arXiv:2603.03787v1 [math.OC] 4 Mar 2026

一种通过随机变分不等式求解一类两阶段非凸非光滑随机锥规划的连续凸差方法

张超[1]，[2*]和王迪[1]

[1]北京交通大学数学与统计学院，中国北京，邮编100044。[2]北京交通大学北京生物大数据与拓扑统计重点实验室，中国北京，邮编100044。

*通讯作者。电子邮箱：chzhang2@bjtu.edu.cn；贡献作者：wangdi@bjtu.edu.cn；

摘要

我们研究一类两阶段非凸非光滑随机锥规划，其中两个阶段的目标函数都可以包含非光滑项，这些项是近端映射易于计算的函数，并进一步与仿射映射复合。这类问题能够用于建模各种应用。然而，由于其两阶段结构可能涉及大量情景，以及非凸、非光滑甚至非利普希茨不连续项和锥约束，求解这类问题颇具挑战性。本文定义了该问题的一个 KKT 点，并证明在温和条件下它是一个必要最优性条件，继而将其转化为等价的非单调非光滑两阶段随机变分不等式（随机变分不等式）。随后，我们利用莫罗包络提出一种连续凸差（SDC）方法来求解该问题，其子问题通过用于求解极大单调两阶段随机变分不等式的渐进对冲方法近似求解。我们证明了在适当假设下，该方法具有严格的收敛性。最后，本文将马科维茨均值-方差模型的一种扩展作为应用，并通过数值结果展示该模型和 SDC 方法的有效性。

关键词： 两阶段非凸非光滑随机锥规划，随机变分不等式，凸差方法，渐进对冲法

MSC 分类： 90C15 , 90C33 , 49J40

1 引言

给定概率空间 $(\Omega, \mathcal{F}, P)$，其有限支撑集为 $\Xi_K := \{\xi_1, \ldots, \xi_K\}$，对任意 $\xi_i \in \Xi_K$，记概率 $p(\xi_i) = p_i$。对每个 $i \in [K] := \{1, 2, \ldots, K\}$，设 $r_1(\cdot): \mathbb{R}^{m_1} \to \mathbb{R}$，$r_{\xi_i}(\cdot): \mathbb{R}^{m_2} \to \mathbb{R}$，$i \in [K]$ 为闭、下半连续（lsc）、下有界函数，定义如下：

$$r_1(x) = P_1(U_1 x + u_1), \quad r_{\xi_i}(y_i) = P_{\xi_i}(U_{\xi_i} y_i + u_{\xi_i}),\ i \in [K], \tag{1}$$

其中，U_1、U_{ξ_i}、$i \in [K]$ 是给定矩阵，u_1、u_{ξ_i}、$i \in [K]$ 是给定向量，P_1、P_{ξ_i}、$i \in [K]$ 是维数适当的给定映射。

我们考虑如下形式的两阶段非凸非光滑随机锥规划（T-NNS-SCP）。

$$\begin{aligned} \min_{x \in \mathbb{R}^{m_1}} \quad & c(x) + r_1(x) + \sum_{k=1}^{K} p_i[\vartheta(x, \xi_i)] \\ \text{s.t.} \quad & Ax = a, \\ & Bx - b \in -C_1, \end{aligned} \tag{2a}$$

其中，$\vartheta(x, \xi_i)$是第二阶段问题的最优值

$$\begin{aligned} \min_{y_i \in \mathbb{R}^{m_2}} \quad & q_{\xi_i}(y_i) + r_{\xi_i}(y_i) \\ \text{s.t.} \quad & A_{1,\xi_i} y_i + A_{2,\xi_i} x = d_{\xi_i}, \\ & W_{\xi_i} y_i + T_{\xi_i} x - h_{\xi_i} \in -C_{2,\xi_i}. \end{aligned} \tag{2b}$$

这里，$A \in \mathbb{R}^{n_1 \times m_1}$、$B \in \mathbb{R}^{s_1 \times m_1}$、$A_{1,\,\xi_i} \in \mathbb{R}^{n_2 \times m_2}$、$A_{2,\,\xi_i} \in \mathbb{R}^{n_2 \times m_1}$、$W_{\xi_i} \in \mathbb{R}^{s_2 \times m_2}$以及 $T_{\xi_i} \in \mathbb{R}^{s_2 \times m_1}$、$i \in [K]$ 是给定矩阵；$a \in \mathbb{R}^{n_1}$、$b \in \mathbb{R}^{s_1}$、$d_{\xi_i} \in \mathbb{R}^{n_2}$以及$h_{\xi_i} \in \mathbb{R}^{s_2}$、$i \in [K]$是给定向量；$C_1 \subseteq \mathbb{R}^{s_1}$、$C_{2,\,\xi_i} \subseteq \mathbb{R}^{s_2}$、$i \in [K]$ 是具有非空内部的对称闭凸锥，其中任意一个都可以是二阶锥、非负锥、半正定矩阵锥，或由上述三种类型的对称锥构成的笛卡尔积。

将第一阶段问题（2a）的可行集记为 X，对每个固定的 $(x, \xi_i) \in X \times \Xi_K$，将第二阶段问题（2b）的可行集记为 $Y(x, \xi_i)$，并定义

$$\Phi := \{(x, y_1, \ldots, y_K) \in \mathbb{R}^{m_1 + m_2 K} \ :\ x \in X,\ y_i \in Y(x, \xi_i),\ i \in [K]\}. \tag{3}$$

全文中我们作如下假设：

A1. P_1、P_{ξ_i}、$i \in [K]$ 均为有下界的闭下半连续函数，且其近端算子易于计算。

A2. $c: \mathbb{R}^{m_1} \to \mathbb{R}$ 和 $q_{\xi_i}: \mathbb{R}^{m_2} \to \mathbb{R}$, $i \in [K]$ 均为二阶连续可微凸函数。

A3. Φ为紧集。

A4. 对于任意的 $(x, \xi_i) \in X \times \Xi_K$, $\sum_{i=1}^{K} p_i |\vartheta(x, \xi_i)| < +\infty$。

A5. 集合 Φ 有一个斯莱特点：存在（$\mathring{x}, \mathring{y}_1, \ldots, \mathring{y}_K$），使得 $A\mathring{x} = a$,$A_{1,\xi_i}\mathring{y}_i + A_{2,\xi_i}\mathring{x} = d_{\xi_i}$, $B\mathring{x} - b \in$ 属于int$(-C_1)$ $W_{\xi_i}\mathring{y}_i + T_{\xi_i}\mathring{x} - h_{\xi_i} \in$ int$(-C_{2,\xi_i})$, $i \in [K]$。

式（2a）-（2b）中的 T-NNS-SCP 模型在不确定性下的重要应用建模方面前景广阔。它继承了两阶段随机规划（SP）处理不确定性的优点，能够在第一阶段做出“即刻决策”，同时明确考虑未来的不确定事件，并在不确定性消除后于第二阶段做出“观望决策”。此外，该模型能够纳入锥约束，因此可以刻画涉及非负锥、二阶锥和半正定锥的复杂约束。最重要的是，它允许第一阶段和第二阶段的目标函数都包含非凸非光滑正则项，这些正则项甚至可以是非利普希茨且不连续的。非光滑项 P_1 和 P_{ξ_i}, $i \in [K]$ 具有易于计算的近端映射，并且还可以与仿射映射复合。因此，复合函数 r_1 和 r_{ξ_i} 可能没有易于计算的近端项，从而扩大了非光滑项的适用范围。

Mehrotra [18]考虑了一类凸且光滑的两阶段随机锥规划，它是式（2a）-（2b）中 T-NNS-SCP 的一个特例，其中 $r(x) \equiv 0$,$r_{\xi_i}(y_i) \equiv 0, i \in [K]$, $c(\cdot)$ 和 $q_{\xi_i}(\cdot)$, $i \in [K]$ 均为凸二次函数。他们开发了一种实用的原始内点分解算法来求解该问题。文中通过一个有趣的马科维茨投资组合选择模型两阶段 SP 扩展的数值实验，研究所提出方法的性能。在许多应用中，非凸、非光滑甚至非利普希茨的不连续正则项有助于诱导所需的稀疏性或其他适当的惩罚项，而且两个阶段的目标函数并非二次函数。这促使我们在本文中考虑新的 T-NNS-SCP 模型。

尽管 T-NNS-SCP 有望用于对重要应用进行建模，但由于其两阶段结构可能包含大量情景 K，且第一阶段和第二阶段目标函数中均存在非凸、非光滑甚至非利普希茨不连续正则项，以及复杂的锥约束，因此如何高效求解并非易事。(2b) 中的约束在y_1、y_2、……、y_K上是分块可分的。因此，如果将两阶段问题写成同时包含第一阶段和第二阶段变量的单阶段问题，那么第一阶段变量与第二阶段变量之间便具有块角结构。对于这种具有特殊结构的问题，Rockafellar 和 Wets 提出的渐进对冲法（PHM）等分解方法 [24] 可能十分高效。然而，正如 [12],所指出的，PHM 的收敛性在很大程度上局限于凸问题及其相关扩展 [8； 31]这一瓶颈限制了 PHM 在一般非凸两阶段随机规划问题中的应用。近年来，求解非凸两阶段随机规划问题已引起广泛关注 [2； 12； 13； 14]。若干研究致力于为每个情景下最新的第一阶段迭代点构造适当的值函数近似，然后利用这些近似生成下一次第一阶段迭代点。当目标函数中没有显式的非光滑项，且凸约束由线性方程和非线性不等式构成时，Borges、Sagastiz´abal 和 Solodov 提出了一种新颖的平滑方法，

可以通过对第二阶段的不等式约束添加 Tikhonov 正则化障碍函数，处理其可能非凸的值函数。近期，Li 和 Cui [12] 利用补救函数隐含的凸凹结构，通过求解所谓的部分 Moreau 包络获得值函数的代理函数；在每次迭代中，该包络都是原第二阶段约束下的一个凸子问题。

从另一个角度来看，随机规划与随机变分不等式密切相关。随机变分不等式为随机优化和随机博弈的最优性条件提供了统一的数学模型，因此在科学、工程、经济学和金融领域有着广泛的应用。 [5] 研究了一类两阶段随机变分不等式，并采用期望残差最小化将两阶段随机变分不等式重新表述为两阶段补救随机规划。Rockafellar 和 Wets 首先在 [26],中给出了多阶段随机变分不等式的定义，随后 Rockafellar 和 Sun 于 [23]提出了用于求解单调多阶段随机变分不等式的 PHM。 [6研究了两阶段随机变分不等式的离散化近似； 7], 所考虑的底层随机数据服从连续分布。 [16], 研究了通过 SVI 方法求解两阶段随机规划的问题，但与本文考虑的情形不同，该问题的目标函数中没有非光滑项；而且， [16]中的两阶段随机规划是凸的，其等价的两阶段随机变分不等式为极大单调，因此可以采用 PHM 求解。 [17]针对流行病期间医疗物资的储存与动态分配，提出了一种两阶段非光滑随机均衡模型。通过对涉及最大值算子的医疗用品不足凸惩罚项进行平滑处理，平滑问题的乐观版本可等价转化为一个含有大量不等式约束的两阶段随机规划。随后，该两阶段随机规划又可等价转化为一个极大单调的光滑两阶段随机变分不等式，并可采用 PHM 对其进行近似求解。

本文设计了一种新颖的 SDC 方法，通过 SVI 方法求解 T-NNS-SCP 模型，并证明在温和假设下，SDC 方法存在聚点，且任一聚点均为原始 T-NNS-SCP 模型的 KKT 点。具体而言，我们首先研究 T-NNS-SCP 的最优性条件，并将其转化为等价的非凸非光滑两阶段随机变分不等式。受 [15], 启发，我们对每个非光滑项使用莫罗包络来构造一个近似问题，该问题可通过近端映射轻松求解。随后，在每次用于近似求解该近似问题的内层迭代中，我们进一步对莫罗包络 DC 重构中的第二项进行线性化，在当前点处添加正则化项，从而得到一个光滑凸两阶段随机规划。该问题等价于一个最大单调两阶段随机变分不等式，并通过著名的 PHM 求解。我们将 SDC 与 PHM 的组合简称为 SDC-PHM。通过为 PHM 子问题的近似解精心设计停止准则、为近似问题制定非精确求解策略，以及为莫罗包络中的参数选择合适的更新规则，我们在适当假设下证明了 SDC 方法的收敛性。与直接求解两阶段随机规划相比，通过 SVI 方法求解具有以下优点。首先，该方法避免了由第二阶段值函数引起的隐式非光滑性

问题。其次，该方法同时处理原变量和对偶变量（乘子），因此 SDC 方法能够生成一个收敛到 KKT 点的子序列，这比收敛到极限驻点更强。第三，在每次迭代中，只需粗略求解与近似问题对应的随机变分不等式，并且允许迭代点不可行，这有助于降低计算成本，尤其是在处理线性方程和一般锥约束等复杂约束时。所提出的 SDC-PHM 方法可视为将著名的 PHM 扩展到相当一般的非凸非光滑两阶段随机规划或随机变分不等式问题。值得注意的是，SDC 框架中的两阶段随机变分不等式子问题也可由其他高效的两阶段随机变分不等式求解器处理，并不局限于 PHM。

本文结构安排如下。第 2 节介绍记号、一些基本定义和结果。第 3 节分析 T-NNS-SCP，在温和假设下定义 T-NNS-SCP 的一个 KKT 点，并证明该点是最优性的必要条件；随后将 KKT 点改写为等价的非单调非光滑两阶段随机变分不等式。第 4 节提出 SDC 方法，并采用 PHM 求解 SDC 中的子问题。我们证明 SDC-PHM 定义良好，且在适当假设下 SDC 方法收敛。第5节将马科维茨均值-方差模型的一个非凸非光滑两阶段扩展作为 T-NNS-SCP 的应用。我们通过数值实验评估所提出的模型和 SDC 方法。

2 记号与预备知识

全文采用以下记号。假设所有向量均为列向量。矩阵 W 的转置记为 W^T，向量 w 的第 i 个分量记为 $(w)_i$，w 的欧几里得范数记为 $\|w\|$。对于$g_i \in \mathbb{R}^{t_i}$，$i \in [K]$,

$$(g_i; i \in [K]) := (g_1; g_2; \ldots; g_K) := (g_1^T, g_2^T, \cdots, g_K^T)^T \in \mathbb{R}^{t_1+t_2+\ldots+t_K}.$$

给定 $w \in \Omega$，Ω 在 w 处的切锥记为 $T_\Omega(w)$ [25, 定义 6.1，第 197], 页；正则（也称 Fr´échet）意义下的法锥和 Mor- ˆ dukhovich（也称极限或基本）法锥分别记为 $N_\Omega(w)$ 和 $N_\Omega(w)$ [25, 定义 6.3，第 199] 页。对于锥 ˆC，$\hat{C}^o = \{w : \langle w, w' \rangle \leq 0$，任意 $w' \in \hat{C}\}$ 都是其极锥。函数 f 的定义域记为 dom $f := \{w \in \mathbb{R}^n : f(w) < +\infty\}$。示性函数 $\delta_\Omega(w)$ 定义为 $\delta_\Omega(w) = 0$ 若 $w \in \Omega$，且 $\delta_\Omega(w) = +\infty$否则。对于可微映射 $L : \mathbb{R}^n \to \mathbb{R}^m$，我们将其在 w 处的雅可比矩阵记为 $\mathcal{J}L(w) \in \mathbb{R}^{m\times n}$，并记 $\nabla L(w) := \mathcal{J}L(w)^T$。我们用 $\mathcal{J}_s L(w)$ 表示映射 $L(w)$ 关于 s 的偏导数，其中 s 的每个分量均来自向量 w。给定锥 ˆ $C_i, i \in [K]$，我们将锥 ˆ $C_i, i \in [K]$ 的笛卡尔积记为 $\prod_{i=1}^K \hat{C}_i$，且 $(\hat{C})^{\otimes K} := \underbrace{\hat{C} \times \cdots \times \hat{C}}_{K}$.

