  $\sigma$  下, 不同算法在体系一上 10 次模拟平均 err\_obj、err\_sce 和 T.

$$\text{err\_obj} := \left| \frac{f_{\text{out}} - f^*}{f^*} \right| \quad (f^* : \text{最优值}), \quad \text{err\_sce} := \frac{\|\hat{\lambda} - \lambda^*\|_\infty}{\|\lambda^*\|_\infty} \quad (\lambda^* : \text{精确 SCE 向量}).$$

KLALM 和 S-KLALM 算法的数值表现对于参数的选择更加鲁棒.

## 算法比较具体设置

- 离散规模  $K \in \{90, 180, \dots, 2880\}$ .
- 终止参数:  $tol = 5 \times 10^{-3}$ .

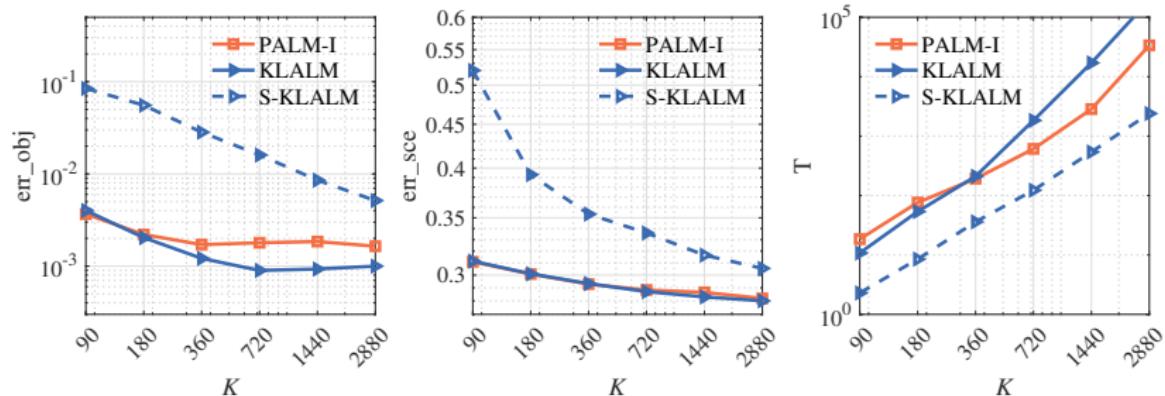


图 9: 在不同  $K$  下, 不同算法在体系一上 10 次模拟平均  $err\_obj$ 、 $err\_sce$  和  $T$ .

KLALM 算法 **误差最小**, S-KLALM 算法 **标度最好**.

## 三维体系模拟具体设置

- 罚参数  $\beta = 1.$
- CMGOPT 框架: 均匀剖分 + 一致加密 + KLALM + S-KLALM.
- 终止参数:

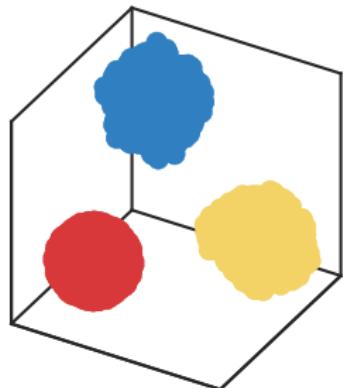
$$tol = \begin{cases} 5 \times 10^{-3}, & \text{初始网格;} \\ 10^{-2} \times (\sqrt{2^d})^{\log_2(K_\ell/K_0)}, & \text{第 } \ell \text{ 层网格,} \end{cases}$$

其中  $K_\ell \in \mathbb{N}$  为第  $\ell \in \mathbb{N}$  层网格的离散规模.

表 6: 三维体系在每层网格上的目标函数值.

