arXiv:2603.01304v1 [cs.LG] 1 Mar 2026

通过对数和与极大极小凹惩罚实现非凸潜在最优分区块稀疏恢复

Takanobu Furuhashi、平 oki Kuroda、Masahiro Yukawa、Qibin Zhao、Hidekata Hontani 和 Tatsuya Yokota

摘要——我们提出两种非凸正则化方法 LogLOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1，用于恢复块划分未知的块稀疏信号。这些方法通过新颖的变分形式，将对数和惩罚及极大极小凹惩罚（MCP）推广至块稀疏域，从而解决现有凸方法中的低估偏差。与依赖平方误差数据保真项的广义莫罗增强（GME）和贝叶斯方法不同，我们提出的方法兼容多种数据保真项。我们针对这些变分形式开发了高效的交替方向乘子法（ADMM）算法，并发现其表现出稳定的经验收敛性。在合成数据、角功率谱估计和纳米孔电流去噪方面进行的数值实验表明，我们的方法在估计精度方面优于最先进的基准方法。

索引术语—正则化、块稀疏性、未知划分、非凸惩罚、交替优化

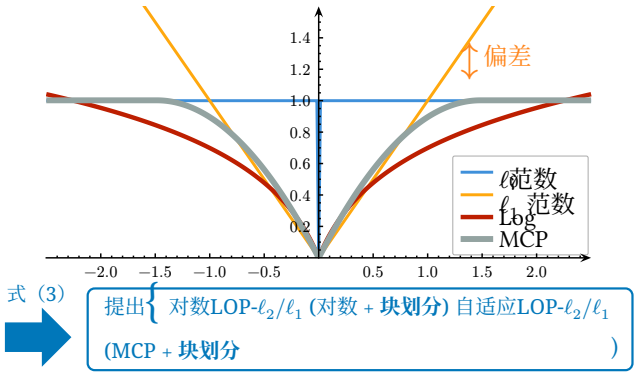


图1：稀疏性诱导惩罚的比较。我们提出的方法通过整合块分区机制来处理未知的块稀疏性，从而扩展了非凸对数和惩罚及MC惩罚的概念。

I. 引言

SPARSE 正则化是信号处理与机器学习中的一项基本技术，它通过利用自然信号固有的稀疏性来实现最优解的恢复 [1–5]。我们考虑根据由线性模型 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}+\boldsymbol{\epsilon}$支配的观测值 $\boldsymbol{y} \in \mathbb{R}^J$ 估计稀疏信号 $\boldsymbol{x} \in \mathbb{R}^N$ 的问题，其中 $\boldsymbol{A} \in \mathbb{R}^{J\times N}$ 是观测矩阵，$\boldsymbol{\epsilon} \in \mathbb{R}^J$ 是观测噪声。

稀疏估计通常被表述为最小化一个由稀疏诱导惩罚正则化的目标函数：

$$\underset{\boldsymbol{x}\in\mathbb{R}^N}{\text{minimize}}\, f(\boldsymbol{A}\boldsymbol{x}) + \lambda\Psi(\boldsymbol{L}\boldsymbol{x}), \tag{1}$$

其中 $f(\boldsymbol{A}\boldsymbol{x})$是数据保真项（例如，针对高斯噪声的平方误差$\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2/2$、针对拉普拉斯噪声的绝对误差，或针对泊松噪声的I-散度）。参数 $\lambda > 0$用于平衡保真度与稀疏性，而 $\boldsymbol{L} \in \mathbb{R}^{K\times N}$ 是一种稀疏化变换，例如傅里叶算子或微分算子 [6–10]。稀疏性最直接的度量是 ℓ_0 伪范数，$\|\boldsymbol{x}\|_0$，用于统计非零元素的数量。然而，ℓ_0 伪范数是不连续且非凸的，其最小化问题是NP难的 [1]。因此，ℓ_1 范数 $\|\boldsymbol{x}\|_1 = \sum_{n=1}^N |x_n|$是 ℓ_0 伪范数的标准凸替代，因为它易于计算。然而，其凸性会导致固有的低估偏差，尤其是对于大振幅分量，这可能会阻碍真实支撑集的恢复 [11, 12]。