¯根据 [27, 第333页，第 366], 页，函数 $f : \mathbb{R}^n \to \mathbb{R} := [-\infty, \infty]$ 在点ˆ$w \in \mathbb{R}^n$ 处称为下半连续（下半连续），如果 $f(\hat{w}) \leq \liminf_{w \to \hat{w}} f(w)$。称

则 f 为下半连续，若其在每一点处均为下半连续，即 $\mathbb{R}^n$。根据 [25, 定义8.3，第 301], 页，对任意¯非凸函数 $f:\ \mathbb{R}^n \to \mathbb{R}$ 以及对任意 $\bar{w} \in$ 定义域f，Fr´echet次微分、极限次微分与地平线次微分分别定义为

$$\hat{\partial} f(\bar{w}) := \left\{ v \in \mathbb{R}^n : \operatorname{liminf}_{w \to \bar{w},\ w \neq \bar{w}} \frac{f(w) - f(\bar{w}) - v^T(w - \bar{w})}{\|w - \bar{w}\|} \geq 0 \right\},$$

$$\partial f(\bar{w}) := \left\{ v \in \mathbb{R}^n : \exists\ w^k \xrightarrow{f} \bar{w},\ v^k \to v \text{ with } v^k \in \hat{\partial} f(w^k) \text{ for each } k \right\},$$

$$\partial^\infty f(\bar{w}) := \left\{ v \in \mathbb{R}^n : \exists\ w^k \xrightarrow{f} \bar{w},\ \alpha^k \downarrow 0,\ \alpha^k v^k \to v \text{ with } v^k \in \hat{\partial} f(w^k) \text{ for each } k \right\},$$

其中，记号 $w^k \xrightarrow{f} \bar{w}$ 表示 $w^k \to \bar{w}$ 和 $f(w^k) \to f(\bar{w})$。我们还定义$\partial f(\bar{w}) = \partial^\infty f(\bar{w}) := \emptyset$，当 $\bar{w} \notin$ dom$f$时。若此外 $f(\bar{w})$ 有限，则根据 [25, 命题 8.7，第 302]页， ∂f 和 $\partial^\infty f$ 是外半连续的，即，

$$\left\{ v \in \mathbb{R}^n : \exists\ w^k \xrightarrow{f} \bar{w},\ v^k \to v \text{ with } v^k \in \partial f(w^k) \text{ for each } k \right\} \subseteq \partial f(\bar{w}),$$

$$\left\{ v \in \mathbb{R}^n : \exists\ w^k \xrightarrow{f} \bar{w},\ \alpha^k \downarrow 0,\ \alpha^k v^k \to v, \text{ with } v^k \in \partial f(w^k) \text{ for each } k \right\} \subseteq \partial^\infty f(\bar{w}).$$

ˆ 众所周知， $\partial f(w) \subseteq \partial f(w)$（[25,定理 8.6，第 302]页）。由 [25,定理 9.13，第 358] 页，可知 f 在 $\bar{w}$ 处利普希茨连续，当且仅当 $\partial^\infty f(\bar{w}) = \{0\}$。此外，若 f 在 $\bar{w}$的某个邻域内连续可微，则根据 [25, 习题 8.8(b)，第 304]页，有 $\partial f(\bar{w}) = \{\nabla f(\bar{w})\}$。

对于任意有下界的实值闭下半连续函数 P_0 以及任意 $\rho > 0$，莫罗包络 [25, 定义1.22，第 20] 页，定义如下:

$$P_{0,\rho}(w) := \inf_{w'} \left\{ \frac{1}{2\rho} \|w' - w\|^2 + P_0(w') \right\}.$$

莫罗包络定义中的下确界在 w处由 ρP_0 的所谓近端映射取得，其定义如下:

$$\operatorname{prox}_{\rho P_0}(w) := \arg\min_{w'} \left\{ \frac{1}{2\rho} \|w' - w\|^2 + P_0(w') \right\}.$$

根据 [25, 定理1.25，第 21]页，由于 P_0 是真闭且有下界，此集合总是非空的。容易看出，对于任意 w,

$$-\infty < \inf_{w'} P_0(w') \leq P_{0,\rho}(w) \leq P_0(w). \tag{4}$$

此外，我们有以下引理，见 [15, 引理 1]。

引理 1 P_0 $P_0 > -\infty$ $w^t \to$ 设为一个下确界为的实值 w^* $\rho_t \downarrow 0$ $\zeta^t \in$ $\rho_t P_0(w^t)$ t $\zeta^t \to w^*$ 闭函数。假设，并为每个选取任意一个近端算子。则有。

莫罗包络可以表示为差分凸函数：

$$P_{0,\rho}(w) = \frac{1}{2\rho}\|w\|^2 - \underbrace{\max_{w'}\left\{\frac{1}{\rho}w^T w' - \frac{1}{2\rho}\|w'\|^2 - P_0(w')\right\}}_{D_{\rho,P_0}(w)}, \tag{5}$$

其中 $D_{\rho,\ P_0}$ 是一个凸函数，因为它是w的仿射函数的最大值。此外，根据$P_{0,\ \rho}$（w）的定义、近端算子ρP_0（w）和（5），我们看到，D_{ρ}中的最大值P_0(w)在近端算子ρP_0(w)中的任意点处取得。

下面的引理给出了一种方法，可以根据 [29, 引理 1]计算$D_{\rho,\ P_0}$(w)的一个次梯度，以及 P(w)的极限次微分中的一个元素。

引理 2 $\frac{1}{\rho}$近端算子ρP_0(w) $\subseteq \partial D_{\rho,\ P_0}$($w$)，以及$\{\frac{1}{\rho}$($w$ − ζ): $\zeta \in$近端算子ρP_0(w)$\} \subseteq \partial P_0$($\zeta$)。

设 U_0 和 u_0 为任意给定的适当大小的矩阵和向量，并设r_0 为一个任意的适当的下半连续复合函数，其形式为

$$r_0(s) = P_0(U_0 s + u_0).$$

记

$$R_{0,\rho}(s) := D_{\rho,P_0}(U_0 s + u_0) \quad \text{and} \quad r_{0,\rho}(s) := P_{0,\rho}(U_0 s + u_0). \tag{6}$$

利用与 [15, 第3节1], 类似的论证，我们可以证明

$$\frac{1}{\rho}U_0^T \operatorname{prox}_{\rho P_0}(U_0 s + u_0) \subseteq U_0^T \partial D_{\rho,P_0}(U_0 s + u_0) = \partial R_{0,\rho}(s). \tag{7}$$

3 T-NNS-SCP 的最优性条件

本节首先将 T-NNS-SCP 描述为大规模锥规划的形式，并证明解的存在性。

我们定义 T-NNS-SCP 的 KKT 点，它是在 Robinson约束资格条件（RCQ）下的一个必要最优性条件。最终，我们将其等价地转化为一个两阶段非凸非光滑随机变分不等式。这促使我们在下一节设计基于随机变分不等式的 SDC 算法。为简洁地表述 RCQ 和随机变分不等式，我们定义如下记号。

命题 1 *The T-NNS-SCP (2a)-(2b) is equivalent to the following nonconvex nonsmooth stochastic conic program*

$$\begin{aligned} &\min_{\mathbf{x_y}} \ \hat{f}(\mathbf{x_y}) + \Psi(\mathbf{x_y}) \\ &\text{s.t. } G(\mathbf{x_y}) \in \mathcal{K}. \end{aligned} \tag{8}$$

The equivalence is understood in the sense that optimal values of problem T-NNS-SCP in (2a)-(2b) and (8) are equal to each other and finite. Moreover, the set of optimal solutions of the T-NNS-SCP in (2a)-(2b) is nonempty and bounded.

记号	说明
$\mathcal{K} := \{0_{n_1}\} \times -C_1 \times (\{0_{n_2}\})^{\otimes K} \times \prod_{i=1}^{K} -C_{2,\ \xi_i}$	与con-相关的凸锥
$D := \mathbb{R}^{m_1} \times (\mathbb{R}^{m_2})^{\otimes K} \times \mathcal{K}^\circ$	两个阶段的约束 与原变量 和对偶变量
$\mathbf{y} := (y_i; i \in [K]) \in (\mathbb{R}^{m_2})^{\otimes K}$	所有情景下第二阶段原变量的向量
$\mathbf{xy} := (x; \mathbf{y}) \in \mathbb{R}^{m_1} \times (\mathbb{R}^{m_2})^{\otimes K}$	两个阶段的原变量向量 两个阶段
$M := (\alpha_1; \alpha_2; \pi_{1,i}; i \in [K]; \pi_{2,i}; i \in [K]) \in \mathcal{K}^\circ$	两个阶段的对偶变量向量 阶段
$\mathbf{z} := (\mathbf{xy}; M) \in D$	两阶段中的原始变量和对偶变量向量
$\hat{f}(\mathbf{xy}) := c(x) + \sum_{i=1}^{K} p_i q_{\xi_i}(y_i)$	两阶段目标函数的光滑部分 两个阶段中的作用
$\Psi(\mathbf{xy}) := r_1(x) + \sum_{i=1}^{K} p_i r_{\xi_i}(y_i)$	非光滑　　目标函数的部分 两个阶段中的函数
$\mathcal{U}(\mathbf{xy}) := (U_1 x + u_1; U_{\xi_i} y_i + u_{\xi_i}; \forall i \in [K])$	向量值　　函数　　for Ψ的内映射
$P(\mathcal{U}(\mathbf{xy})) := P_1(U_1 x + u_1) + \sum_{i=1}^{K} p_i P_{\xi_i}(U_{\xi_i} y_i + u_{\xi_i})$	$\Psi(\mathbf{xy})$的复合表达式
$G(\mathbf{xy}) :\ = \begin{pmatrix} G^1(\mathbf{xy}) :\ = Ax - a \\ G^2(\mathbf{x_y}) := Bx - b \\ G^{1,\ \xi_i}(\mathbf{xy}) :\ = A_{1,\ \xi_i} y_i + A_{2,\ \xi_i} x - d_{\xi_i};\ \ \forall i \in [K] \\ G^{2,\ \xi_i}(\mathbf{xy}) :\ = W_{\xi_i} y_i + T_{\xi_i} x - h_{\xi_i};\ \ \forall i \in [K] \end{pmatrix}$	向量值函数，用于关联 与两个阶段的约束

表 1 符号说明

Proof 我们容易得到问题 T-NNS-SCP 在 (2a)-(2b) 与 (8) 中的最优值彼此相等。由于 (8) 中的目标函数下有界且下半连续，并根据假设 A3 和假设 A4，我们可以得出 (8) 的最优解集非空且有界。由假设 A4，(8) 的最优值是有限的。因此，问题 (2a)-(2b) 与 (8) 的共同最优值是有限的。显然，若 $\bar{\mathbf{x}}_{\bar{\mathbf{y}}}$ 是问题 (8) 的最优解，则 $\bar{x}$ 是 (2a) 的最优解，且对所有 $i \in [K]$，$\bar{y}_i$ 是第二阶段问题 (2b) 的最优解。因此，(2a)-(2b) 中 T-NNS-SCP 的解集非空。解集的有界性来自假设 A3 中 $\Phi = \{\mathbf{xy} :\ G(\mathbf{xy}) \in \mathcal{K}\}$ 的紧性。□

设$\mathbf{x}^*_{\mathbf{y}} \in \Phi$ 为给定点。显然，$N_{\{0_{n_1}\}}(G^1(\mathbf{x}^*_{\mathbf{y}})) = \mathbb{R}^{n_1}$，且$N_{\{0_{n_2}\}}(G^{1,\xi_i}(\mathbf{x}^*_{\mathbf{y}})) = \mathbb{R}^{n_2}$ 对任意 $i \in [K]$均成立。根据 [19, 命题 1.4，pp. 4], ，我们有

$$N_{\mathcal{K}}(G(\mathbf{x}^*_{\mathbf{y}})) = \mathbb{R}^{n_1} \times N_{\{-C_1\}}(G^2(\mathbf{x}^*_{\mathbf{y}})) \times (\mathbb{R}^{n_2})^{\otimes K} \times \prod_{i=1}^{K} N_{\{-C_{2,\xi_i}\}}\left(G^{2,\xi_i}(\mathbf{x}^*_{\mathbf{y}})\right). \quad (9)$$

根据 [1, (3.9),pp. 147],可得

$$\begin{aligned} M^* \in N_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*)) &\Longleftrightarrow G(\mathbf{x}_{\mathbf{y}}^*) \in N_{\mathcal{K}^\circ}(M^*) \\ &\Longleftrightarrow G(\mathbf{x}_{\mathbf{y}}^*) \in \mathcal{K},\ M^* \in \mathcal{K}^\circ,\ {M^*}^T G(\mathbf{x}_{\mathbf{y}}^*) = 0. \end{aligned} \tag{10}$$

以下定义中的 RCQ 是锥规划中使用的一种经典约束资格条件；例如参见 [32]。

定义1 我们称 RCQ 在 $\mathbf{x}_{\mathbf{y}}^* \in \Phi$ 处成立，如果梯度 $\nabla(G^1(\mathbf{x}_{\mathbf{y}}^*))_{j_1}$ 和 $\nabla(G^{1,\xi_i}(\mathbf{x}_{\mathbf{y}}^*))_{j_2}$，$i \in [K]$，$j_1 \in [n_1]$，$j_2 \in [n_2]$ 线性无关，且存在一个向量 $d \in \mathbb{R}^{m_1+m_2K}$ 使得

$$\begin{aligned} &\mathcal{J}G^1(\mathbf{x}_{\mathbf{y}}^*)d = 0, \\ &\mathcal{J}G^{1,\xi_i}(\mathbf{x}_{\mathbf{y}}^*)d = 0,\ i \in [K], \\ &G^2(\mathbf{x}_{\mathbf{y}}^*) + \mathcal{J}G^2(\mathbf{x}_{\mathbf{y}}^*)d \in \mathrm{int}(-C_1), \\ &G^{2,\xi_i}(\mathbf{x}_{\mathbf{y}}^*) + \mathcal{J}G^{2,\xi_i}(\mathbf{x}_{\mathbf{y}}^*)d \in \mathrm{int}(-C_{2,\xi_i}),\ i \in [K]. \end{aligned} \tag{11}$$

假设A5 意味着约束 $G(\mathbf{xy}) \in \mathcal{K}$ 存在一个斯莱特点；根据 [1, 命题2.104，第 73] 页，这确实保证了 RCQ 在 $\mathbf{x_y} \in \Phi$ 的任意可行点处成立。

引理3 在任意点 $\mathbf{x}_{\mathbf{y}}^* \in \Phi$ 处，我们有

$$\begin{aligned} T_\Phi(\mathbf{x}_{\mathbf{y}}^*) &= \left\{ d \in \mathbb{R}^{m_1+m_2K} : \mathcal{J}G(\mathbf{x}_{\mathbf{y}}^*)d \in T_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*)) \right\}, \\ N_\Phi(\mathbf{x}_{\mathbf{y}}^*) &= \hat{N}_\Phi(\mathbf{x}_{\mathbf{y}}^*) = T_\Phi^\circ(\mathbf{x}_{\mathbf{y}}^*) = \nabla G(\mathbf{x}_{\mathbf{y}}^*) N_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*)). \end{aligned}$$

Proof 切锥的公式由 [1, 推论2.91，第 66] 页得出，这是利用了 RCQ 对系统 $G(\mathbf{xy}) \in \mathcal{K}$ 中任意 $\mathbf{x}_{\mathbf{y}}^*$ 均成立这一事实。此外，由于 Φ 是凸的，根据 [25, 定理6.9，第 203] 页，可得法锥的第一个等式；根据 [25, 定理6.28，第 220]，页，可得第二个等式；最后一个等式则由 [22, (4)，第 107] 页得出。□

定义2 (KKT 点) 若 $\mathbf{x}_{\mathbf{y}}^*$ 是问题 (8) 的 KKT 点，则 $\mathbf{x}_{\mathbf{y}}^* \in \Phi$ 且存在向量 $\lambda^* \in \partial P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*))$ 和 $M^* \in N_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*))$，使得

$$0 = \nabla \hat{f}(\mathbf{x}_{\mathbf{y}}^*) + \nabla \mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)\lambda^* + \nabla G(\mathbf{x}_{\mathbf{y}}^*)M^*. \tag{12}$$

设 $\mathbf{x}_{\mathbf{y}}^* \in \Phi$。为了证明，(8) 的 KKT 点是 (8) 的必要最优性条件，我们需要以下两个假设。

A6. $-\partial^\infty \Psi(\mathbf{x}_{\mathbf{y}}^*) \cap N_\Phi(\mathbf{x}_{\mathbf{y}}^*) = \{0\}$.**A7.** $\nabla \mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)\lambda = 0, \lambda \in \partial^\infty P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)) \implies \lambda = 0.$

假设**A6** 中的条件称为基本资格条件（BQ），并广泛应用于非凸非光滑优化中；参见 [3,定义2.3]。假设 **A7** 中的条件也是非凸非光滑优化问题的一种经典约束资格条件；参见 [15, (12)]。如果 P_1 和 P_{ξ_i} 对所有 $i \in [K]$ 都是局部利普希茨连续的，那么 P和 Ψ也都是局部利普希茨连续的。这保证了 $\partial^\infty \Psi(\mathbf{x}_{\mathbf{y}}^*) = \{0\}$ 和 $\partial^\infty P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)) = \{0\}$。因此，在这种情况下，假设 **A6** 和 **A7** 显然成立。如果 U_1 和 U_{ξ_i}, $i \in [K]$ 满行秩，那么假设 **A7** 成立。

定理 2 *Let* $\mathbf{x}_{\mathbf{y}}^*$ *be a local optimal solution of problem* (8). *Suppose assumptions* **A6** *and* **A 7** *hold, then* $\mathbf{x}_{\mathbf{y}}^*$ *is a KKT point of problem* (8)。

Proof 容易看出，8）等价于

$$\min \hat{f}(\mathbf{x_y}) + \Psi(\mathbf{x_y}) + \delta_\Phi(\mathbf{x_y}).$$

由费马法则 [25,定理10.1，第 422]页可知

$$0 \in \nabla \hat{f}(\mathbf{x}_{\mathbf{y}}^*) + \partial \left(\Psi(\mathbf{x}_{\mathbf{y}}^*) + \delta_\Phi(\mathbf{x}_{\mathbf{y}}^*)\right). \tag{13}$$

利用假设 **A6**， $\partial\delta_\Phi(\mathbf{x}_{\mathbf{y}}^*) = \partial^\infty\delta(\mathbf{x}_{\mathbf{y}}^*) = N_\Phi(\mathbf{x}_{\mathbf{y}}^*)$ 以及极限次微分的和法则 [25, 推论10.9，第 430], 页，可得

$$\partial \left(\Psi(\mathbf{x}_{\mathbf{y}}^*) + \delta_\Phi(\mathbf{x}_{\mathbf{y}}^*)\right) \subseteq \partial\Psi(\mathbf{x}_{\mathbf{y}}^*) + \partial\delta_\Phi(\mathbf{x}_{\mathbf{y}}^*) = \partial\Psi(\mathbf{x}_{\mathbf{y}}^*) + N_\Phi(\mathbf{x}_{\mathbf{y}}^*), \tag{14}$$

根据 [25, 练习8.14，第 310]页。此外，由假设 **A7** 和 [25, 定理10.6，第 427], 页可知

$$\partial\Psi(\mathbf{x}_{\mathbf{y}}^*) \subseteq \nabla\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)\partial P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)). \tag{15}$$

由于 RCQ 在$\mathbf{x}_{\mathbf{y}}^*$ 处对系统 $G(\mathbf{xy}) \in \mathcal{K}$成立，根据引理 3 可知

$$N_\Phi(\mathbf{x}_{\mathbf{y}}^*) = \nabla G(\mathbf{x}_{\mathbf{y}}^*) N_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*)). \tag{16}$$

结合（13）-（16），我们发现 $\mathbf{x}_{\mathbf{y}}^* \in \Phi$ 满足（12），从而

$$0 \in \nabla \hat{f}(\mathbf{x}^*) + \nabla\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)\partial P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)) + \nabla G(\mathbf{x}_{\mathbf{y}}^*) N_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*)).$$