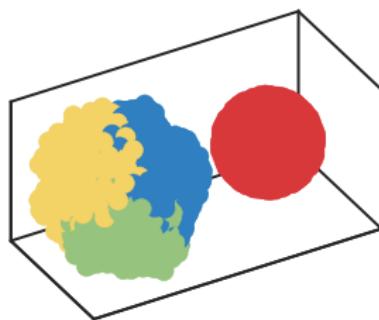
$\ell$	体系二		体系三	
	$K$	$f_{\text{out}}$	$K$	$f_{\text{out}}$
0	1728	1.0202	1000	4.6193
1	13824	1.0209	8000	4.6716
2	110592	1.0209	64000	4.6833

# 数值实验—强关联电子体系计算 (续)



(a) 体系二

●  $\tilde{\Omega}$   
●  $\hat{T}_2(\tilde{\Omega})$   
●  $\hat{T}_3(\tilde{\Omega})$



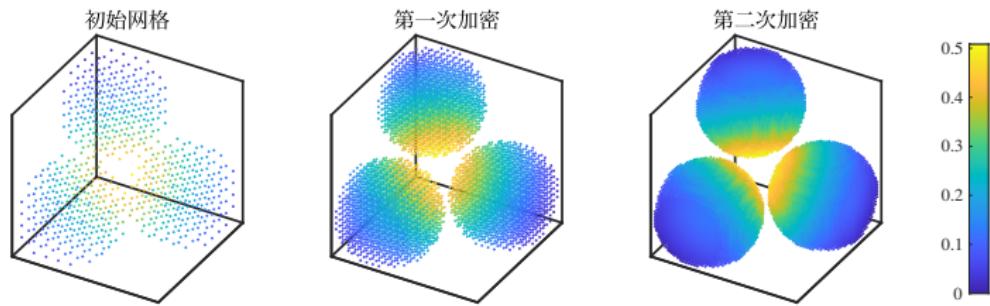
(b) 体系三

●  $\tilde{\Omega}$   
●  $\hat{T}_2(\tilde{\Omega})$   
●  $\hat{T}_3(\tilde{\Omega})$   
●  $\hat{T}_4(\tilde{\Omega})$

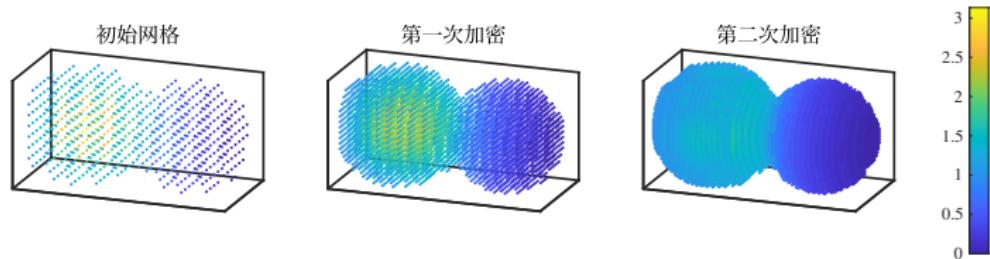
图 10: 三维体系在最后一层网格上的近似映射  $\{\hat{\mathcal{T}}_i\}_{i=2}^N$  切片.

首次可视化了三维情形下电子位置之间的近似映射.

# 数值实验—强关联电子体系计算 (续)



(a) 体系二



(b) 体系三

图 11: 三维体系在每层网格上的  $\hat{\lambda}$ .

## 二、第一性原理晶体结构弛豫中的 优化问题、理论与算法

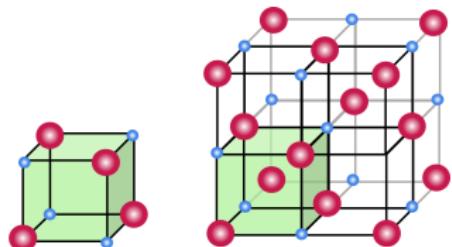


图 12: 晶体结构与单胞.<sup>(3)</sup>

- “自然界的固体物质绝大多数是晶体”.  
[Huang-Han 1998]
- 原子  $\xrightarrow{\text{周期排列}}$  晶体结构.
- 单胞: 晶体结构最小代表元.

## 描述晶体结构所需自由度

- 单胞中  $M \in \mathbb{N}$  个原子的位置  $R := [R_1, \dots, R_M] \in \mathbb{R}^{3 \times M}$ .
- 单胞的三个晶格基矢  $A := [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] \in \mathbb{R}^{3 \times 3}$ .

$$R_m + \mathbb{Z}\mathbf{a}_1 + \mathbb{Z}\mathbf{a}_2 + \mathbb{Z}\mathbf{a}_3, \quad m = 1, \dots, M.$$

<sup>(3)</sup> 图片来源: <https://www.expii.com/t/crystal-lattice-structure-formation-7999>.