非凸惩罚项弥合了 ℓ_0 与ℓ_1 惩罚项之间的差距（图1）。通过部分牺牲凸性，这些惩罚项在保持计算可行的同时，能够实现近乎无偏的恢复。对数和惩罚项 [13] 和极小极大凹惩罚（MCP） [11, 14]是广为人知的例子，可抑制对大系数的过度惩罚。MCP可解释为一种经过最优加权的 ℓ_1 惩罚项，称为等效MCP（EMCP） [15, 16],；其加权方案已被证明有效，能够为较大的系数分配较小的权重 [17–19]。

尽管逐元素稀疏正则项效果显著，但自然信号通常具有块稀疏性，其非零系数会聚集成组 [20–26]。然而在实践中，此类分区通常无法预先获得，因此标准块稀疏正则项难以奏效。

为解决这一问题，基于潜在最优分区（LOP）的方法会同时估计信号及其潜在块分区 [27–34]。值得注意的是，凸 LOP-ℓ_2/ℓ_1 惩罚项 [35–37]支持高效的近端优化，但仍存在低估偏差。其非凸扩展——广义莫罗增强（GME）-LOP-ℓ_2/ℓ_1惩罚项 [38],——能够缓解这一偏差，但从理论上讲仅适用于最小二乘问题（即满足高斯噪声假设）。

类似地，稀疏贝叶斯学习（SBL） [28, 33] 能够有效恢复未知块分区，但为了便于解析求解（例如利用共轭性），主要依赖高斯噪声。此外，SBL 通常需要显式进行矩阵求逆，这在高维场景中代价过高。

本文提出了两种新颖的基于非凸 LOP 的方

本研究部分获得日本学术振兴会（JSPS）科学研究费助成事业（KAKENHI）项目（资助编号：23K28109）的支持。

Takanobu Furuhashi、Hidekata Hontani 和 Tatsuya Yokota 任职于日本名古屋的名古屋工业大学（电子邮箱：clz14116@nitech.jp，{hontani.hidekata / t.yokota}@nitech.ac.jp）。Qibin Zhao 和 Tatsuya Yokota 也任职于日本东京的理化学研究所先进智能项目中心（电子邮箱：qibin.zhao@riken.jp）。Hiroki Kuroda 任职于日本新潟的长冈技术科学大学（电子邮箱：kuroda@vos.nagaokaut.ac.jp）。Masahiro Yukawa 任职于日本横滨的庆应义塾大学（电子邮箱：yukawa@elec.keio.ac.jp）。

本作品已提交至电气与电子工程师协会，拟供其发表。版权可能在未另行通知的情况下转让，届时本版本可能不再可访问。

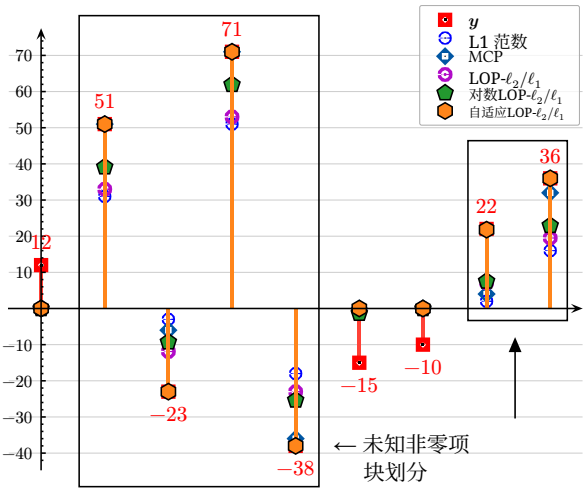


图 2：近端算子比较。与 LOP-ℓ_2/ℓ_1 不同，所提出的对数LOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1 在保留块划分的同时，减轻了大幅值（例如，$y_2 = 51$，$y_4 = 71$）的低估偏差，其性能优于逐元素惩罚项。

法，旨在提高准确性，同时保持其在块稀疏信号恢复中的适用性。如图 2 所示，与凸方法相比，这些方法能更好地保留较大的信号分量。

对数LOP-ℓ_2/ℓ_1 (对数LOP-ℓ_2/ℓ_1)：该方法将对数和惩罚项融入 LOP 框架，以减轻自适应识别的块内的低估偏差。这是首个采用对数惩罚项实现块稀疏信号近乎无偏恢复的 LOP 方法。