这表明存在 $\lambda^* \in \partial P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*))$ 和 $M^* \in N_{\mathcal{K}}(G(\mathbf{x}_{\mathbf{y}}^*))$ 满足（12）。因此 $\mathbf{x}_{\mathbf{y}}^*$ 是一个 KKT 点。□

对于特殊的非利普希茨函数 $\Psi(\mathbf{xy})$，问题（8）的局部最优解 $\mathbf{x}_{\mathbf{y}}^*$ 是（8）的驻点，无需假设 **A6** 和 **A7**。与 [3,定理2.2], 类似，我们可以不难地证明如下推论。

推论 1 *Suppose* $\Psi \quad \mathbf{y} = \sum_{i=1}^{m_1+m_2K} \psi_i \quad \mathbf{y}\ i$ *for some lsc functions* $\psi_i \quad \mathbb{R} \to \mathbb{R}$. *Let*(x) ((x)) : : ((x)) $\mathcal{I} = \{i \quad \partial^\infty\psi_i \quad {}^*_{\mathbf{y}}\ i = \{0\}\}$ *and* $\mathcal{I}^c$ *be the complement of* $\mathcal{I}$ *with respect to* $\{1, \ldots, m_1 + ,((x)) ,(x) m_2K\}$. *Assume further that for any* $i \in \mathcal{I}^c \quad \partial\psi_i \quad {}^*_{\mathbf{y}}\ i = \mathbb{R}$ *and the RCQ holds at* ${}^*_{\mathbf{y}}\ \mathcal{I}$ *for the system* $G \quad \mathbf{y}\ \mathcal{I} \quad {}^*_{\mathbf{y}}\ \mathcal{I}^c \in \mathcal{K}$. *Then* ${}^*_{\mathbf{y}}$ *is a KKT point of problem* 8 . ((x),(x)) x( )

显然，稀疏性诱导惩罚项 $\Psi(\mathbf{xy}) := \|x\|_p^p + \sum_{i=1}^K p_i\|y_i\|_p^p$ 对于 $p \in (0, \mathbf{1})$以及 $\Psi(\mathbf{xy}) := \|x\|_0 + \sum_{i=1}^K p_i\|y_i\|_0$ 满足推论 Ψ 1中的要求；它们均为非利普希茨函数，且后者甚至是不连续的。

Remark 1 当锥约束可以由有限个等式约束和不等式约束定义时，根据 [1, 第 71]页，RCQ等价于Mangasarian-Fromovitz约束资格条件（MFCQ）。利用 [3, 命题2。2], 弱于RCQ的松弛常数正线性相关性（RCPLD）足以推导出引理 $N_\Phi(\mathbf{x}_{\mathbf{y}}^*)$中的表示式3。因此，利用RCPLD可以得出与定理2 和推论1 中相同的结论。

我们进一步将（12）中的 KKT条件重新表述为一个等价的非单调非光滑两阶段变分不等式问题，为下一节通过 SVI 方法开发结合 PHM 的 SDC 方法提供思路。

令

$$H(\mathbf{z};\lambda^*) := \begin{pmatrix} \nabla \hat{f}(\mathbf{x_y}) + \nabla G(\mathbf{x_y})M + \nabla \mathcal{U}(\mathbf{x_y})\lambda^* \\ -G(\mathbf{x_y}) \end{pmatrix},$$

条件（12）等价于存在 $\mathbf{z}^* = (\mathbf{x}_{\mathbf{y}}^*;\ M^*) \in D$和 $\lambda^* \in \partial P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*))$，使得 $\mathbf{z}^*$ 是以下两阶段变分不等式的一个解

$$0 \in H(\mathbf{z};\lambda^*) + N_D(\mathbf{z}). \tag{17}$$

定义自然残差函数如下:

$$H_D^{\mathrm{nat}}(\mathbf{z};\lambda^*) := \mathbf{z} - \mathbf{Proj}_D\left(\mathbf{z} - H(\mathbf{z};\lambda^*)\right), \tag{18}$$

其中，$\mathbf{Proj}_D(\hat{\mathbf{z}})$ 是 $\hat{\mathbf{z}}$ 在 D上的投影。记（17）的解集为 $\mathrm{SOL}(D,H)$。根据 [10, 命题1.5.8，第 83], 页，$\mathbf{z}^* \in \in \mathrm{SOL}(D,H)$ 当且仅当

$$H_D^{\mathrm{nat}}(\mathbf{z}^*;\lambda^*) = 0.$$

（8）的求解方法

4.1 逐次差凸方法

我们首先通过莫罗包络近似非凸非光滑项，并考虑由 Moreau 包络诱导的近似问题（MEAP）：8，其定义如下:

$$\begin{aligned} \min_{\mathbf{x_y}}\ & F_\rho(\mathbf{x_y}) := \hat{f}(\mathbf{x_y}) + r_{1,\rho}(x) + \sum_{i=1}^{K} p_i r_{\xi_i,\rho}(y_i) \\ \text{s.t. }& G(\mathbf{x_y}) \in \mathcal{K}. \end{aligned} \tag{MEAP}$$

其中，$r_{1,\ \rho}$ 和$r_{\xi_i,\ \rho}$、$i \in [K]$ 分别是 r_1 和 r_{ξ_i}、$i \in [K]$的莫罗包络。它们的差凸分解为:

$$\begin{aligned} r_{1,\rho}(x) &= \frac{1}{2\rho}\|U_1 x + u_1\|^2 - R_{1,\rho}(x), \\ r_{\xi_i,\rho}(y_i) &= \frac{1}{2\rho}\|U_{\xi_i} y_i + u_{\xi_i}\|^2 - R_{\xi_i,\rho}(y_i),\ i \in [K], \end{aligned}$$

其中

$$R_{1,\rho}(x) = D_{\rho,P_1}(U_1 x + u_1) \quad \text{and} \quad R_{\xi_i,\rho}(y_i) = D_{\rho,P_{\xi_i}}(U_{\xi_i} y_i + u_{\xi_i}).$$

莫罗包络$r_{1,\ \rho}$，$r_{\xi_i,\ \rho}$，$i \in [K]$ 的 r_1 和 r_{ξ_i}，$i \in [K]$，以及用于处理非凸非光滑函数的 DC 分解，灵感来源于 [15]。

我们进一步为（MEAP）设计了一种带正则化项的线性化近似。给定 ρ_t、$\mathbf{x}_\mathbf{y}^t = (x^t; y_i^t;\ i \in [K])$，以及

$$\zeta_1^t \in \operatorname{prox}_{\rho_t P_1}(U_1 x^t + u_1), \quad \zeta_{\xi_i}^t \in \operatorname{prox}_{\rho_t P_{\xi_i}}(U_{\xi_i} y_i^t + u_{\xi_i}),\ i \in [K],$$

由（7）可知，$\varpi^t := \left(\varpi_1^t;\ \varpi_{\xi_i}^t;\ \forall i \in [K]\right)$ 满足

$$\varpi_1^t := \frac{1}{\rho_t} U_1^T \zeta_1^t \in \partial R_{1,\rho_t}(x^t), \quad \varpi_{\xi_i}^t := \frac{1}{\rho_t} U_{\xi_i}^T \zeta_{\xi_i}^t \in \partial R_{\xi_i,\rho_t}(y_i^t),\ i \in [K].$$

对于给定的 $\tau := (\tau_1; \tau_2) > 0$，我们将 $r_{1,\ \rho_t}(x)$ 和$r_{\xi_i,\ \rho_t}(y_i)$ 的凸光滑替代函数定义为 $i \in [K]$

$$\begin{aligned}
\hat{r}_{1,\rho_t,\tau_1}(x) &:= \frac{1}{2\rho_t}\|U_1 x + u_1\|^2 - R_{1,\rho_t}(x^t) - \langle \varpi_1^t, x - x^t\rangle + \frac{\tau_1}{2}\|x - x^t\|^2, \\
\hat{r}_{\xi_i,\rho_t,\tau_2}(y_i) &:= \frac{1}{2\rho_t}\|U_{\xi_i} y_i + u_{\xi_i}\|^2 - R_{\xi_i,\rho_t}(y_i^t) - \langle \varpi_{\xi_i}^t, y_i - y_i^t\rangle + \frac{\tau_2}{2}\|y_i - y_i^t\|^2,
\end{aligned}$$

分别使用 DC 分解中第二个凸项的线性化，以及在 $\mathbf{x}_\mathbf{y}^t$ 处添加的二次正则化项。接着，我们定义

$$F^L_{\rho_t,\mathbf{x}_\mathbf{y}^t,\varpi^t,\tau}(\mathbf{x}_\mathbf{y}) := \ \hat{f}(\mathbf{x}_\mathbf{y}) + \hat{r}_{1,\rho_t,\tau_1}(x) + \sum_{i=1}^K p_i \hat{r}_{\xi_i,\rho_t,\tau_2}(y_i),$$

并发现

$$F_{\rho_t}(\mathbf{x}_\mathbf{y}) \ \le \ F^L_{\rho_t,\mathbf{x}_\mathbf{y}^t,\varpi^t,\tau}(\mathbf{x}_\mathbf{y}) - \frac{\tau_1}{2}\|x - x^t\|^2 - \frac{\tau_2}{2}\sum_{i=1}^K p_i \|y_i - y_i^t\|^2, \tag{19}$$

$$F_{\rho_t}(\mathbf{x}_\mathbf{y}^t) \ = \ F^L_{\rho_t,\mathbf{x}_\mathbf{y}^t,\varpi^t,\tau}(\mathbf{x}_\mathbf{y}^t). \tag{20}$$

在 $\mathbf{x}_\mathbf{y}^t$ 处的 (MEAP) 线性化近似问题可写为

$$\begin{aligned}
&\min_{\mathbf{x}_\mathbf{y}} \ F^L_{\rho_t,\mathbf{x}_\mathbf{y}^t,\varpi^t,\tau}(\mathbf{x}_\mathbf{y}) \\
&\text{s.t. } G(\mathbf{x}_\mathbf{y}) \in \mathcal{K}.
\end{aligned} \qquad \text{(L-MEAP-}\rho_t\text{-}\mathbf{x}_\mathbf{y}^t\text{-}\varpi^t\text{-}\tau\text{)}$$

上述线性化近似问题等价于一个极大单调随机变分不等式，因此可以应用 PHM 求解，如下所示。显然，

该问题是凸的。将 T-NNS-SCP (2a)-(2b) 中的 $r_1(x)$替换为 $\hat{r}_{1,\rho_t,\tau_1}(x)$，并将 $r_{\xi_i}(y_i)$替换为 $\hat{r}_{\xi_i,\rho_t,\tau_2}(y_i)$，再将$i \in [K]$ 用于 Ψ，并采用与命题1中类似的重新表述过程，可以得到问题 (L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$) 必定有解且目标值有限。此外，线性化近似问题有解，且其目标值有限，并与具有 $\hat{r}_{1,\rho_t,\tau_1}(x)$ 和 $\hat{r}_{\xi_i,\rho_t,\tau_2}(y_i)$ 的 T-NNS-SCP 的目标值相同，$i \in [K]$。下述定理给出了线性化近似问题的解与极大单调光滑两阶段随机变分不等式的解之间的关系。

定义

$$\lambda^L_{\rho_t,\mathbf{x}_{\mathbf{y}}^t,\varpi^t,\tau}(\mathbf{x}_{\mathbf{y}}) := \begin{pmatrix} \frac{1}{\rho_t}(U_1 x + u_1) - \frac{1}{\rho_t}\zeta_1^t + \tau_1(x - x^t) \\ p_i\left[\frac{1}{\rho_t}(U_{\xi_i} y_i + u_{\xi_i}) - \frac{1}{\rho_t}\zeta_{\xi_i}^t\right] + p_i\tau_2(y_i - y_i^t); \forall i \in [K] \end{pmatrix}, \tag{21}$$

$$H^L_{\rho_t,\mathbf{x}_{\mathbf{y}}^t,\varpi^t,\tau}(\mathbf{z}) := \begin{pmatrix} \nabla \hat{f}(\mathbf{x}_{\mathbf{y}}) + \nabla G(\mathbf{x}_{\mathbf{y}})M + \nabla \mathcal{U}(\mathbf{x}_{\mathbf{y}})\lambda^L_{\rho_t,\mathbf{x}_{\mathbf{y}}^t,\varpi^t,\tau}(\mathbf{x}_{\mathbf{y}}) \\ -G(\mathbf{x}_{\mathbf{y}}) \end{pmatrix}. \tag{22}$$

定理 3 *The problem (L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$) is sure to have a solution. Moreover, thepoint* $\bar{\mathbf{x}}_{\mathbf{y}}$ *is an optimal solution of problem (L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$)* 当且仅当 $\bar{\mathbf{x}}_{\mathbf{y}}$ *is a solution of the following SVI*

$$0 \in \nabla F^L_{\rho_t,\mathbf{x}_{\mathbf{y}}^t,\varpi^t,\tau}(\mathbf{x}_{\mathbf{y}}) + N_\Phi(\mathbf{x}_{\mathbf{y}}). \tag{23}$$

Furthermore, $\bar{\mathbf{x}}_{\mathbf{y}}$ *is an optimal solution of the SVI* (23) *if and only if there exist* $\bar{\mathbf{z}} = (\bar{\mathbf{x}}_{\mathbf{y}}; \bar{M}) \in D$ *such that* $\bar{\mathbf{z}}$ *is a solution of the two-stage SVI*

$$0 \in H^L_{\rho_t,\mathbf{x}_{\mathbf{y}}^t,\varpi^t,\tau}(\mathbf{z}) + N_D(\mathbf{z}). \tag{24}$$

Proof 问题（L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$）等价于如下光滑凸规划

$$\min\ F^L_{\rho_t,\mathbf{x}_{\mathbf{y}}^t,\varpi^t,\tau}(\mathbf{x}_{\mathbf{y}}) \quad \text{s.t.} \quad \mathbf{x}_{\mathbf{y}} \in \Phi,$$

根据 [25, 定理6.12，第 207]页， $\bar{\mathbf{x}}_{\mathbf{y}}$ 是问题（L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$- ϖ^t-τ）的全局最优解，当且仅当 $\bar{\mathbf{x}}_{\mathbf{y}}$ 是随机变分不等式（23）的一个解。此外，根据假设A5，我们知道，对于系统 $G(\mathbf{x}_{\mathbf{y}}) \in \mathcal{K}$，RCQ在任意点 $\bar{\mathbf{x}}_{\mathbf{y}} \in \Phi$ 处成立。基于引理3和KKT条件可知， $\bar{\mathbf{x}}_{\mathbf{y}}$ 是问题（L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$）的一个最优解，当且仅当存在 $\bar{\mathbf{z}} = (\bar{\mathbf{x}}_{\mathbf{y}}; \bar{M}) \in D$，使得 $\bar{\mathbf{z}}$是（24）中的两阶段变分不等式的一个解。□

引理4 (24) 中的两阶段变分不等式属于极大单调型。

Proof 利用 T-NNS-SCP 的约束（2a）-（2b）中的所有函数均为仿射函数这一事实，按照 [16], 中定理3.2和推论3.1的证明，我们可以轻松证明该引理。□

定义自然残差函数

$$H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau,D}(\mathbf{z}) := \mathbf{z} - \mathbf{Proj}_D\left(\mathbf{z} - H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(\mathbf{z})\right). \tag{25}$$

通过直接计算，我们有

$$\left\|H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,0}(\mathbf{z}) - H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(\mathbf{z})\right\| = \tau_1\left\|x - x^t\right\| + \tau_2\sum_{i=1}^K p_i\left\|y_i - y_i^t\right\|. \tag{26}$$

由投影算子的非扩张性质，我们得到

$$\begin{aligned}
&\left\|H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau,D}(\mathbf{z}) - H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,0,D}(\mathbf{z})\right\| \\
=& \left\|\mathbf{Proj}_D\left(\mathbf{z} - H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,0}(\mathbf{z})\right) - \mathbf{Proj}_D\left(\mathbf{z} - H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(\mathbf{z})\right)\right\| \\
\le& \left\|H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(\mathbf{z}) - H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,0}(\mathbf{z})\right\| \\
=& \tau_1\left\|x - x^t\right\| + \tau_2\sum_{i=1}^K p_i\left\|y_i - y_i^t\right\|.
\end{aligned} \tag{27}$$

我们已准备好描述用于求解（8）的 SDC 方法，见算法1。

如何高效地近似求解（28）（见算法1 ）中的两阶段变分不等式是一个重要问题。两阶段变分不等式已在 [5； 26]中得到研究。在 [23] 中提出的PHM能够高效求解两阶段变分不等式，因为它在面对大量情景时可以使用并行计算，并且可以采用高效算法，例如外梯度法 [10, 算法12.1。9], 半光滑牛顿法[21； 11], 或平滑方法 [28； 20] 可用于求解PHM中的子问题。

由于PHM是一种分解法，并且在每次迭代中考虑一个对应于第一阶段以及关于单个情景的第二阶段的随机变分不等式，我们用

$$\mathcal{K}_i, D_i, (x_i; y_i), M^i, \mathbf{z}_i, \hat{f}^i(x_i, y_i), \mathcal{U}^i(x_i, y_i), G^i(x_i, y_i), \lambda^{i,L}_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(x_i, y_i), H^{i,L}_{\rho_t,\mathbf{x}^t_{\mathbf{y}}\varpi^t,\tau}(\mathbf{z}_i)$$

其对应量

$$\mathcal{K}, D, \mathbf{x}_{\mathbf{y}}, M, \mathbf{z}, \hat{f}(\mathbf{x}_{\mathbf{y}}), \mathcal{U}(\mathbf{x}_{\mathbf{y}}), G(\mathbf{x}_{\mathbf{y}}), \lambda^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(\mathbf{x}_{\mathbf{y}}), H^L_{\rho_t,\mathbf{x}^t_{\mathbf{y}},\varpi^t,\tau}(\mathbf{z})$$

分别见表1、(21)和(22)，删去与 $j \neq i$、$i \in [K]$对应的部分即可。例如，

$$\mathcal{K}_i := \{0_{n_1}\} \times -C_1 \times \{0_{n_2}\} \times -C_{2,\xi_i}, \quad D_i := \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times (\mathcal{K}_i)^\circ,\ i \in [K].$$