# 工作四：求解行列式约束优化问题的算法

## 第一性原理固定晶格体积晶体结构弛豫

$$\min_{R, A} E(R, A), \text{ s. t. } \det(A) = V.$$

- $E : \mathbb{R}^{3 \times M} \times \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$ : 第一性原理能量泛函.  $V > 0$ : 预设体积.
- 应用: 材料物态方程计算  $\Rightarrow$  热力学性质. [Kittel 2005; Giannozzi et al. 2009; ...]
- 常用算法: 效率较低、无理论保证.
  - 非线性共轭梯度 (CG) 算法.  
[Hestenes-Stiefel 1952; Fletcher-Reeves 1964; Polak-Ribiere 1969; Polyak 1969; Dai-Yuan 1999; ...]
  - 拟 Newton (QN) 算法.  
[Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970; Pulay 1982; ...]

缩放算子:  $\mathcal{P}_V(X) := \sqrt[n]{V / \det(X)} X$  保持可行性.

## 一般行列式约束优化问题

$$\min_X f(X), \text{ s. t. } \det(X) = V.$$

主要贡献: 使用缩放算子的高效算法及其理论性质.

# 投影梯度下降 (PGD) 算法

步 1. 初始化  $X^{(0)} \in \mathbb{R}^{n \times n}$ :  $\det(X^{(0)}) = V$ ,  $\bar{\alpha}, \underline{\alpha}, \eta > 0$ ,  $\delta \in (0, 1)$ . 置  $k := 0$ ,  $\bar{f}_0 := f(X^{(0)})$ ,  $q_0 := 1$ .

步 2. 计算负梯度到切空间  $\mathcal{T}^{(k)}$  上的正交投影

$$D^{(k)} := \mathcal{P}_{\mathcal{T}^{(k)}}(-\nabla f(X^{(k)})) \in \mathbb{R}^{n \times n}.$$

步 3. 选取初始试探步长  $\alpha_{k-1,0} \in [\underline{\alpha}, \bar{\alpha}]$ . 置  $l := 0$ .

步 4. 更新  $\tilde{X}^{(k,l)} := X^{(k)} + \alpha_{k-1,l} D^{(k)}$ ,  $X^{(k,l)} := \mathcal{P}_V(\tilde{X}^{(k,l)})$ .

步 5. 检查是否满足非单调线搜索准则 [Zhang-Hager 2004]

$$f(X^{(k,l)}) \leq \bar{f}_k - \eta \cdot \alpha_{k-1,l} \|D^{(k)}\|^2.$$

若是, 转步 6; 若否, 置  $\alpha_{k-1,l+1} := \delta \cdot \alpha_{k-1,l}$ ,  $l := l + 1$ , 转步 4.

步 6. 更新  $X^{(k+1)} := X^{(k,l)}$  以及代理序列

$$\bar{f}_{k+1} := \frac{\bar{f}_k + \mu_k q_k f(X^{(k+1)})}{1 + \mu_k q_k}, \quad q_{k+1} := 1 + \mu_k q_k, \quad \mu_k \in [0, 1].$$

步 7. 检查终止准则是否满足. 若是, 停机; 若否, 置  $k := k + 1$ , 转步 2.

主要创新点:  $f(\mathcal{P}_V(\tilde{X}^{(k,l)})) - f(\tilde{X}^{(k,l)}) \leq \mathcal{O}(\|\tilde{X}^{(k,l)} - X^{(k)}\|^2)$ .

## 条件 13 (目标函数的光滑性)

目标函数  $f$  局部 Lipschitz 连续可微且在可行域上是强制的, 即

$$\lim_{\substack{\|X\|_2 \rightarrow \infty \\ \det(X) = V}} f(X) = \infty.$$

## 定理 14 (PGD 算法的全局依子列收敛性)

假设条件 13 成立. 令  $\{X^{(k)}\}$  为 PGD 算法产生的迭代点序列, 其中  $\{\mu_k\} \subseteq (0, 1]$  且  $\inf_k \mu_k > 0$ . 则  $\{X^{(k)}\}$  至少有一个聚点且其每个聚点都是问题的 KKT 点.