自适应加权 LOP-ℓ_2/ℓ_1（自适应LOP-ℓ_2/ℓ_1)：该方法采用一种受 EMCP 启发的自适应加权方案。它根据已识别块中的信号幅值更新权重，从而更好地区分显著系数。

本研究的主要贡献如下：

- 我们建立了一个通用的 LOP 型惩罚框架，统一了凸正则化项和本文提出的非凸 LOP 型正则化项。我们证明，渐近水平稳定（als）性质可保证极小值点的存在。
- 我们提出了 LogLOP-ℓ_2/ℓ_1 和 AdaLOP-ℓ_2/ℓ_1 惩罚项，这是首批将对数惩罚和自适应加权引入 LOP 框架的方法。这些方法显著降低了低估偏差。
- 我们针对这些形式提出了高效的、基于交替方向乘子法（ADMM）的算法。尽管由于非凸性而缺乏理论保证（详见注2），但实验表明，这些算法能够稳定收敛到所需解。
- 与 GME-LOP-ℓ_2/ℓ_1 和基于SBL的方法不同，所提出的方法兼容任何近端算子可计算的数据保真项 [39, 40]，因此可应用于多种噪声模型和应用场景。

本文其余部分安排如下。第二节回顾现有方法。第三节详细介绍所提出的对数LOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1 惩罚。第四节介绍优化算法。第五节给出数值结果，第六节总结全文。

II. 现有方法综述

A. 数学符号

向量和矩阵分别用黑体小写字母（$\boldsymbol{a}$）和大写字母（$\boldsymbol{A}$）表示。$\boldsymbol{x} \in \mathbb{R}^N$ 表示一个 N维向量，其第 n个分量为 x_n；$\boldsymbol{X} \in \mathbb{R}^{I\times J}$表示一个 $I \times J$ 矩阵，其第（i，j）个元素为$X_{i,\ j}$。$(\cdot)^\top$ 和 $(\cdot)^{\mathsf{H}}$。$\mathbb{N}$、$\mathbb{R}$、$\mathbb{R}_{>0}$和 $\mathbb{R}_{\geq 0}$ 分别表示正整数、实数、正实数和非负实数集合。$\boldsymbol{O}$ 和 $\boldsymbol{I}$ 分别表示零矩阵和单位矩阵。集合 $[N]$ 表示 $\{1,\ldots,N\}$，而 $|S|$表示集合 S的基数。对于向量 $\boldsymbol{x}$ 和指标集 S，$\boldsymbol{x}_S$表示由 S索引的子向量。$|x|$表示绝对值，$\mathrm{sign}(x)$表示符号函数。$|\boldsymbol{x}|$ 表示逐元素绝对值向量。

ℓ_1 ℓ_2 $\|\boldsymbol{x}\|_1 =$ 范数 $\sum_{n=1}^N |x_n|$ $\|\boldsymbol{x}\|_2 = \sum_{n=1}^N x_n^2\ ^{1/2}$ $\boldsymbol{x}\odot\boldsymbol{y}$ 和范数定义为和（）。表示逐元素乘积。对于向量 $\boldsymbol{w}$ 和标量函数 h $\boldsymbol{w}\odot$，h $\boldsymbol{x}$ 我们还用 $\boldsymbol{w}\odot h$ 表示由（） （）$=\sum_{n=1}^N w_n h(x_n)$定义的函数。若函数：$\mathbb{R}^N \to \mathbb{R}\cup\{\infty\}$ 的定义域 $\mathrm{dom}(h)$：$=\{\boldsymbol{x}\in\mathbb{R}^N \mid h(\boldsymbol{x}) < \infty\}$ 非空，则称其为真函数。若对每个 $a\in\mathbb{R}$，其下水平集 $\{\boldsymbol{x}\in\mathbb{R}^N \mid h(\boldsymbol{x})\leq a\}$ 均为闭集，则称其为下半连续（lsc）。若 $\lim_{\|\boldsymbol{x}\|_2\to\infty} h(\boldsymbol{x}) = \infty$，则称其为强制。具有 $\beta>0$ 的真下半连续凸函数 h 的近端算子为

$$\mathrm{prox}_{\beta h}(\boldsymbol{v}) \coloneqq \arg\min_{\boldsymbol{x}\in\mathbb{R}^N}\left\{h(\boldsymbol{x}) + \frac{1}{2\beta}\|\boldsymbol{x}-\boldsymbol{v}\|_2^2\right\}.$$