设 $\mathbf{\Lambda}_i := (u_i,\ c_{1,\ i},\ c_{2,\ i}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{s_1}$，其中元素 u_i、$c_{1,\ i}$ 和 $c_{2,\ i}$充当对应于非预见性的拉格朗日乘子，针对

算法1 用于求解（8）的逐次DC近似方法（SDC-PHM）

1: 步骤 0。选取 $\eta = (\eta_1, \eta_2, \eta_3) > 0$ 和 $\tau = (\tau_1, \tau_2) > 0$，满足 ρ_t 为正数序列且 $\rho_t \downarrow 0$。初始化 $\mathbf{z}^0 \in D$ 并设置 $t = 0$。2: 步骤 1。使用逐次PHM近似求解带有 $\rho = \rho_t$的 (MEAP)。设置 $\mathbf{z}^{t,0} = \mathbf{z}^0$, 和 $l = 0$。3: 步骤 1-1。选取 $\zeta_1^{t,l} \in$ 近端算子 $\rho_t P_1(U_1 x^{t,l} + u_1)$ 以及 $\zeta_{\xi_i}^{t,l} \in$ 近端算子 $\rho_t P_{\xi_i}(U_{\xi_i} y_i^{t,l} + u_{\xi_i}), i \in [K]$。计算

$$\varpi_1^{t,l} = \frac{1}{\rho_t} U_1^T \zeta_1^{t,l} \in \partial R_{1,\rho_t}(x^{t,l}) \quad \text{and} \quad \varpi_{\xi_i}^{t,l} = \frac{1}{\rho_t} U_{\xi_i}^T \zeta_{\xi_i}^{t,l} \in \partial R_{\xi_i,\rho_t}(y_i^{t,l}),\ i \in [K].$$

4: **步骤 1-2**。使用 PHM（算法2）近似求解以下两阶段变分不等式

$$0 \in H^L_{\rho_t, \mathbf{x}_\mathbf{y}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{z}) + N_D(\mathbf{z}), \tag{28}$$

并得到满足以下条件的 $\tilde{\mathbf{z}}$

$$\left\| H^{L,\mathrm{nat}}_{\rho_t, \mathbf{x}_\mathbf{y}^{t,l}, \varpi^{t,l}, \tau, D}(\tilde{\mathbf{z}}) \right\| \le \eta_1 \rho_t, \tag{29}$$

and

$$F^L_{\rho_t, \mathbf{x}_\mathbf{y}^{t,l}, \varpi^{t,l}, \tau}(\tilde{\mathbf{x}}_\mathbf{y}) \le F^L_{\rho_t, \mathbf{x}_\mathbf{y}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{x}_\mathbf{y}^{t,l}) + \frac{\eta_2}{(l+1)^2}. \tag{30}$$

5: **步骤 1-3**。设置 $\mathbf{z}^{t,l+1} := \tilde{\mathbf{z}}$。如果 $\tilde{\mathbf{z}}$ 不满足

$$\tau_1 \left\| \tilde{x} - x^{t,l} \right\| + \tau_2 \sum_{i=1}^K p_i \left\| \tilde{y}_i - y_i^{t,l} \right\| \le \eta_3 \rho_t^2, \tag{31}$$

令 $l := l + 1$ 并转到 **步骤 1-1**。否则，令 $l_t := l + 1$ 并转到 **步骤 2**。6: **步骤 2.** 更新 $\mathbf{z}^{t+1} := \mathbf{z}^{t,l_t}$, $t := t + 1$ 并转到 **步骤 1**。

第一阶段向量 x, α_1 和 α_2，并定义

$$C_2^i(\breve{\mathbf{\Lambda}}_i^\nu) := (\breve{u}_i^\nu; 0_{\{m_2\}}; \breve{c}_{1,i}^\nu; \breve{c}_{2,i}^\nu; 0_{\{n_2+s_2\}}).$$

我们准备给出渐进对冲法（PHM）的框架，以近似求解（28），见算法2 如下。

算法2 用于近似求解（28）的 PHM

1: 给定初始点 $\breve{\mathbf{z}}^0 \in \mathcal{N}$ 与 $\breve{\Lambda}^0 \in \mathcal{M}$，其中 $\sum_{i=1}^K p_i \breve{u}_i^0 = 0$，$\sum_{i=1}^K p_i \breve{c}_{1,\,i}^0 = 0$，$\sum_{i=1}^K p_i \breve{c}_{2,\,i}^0 = 0$，$\breve{v}_i^0 = \breve{s}_{1,\,i}^0 = \breve{s}_{2,\,i}^0 = 0$，$i = 1, \ldots\ldots, K$。选取步长 $\sigma > 0$。置 $\nu = 0$。

2: **步骤1.** 对于 $i \in [K]$，求解变分不等式

$$0 \in H^{i,L}_{\rho_t, x_i^t, y_i^t, \varpi^t, \tau}(\mathbf{z}_i) + C_2^i(\breve{\mathbf{\Lambda}}_i^\nu) + \sigma(\mathbf{z}_i - \breve{\mathbf{z}}_i^\nu) + N_{D_i}(\mathbf{z}_i), \tag{32}$$

以获得一个解 $\hat{\mathbf{z}}_i^\nu$。

3: **步骤2.** 对于 $i \in [K]$，令

$$\begin{aligned}
&\breve{x}^{\nu+1} = \sum_{i=1}^K p_i \hat{x}_i^\nu,\ \breve{\alpha}_1^{\nu+1} = \sum_{i=1}^K p_i \hat{\alpha}_{1,i}^\nu,\ \breve{\alpha}_2^{\nu+1} = \sum_{i=1}^K p_i \hat{\alpha}_{2,i}^\nu,\\
&\breve{y}_i^{\nu+1} = \hat{y}_i^\nu,\ \breve{\pi}_{1,i}^{\nu+1} = \hat{\pi}_{1,i}^\nu,\ \breve{\pi}_{2,i}^{\nu+1} = \hat{\pi}_{2,i}^\nu,\ \breve{u}_i^{\nu+1} = \breve{u}_i^\nu + \sigma(\hat{x}_i^\nu - \breve{x}^{\nu+1}),\\
&\breve{c}_{1,i}^{\nu+1} = \breve{c}_{1,i}^\nu + \sigma(\hat{\alpha}_{1,i}^\nu - \breve{\alpha}_1^{\nu+1}),\ \breve{c}_{2,i}^{\nu+1} = \breve{c}_{2,i}^\nu + \sigma(\hat{\alpha}_{2,i}^\nu - \breve{\alpha}_2^{\nu+1}).
\end{aligned}$$

如果（29）和（30）在 $\breve{\mathbf{z}}$ 处满足，则停止并令 $\tilde{\mathbf{z}} := \breve{\mathbf{z}}$。否则，令 $\nu := \nu + 1$，并转到**步骤1**。

4.2 全局收敛的理论保证

在本小节中，我们首先在以下较弱假设**A8** 或**A9**下，于定理4 证明算法1中的SDC定义良好，其中两阶段随机变分不等式（28）由算法2中的PHM方法近似求解。

A8.函数 $c(\cdot)$在 $\mathbb{R}^{m_1}$ 上有下界，并且对每个 $i \in [K]$，函数 $q_{\xi_i}(\cdot)$在 $\mathbb{R}^{m_2}$ 上有下界。

A9. 集合 $\{\mathbf{x_y} : \hat{f}(\mathbf{x_y}) \le \upsilon\}$ 对任意 $\upsilon \in \mathbb{R}$都有界。

定理 4 *Suppose that the successive PHM in* **1** *of Algorithm 1 is applied to approximately solve (MEAP) with* $\rho = \rho_t$ *in the* $t+1$ **步**() *-th iteration of SDC. Then the following statements hold.*

(i) For each fixed l *and the corresponding two-stage SVI in* **1 2** *of Algorithm 1 the* **步 -**, () () *criteria* 29 *and* 30 *can be satisfied by PHM in Algorithm 2 within finite iterations.*

(ii) If in addition assumption **8** *or* **9** *holds then the criterion* 31 *in* **1 3** *of ,,***A A**,() **步骤 - 步骤 - 步骤 -** *Algorithm 1 is satisfied after finite calls of* **1 1** *and* **1 2**.

Proof （i）令 $\{\breve{\mathbf{z}}^\nu\}$ 为用于求解（28 ）的算法 2生成的序列。根据 PHM 的收敛结果，当 ν 趋于无穷大时，$\breve{\mathbf{z}}^\nu$ 收敛到一点 $\breve{\mathbf{z}}^*$，该点是（28）中两阶段随机变分不等式的一个解。因此，

$$\lim_{\nu \to +\infty} \left\| H^{L,\mathrm{nat}}_{\rho_t, \mathbf{x}_\mathbf{y}^t, \varpi^t, \tau, D}(\breve{\mathbf{z}}^\nu) \right\| = \left\| H^{L,\mathrm{nat}}_{\rho_t, \mathbf{x}_\mathbf{y}^t, \varpi^t, \tau, D}(\breve{\mathbf{z}}^*) \right\| = 0.$$

因此，对于常数 $\varepsilon_1 := \eta_1\rho_t > 0$，存在 $\Gamma(\varepsilon_1) \in \mathbb{N} := \{1, 2, \ldots\}$，使得对所有$\nu > \Gamma(\varepsilon_1)$

$$\left\| H^{L,\mathrm{nat}}_{\rho_t, \mathbf{x}_{\mathbf{y}}^t, \varpi^t, \tau, D}(\check{\mathbf{z}}^\nu) \right\| \le \eta_1\rho_t.$$

根据定理 3，我们知道，$\check{\mathbf{x}}^*_{\mathbf{y}}$ 是$\check{\mathbf{z}}^*$ 的一个分块，且为（L-MEAP-ρ_t-$\mathbf{x}^t_{\mathbf{y}}$-$\varpi^t$-$\tau$）的最优解。由于$F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}$连续，我们有

$$\lim_{\nu \to +\infty} F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\check{\mathbf{x}}^\nu_{\mathbf{y}}) = F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\check{\mathbf{x}}^*_{\mathbf{y}}) = \min_{\mathbf{x}_{\mathbf{y}} \in \Phi} F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{x}_{\mathbf{y}}). \tag{33}$$

那么，对于常数 $\varepsilon_2 := \frac{\eta_2}{(l+1)^2} > 0$，存在 $\Gamma_1(\varepsilon_2) \in \mathbb{N}$，使得对于任意 $\nu > \Gamma_1(\varepsilon_2)$，

$$\left| F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\check{\mathbf{x}}^\nu_{\mathbf{y}}) - F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\check{\mathbf{x}}^*_{\mathbf{y}}) \right| \le \varepsilon_2, \tag{34}$$

因此，对于任意 $\nu > \Gamma_1(\varepsilon_2)$，

$$F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\check{\mathbf{x}}^\nu_{\mathbf{y}}) \le F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\check{\mathbf{x}}^*_{\mathbf{y}}) + \varepsilon_2 \le F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{x}_{\mathbf{y}}^{t,l}) + \varepsilon_2. \tag{35}$$

显然，对于 $\bar{\nu} = \max\{\Gamma(\varepsilon_1)\ \Gamma_1(\varepsilon_2)\} + 1$，$\check{\mathbf{z}} = \check{\mathbf{z}}^\nu$ 满足（29）和（30）。这保证了算法 2 经过有限次迭代后满足条件（29）和（30）。

(ii) 根据（19），结合$\mathbf{x}_{\mathbf{y}}^{t,l+1}$，（30）中的算法 1，以及（20），可得

$$\begin{aligned}
& F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t,l+1}) \\
\le\ & F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{x}_{\mathbf{y}}^{t,l+1}) - \frac{\tau_1}{2} \left\| x^{t,l+1} - x^{t,l} \right\|^2 - \frac{\tau_2}{2} \sum_{i=1}^K p_i \left\| y_i^{t,l+1} - y_i^{t,l} \right\|^2 \\
\le\ & F^L_{\rho_t, \mathbf{x}_{\mathbf{y}}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{x}_{\mathbf{y}}^{t,l}) + \frac{\eta_2}{(l+1)^2} - \frac{\tau_1}{2} \left\| x^{t,l+1} - x^{t,l} \right\|^2 - \frac{\tau_2}{2} \sum_{i=1}^K p_i \left\| y_i^{t,l+1} - y_i^{t,l} \right\|^2 \\
=\ & F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t,l}) - \frac{\tau_1}{2} \left\| x^{t,l+1} - x^{t,l} \right\|^2 - \frac{\tau_2}{2} \sum_{i=1}^K p_i \left\| y_i^{t,l+1} - y_i^{t,l} \right\|^2 + \frac{\eta_2}{(l+1)^2}.
\end{aligned} \tag{36}$$

将上述不等式从 l 0 求和至 L，令 L 趋于 $+\infty$，并利用事实 $\sum_{j=0}^{+\infty} \frac{\eta_2}{j+1)_2} = \frac{\pi^2}{6}\eta_2$，我们得到

$$\begin{aligned}
& \sum_{l=0}^{+\infty} \frac{\tau_1}{2} \left\| x^{t,l+1} - x^{t,l} \right\|^2 + \frac{\tau_2}{2} \sum_{l=0}^{+\infty} \sum_{i=1}^K p_i \left\| y_i^{t,l+1} - y_i^{t,l} \right\|^2 + \liminf_{\hat{L} \to +\infty} F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t,\hat{L}+1}) \\
\le\ & F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t,0}) + \sum_{j=0}^{+\infty} \frac{\eta_2}{(j+1)^2} < +\infty.
\end{aligned} \tag{37}$$

接下来，我们证明在假设 $\hat{L} \to +\infty$ $F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t\ L+1})$ 下，或者在假设 **A8** 或 **A9**. 根据（4），对于任意 $\rho > 0$,

$$F_\rho(\mathbf{x}_{\mathbf{y}}) = \hat{f}(\mathbf{x}_{\mathbf{y}}) + r_{1,\rho}(x) + \sum_{i=1}^K p_i r_{\xi_i,\rho}(y_i), \tag{38}$$

并且存在 $r(\rho) \in \mathbb{R}$ 使得

$$\inf_{\mathbf{x}_{\mathbf{y}} \in \mathbb{R}^{m_1 + m_2 K}} \left\{ r_{1,\rho}(x) + \sum_{i=1}^K p_i r_{\xi_i,\rho}(y_i) \right\} \ge \underline{r}(\rho),$$

根据假设 **A1**, (1), and(4) 中的 P_0 分别替换为 P_1 和 P_{ξ_i}。

因此，如果假设 **A8**成立，我们知道 $f(\mathbf{xy})$ 也下有界，ˆ因此 $\liminf_{\hat{L}\to+\infty} F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t\,L+1},)$ 下有界。此外，由（36），$\sum_{j=0}^{+\infty} \eta_{2_{(j+1)_2}} = \frac{\pi^2}{6}\eta_2$，以及（38），我们发现，对于任意 $l \geq 0$,

$$\hat{f}(\mathbf{x}_{\mathbf{y}}^{t,l+1}) + \underline{r}(\rho_t) \leq F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t,l+1}) \leq F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t,0}) + \frac{\pi^2}{6}\eta_2. \tag{39}$$

因此，如果假设 **A9** 成立，则有 $\{\mathbf{x}_{\mathbf{y}}^{t\,l+1}\}_{l\geq 0}$,有界，因此我们ˆ发现 $\liminf_{\hat{L}\to+\infty} F_{\rho_t}(\mathbf{x}_{\mathbf{y}}^{t\,L+1},)$ 也下有界。

因此，在假设 **A8** 或假设 **A9**成立的情况下，（37）表明

$$\lim_{l\to+\infty} \left\|\mathbf{x}_{\mathbf{y}}^{t,l+1} - \mathbf{x}_{\mathbf{y}}^{t,l}\right\| = 0. \tag{40}$$

那么，对于常数 $\varepsilon_3 := \eta_3\rho_t^2 > 0$，存在ˆ $L_1(\varepsilon_3) \in \mathbb{N}$，使得对于任意 $l > \hat{}L_1(\varepsilon_3)$,

$$\tau_1 \left\|x^{t,l+1} - x^{t,l}\right\| + \tau_2 \sum_{i=1}^{K} p_i \left\|y_i^{t,l+1} - y_i^{t,l}\right\| \leq \varepsilon_3.$$

因此，步骤 1-3中的判据（**31**）在有限次调用**步骤 1-1 和步骤 1-2 后得到满足**。□

Remark 2 事实上，任何用于求解极大单调光滑两阶段变分不等式的算法，都可以在**步骤 1-2** 中使用，以替代算法2中的 PHM方法，只要该算法具有如下收敛性结果：存在该算法的一个聚点 $\check{}\mathbf{z}$，使其为两阶段变分不等式的解。也就是说，存在一个无限子序列 $\mathcal{N}^\sharp \subseteq \mathbb{N}$使得

$$\lim_{\nu\to+\infty,\ \nu\in\mathcal{N}^\sharp} \mathbf{z}^\nu = \check{\mathbf{z}}.$$

算法1的定义良好性，即定理4 中的相同结论，可以通过类似的论证得到，不同之处仅在于，在关于陈述（i）的 $\mathbb{N}$ 证明中，应将其替换为 $\mathcal{N}^\sharp$ 。

接下来，我们将建立算法1中 SDC 方法的收敛性。我们需要以下假设。

A10. P_1 和 P_{ξ_i}、 $i \in [K]$ 是连续函数 .

A11. $-\nabla\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)\partial^\infty P(\mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)) \cap N_\Phi(\mathbf{x}_{\mathbf{y}}^*) = \{0\}.$

定理 5 (SDC 的收敛性) *Let* $\{\mathbf{z}^{t,\ l_t}\}$ *be the sequence generated by the SDC method in Algorithm 1 and let* $^*_{\mathbf{y}}$ *be an arbitrary accumulation point of* $\{^{t\,l_t}\}$. *Suppose that* , $\mathbf{xx}$, $\mathbf{y}$ **AAA**, **Ax( )** *assumptions* **5**, **7**, **10** *and* **11** *hold. Then* $^*_{\mathbf{y}}$ *is a KKT point of* 8 .