# 求解第一性原理固定晶格体积晶体结构弛豫的 PGD 算法

$$F_{\text{atom}}^{(k)} := -\nabla_R E(R^{(k)}, A^{(k)}), F_{\text{latt}}^{(k)} := -\nabla_A E(R^{(k)}, A^{(k)}).$$

步 1. 初始化.....

步 2. 计算搜索方向

$$D_{\text{atom}}^{(k)} := F_{\text{atom}}^{(k)} \in \mathbb{R}^{3 \times M}, \quad D_{\text{latt}}^{(k)} := \mathcal{P}_{\mathcal{T}^{(k)}}(F_{\text{latt}}^{(k)}) \in \mathbb{R}^{3 \times 3}.$$

步 3. 选取初始试探步长  $\alpha_{\text{atom}}^{(k-1,0)}$  和  $\alpha_{\text{latt}}^{(k-1,0)}$ , 置  $l := 0$ .

步 4. 更新

$$R^{(k,l)} := R^{(k)} + \alpha_{\text{atom}}^{(k-1,l)} D_{\text{atom}}^{(k)}, \quad A^{(k,l)} := \mathcal{P}_V(A^{(k)} + \alpha_{\text{latt}}^{(k-1,l)} D_{\text{latt}}^{(k)}).$$

步 5. 检查是否满足非单调搜索准则

$$E(R^{(k,l)}, A^{(k,l)}) \leq \bar{E}_k - \eta \left( \alpha_{\text{atom}}^{(k-1,l)} \left\| D_{\text{atom}}^{(k)} \right\|^2 + \alpha_{\text{latt}}^{(k-1,l)} \left\| D_{\text{latt}}^{(k)} \right\|^2 \right).$$

若是, 转步 6; 若否, 更新  $\alpha_{\text{atom}}^{(k-1,l+1)}$  和  $\alpha_{\text{latt}}^{(k-1,l+1)}$ , 置  $l := l + 1$ , 转步 4.

步 6. 更新  $R^{(k+1)} := R^{(k,l)}$  和  $A^{(k+1)} := A^{(k,l)}$  以及代理序列.....

步 7. 检查终止准则是否满足. 若是, 停机; 若否, 置  $k := k + 1$ , 转步 2.

## 比较算法及其设置<sup>(4)</sup>

- CG 算法: Polak-Ribiere-Polyak 共轭参数 + Wolfe 线搜索准则.
- QN 算法: 第二类 Broyden 方法.
- PGD 算法: 截断交替 Barzilai-Borwein (ABB) 步长 [Barzilai-Borwein 1988; Dai-Fletcher 2005]

$$\alpha_{\text{atom}}^{(k-1,0)} = \max \left\{ \min \left\{ \left| \alpha_{\text{atom}, \text{ABB}}^{(k)} \right|, 10 \right\}, 10^{-5} \right\},$$

$$\alpha_{\text{latt}}^{(k-1,0)} = \max \left\{ \min \left\{ \left| \alpha_{\text{latt}, \text{ABB}}^{(k)} \right|, 10^{-1} \right\}, 10^{-7} \right\},$$

$$\alpha_{\text{atom}, \text{ABB}}^{(k)} := \begin{cases} \frac{\| R^{(k)} - R^{(k-1)} \|^2}{\langle R^{(k)} - R^{(k-1)}, D_{\text{atom}}^{(k-1)} - D_{\text{atom}}^{(k)} \rangle}, & \text{若 } \text{mod}(k, 2) = 0; \\ \frac{\langle R^{(k)} - R^{(k-1)}, D_{\text{atom}}^{(k-1)} - D_{\text{atom}}^{(k)} \rangle}{\| D_{\text{atom}}^{(k-1)} - D_{\text{atom}}^{(k)} \|^2}, & \text{若 } \text{mod}(k, 2) = 1, \end{cases}$$

$\alpha_{\text{latt}, \text{ABB}}^{(k)}$  定义类似,  $\eta = 10^{-4}$ ,  $\delta_{\text{atom}} = 0.1$ ,  $\delta_{\text{latt}} = 0.5$ .