B. 非凸稀疏正则化

尽管 ℓ_1 范数是稀疏估计的标准工具，但其凸性和强制性会对大系数产生低估偏差。非凸正则化能够更准确地衡量稀疏性 [19, 41, 42]。

对数和惩罚 [13]可以缓解这种偏差：

$$\Omega_\epsilon(\boldsymbol{x}) = \sum_{n=1}^N \log\left(\frac{|x_n|}{\epsilon}+1\right),$$

其中，$\epsilon>0$ 是曲率参数。当 $\epsilon\to 0$ 时，它趋近于 ℓ_0伪范数；当 $\epsilon\to\infty$时，则趋近于 ℓ_1 范数。与 ℓ_1 范数相比，其对数增长对大幅值的惩罚较弱。然而，由于该惩罚项具有强制性且没有上界，偏差虽有所减小，但并未消除。

极大极小凹惩罚（MCP） [11] 通过采用有界惩罚项来解决这一问题：

$$\Omega_{\gamma,w}(x) = \begin{cases} w|x| - \frac{x^2}{2\gamma}, & \text{if } |x|\leq\gamma w; \\ \frac{\gamma w^2}{2}, & \text{if } |x|>\gamma w, \end{cases}$$

其中，$w>0$ 是稀疏性参数，$\gamma>1$ 是曲率参数。它趋近于 ℓ_0（$\gamma\to 1$）范数和 ℓ_1（$\gamma\to\infty$）范数。当 $|x|>\gamma w$时变为常数，MCP 对大系数实现了无偏性。MCP 可以扩展到向量，表示为 $\Omega_{\gamma,\ \boldsymbol{w}}(\boldsymbol{x})\ =\sum_n \Omega_{\gamma,\ w_n}(x_n)$，其中 $\boldsymbol{w}\in\mathbb{R}^N_{\geq 0}$。

关键在于，MCP 存在一种称为向量等效MCP（EMCP）的变分表示 [15]:

$$\Omega_{\gamma,\boldsymbol{w}}(\boldsymbol{x}) = \min_{\boldsymbol{\omega}\in\mathbb{R}^N_{\geq 0}}\left\{\|\boldsymbol{x}\|_{\boldsymbol{\omega},1} + \frac{\gamma}{2}\|\boldsymbol{\omega}-\boldsymbol{w}\|_2^2\right\},$$

其中 $\|\boldsymbol{x}\|_{\boldsymbol{\omega},\,1} = \sum_n \omega_n |x_n|$。这将 MCP 解释为一种经过最优加权的 ℓ_1 惩罚项。 [15] 中的方法提出自适应地更新 $\boldsymbol{w}$，以提高估计精度。这种自适应能力是我们所提方法的核心。

C. 潜在最优划分（*LOP*）-ℓ_2/ℓ_1

标准的块稀疏正则化依赖于给定的块划分。为应对划分未知的情形，凸 LOP-ℓ_2/ℓ_1 惩罚项 [35] 会同时估计信号及其最优块划分。

该方法旨在确定一个块划分 $(\mathcal{B}_k)_{k=1}^{j}$，使凸混合 ℓ_2/ℓ_1 范数最小：

$$\psi_K(\boldsymbol{x}) \coloneqq \min_{j\in[K]} \left[\min_{(\mathcal{B}_k)_{k=1}^{j}\in\mathcal{P}_j} \sum_{k=1}^{j} \sqrt{|\mathcal{B}_k|}\|\boldsymbol{x}_{\mathcal{B}_k}\|_2 \right],$$

其中，$\mathcal{P}_j$ 由所有可能的、互不重叠的块划分组成，且包含 j 个块 [35, 见式 3]。由于这是一个难以求解的组合优化问题，LOP-ℓ_2/ℓ_1 采用基于 ℓ_2 范数变分形式的紧凸松弛。变分函数 $\phi\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}_{\geq 0}\cup\{\infty\}$ 定义如下：

$$\phi(x,\tau) \coloneqq \begin{cases} \dfrac{|x|^2}{2\tau} + \dfrac{\tau}{2}, & \text{if } \tau > 0; \\ 0, & \text{if } x = 0 \text{ and } \tau = 0; \\ \infty, & \text{otherwise.} \end{cases} \tag{2}$$