Proof 设 $\mathcal{T}$ 是 $\mathbb{N}$ 的一个无限子序列，使得 $\lim_{t\in\mathcal{T},\ t\to+\infty} \mathbf{x}^{t,\ l_t}\mathbf{y} = \mathbf{x}_{\mathbf{y}}^*$。根据算法 27、29和31中的式 (1，我们有

$$\left\|H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l_t-1},\varpi^{t,l_t-1},0,D}(\mathbf{z}^{t,l_t})\right\|$$

$$\leq \left\| H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l_t-1},\varpi^{t,l_t-1},\tau,D}(\mathbf{z}^{t,l_t}) - H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l_t-1},\varpi^{t,l_t-1},0,D}(\mathbf{z}^{t,l_t}) \right\| + \left\| H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l_t-1},\varpi^{t,l_t-1},\tau,D}(\mathbf{z}^{t,l_t}) \right\|$$

$$\leq \tau_1 \left\| x^{t,l_t} - x^{t,l_t-1} \right\| + \tau_2 \sum_{i=1}^{K} p_i \left\| y_i^{t,l_t} - y_i^{t,l_t-1} \right\| + \left\| H^{L,\mathrm{nat}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l_t-1},\varpi^{t,l_t-1},\tau,D}(\mathbf{z}^{t,l_t}) \right\|$$

$$\leq \eta_3 \rho_t^2 + \eta_1 \rho_t. \tag{41}$$

令 $t \in \mathcal{T}$， $t \to +\infty$ 代入式 (41)，并利用公式 (21)-(22、25)，可得

$$\lim_{t\in\mathcal{T},\ t\to+\infty} \left[\nabla \hat{f}(\mathbf{x}_{\mathbf{y}}^*) + \nabla G(\mathbf{x}_{\mathbf{y}}^*) M^{t,l_t} + \nabla \mathcal{U}(\mathbf{x}_{\mathbf{y}}^*) \lambda^{t,l_t-1} \right] = 0, \tag{42}$$

$$\lim_{t\in\mathcal{T},\ t\to+\infty} \left[M^{t,l_t} - \mathbf{Proj}_{\mathcal{K}^\circ}(M^{t,l_t} + G(\mathbf{x}_{\mathbf{y}}^*)) \right] = 0, \tag{43}$$

其中

$$M^{t,l_t} := (\alpha_1^{t,l_t}; \alpha_2^{t,l_t}; \pi_{1,1}^{t,l_t}; \ldots; \pi_{1,K}^{t,l_t}; \pi_{2,1}^{t,l_t}; \ldots; \pi_{2,K}^{t,l_t}),$$

$$\lambda^{t,l_t-1} := \begin{pmatrix} \lambda_1^{t,l_t-1} := \frac{1}{\rho_t}(U_1 x^{t,l_t} + u_1) - \frac{1}{\rho_t}\zeta_1^{t,l_t-1} \\ \lambda_{\xi_i}^{t,l_t-1} := p_i \left[\frac{1}{\rho_t}(U_{\xi_i} y_i^{t,l_t} + u_{\xi_i}) - \frac{1}{\rho_t}\zeta_{\xi_i}^{t,l_t-1} \right]; \forall i \in [K] \end{pmatrix}.$$

Let

$$\delta^{t,l_t} = \begin{pmatrix} \delta_1^{t,l_t} := \frac{1}{\rho_t} U_1 (x^{t,l_t} - x^{t,l_t-1}) \\ \delta_{\xi_i}^{t,l_t} := p_i \frac{1}{\rho_t} U_{\xi_i} (y_i^{t,l_t} - y_i^{t,l_t-1}); \forall i \in [K] \end{pmatrix}. \tag{44}$$

根据（31），我们可以推得 $\{\delta^{t,l_t}\}$ 有界且

$$\lim_{t\in\mathcal{T},\ t\to+\infty} \delta^{t,l_t} = 0. \tag{45}$$

鉴于 $\zeta^{t,\ l_t-1}_{,\ 1} \in$ 近端算子$\rho_t P_1$（$U_1 x^{t,\ l_t-1} + u_1$），$\zeta^{t,\ l_t-1}_{,\ \xi_i} \in$ 近端算子$\rho_t P_{\xi_i}$（$U_{\xi_i} y^{t,}_{,\ i}{}^{l_t-1} + u_{\xi_i}$），以及引理2，可得

$$\lambda_1^{t,l_t-1} - \delta_1^{t,l_t} = \frac{1}{\rho_t}(U_1 x^{t,l_t-1} + u_1) - \frac{1}{\rho_t}\zeta_1^{t,l_t-1} \in \partial P_1(\zeta_1^{t,l_t-1}),$$

$$\lambda_{\xi_i}^{t,l_t-1} - \delta_{\xi_i}^{t,l_t-1} = p_i \left[\frac{1}{\rho_t}(U_{\xi_i} y_i^{t,l_t-1} + u_{\xi_i}) - \frac{1}{\rho_t}\zeta_{\xi_i}^{t,l_t-1} \right] \in \partial \left(p_i P_{\xi_i}(\zeta_{\xi_i}^{t,l_t-1}) \right),\ i \in [K].$$

上述两个包含关系结合表 P（$\mathcal{U}$（$\mathbf{xy}$））的定义，即可得到1，即

$$\lambda^{t,l_t-1} - \delta^{t,l_t} \in \partial P(\zeta^{t,l_t-1}). \tag{46}$$

接下来，我们断言

$$\{\Gamma_{t,l_t} := \|M^{t,l_t}\| + \|\lambda^{t,l_t-1}\|\}_{t\in\mathcal{T}} \text{ is bounded.} \tag{47}$$

否则，从两边减去 $\nabla \mathcal{U}(\mathbf{x}_{\mathbf{y}}^*)\delta^{t,\ l_t}$，从式（42）两边减去该项，然后除以$\Gamma_{t,\ l_t}$并令 $t \in \mathcal{T}$趋于 $+\infty$，可得

$$\lim_{t\in\mathcal{T},\ t\to+\infty} \nabla G(\mathbf{x}_{\mathbf{y}}^*) \frac{M^{t,l_t}}{\Gamma_{t,l_t}} + \nabla \mathcal{U}(\mathbf{x}_{\mathbf{y}}^*) \frac{\lambda^{t,l_t-1} - \delta^{t,l_t}}{\Gamma_{t,l_t}} = 0. \tag{48}$$

由于 $\left\{ \frac{M^{t}{}^{,l_t}_{,\Gamma_t}}{} \right\}_{t\in\mathcal{T}}$和 $\left\{ \frac{\lambda^{t}{}^{,l_t-1}_{,l_t}-\delta^{t}{}^{,l_t}_{,l_t}}{\Gamma_t} \right\}_{t\in\mathcal{T}}$有界，因此存在 $\mathcal{T}_1 \subseteq \mathcal{T}$使得

$$\lim_{t\in\mathcal{T}_1,\ t\to+\infty} \Gamma_{t,l_t} = +\infty, \quad \lim_{t\in\mathcal{T}_1,\ t\to+\infty} \frac{M^{t,l_t}}{\Gamma_{t,l_t}} = \tilde{M}, \quad \lim_{t\in\mathcal{T}_1,\ t\to+\infty} \frac{\lambda^{t,l_t-1} - \delta^{t,l_t}}{\Gamma_{t,l_t}} = \tilde{\lambda}. \tag{49}$$

容易看出 $\|\tilde{M}\| + \|\tilde{\lambda}\| = 1$. 此外，将（48）除以 Γ_{t,l_t} 并令 $t \in \mathcal{T}_1, t \to +\infty$，可得

$$\tilde{M} = (\tilde{\alpha}_1; \tilde{\alpha}_2; \tilde{\pi}_{1,1}; \ldots; \tilde{\pi}_{1,K}; \tilde{\pi}_{2,1}; \ldots; \tilde{\pi}_{2,K}) \in \mathcal{K}^\circ, \ G(\mathbf{x}_\mathbf{y}^*) \in \mathcal{K}, \ \tilde{M}^T G(\mathbf{x}_\mathbf{y}^*) = 0. \tag{50}$$

也就是说，

$$\tilde{M} \in N_{\mathcal{K}}(G(x^*)).$$

由于 $\mathcal{U}(\mathbf{x}^{t,\ l_t-1}_\mathbf{y}) \to \mathcal{U}(\mathbf{x}^*_\mathbf{y})$，$\zeta^{t,\ l_t-1} \in$ 近端算子$\rho_t P(\mathcal{U}(\mathbf{x}^{t,\ l_t-1}_\mathbf{y}))$，且假设 A1 保证 P 下有界，由引理1 可知 $\zeta^{t,\ l_t-1} \to \mathcal{U}(\mathbf{x}^*_\mathbf{y})$。因此我们还有

$$\tilde{\lambda} \in \partial^\infty P(\mathcal{U}(\mathbf{x}^*_\mathbf{y})), \tag{51}$$

由此可得（46），$\Gamma^{1}_{t,\ l_t} \downarrow 0$；再由（49）中的最后一个等式、假设 A10 以及 $\partial^\infty P$ 的外半连续性可得 $\zeta^{t,\ l_t-1} \xrightarrow{P} \mathcal{U}(\mathbf{x}^*_\mathbf{y})$。

$\{\Gamma_{t,l_t}\}_{t\in\mathcal{T}}$ 无界仅有以下三种可能情况。

(i) $\tilde{\lambda} = 0$ and $\|\tilde{M}\| = 1$; (ii) $\tilde{M} = 0$ and $\|\tilde{\lambda}\| = 1$; (iii) $\tilde{M} \neq 0$ and $\tilde{\lambda} \neq 0$.

下面我们将证明这三种情况确实都不可能发生。

若 (i) 成立，则 (48)-(49) 意味着

$$0 = \nabla G^1(\mathbf{x}^*_\mathbf{y})\tilde{\alpha}_1 + \nabla G^2(\mathbf{x}^*_\mathbf{y})\tilde{\alpha}_2 + \sum_{i=1}^{K}\left(\nabla G^{1,\xi_i}(\mathbf{x}^*_\mathbf{y})\tilde{\pi}_{1,i} + \nabla G^{2,\xi_i}(\mathbf{x}^*_\mathbf{y})\tilde{\pi}_{2,i}\right). \tag{52}$$

根据假设 $\mathbf{x}^*_\mathbf{y}$ 处的系统 $G(\mathbf{xy}) \in \mathcal{K}$ 满足 RCQ **A5**，存在$\bar{d} \in \mathbb{R}^{m_1+m_2K}$ 满足（11）。现在我们考虑以下仅有的两种可能情形。

情形 1。当 $\alpha_2 = 0$ 且 $\pi_{2,\ i} = 0$ 对所有 $i \in [K]$，（52）变为

$$0 = \nabla G^1(\mathbf{x}^*_\mathbf{y})\tilde{\alpha}_1 + \sum_{i=1}^{K}\nabla G^{1,\xi_i}(\mathbf{x}^*_\mathbf{y})\tilde{\pi}_{1,i}, \tag{53}$$

其中 $\|(\tilde{\alpha}_1; \tilde{\pi}_{1,\ 1}; \ldots; \tilde{\pi}_{1,\ K})\| = 1$ 这与以下事实矛盾：梯度 $\nabla(G^1(\mathbf{x}^*_\mathbf{y}))_{j_1}$ 和 $\nabla(G^{1,\ \xi_i}(\mathbf{x}^*_\mathbf{y}))_{j_2}$，$i \in [K], j_1 \in [n_1], j_2 \in [n_2]$ 根据系统 $G(\mathbf{xy}) \in \mathcal{K}$ 在 $\mathbf{x}^*_\mathbf{y}$ 处满足的 RCQ 线性无关。

情形 2。否则，$\tilde{\alpha}_2 \neq 0$ 或 $(\tilde{\pi}_{2,\ 1}; \ldots; \tilde{\pi}_{2,\ K}) \neq 0$。根据 RCQ 的定义，由于系统 $G(\mathbf{xy}) \in \mathcal{K}$ 在 $\mathbf{x}^*_\mathbf{y}$ 处满足 RCQ，我们知道存在一个向量 $\bar{d} \in \mathbb{R}^{m_1+m_2K}$，使得

$$\tilde{\alpha}_1^T \mathcal{J}G^1(\mathbf{x}^*_\mathbf{y})\bar{d} = 0, \ \tilde{\pi}_{1,i}^T \mathcal{J}G^{1,\xi_i}(\mathbf{x}^*_\mathbf{y})\bar{d} = 0, \ i \in [K].$$

利用这些事实，并与方向 d 对（52）的两边取内积，我们得到

$$0 = \tilde{\alpha}_2^T \mathcal{J}G^2(\mathbf{x}^*_\mathbf{y})\bar{d} + \sum_{i=1}^{K}\tilde{\pi}_{2,i}^T \mathcal{J}G^{2,\xi_i}(\mathbf{x}^*_\mathbf{y})\bar{d}. \tag{54}$$

由（50），我们有

$$\tilde{\alpha}_2^T G^2(\mathbf{x}^*_\mathbf{y}) = 0 \quad \text{and} \quad \tilde{\pi}_{2,i}^T G^{2,\xi_i}(\mathbf{x}^*_\mathbf{y}) = 0, \ i \in [K]. \tag{55}$$

若 $\tilde{\alpha}_2 \neq 0$，利用式（55）中的第一个方程可得

$$\tilde{\alpha}_2^T \mathcal{J}G^2(\mathbf{x}^*_\mathbf{y})\bar{d} = \tilde{\alpha}_2^T\left(G^2(\mathbf{x}^*_\mathbf{y}) + \mathcal{J}G^2(\mathbf{x}^*_\mathbf{y})\bar{d}\right).$$

由于系统 $G(\mathbf{xy}) \in \mathcal{K}$ 在 $\mathbf{x}_\mathbf{y}^*$ 处满足 RCQ，因此可得

$$G^2(\mathbf{x}_\mathbf{y}^*) + \mathcal{J}G^2(\mathbf{x}_\mathbf{y}^*)\bar{d} \in \text{int}(-C_1).$$

这与 $\tilde{\alpha}_2 \in -C_1^\circ$ 一起得到

$$\tilde{\alpha}_2^T \mathcal{J}G^2(\mathbf{x}_\mathbf{y}^*)\bar{d} = \tilde{\alpha}_2^T \left(G^2(\mathbf{x}_\mathbf{y}^*) + \mathcal{J}G^2(\mathbf{x}_\mathbf{y}^*)\bar{d} \right) < 0. \tag{56}$$

否则（$\tilde{\pi}_{2,\ 1}$；...；$\tilde{\pi}_{2,\ K}$）$\neq 0$。则存在 $\bar{i} \in [K]$ 使得 $\tilde{\pi}_{2,\ \bar{i}} \neq 0$。利用（55）中的第二个方程，$\tilde{\pi}_{2,\ i} \in -C_{2,\ \xi_i}^\circ$ 和 $G^{2,\ \xi_i}(\mathbf{x}_\mathbf{y}^*) + \mathcal{J}G^{2,\ \xi_i}(\mathbf{x}_\mathbf{y}^*)\ \bar{d} \in \text{int}(-C_{2,\ \xi_i})$，可得

$$\tilde{\pi}_{2,i}^T \mathcal{J}G^{2,\xi_i}(\mathbf{x}_\mathbf{y}^*)\bar{d} = \tilde{\pi}_{2,i}^T \left(G^{2,\xi_i}(\mathbf{x}_\mathbf{y}^*) + \mathcal{J}G^{2,\xi_i}(\mathbf{x}_\mathbf{y}^*)\bar{d} \right) \leq 0, \quad \forall i \in [K], \tag{57}$$

and

$$\tilde{\pi}_{2,\bar{i}}^T \mathcal{J}G^{2,\xi_{\bar{i}}}(\mathbf{x}_\mathbf{y}^*)\bar{d} = \tilde{\pi}_{2,\bar{i}}^T \left(G^{2,\xi_{\bar{i}}}(\mathbf{x}_\mathbf{y}^*) + \mathcal{J}G^{2,\xi_{\bar{i}}}(\mathbf{x}_\mathbf{y}^*)\bar{d} \right) < 0. \tag{58}$$

鉴于（56）、（57）和（58），可知

$$\tilde{\alpha}_2^T \mathcal{J}G^2(\mathbf{x}_\mathbf{y}^*)\bar{d} + \sum_{i=1}^{K} \tilde{\pi}_{2,i}^T \mathcal{J}G^{2,\xi_i}(\mathbf{x}_\mathbf{y}^*)\bar{d} < 0, \tag{59}$$

这与（54）矛盾。

因此，系统 G（$\mathbf{xy}$）$\in \mathcal{K}$ 在 $\mathbf{x}_\mathbf{y}^*$ 处的 RCQ 保证情况（i）不存在。

若（ii）成立，则（48）-（51）推出

$$\nabla \mathcal{U}(\mathbf{x}_\mathbf{y}^*)\tilde{\lambda} = 0, \quad \tilde{\lambda} \in \partial^\infty P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*)).$$

然而，根据**假设A7**，我们可以推出 $\tilde{\lambda} = 0$，这与事实 $\|\tilde{\lambda}\| = 1$相矛盾。因此，情形(ii)也不存在。

如果 (i ii) holds, then (48)-(51) yield

$$\nabla G(\mathbf{x}_\mathbf{y}^*)\tilde{M} = -\nabla \mathcal{U}(\mathbf{x}_\mathbf{y}^*)\tilde{\lambda}, \ \tilde{M} \in N_\mathcal{K}(G(\mathbf{x}_\mathbf{y}^*)), \ \tilde{\lambda} \in \partial^\infty P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*)).$$

根据(16)，我们知道 $\nabla G(\mathbf{x}_\mathbf{y}^*)M \in N_\Phi(\mathbf{x}_\mathbf{y}^*)$利用 $\lambda \in \partial^\infty P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*))$和假设A11，可得 $\nabla \mathcal{U}(\mathbf{x}_\mathbf{y}^*)$ $\tilde{\lambda} = 0$，因为 $\tilde{\lambda} \in \partial^\infty P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*))$。根据假设A7，我们进而得到 $\tilde{\lambda} = 0$。这与(iii)中的事实 $\tilde{\lambda} \neq 0$ 相矛盾。因此，情形(iii)也不存在。