- 终止准则:
  - 原子受力不超过  $0.01 \text{ eV}/\text{\AA}$  且晶格偏应力不超过  $M \times 0.01 \text{ eV}/\text{\AA}$ .
  - KSDFT 求解次数 (#KS) 超过 1000.

---

<sup>(4)</sup> 我们直接调用软件 Correlated Electron System Simulation Package [Fang et al. 2016; Gao et al. 2017; Zhou et al. 2018; Fang et al. 2019] 中实现的 CG 与 QN 算法. 第一性原理计算均采用默认设置.

## 数值实验—算法比较 (续)

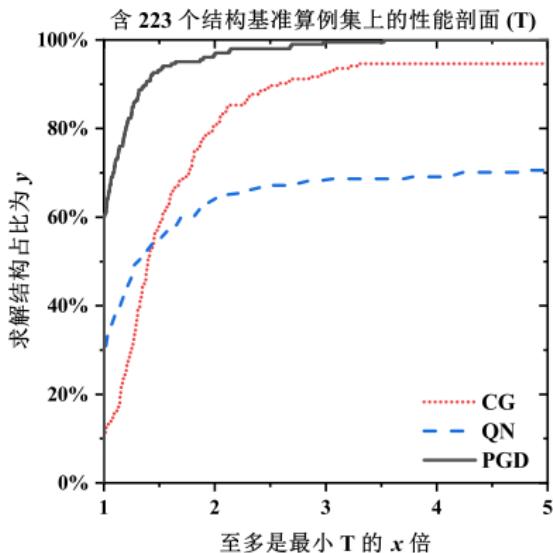
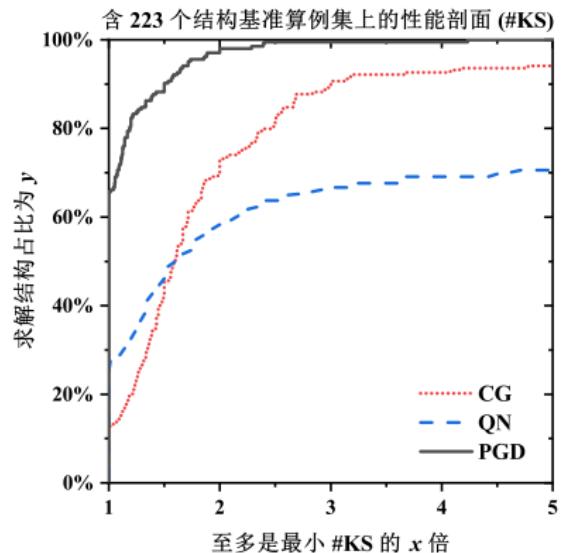


图 13: CG、QN、PGD 算法在基准算例集上的性能剖面 [Dolan-Moré 2002].

- 运行时间加速比: PGD:CG = 1.41, PGD:QN = 1.45.
- 无法正常终止结构数: CG:11, QN:56, PGD: 0.

# 数值实验—高熵合金物态方程计算

## 静态三阶 Birch-Murnaghan (BM3) 物态方程 (能量—体积关系) [Birch 1947]

$$E(V) = E_0 + \frac{9V_0B_0}{16} \left\{ \left[ \left( \frac{V_0}{V} \right)^{2/3} - 1 \right]^3 B'_0 + \left[ \left( \frac{V_0}{V} \right)^{2/3} - 1 \right]^2 \left[ 6 - 4 \left( \frac{V_0}{V} \right)^{2/3} \right] \right\},$$

$V_0, E_0, B_0, B'_0$ : 平衡体积及其对应的能量、**体模量**与体模量对压强的导数(待拟合参数).

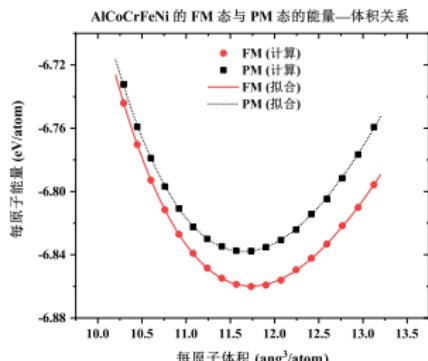


图 14: AlCoCrFeNi 的铁磁 (FM) 态与顺磁 (PM) 态的能量—体积关系.

表 7: 拟合的静态 BM3 物态方程中的参数.