据此，我们将 LOP-ℓ_2/ℓ_1 惩罚项定义如下：

$$\Psi_\alpha(\boldsymbol{x}) \coloneqq \min_{\substack{\boldsymbol{\sigma}\in\mathbb{R}^N \\ \|\boldsymbol{D}\boldsymbol{\sigma}\|_1\leq\alpha}} \sum_{n=1}^{N} \phi(x_n,\sigma_n).$$

潜在向量 $\boldsymbol{\sigma}$ 表示块划分，并受全变差（TV） $\|\boldsymbol{D}\boldsymbol{\sigma}\|_1 \leq \alpha$ 的约束。参数 $\alpha \geq 0$控制划分粒度，而 $\boldsymbol{D}$ 表示一阶差分矩阵。尽管 LOP-ℓ_2/ℓ_1 能够有效发现块划分，但它仍具有凸性，因此对于大幅值会产生相同的低估偏差（图 2）。

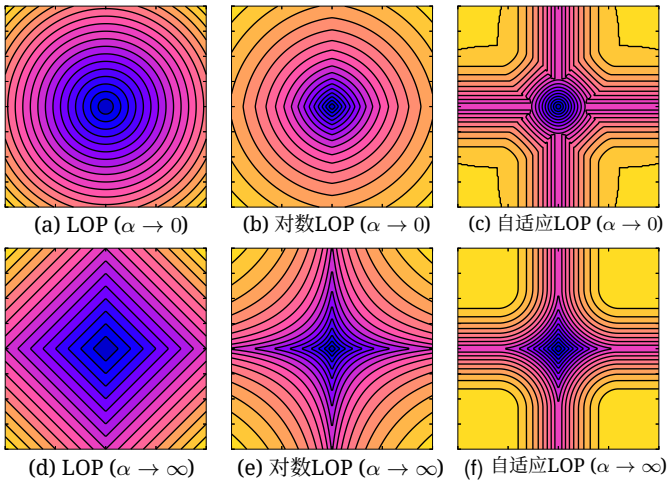
D. 广义莫罗增强（*GME*）-*LOP*-ℓ_2/ℓ_1

为减轻这种偏差，GME-LOP-ℓ_2/ℓ_1 [38] 将广义莫罗增强（GME）框架 [43–45]应用于 LOP-ℓ_2/ℓ_1。它通过从 $\Psi_\alpha(\boldsymbol{x})$ 中减去广义莫罗包络，构造非凸惩罚项 $\Psi_{\mathrm{GL}}(\boldsymbol{x})$：

$$\Psi_{\mathrm{GL}}(\boldsymbol{x}) = \Psi_\alpha(\boldsymbol{x}) - \min_{\boldsymbol{v}\in\mathbb{R}^N} \left\{ \Psi_\alpha(\boldsymbol{v}) + \frac{1}{2}\|\boldsymbol{B}(\boldsymbol{x}-\boldsymbol{v})\|_2^2 \right\}.$$

GME-LOP-ℓ_2/ℓ_1 的一个关键特点是，在满足 GME 参数矩阵 $\boldsymbol{B}$ 的特定条件（例如 $\boldsymbol{A}^\top\boldsymbol{A} - \lambda\boldsymbol{L}^\top\boldsymbol{B}^\top\boldsymbol{B}\boldsymbol{L} \succeq \boldsymbol{O}$）时，整体正则化最小二乘目标函数仍保持凸性。因此，可以保证其收敛至全局最小值。

然而，这种凸性保证严格依赖于平方误差损失，因此限制了其在其他数据保真项中的适用性，也促使我们提出更灵活的基于 LOP 的非凸方法。需要注意的是，尽管我们提出的方法不具备与 GME-LOP-ℓ_2/ℓ_1 （见注2）相同的理论收敛保证，但实验结果表明，即使采用非平方误差损失，它们也表现出稳定的收敛行为。



(a) LOP ($\alpha\to 0$) (b) 对数LOP ($\alpha\to 0$) (c) 自适应LOP ($\alpha\to 0$)
(d) LOP ($\alpha\to\infty$) (e) 对数LOP ($\alpha\to\infty$) (f) 自适应LOP ($\alpha\to\infty$)