到目前为止，我们已证明(47)成立，即，序列 $\{\Gamma_{t,l_t}\}_{t\in\mathcal{T}}$ 有界。因此，显然序列 $\{\mathbf{z}^{t,l_t}\}_{t\in\mathcal{T}}$ 由算法1 生成的有界。因此，存在一个无限子序列 $\mathcal{T}_2 \subseteq \mathcal{T}_1$ 以及向量$\mathbf{z}^*$, M^* 和 λ^*，使得

$$\mathbf{z}^* = \lim_{t\in\mathcal{T}_2,\ t\to+\infty} \mathbf{z}^{t,l_t}, \quad M^* = \lim_{t\in\mathcal{T}_2,\ t\to+\infty} M^{t,l_t},$$

and

$$\lambda^* = \lim_{t\in\mathcal{T}_2,\ t\to+\infty} \lambda^{t,l_t-1} = \lim_{t\in\mathcal{T}_2,\ t\to+\infty} (\lambda^{t,l_t-1} - \delta^{t,l_t}) \in \partial P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*))$$

由（45）、（46）、$\mathbf{x}^{t,\ l_t-1}{}_\mathbf{y} P\ (\mathcal{U}\ (\cdot)\ ) \longrightarrow \mathbf{x}_\mathbf{y}^*$ 当 $t \in \mathcal{T}_2 \subseteq \mathcal{T}_1$时，以及 $\partial P\ (\mathcal{U}\ (\cdot)\ )$ 的外半连续性。利用（42）、（43），以及上述三个等式，我们得到

$$0 = \nabla \hat{f}(\mathbf{x}_\mathbf{y}^*) + \nabla G(\mathbf{x}_\mathbf{y}^*)M^* + \nabla \mathcal{U}(\mathbf{x}_\mathbf{y}^*)\lambda^*, \ \lambda^* \in \partial P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*)),$$
$$M^* = \text{Proj}_{\mathcal{K}^\circ}(M^* + G(\mathbf{x}_\mathbf{y}^*)),$$

其中最后一个等式等价于 $M^* \in N_\mathcal{K}(G(\mathbf{x}_\mathbf{y}^*))$。因此，存在 $\lambda^* \in \partial P(\mathcal{U}(\mathbf{x}_\mathbf{y}^*))$和 $M^* \in N_\mathcal{K}(G(\mathbf{x}_\mathbf{y}^*))$，使得

$$0 = \nabla \hat{f}(\mathbf{x}_\mathbf{y}^*) + \nabla \mathcal{U}(\mathbf{x}_\mathbf{y}^*)\lambda^* + \nabla G(\mathbf{x}_\mathbf{y}^*)M^*,$$

因此$\mathbf{x}_\mathbf{y}^*$ 是（8）的 KKT 点 . □

定理 6 *Suppose* $\Psi(\mathbf{x}_\mathbf{y}) = \sum_{i=1}^{m_1+m_2K} \psi_i(\mathbf{x}_\mathbf{y})$ *for some lsc functions* $\psi_i : \mathbb{R} \to \mathbb{R}$. *Let* $\mathcal{I} = \{i : \partial^\infty \psi_i(\mathbf{x}^*_\mathbf{y}) = \{0\}\}$ *and* $\mathcal{I}^c$ *be the complement of* $\mathcal{I}$ *with respect to* $\{1, \ldots, m_1 + m_2K\}$. *Assume further that for any* $i \in \mathcal{I}^c$ $\partial\psi_i(\mathbf{x}^*_\mathbf{y}) = \mathbb{R}$ *and the RCQ holds at* $\mathbf{x}^*_\mathbf{y}$ *for* $\mathbf{x}$ *the system* $G(\mathbf{x}_\mathcal{I}, (\mathbf{x}^*_\mathbf{y})_{\mathcal{I}^c}) \in \mathcal{K}$. *Let* $\{\mathbf{z}^t, l_t\}$ *be the sequence generated by the SDC method in Algorithm 1 and* $\mathbf{x}^*_\mathbf{y}$ *be an arbitrary accumulation point of* $\{\mathbf{x}^{t,l_t}\}$. *Then* $\mathbf{x}^*_\mathbf{y}$ *is a KKT* $\mathbf{x}$ *point of* 8 .

Proof 利用 Ψ 的可分性，我们知道对于任意 $\mathbf{x}\mathbf{y}$, $\mathcal{U}(\mathbf{x}\mathbf{y}) = \mathbf{x}\mathbf{y}$,

$$\partial P(\mathcal{U}(\mathbf{x}_\mathbf{y})) = \partial P(\mathbf{x}_\mathbf{y}) = \partial\Psi(\mathbf{x}_\mathbf{y}),$$

并且根据 [25,命题10。5],

$$\partial\Psi(\mathbf{x}_\mathbf{y}) = \partial_{(\mathbf{x}_\mathbf{y})_\mathcal{I}}\Psi(\mathbf{x}_\mathbf{y}) \times \partial_{(\mathbf{x}_\mathbf{y})_{\mathcal{I}^c}}\Psi(\mathbf{x}_\mathbf{y}).$$

根据 $\partial^\infty\psi_i((\mathbf{x}^*_\mathbf{y})_i) = \{0\}$ 对所有 $i \in \mathcal{I}$ 均成立，我们有

$$\partial^\infty\Psi_\mathcal{I}(\mathbf{x}^*_\mathbf{y}) \cap N_{\Phi_\mathcal{I}}(\mathbf{x}^*_\mathbf{y}) = \{0\}, \text{ with } \Phi_\mathcal{I} := \{\mathbf{x}_\mathbf{y} : G((\mathbf{x}_\mathbf{y})_\mathcal{I}, (\mathbf{x}^*_\mathbf{y})_{\mathcal{I}^c}) \in \mathcal{K}\},$$
$$\lambda \in \partial^\infty\Psi_\mathcal{I}(\mathbf{x}^*_\mathbf{y}) \Longrightarrow \lambda = 0.$$

与 (42)-(43) 类似，算法1 保证

$$\lim_{t\in\mathcal{T},\ t\to+\infty} \left[\nabla_{(\mathbf{x}_\mathbf{y})_\mathcal{I}}\hat{f}(\mathbf{x}^*_\mathbf{y}) + \nabla_{(\mathbf{x}_\mathbf{y})_\mathcal{I}}G(\mathbf{x}^*_\mathbf{y})M^{t,l_t} + \lambda^{t,l_t-1}\right] = 0,$$
$$\lim_{t\in\mathcal{T},\ t\to+\infty} \left[M^{t,l_t} - \mathbf{Proj}_{\mathcal{K}^\circ}(M^{t,l_t} + G(\mathbf{x}^*_\mathbf{y}))\right] = 0.$$

由于系统 $G((\mathbf{x}\mathbf{y})_\mathcal{I}, (\mathbf{x}^*_\mathbf{y})_{\mathcal{I}^c}) \in \mathcal{K}$在 $(\mathbf{x}^*_\mathbf{y})_\mathcal{I}$ 处满足 RCQ，运用与定理5类似的论证，并由上述两个方程可推得

$$0 \in \nabla_{(\mathbf{x}_\mathbf{y})_\mathcal{I}}\hat{f}(\mathbf{x}^*_\mathbf{y}) + \partial_{(\mathbf{x}_\mathbf{y})_\mathcal{I}}\Psi(\mathbf{x}^*_\mathbf{y}) + \nabla_{(\mathbf{x}_\mathbf{y})_\mathcal{I}}G(\mathbf{x}^*_\mathbf{y})N_\mathcal{K}(G(\mathbf{x}^*_\mathbf{y})).$$

根据假设， $\partial\psi_i((\mathbf{x}^*_\mathbf{y})_i) = \mathbb{R}$ 对所有 $i \in \mathcal{I}^c$ 均成立，显然有

$$0 \in \nabla_{(\mathbf{x}_\mathbf{y})_{\mathcal{I}^c}}\hat{f}(\mathbf{x}^*_\mathbf{y}) + \partial_{(\mathbf{x}_\mathbf{y})_{\mathcal{I}^c}}\Psi(\mathbf{x}^*_\mathbf{y}) + \nabla_{(\mathbf{x}_\mathbf{y})_{\mathcal{I}^c}}G(\mathbf{x}^*_\mathbf{y})N_\mathcal{K}(G(\mathbf{x}^*_\mathbf{y})).$$

结合上述两个包含关系，我们有

$$0 \in \nabla\hat{f}(\mathbf{x}^*_\mathbf{y}) + \partial\Psi(\mathbf{x}^*_\mathbf{y}) + \nabla G(\mathbf{x}^*_\mathbf{y})N_\mathcal{K}(G(\mathbf{x}^*_\mathbf{y})).$$

因此，$\mathbf{x}^*_\mathbf{y}$ 是 (8) 的 KKT 点。 □

4.3 有用的观察与扩展

在本小节中，我们给出一些有用的观察与扩展。发展 SDC 方法的思想以及假设的作用在图 1 中作了简要说明。

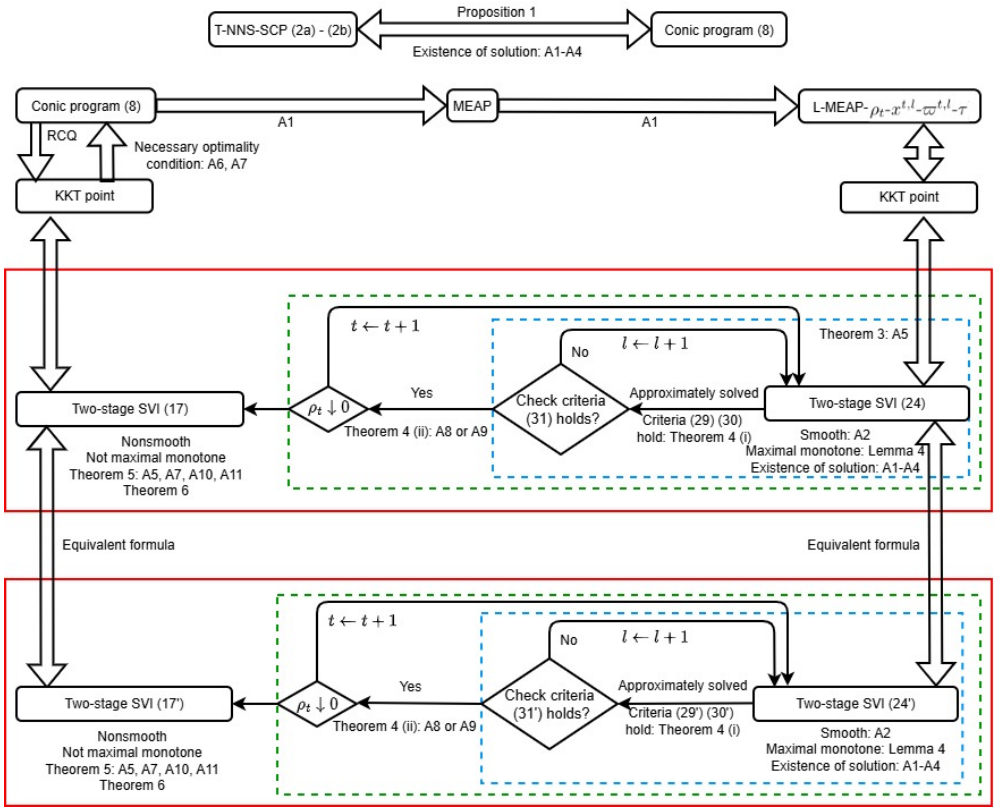

图1 所提出的 SDC-PHM 总体框架及其理论依据

观察与扩展

（I）第一个红框上方的前两行给出了将原始非凸非光滑 T-NNS-SCP 简化为凸光滑两阶段 SP 的思想。算法 1 在图 1 的第一个红框中加以说明。也就是说，通过 KKT 点，非光滑 T-NNS-SCP 由极大单调两阶段 SVI 出发，采用 SVI 方法求解。

（II）在实践中，8）的 KKT 点可能有不同但等价的公式，由此得到一个不同于（17）的等价两阶段变分不等式。对于（L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$）的 KKT 点及相应的两阶段变分不等式，采用同样的公式而非（24），仍可使用算法 1 的相同框架，并得到收敛性结果：任一聚点都是与原问题对应的两阶段变分不等式的解，且为（8）提供一个 KKT 点。图1底部的红框展示了这一扩展的示例。该扩展示例见第5节。

(III)（2a）和（2b）中的锥约束函数均为仿射函数。此要求仅用于引理4的证明，以确保与 L-MEAP 对应的（24）中的两阶段随机变分不等式是极大单调的。因此，如果存在

锥约束中的函数不是仿射函数，但（24）中的随机变分不等式仍为极大单调，且能够保证解的存在，那么算法 1 以及定理 4-6 中的收敛性结果仍然成立。对于这一推广，本文通篇以 $\nabla G(\mathbf{x}\boldsymbol{y})$ 替代常数矩阵 ∇G 。

(IV) 停止准则中的自然残差函数 $H^{L,\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau,D}_{\text{nat}}(\tilde{\mathbf{z}})$ 可以替换为任意其他残差函数 $H^{L,\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau,D}_{\text{res}}(\tilde{\mathbf{z}})$，只要它满足

$$H^{L,\text{res}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau,D}(\tilde{\mathbf{z}}) = 0 \Longleftrightarrow \tilde{\mathbf{z}} \in \text{SOL}\left(D, H^{L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau,D}(\mathbf{z})\right), \tag{60}$$

and

$$\begin{aligned}&\left\|H^{L,\text{res}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau,D}(\tilde{\mathbf{z}}) - H^{L,\text{res}}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},0,D}(\tilde{\mathbf{z}})\right\| \\ \leq\ & \tau_1\|x - x^{t,l}\| + \tau_2\sum_{i=1}^{K} p_i\|y_i - y_i^{t,l}\| \quad \text{for some constants } \tau_1, \tau_2 > 0. \end{aligned} \tag{61}$$

(V) 当只有第一阶段/第二阶段目标函数包含非光滑项时，我们只需相应地省略本文中针对第二阶段/第一阶段目标函数中的非光滑项所描述的过程。此时，算法 1 以及定理 4-6 中的收敛性结果仍然成立。

(VI) 假设 **A5** 要求 T-NNS-SCP (2a)- (2b) 存在一个斯莱特点，因此，针对系统$G(\mathbf{x}\boldsymbol{y}) \in \mathcal{K}$，RCQ 在任意可行点处成立。在假设 **A5**下，定理3 证明了凸规划 (L-MEAP-ρ_t-$\mathbf{x}_{\mathbf{y}}^t$-$\varpi^t$-$\tau$) 的最优性条件等价于光滑的极大单调两阶段 SVI (24)。如果两阶段 SVI (24) 具有特殊结构，例如它是一个 P矩阵线性互补问题 [4], ，那么即使不要求其为极大单调问题，我们仍然可以求解它。在这种情况下，我们不需要 **A5**，只需要求针对 $\{\mathbf{x}_{\mathbf{y}}^t\}$ 的任意聚点，RCQ 均成立，即可获得与本文相同的收敛性结果。

5 马科维茨均值-方差模型的非凸非光滑两阶段扩展

本节采用扩展的马科维茨均值-方差模型 [18] 作为基准模型。为使模型更符合实际资产管理需求，我们对投资组合权重施加非负约束以禁止卖空，并在目标函数中加入 ℓ_0-范数正则化项，以诱导

稀疏性，从而得到如下非凸非光滑两阶段随机规划：

$$\begin{aligned}\min_{x\in\mathbb{R}^n}\ & x^TQ_1x+\gamma\|x\|_0+\sum_{i=1}^{K}\frac{1}{K}\vartheta_i(x,\xi_i)\\ \text{s.t.}\ & e^Tx=1,\\ & \bar{r}_1^Tx\geq\bar{r}_1^{\min},\\ & x\geq 0,\end{aligned}\tag{62}$$

其中，$\vartheta_i(x,\xi_i)$，$i=1,...,K$ 是第二阶段问题的最优值

$$\begin{aligned}\min_{y_i\in\mathbb{R}^n}\ & y_i^TQ_{2,i}y_i\\ \text{s.t.}\ & e^Ty_i=1,\\ & \bar{r}_{2,i}^Ty_i\geq\bar{r}_{2,i}^{\min},\\ & y_i\geq 0,\\ & \|x-y_i\|\leq\tau_{2,i}.\end{aligned}\tag{63}$$

这里，

- x：第一阶段投资组合头寸向量；
- e：全1向量；
- $r_1,\bar{r}_1,\bar{r}_{\min 1}$：第一期的资产收益率、预期资产收益率和预期收益率下界，分别为；
- Q_1：第一期的资产收益率协方差；
- y_i：情景 i 下第二阶段投资组合头寸向量；
- $r_{2,i}$，$\bar{r}_{2,i}$，$\bar{r}_{\min 2,i}$：分别表示情景 i 下第二期的资产收益率向量、预期资产收益率向量和最低要求收益率；
- $Q_{2,i}$：情景 i 下第二期资产收益率的协方差；
- $\tau_{2,i}$：第一阶段决策与第二阶段决策之间距离的上界。

约束 $\|x-y_i\|\leq\tau_{2,i}$ 可通过引入辅助变量 t，视为一个二阶锥（SOC）约束，即二阶锥中的一个线性函数

$$(x-y_i;t)\in\{(s_1;s_2)\in\mathbb{R}^{n+1}\ :\ \|s_1\|\leq s_2\},$$

以及等式约束 $t=\tau_{2,i}$。它也可以等价地转换为一个非线性约束

$$\tau_{2,i}^2-\|x-y_i\|^2\geq 0,$$

这是非负锥中的一个非线性函数。为便于计算，这里我们采用第二种。假设**A5**成立。于是容易看出，假设**A1**-**A4**以及**A6** 成立。则（62）-（63）可重新表述为