数据来源	$V_0$ ( $\text{\AA}^3/\text{atom}$ )	$E_0$ (eV/atom)	$B_0$ (GPa)	$B'_0$
FM ([Wu et al. 2020])	11.77	-6.85	157.32	8.9
FM (PGD)	11.74	-6.86	158.49	5.3
PM ([Wu et al. 2020])	11.70	-6.82	161.64	5.3
PM (PGD)	11.65	-6.84	167.59	4.7

拟合的力学参数与已有结果相合.

# 数值实验—高熵合金物态方程计算 (续)

## 静态 BM3 物态方程 (压强—体积关系) [Birch 1947]

$$P(V) = \frac{3B_0}{2} \left[ \left( \frac{V_0}{V} \right)^{7/3} - \left( \frac{V_0}{V} \right)^{5/3} \right] \left\{ 1 + \frac{3}{4}(B'_0 - 4) \left[ \left( \frac{V_0}{V} \right)^{2/3} - 1 \right] \right\}.$$

静态 BM3 物态方程 + 修正平均场势方法 [Song-Liu 2007]  $\Rightarrow$  有限温 BM3 物态方程.

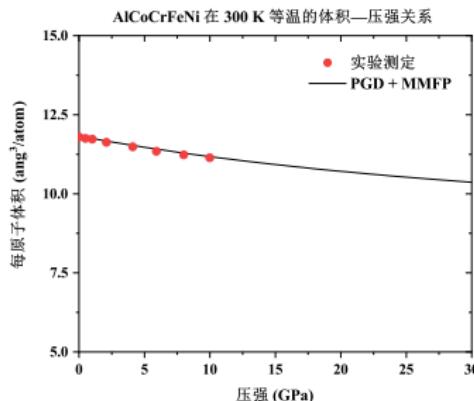


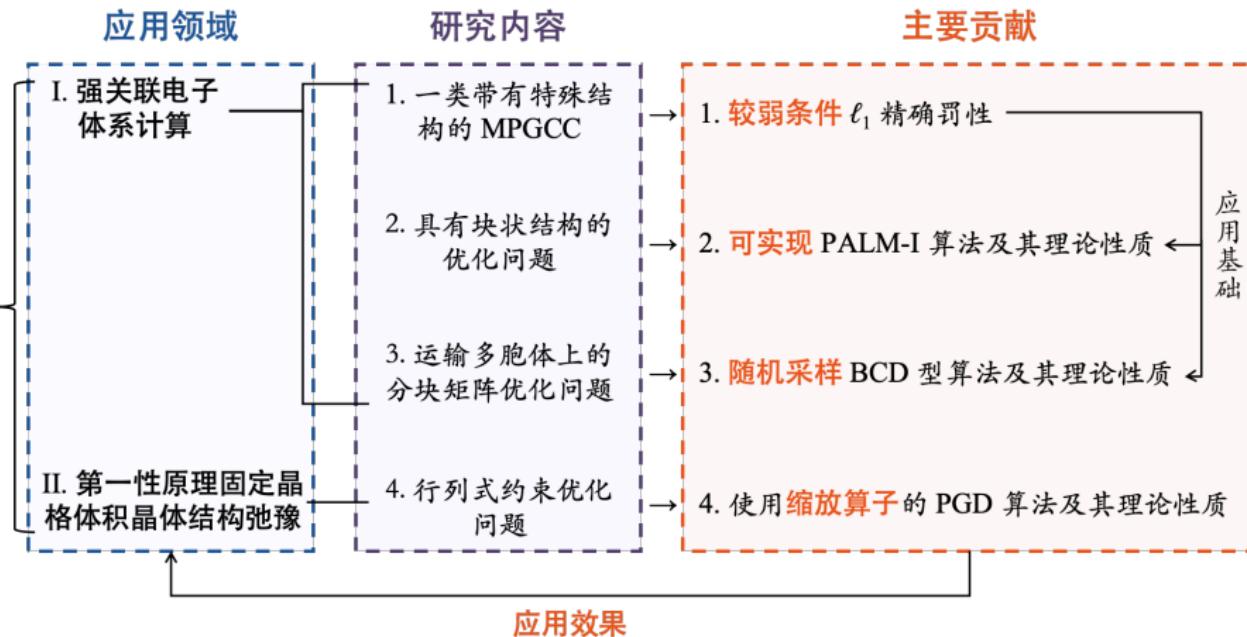
图 15: AlCoCrFeNi 在 300 K 等温的体积—压强关系, 其中“实验测定”表示使用 X 射线衍射与金刚石压砧实验测定的结果 [Cheng et al. 2019].