以下等价的两阶段 SP：

$$
\begin{aligned}
\min_{\mathbf{x_y}}\ & x^T Q_1 x + \gamma\|x\|_0 + \sum_{i=1}^{K} \frac{1}{K} y_i^T Q_{2,i} y_i \\
\text{s.t. } & e^T x - 1 = 0, && \text{[Multiplier: } \alpha_1 \in \mathbb{R}] \\
& \bar{r}_1^T x - \bar{r}_1^{\min} \geq 0, && \text{[Multiplier: } \alpha_2^r \geq 0] \\
& x \geq 0, && \text{[Multiplier: } \alpha^x \geq 0] \\
& e^T y_i - 1 = 0,\ i \in [K], && \text{[Multiplier: } \pi_{1,i} \in \mathbb{R}] \\
& \bar{r}_{2,i}^T y_i - \bar{r}_{2,i}^{\min} \geq 0,\ i \in [K], && \text{[Multiplier: } \pi_{2,i}^r \geq 0] \\
& y_i \geq 0,\ i \in [K], && \text{[Multiplier: } \pi_{2,i}^y \geq 0] \\
& \tau_{2,i}^2 - \|x - y_i\|^2 \geq 0,\ i \in [K], && \text{[Multiplier: } \pi_{2,i}^\tau \geq 0].
\end{aligned} \tag{64}
$$

对于该问题，仅第一阶段涉及由 ℓ_0-范数构成的非凸非光滑项。通过使用莫罗包络，对其DC 公式的第二项进行线性化，并在当前迭代点 x^t、^{l}处添加正则化项，我们得到（L-MEAP-ρ_t-$\mathbf{x}_\mathbf{y}^t$-ϖ^t-τ）的目标函数

$$
\begin{aligned}
F^L_{\rho_t, x^{t,l}, \varpi^{t,l}, \tau}(\mathbf{x_y}) := x^T Q_1 x + \sum_{i=1}^{K} \frac{1}{K} y_i^T Q_{2,i} y_i + \frac{1}{2\rho_t}\|x\|^2 - R_{\rho_t}(x^{t,l}) - \langle x - x^{t,l}, \varpi^{t,l}\rangle \\
+ \frac{\tau}{2}\|x - x^{t,l}\|^2,
\end{aligned}
$$

根据引理2 的推导， [30]， $R_{\rho_t}(x^{t,l})$的次微分为

$$
(\partial R_{\rho_t}(x^{t,l}))_j = \begin{cases} \{0\}, & |(x^{t,l})_j| < \sqrt{2\gamma\rho_t}, \\ \{\frac{(x^{t,l})_j}{\rho_t}\}, & |(x^{t,l})_j| > \sqrt{2\gamma\rho_t}, \\ \left\{0, \frac{(x^{t,l})_j}{\rho_t}\right\}, & |(x^{t,l})_j| = \sqrt{2\gamma\rho_t}, \end{cases} \quad j \in [n].
$$

我们选取一个特定元素 $\varpi^{t,l} \in \partial R_{\rho_t}(x^{t,l})$：

$$
(\varpi^{t,l})_j = \begin{cases} 0, & |(x^{t,l})_j| < \sqrt{2\gamma\rho_t}, \\ \frac{(x^{t,l})_j}{\rho_t}, & |(x^{t,l})_j| \geq \sqrt{2\gamma\rho_t}, \end{cases} \quad j \in [n]. \tag{65}
$$

应用定理 3，我们发现，针对（L-MEAP-ρ_t-$\mathbf{x}_\mathbf{y}^t$-ϖ^t-τ）的（64）的 KKT条件等价于以下两阶段变分不等式，事实上这是一个两阶段混合

非线性互补问题：

$$\begin{cases} 2Q_1x+\dfrac{1}{\rho_t}x-\varpi^{t,l}+\tau(x-x^{t,l})+\alpha_1 e-\alpha_2^r\bar{r}_1-\alpha_2^x+\displaystyle\sum_{i=1}^{K}2\pi_{2,i}^{\tau}(x-y_i)=0, \\ \dfrac{2}{K}Q_{2,i}y_i+\pi_{1,i}e-\pi_{2,i}^r\bar{r}_{2,i}-\pi_{2,i}^{y_i}+2\pi_{2,i}^{\tau}(y_i-x)=0, \quad i\in[K], \\ -e^Tx+1=0, \\ 0\le\left(\bar{r}_1^Tx-\bar{r}_1^{\min}\right)\perp\alpha_2^r\ge 0, \\ 0\le x\perp\alpha_2^x\ge 0, \\ -e^Ty_i+1=0, i\in[K], \\ 0\le\left(\bar{r}_{2,i}^Ty_i-\bar{r}_{2,i}^{\min}\right)\perp\pi_{2,i}^r\ge 0, \quad i\in[K], \\ 0\le y_i\perp\pi_{2,i}^{y_i}\ge 0, \quad i\in[K], \\ 0\le\left(\tau_{2,i}^2-\|x-y_i\|^2\right)\perp\pi_{2,i}^{\tau}\ge 0, \quad \forall i\in[K]. \end{cases} \tag{66}$$

这里，$d_1\perp d_2$ 指 $d_1^Td_2=0$。显然，上述两阶段变分不等式等价于以下简洁形式的两阶段变分不等式：

$$\begin{cases} 0\le\left(2Q_1x+\frac{1}{\rho_t}x-\varpi^{t,l}+\tau(x-x^{t,l})+\alpha_1e-\alpha_2^r\bar{r}_1+\sum_{i=1}^{K}2\pi_{2,i}^{\tau}(x-y_i)\right)\perp x\ge 0, \\ 0\le\left(\frac{2}{K}Q_{2,i}y_i+\pi_{1,i}e-\pi_{2,i}^r\bar{r}_{2,i}+2\pi_{2,i}^{\tau}(y_i-x)\right)\perp y_i\ge 0,\ i\in[K], \\ -e^Tx+1=0, \\ 0\le\left(\bar{r}_1^Tx-\bar{r}_1^{\min}\right)\perp\alpha_2^r\ge 0, \\ -e^Ty_i+1=0,\ i\in[K], \\ 0\le\left(\bar{r}_{2,i}^Ty_i-\bar{r}_{2,i}^{\min}\right)\perp\pi_{2,i}^r\ge 0,\ i\in[K], \\ 0\le\left(\tau_{2,i}^2-\|x-y_i\|^2\right)\perp\pi_{2,i}^{\tau}\ge 0,\ i\in[K]. \end{cases} \tag{67}$$

在计算中，我们将使用式（67）中的简洁随机变分不等式，由此通过算法 1，可得到原问题（64）的 KKT 条件所对应的简洁两阶段变分不等式。容易看出，式（67）可写为

$$0\in H^L_{\rho_t,x^{t,l},\varpi^{t,l},\tau}(\mathbf{z})+N_D(\mathbf{z}),$$

其中

$$\mathbf{z}:=(x;y_1;\ldots;y_K;\alpha;\pi_1;\ldots;\pi_K)\in D:=\mathbb{R}^n_+\times(\mathbb{R}^n_+)^{\otimes K}\times(\mathbb{R}\times\mathbb{R}_+)\times(\mathbb{R})^{\otimes K}\times(\mathbb{R}^2_+)^{\otimes K},$$

其中 $\alpha := (\alpha_1; \alpha_2^r)$，且 $\pi_i := (\pi_{1,i}; \pi_{2,i}^r; \pi^{\tau_{2,i}})$ 对于 $i \in [K]$，并且

$$H^L_{\rho_t, x^{t,l}, \varpi^{t,l}, \tau}(\mathbf{z}) := \begin{pmatrix} 2Q_1 x + \frac{1}{\rho_t} x - \varpi^{t,l} + \tau(x - x^{t,l}) + \alpha_1 e - \alpha_2^r \bar{r}_1 + \sum_{i=1}^K 2\pi_{2,i}^{\tau_{2,i}}(x - y_i) \\ \frac{2}{K} Q_{2,i} y_i + \pi_{1,i} e - \pi_{2,i}^r \bar{r}_{2,i} + 2\pi_{2,i}^{\tau_{2,i}}(y_i - x); \quad i \in [K] \\ -e^T x + 1 \\ \bar{r}_1^T x - \bar{r}_1^{\min} \\ -e^T y_i + 1; \quad i \in [K] \\ \bar{r}_{2,i}^T y_i - \bar{r}_{2,i}^{\min}; \quad i \in [K] \\ \tau_{2,i}^2 - \|x - y_i\|^2; \quad i \in [K] \end{pmatrix}.$$

PHM 在算法 1 中的成功实现取决于能否确立下述命题所述的结果。

命题 7 *The two-stage SVI (67) is of maximal monotone type.*

Proof 根据 [26, 定理 3。5], 两阶段变分不等式（67）的单调性等价于每个情景下分解变分不等式的单调性 $i \in [K]$:

$$\begin{cases} 0 \le \left(2Q_1 x + \frac{1}{\rho_t} x - \varpi^{t,l} + \tau(x - x^{t,l}) + \alpha_1 e - \alpha_2^r \bar{r}_1 + 2\pi_{2,i}^{\tau}(x - y_i)\right) \perp x \ge 0, \\ 0 \le \left(\frac{2}{K} Q_{2,i} y_i + \pi_{1,i} e - \pi_{2,i}^r \bar{r}_{2,i} + 2\pi_{2,i}^{\tau}(y_i - x)\right) \perp y_i \ge 0, \\ -e^T x + 1 = 0, \\ 0 \le \left(\bar{r}_1^T x - \bar{r}_1^{\min}\right) \perp \alpha_2^r \ge 0, \\ -e^T y_i + 1 = 0, \\ 0 \le \left(\bar{r}_{2,i}^T y_i - \bar{r}_{2,i}^{\min}\right) \perp \pi_{2,i}^r \ge 0, \\ 0 \le \left(\tau_{2,i}^2 - \|x - y_i\|^2\right) \perp \pi_{2,i}^{\tau} \ge 0. \end{cases} \tag{68}$$

设 $\mathbf{z_P} := (x;\ y_i;\ \alpha_1;\ \alpha_2^r; \pi_{1,i}; \pi_{2,i}) \in D_i := \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^2_+$。我们只需证明与变分不等式 (68) 相关联的函数 $H^{i}_{\rho_t, \mathbf{x}_\mathbf{y}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{z_P})$ 在集合 D_i 上单调，其中

$$H^{i,L}_{\rho_t, \mathbf{x}_\mathbf{y}^{t,l}, \varpi^{t,l}, \tau}(\mathbf{z}_\mathrm{P}) := \begin{pmatrix} 2Q_1 x + \left(\frac{1}{\rho_t} + \tau\right) x - \varpi^{t,l} - \tau x^{t,l} + \alpha_1 e - \alpha_2^r \bar{r}_1 + 2\pi_{2,i}^{\tau}(x - y_i) \\ \frac{2}{K} Q_{2,i} y_i + \pi_{1,i} e - \pi_{2,i}^r \bar{r}_{2,i} + 2\pi_{2,i}^{\tau}(y_i - x) \\ -e^T x + 1 \\ \bar{r}_1^T x - \bar{r}_1^{\min} \\ -e^T y_i + 1 \\ \bar{r}_{2,i}^T y_i - \bar{r}_{2,i}^{\min} \\ \tau_{2,i}^2 - \|x - y_i\|^2 \end{pmatrix}.$$

通过直接计算，$H^{i,L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau}(\mathbf{z}_{\mathrm{P}})$ 的雅可比矩阵为 $\mathcal{J}H^{i,L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau}(\mathbf{z}_{\mathrm{P}})$

$$\mathcal{J}H^{i,L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau}(\mathbf{z}_{\mathrm{p}}) = \begin{pmatrix} 2Q_1 + (2\pi^{\tau}_{2,i} + \tau + \frac{1}{\rho_t})I_n & -2\pi^{\tau}_{2,i}I_n & e & -\bar{r}_1 & 0 & 0 & 2(x-y_i) \\ -2\pi^{\tau}_{2,i}I_n & \frac{2}{K}Q_{2,i} + 2\pi^{\tau}_{2,i}I_n & 0 & 0 & e & -\bar{r}_{2,i} & -2(x-y_i) \\ -e^T & 0 & 0 & 0 & 0 & 0 & 0 \\ \bar{r}_1^T & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -e^T & 0 & 0 & 0 & 0 & 0 \\ 0 & \bar{r}_{2,i}^T & 0 & 0 & 0 & 0 & 0 \\ -2(x-y_i)^T & 2(x-y_i)^T & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

令 $\mathbf{u} = (u_x;\ u_y;\ u_{\alpha_1};\ u_{\alpha_2^r};\ u_{\pi_1};\ u_{\pi_r};\ u_{\pi_\tau}) \in \mathbb{R}^{2n+5}$ 为一个任意向量，其分块方式与 $\mathcal{J}H^{i,L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t},\varpi^{t},\tau}(\mathbf{z}_{\mathrm{P}})$ 的列分块方式相同。于是得到二次型

$$\mathbf{u}^T\mathcal{J}H^{i,L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau}(\mathbf{z}_{\mathrm{p}})\mathbf{u} = u_x^T\left(2Q_1 + (\tau + \frac{1}{\rho_t})I_n\right)u_x + \frac{2}{K}u_y^TQ_{2,i}u_y + 2\pi^{\tau}_{2,i}\|u_x - u_y\|^2 \geq 0,$$

由于 $Q_1, Q_{2,i}$ 是半正定矩阵，$\rho_t > 0, \tau > 0$ 和 $\pi^{\tau}_{2,i} \geq 0$。因此，根据 [10, 命题 2.3.2]，与变分不等式 (68) 相关的函数 $H^{i,L}_{\rho_t,\mathbf{x}_{\mathbf{y}}^{t,l},\varpi^{t,l},\tau}(\mathbf{z}_{\mathrm{P}})$ 在 D_i 上是单调的。根据 [25, 例 12.48]，两阶段变分不等式 (67) 的单调性与所涉及函数的连续性相结合，意味着 (67) 属于极大单调型。

□

在下一节的数值实验中，我们使用自然残差函数，并通过高效的半光滑牛顿法求解 PHM 中每个情景对应的变分不等式 (68) [9]。

6 数值实验

本节针对第5节所示的应用开展数值实验，从多个角度评估所提出的模型和 SDC-PHM 算法。所有实验均在 Windows 10 系统下，使用配备 64 GB 内存、主频为 3.70 GHZ 的 Intel Core 10 CPU，并通过 MATLAB R2021a 完成。

为构建所提出的两阶段非凸非光滑随机模型（62）-（63）的数据参数，我们使用从2023年1月5日至2025年7月1日期间标普500指数中选取的 $n = 40$ 只股票的历史日收益率，共包含 $T = 500$ 个每日观测值。令 $P_{i,t}$ 表示资产 i 在第 t 天的收盘价。第 t 天的日收益率向量定义为 $R_t := (R_{1,t};\ \ldots;\ R_{n,t}) \in \mathbb{R}^n$，其中 $R_{i,t} = (P_{i,t} - P_{i,t-1})/P_{i,t-1}$。利用这些历史数据，我们通过以下表达式估计第一阶段的预期收益率向量 $\bar{r}_1$ 和协方差矩阵 Q_1。

$$\bar{r}_1 = \frac{1}{T}\sum_{t=1}^{T} R_t, \quad Q_1 = \frac{1}{T-1}\sum_{t=1}^{T}(R_t - \bar{r}_1)(R_t - \bar{r}_1)^\top + \epsilon I, \tag{69}$$

其中 $\epsilon = 10^{-9}$ 是为确保数值稳定性而引入的正则化参数。第二阶段资产收益率 $\bar{r}_{2,i}$ 以及协方差矩阵 $Q_{2,i}$ 对于所有 $i \in [K]$

由多元 GARCH 模型生成，其选取使得资产收益率落入 $[0.5, 1.5]$ 范围内，波动率在10%左右变化，且相关性介于 -0.5 与0.5，如 [18]中所采用。我们考虑等权投资组合 $\bar{x}: = \frac{1}{n}e$ 并设定

$$\bar{r}_1^{\min} := \bar{r}_1^T \bar{x} - 0.05|\bar{r}_1^T \bar{x}|, \quad \bar{r}_{2,i}^{\min} := \bar{r}_{2,i}^T \bar{x} - 0.05|\bar{r}_{2,i}^T \bar{x}|, \ i \in [K].$$

于是，假设A5成立，因为 $\bar{x}$ 是（64）可行域的斯莱特点。

我们测试不同的 $K \in \{1000, 3000, 5000\}$。对于每种 K选择，我们通过独立生成数据参数 $\bar{r}_{2、i}$和 $Q_{2、i}$、 $i \in [K]$，生成 20 个测试问题，每个测试问题的数据参数重复试验 20 次。为确保结果可复现，每次重复试验均分配不同的随机种子。表 2汇总了实验所用测试问题的维度。

表 2 测试问题的维度：约束数量（♯）和变量数量。

	问题规模	
K	♯约束（42+43K）	♯ 变量（40+40K）
1,000	43,042	40,040
3,000	129,042	120,040
5,000	215,042	200,040

为了分析诱导稀疏正则化项 $\gamma\|x\|_0$ 和 SOC 约束的影响，我们考虑四种模型，分别标记为 A、B、C 和 D。模型A是所提出的模型（62)-(63)。基于模型A，模型B去除了 SOC 约束；模型C去除了 l_0-范数正则化项；模型D则同时去除了 SOC 约束和 l_0-范数正则化项。由于模型A和模型B都涉及非凸、非光滑的 l_0-范数，因此我们使用所提出的 SDC-PHM 求解它们。模型C和模型D是凸的，因此我们直接使用 PHM 求解。

我们使用以下准则来衡量计算得到的解 ˆ**z**ˆ 与 $\lambda \in \partial\gamma\|\hat{}x\|_0$ 的性能，其中这些量基于引理 2 和 [30] 构造如下：

$$\hat{\lambda} := \frac{1}{\hat{\rho}}\hat{x} - \hat{\varpi}, \tag{70}$$

其中，ˆρ 是与计算得到的解ˆx相对应的平滑参数，ˆϖ 的各分量定义为

$$(\hat{\varpi})_j = \begin{cases} 0, & |(\hat{x})_j| < \sqrt{2\gamma\hat{\rho}}, \\ \frac{(\hat{x})_j}{\hat{\rho}}, & |(\hat{x})_j| \geq \sqrt{2\gamma\hat{\rho}}, \end{cases} \quad j \in [n]. \tag{71}$$

根据KKT系统定义的可行性误差和最优性误差（KKT_∞和$\mathrm{KKT}_{\mathrm{rel}}$）、与SOC约束相关的值（soc），以及与原变量相关的可行性误差（FeasError），分别定义为

$$\mathrm{KKT}_\infty := \|H_D^{\mathrm{nat}}(\hat{\mathbf{z}};\hat{\lambda})\|_\infty = \max_i |(H_D^{\mathrm{nat}}(\hat{\mathbf{z}};\hat{\lambda}))_i|, \tag{72a}$$

$$\mathrm{KKT}_{\mathrm{rel}} := \frac{\|H_D^{\mathrm{nat}}(\hat{\mathbf{z}};\hat{\lambda})\|}{1+\|\hat{\mathbf{z}}\|}, \tag{72b}$$

$$\mathrm{soc}(\hat{\mathbf{z}}) := \frac{1}{K}\sum_{i=1}^{K}\|x-y_i\|, \tag{72c}$$

$$\mathrm{FeasErr}(\hat{\mathbf{z}}) := \|G_{\mathrm{eq}}(\hat{\mathbf{z}})\|^2 + \|\max(0, G_{\mathrm{ineq}}(\hat{\mathbf{z}}))\|^2, \tag{72d}$$

其中 G_{eq} 和 G_{ineq} 分别是 G中对应于等式约束和不等式约束的子向量。

在数值实验中，我们将正则化参数 $\gamma = 10^{-5}$ 设置于模型A和模型B，$\tau_{2,\,i} \equiv 0.2$ 对所有 $i \in [K]$ 设置于模型A和模型C，并将SDC-PHM在算法1和算法2中的参数设置如下。

$$\tau = 10^{-4},\ \eta_1 = \eta_2 = \eta_3 = K/5,\ \sigma = 1,$$

以及平滑参数

$$\rho_0 = 1; \quad \rho_{t+1} = 0.8\rho_t \text{ for } t \geq 0.$$

为了终止SDC-PHM 方法或PHM方法，我们首先要求

$$|\mathrm{obj}_{t-1} - \mathrm{obj}_t| \leq 10^{-3},$$

其中“obj”指所关注的每个模型的目标值。此外，对于SDC-PHM，我们还要求平滑参数

$$\rho_t \leq 10^{-4}.$$

ˆ我们评估计算得到的解ˆ$\mathbf{z}$ 连同 λ，这些解由不同模型得到，评估指标包括ˆx 的非零元素数（nnz）、（72）中的准则、CPU时间以及PHM迭代次数（PHM iter），其指求解变分不等式（32）的总次数，位于算法2中，这是SDC-PHM的主要计算成本。表3 所报告的结果表示上述性能指标在这20次重复试验上的表现。

从表 3的“nnz”列可以发现，采用 ℓ_0-范数正则化的模型A和模型B始终具有较高的稀疏性，非零元素数平均为14。相比之下，缺少基数约束的模型C和模型D会产生更加稠密的投资组合（nnz在24至32之间）。这证实了所提出的 ℓ_0-正则化项在筛选重要资产方面的有效性。从表 3的“soc”列中，比较模型A与模型B（同理，比较模型C

表3 使用20次重复试验的均值（不带括号）和标准差（括号内）比较模型 A、B（无 SOC）、C（无 ℓ_0）和 D（无 SOC、无 ℓ_0）的计算得到的解。

K	模型	nnz	KKT_∞	$\mathrm{KKT}_{\mathrm{rel}}$	soc	可行性误差	CPU（秒）	PHM 迭代次数
1,000	A	14 (0)	7.2e-3 (1.2e-3)	2.2e-4 (3.8e-5)	0.17 (1.3e-3)	2.1e-5 (7.2e-6)	22 (3)	55 (6)
	B	14 (0)	3.9e-3（4.6e-4）	2.9e-2 (4.3e-3)	0.31 (1.4e-4)	2.4e-6 (8.3e-7)	16 (1)	42 (0)
	C	32 (2)	4.7e-2 (1.8e-2)	1.7e-2（9.5e-3）	0.20 (4.5e-3)	1.3e-5 (4.4e-6)	117 (18)	269 (41)
	D	24 (1)	5.0e-2 (2.1e-2)	1.7e-2 (9.2e-3)	0.28 (2.1e-2)	9.0e-6 (4.4e-6)	128 (19)	301 (43)
3,000	A	14 (0)	2.2e-2 (6.4e-3)	1.5e-4 (2.0e-5)	0.17 (1.4e-3)	7.8e-5 (2.4e-5)	36 (5)	51 (7)
	B	14 (0)	5.0e-3 (6.8e-4)	5.0e-2 (7.6e-3)	0.31（9.1e-5）	6.9e-6 (1.7e-6)	30 (2)	42 (0)
	C	32 (2)	5.0e-2 (1.7e-2)	9.9e-3 (5.5e-3)	0.20（4.5e-3）	4.0e-5 (1.3e-5)	218 (32)	270 (40)
	D	24 (1)	5.4e-2 (2.0e-2)	9.8e-3（5.3e-3）	0.28（2.1e-2）	2.7e-5 (1.3e-5)	243 (37)	301 (43)
5,000	A	14 (0)	3.7e-2 (1.1e-2)	1.2e-4 (1.4e-5)	0.17 (1.4e-3)	1.4e-4 (4.8e-5)	56 (9)	50 (8)
	B	14 (0)	5.0e-3 (6.8e-4)	6.3e-2 (9.6e-3)	0.31 (8.4e-5)	1.2e-5 (2.9e-6)	46 (2)	42 (0)
	C	32 (1)	5.2e-2 (1.8e-2)	7.7e-3 (4.3e-3)	0.20 (4.4e-3)	6.7e-5 (2.2e-5)	357 (56)	270 (40)
	D	24 (1)	5.6e-2 (2.0e-2)	7.6e-3 (4.1e-3)	0.28 (2.1e-2)	4.6e-5 (2.2e-5)	384（56）	302（43）

与模型D），可以看到，加入SOC约束显著降低了SOC指标（表示第一阶段决策与第二阶段决策之间的距离）。例如，模型A的SOC平均值约为0.17，而模型B则升至约0.31，对所有 $i \in [K]$均超过$\tau_{2,\,i} \equiv 0.2$ 。这一结果表明，SOC约束能够有效确保第二阶段调整后的决策不会与第一阶段决策相差过大。

从表$_\infty$中“$\mathrm{KKT}_{\mathrm{rel}}$”和“FeasError”这几列可以看出，SDC-PHM 成功为模型 A 和 B 生成了近似 KKT 点，而 PHM 成功为模型 C 和 D 生成了近似 KKT 点。有趣的是，如表3最后两列所示，对于非凸不连续模型 A 和 B，所提出的 SDC-PHM 比用于凸模型 C 和 D 的 PHM 收敛快得多。这似乎与我们的预期不同，因为 SDC-PHM 将 PHM 用作子程序，而 ℓ_0-正则化项是非凸且不连续的，通常被认为难以处理。为了解释这一现象，我们在图2 中绘制了四个模型在一次具有代表性的运行中的目标值随 CPU 时间变化的曲线（蓝线），以及第一阶段决策向量的基数随 CPU 时间变化的曲线（红线）；该次运行的情景规模为 $K = 5000$。图2 显示，由于存在 ℓ_0-正则化项，模型 A 和 B 的基数下降得快得多，从而使模型 A 和 B 的目标值也下降得更快，并使 SDC-PHM 比用于模型 C 和 D 的 PHM 收敛得更快。我们还在图3 中绘制了相应的可行性误差随 CPU 时间变化的曲线，结果表明，模型 A 和 B 的 SDC-PHM 达到了与模型 C 和 D 的 PHM 方法相当的可接受可行性误差水平。

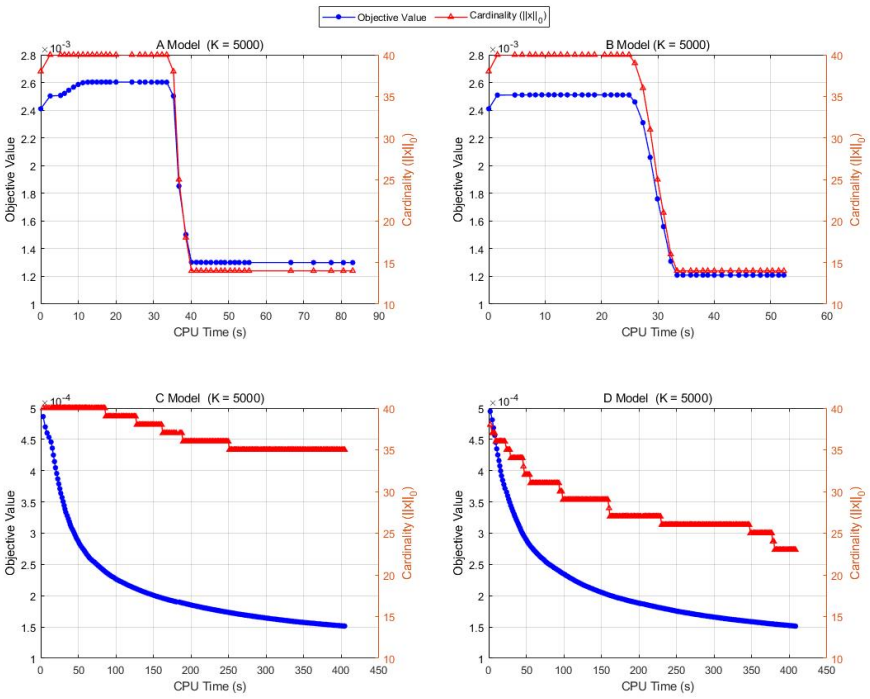

图2 模型 A–D 在 K =5000 条件下的目标值（蓝色）和基数（红色）随 CPU时间的变化

7 结论

我们提出 SDC 方法，通过 SVI 方法求解一类两阶段非凸非光滑随机锥规划。我们证明，在温和假设下，SDC 方法存在一个聚点，且任何聚点都是 KKT 点。SDC 方法为处理两阶段非凸非光滑随机锥规划提供了一种新方法，并且具有灵活性，可进行各种扩展，以拓宽其适用范围或提高计算效率，如第4.3小节所示。SDC 方法有望用于求解第二阶段目标函数与第一阶段决策变量相互交织的问题，这将是我们未来的研究方向。

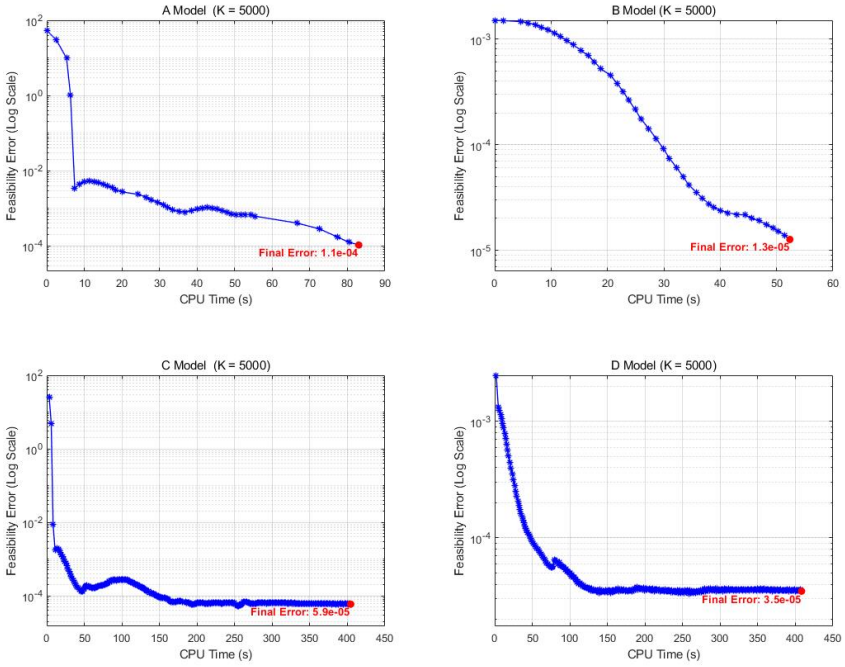

图3 模型 A–D 在 $K = 5000$ 下的可行性误差与 CPU时间

致谢。 本工作得到国家自然科学基金委员会（项目编号：12171027；12571316）和国家留学基金管理委员会（CSC）（项目编号：202507090144）的资助。

参考文献

[1]Bonnas, J.F., Shapiro, A.：《优化问题的扰动分析》。施普林格，纽约（2000）[2] Borges, P., Sagastiz´abal, C., Solodov, M.：一种用于完全参数化凸问题的正则化平滑方法及其在凸与非凸两阶段随机规划中的应用。数学规划 **189(1)**，117–149（2021）[3] Chen, X., Guo, L., Lu, Z., Ye, J.J.：一种用于非 Lipschitz 非凸规划的增广拉格朗日方法。SIAMJ. Numer.Anal. **55(1)**，168–193（2017）[4] Cottle, R.W., Pang, J.S., Stone, R.E.：《线性互补问题》。SIAM，费城（2009）[5] Chen, X., Pong, T.K., Wets, R.J.：两阶段随机变分不等式：

一种erm-解程序。数学规划 **165**，71–111（2017）

[6] Chen, X., Shapiro, A., Sun, H.：两阶段随机广义方程样本均值近似的收敛性分析。SIAM优化学报 **29(1)**，135–161（2019年）[7] Chen, X., Sun, H., Xu, H.：两阶段随机及分布鲁棒线性互补问题的离散近似。数学规划 **177**，255–289（2019年）[8] Cui, X., Zhang, L.：一种求解非单调随机变分不等式的局部渐进对冲算法。运筹学数学 **49**，1915–1937（2024）[9]De, L., Facchinei, F., Kanzow, C.：一种求解非线性互补问题的半光滑方程方法。数学规划 **75**，407–439（1996）[10] Facchinei, F., Pang, J.S.：《有限维变分不等式与互补问题》。施普林格，纽约（2003）[11] Kanzow, C., Ferenczi, I., Fukushima, M.：关于无严格互补条件下线性与非线性二阶锥规划半光滑牛顿法的局部收敛性。SIAM优化学报 **20(1)**，297–320（2009）[12] Li, H., Cui, Y.：一种求解具有非凸补救函数的两阶段随机规划的分解算法。SIAM优化学报 **34(1)**，306–335（2024）[13] Liu, J., Cui, Y., Pang, J.S., Sen, S.：具有线性双参数化二次补救的两阶段随机规划。SIAM优化学报**30(3)**，2530–2558（2020）[14] Liu, A., Lau, V.K.N., Zhao, M.J.：两阶段随机非凸优化的在线逐次凸近似。IEEE Trans. SignalProcess. **66(22)**，5941–5955（2018）[15] Liu, T., Pong, T.K., Takeda, A.：一种针对一类非凸非光滑优化问题的逐次差分凸近似方法。数学规划 **176**，339–367（2019）[16]Li, M., Zhang, C.：源于随机规划的两阶段随机变分不等式。优化理论与应用杂志 **186(1)**，324–343（2020）[17]Li, M., Zhang, C., Ding, M.X., Lv, R.P.：疫情管理中医疗物资储存与动态分配的两阶段随机变分不等式模型。应用数学建模 **102**，35–61（2022）

.[18] Mehrotra, S., Ozevin, M.G.：关于两阶段随机锥规划内点分解算法的实现。SIAM优化学报 **19(4)**, 1846–1880（2009年）[19]Mordukhovich, B.S.：《变分分析及其应用》。施普林格，沙姆（2018）[20] Narushima, Y., Sagara, N., Ogasawara, H.：一种用于二阶锥互补问题的Fischer-Burmeister函数光滑牛顿法。优化理论与应用杂志 **149(1)**, 79–101（2011）[21] Pan, S., Chen, J.S.：一种基于单参数二阶锥互补函数类的二阶锥互补问题半光滑牛顿法。计算优化与应用 **45(1)**, 59–88（2010）[22] Pan, L.L., Luo, Z., Xiu, N.：基数约束锥规划的限制Robinson约束规范与最优性。优化理论与应用杂志**175**, 104–118（2017）[23] Rockafellar, R.T., Sun, J.：通过渐进对冲求解单调随机变分不等式和互补问题。数学规划 **174(1)**, 453–471（2019年）[24] Rockafellar, R.T., Wets, R.J.：不确定性下优化中的情景与策略聚合。运筹学数学 **16(1)**, 119–147（1991）[25] Rockafellar, R.T., Wets, R.J.：《变分分析》。施普林格，柏林（1998）[26] Rockafellar, R.T., Wets, R.J.：随机变分不等式：从单阶段到多阶段。数学规划 **165(1)**, 331–360（2017）[27] Shapiro, A., Dentcheva, D., Ruszczynski, A.：《随机规划讲义：建模与理论》。SIAM，费城（2014）[28] Tang, J., He,G., Dong, L., Fang, L., Zhou, J.：一种用于二阶锥互补问题的光滑牛顿法。应用数学 **58(2)**, 223–247（2013）[29] Xu, Y., Qi, Q., Lin, Q., Jin, R., Yang, T.：针对差分凸函数和非光滑非凸正则项的随机优化及其非渐近收敛性。载于：国际机器学习会议（ICML）论文集，第6942–6951页（2019）[30]Zhang, S., Qian, H., Gong, X.：一种用于非凸结构化稀疏优化的全局收敛交替近端分裂方法。载于：AAAI人工智能会议论文集，第30卷（2016）[31] Zhang, M., Sun, J., Xu, H.：不确定性下的两阶段二次博弈及其渐进对冲算法求解。SIAM优化学报 **29**, 1799–1818

（2019
年）
[32] Zhang, Y., Wu, J., Zhang, L.：含二阶锥互补约束的数学规划的一阶必要最优性条件。Journal of Global Optimization **63**(2), 253–279（2015年）