估计体积与实验测定值**最大偏差不超过 0.7%**.

### 三、总结与展望

# 总结

计算材料科学



## 应用结果

- 首次可视化了二维、三维情形电子位置间映射.
- 计算了高熵合金的物态方程, 拟合的热力学参数符合实验测定值.

- 优化理论与算法的改进和推广.
- 计算材料科学中的其他优化问题(不同模型、跨尺度).

# 在学期间的研究成果

## 学术论文

- [1] **Hu Y**, Chen H, Liu X. A global optimization approach for multimarginal optimal transport problems with Coulomb cost [J]. *SIAM Journal on Scientific Computing*, 2023, 45(3): A1214-A1238.
- [2] **Hu Y**, Liu X. The convergence properties of infeasible inexact proximal alternating linearized minimization [J]. *Science China Mathematics*, 2023, 66(10): 2385-2410.
- [3] **Hu Y**, Liu X. The exactness of the  $\ell_1$  penalty function for a class of mathematical programs with generalized complementarity constraints [J]. *Fundamental Research*, 2023, published online. DOI: 10.1016/j.fmre.2023.04.006.
- [4] **Hu Y**, Li M, Liu X, Meng C. Sampling-based methods for multi-block optimization problems over transport polytopes [J]. *Mathematics of Computation*, accepted.
- [5] **Hu Y**, Gao X, Zhao Y, Liu X, Song H. Force-based gradient descent method for *ab initio* atomic structure relaxation [J]. *Physical Review B*, 2022, 106(10): 104101.
- [6] **Hu Y**, Yin J, Gao X, Liu X, Song H. Projected gradient descent algorithm for *ab initio* crystal structure relaxation under a fixed unit cell volume. arXiv preprint, 2024, arXiv:2405.02934.

## 知识产权

- [1] 高兴誉, 刘歆, 宋海峰, 胡雨宽, 方俊, 杨真, 赵亚帆, 王丽芳, 刘海风. 原子结构弛豫的非单调线搜索方法及装置: 202111534901.5 [P]. 2022-12-02.
- [2] 高兴誉, 刘歆, 宋海峰, 胡雨宽, 陈欣, 王越超, 方俊, 王丽芳, 张乐. 固定晶格体积晶体结构弛豫的计算方法及装置: 202211210741.3 [P]. 2023-07-21.
- [3] 高兴誉, 刘歆, 胡雨宽, 宋海峰, 陈欣, 王越超, 王丽芳. 基于非单调梯度型算法的晶体结构弛豫软件 ProME-SuRe: 2023SR1558824 [软件]. 2023-12-04.

# 感谢各位专家!

电子邮件: ykhu@lsec.cc.ac.cn

个人主页: <https://huyukuan.github.io>

## 与 PALM-I、(S-)ERALM、(S-)KLALM 算法相关的数值实验—工作站

- CPU: Intel Xeon Gold 6242R ( $3.10\text{ GHz} \times 20 \times 2$ ), 510 GB 运行内存.
- 操作系统: Ubuntu 20.04.5.
- 软件: MATLAB R2018b/R2019b.

## 与 PGD 算法相关的数值实验—计算机集群 LSSC-IV<sup>(5)</sup>

- CPU: Intel Xeon Gold 6140 ( $2.30\text{ GHz} \times 18 \times 2$ ), 192 GB 运行内存.
- 操作系统: Red Hat Enterprise Linux Server 7.3.
- 软件: Correlated Electron System Simulation Package (CESSP).

[Fang et al. 2016; Gao et al. 2017; Zhou et al. 2018; Fang et al. 2019]

- 语言与编译器: Fortran 90, Intel oneAPI.

---

<sup>(5)</sup> 集群中文简介: <http://lsec.cc.ac.cn/chinese/lsec/LSSC-IVintroduction.pdf>.