图 3： ℓ_2/ℓ_1 正则化中 LOP、对数LOP（定理1）和自适应LOP（定理2）变体的渐近行为。

III. 所提出正则化的设计

现有的基于 LOP 的方法面临两难：凸的 LOP-ℓ_2/ℓ_1 惩罚项存在低估偏差，而非凸的 GME-LOP-ℓ_2/ℓ_1 惩罚项则仅适用于平方误差损失。为克服这些局限，我们提出了两种新颖的非凸框架，将 LOP-ℓ_2/ℓ_1 的块发现能力与更强的偏差缓解相结合，同时不限制数据保真项。

对数 LOP-ℓ_2/ℓ_1 (对数LOP-ℓ_2/ℓ_1)： 通过将块幅值上的线性惩罚替换为对数函数，减轻低估问题。

自适应加权 LOP-ℓ_2/ℓ_1 (自适应LOP-ℓ_2/ℓ_1)： 联合优化块划分和正则化权重，从而自动降低显著分量的权重。

如图 3 所示，对数LOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1（定理 1 和定理 2 中讨论）的渐近形式均呈现出明显的非凸特征。对于较大的信号幅度，惩罚项的增长会减弱，因此与凸 LOP-ℓ_2/ℓ_1 惩罚项相比，估计偏差得以减轻。

A. 最优划分的一般变分框架

为系统地解决偏差问题，我们首先将 LOP-ℓ_2/ℓ_1 机制抽象为一个一般变分框架。LOP-ℓ_2/ℓ_1 的核心思想是通过潜变量 σ_n 来评估信号分量x_n，该潜变量能够捕获最优块划分。不同于 GME-LOP-ℓ_2/ℓ_1依赖 GME 进行非凸化，我们通过直接采用非凸变分函数φ 来定义一般的 LOP 型惩罚项$\Psi_{\alpha,\,\varphi}$ 具体如下： $\mathbb{R}\times\mathbb{R}\to\mathbb{R}_{\geq 0}\cup\{\infty\}$

定义 1 （一般 LOP 型惩罚）。 设 $\varphi\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}_{\geq 0}\cup\{\infty\}$ 是一个适当下半连续函数。假设 φ满足渐近水平稳定（*als*）性质 *[46, 定义 3.3.2]*。对于给定的分区参数 $\alpha \geq 0$，一般 *LOP* 型惩罚项 $\Psi_{\alpha,\,\varphi}\colon \mathbb{R}^N\to\mathbb{R}_{\geq 0}\cup\{\infty\}$ 定义如下：

$$\Psi_{\alpha,\varphi}(\boldsymbol{x}) \coloneqq \min_{\substack{\boldsymbol{\sigma}\in\mathbb{R}^N \\ \|\boldsymbol{D}\boldsymbol{\sigma}\|_1\leq\alpha}} \sum_{n=1}^{N} \varphi(x_n,\sigma_n). \tag{3}$$

als 性质刻画了水平集在无穷远处的稳定性，并推广了强制性的假设，从而确保 $\Psi_{\alpha,\ \varphi}(\boldsymbol{x})$ 存在极小值点（详情见附录 A）。此外，当最小化由该惩罚项与一个适当、下半连续、强制且凸的函数 f 组成的目标函数（例如 (8)）时，该性质还确保存在一个极小值点 $(\hat{\boldsymbol{x}},\ \hat{\boldsymbol{\sigma}})$。

变分函数 $\varphi(x,\ \sigma)$ 被设计为基础惩罚项 $\Omega(x)$ 的变分上界（例如，ℓ_1范数 [35, 定理 2], 对数和和 MC 惩罚项），满足 $\min_{\sigma\in\mathbb{R}}\varphi(x,\ \sigma)\ =\Omega(x)$。在 TV 约束下最小化该上界，可以使框架在应用 $\Omega(x)$ 的性质的同时发现最优块划分。LOP-ℓ_2/ℓ_1 使用式（2）中的函数，产生绝对惩罚项 $|x|$。为减少估计偏差，对数LOP-ℓ_2/ℓ_1和自适应 LOP-ℓ_2/ℓ_1 采用满足 als 条件的非凸变分函数，从而确保惩罚项定义良好并保证极小值点的存在。

B. 对数*LOP-*ℓ_2/ℓ_1 （对数*LOP-*ℓ_2/ℓ_1)

1) 变分形式：我们通过选择一个诱导对数增长的变分函数 ϕ_ϵ 来实现对数LOP-ℓ_2/ℓ_1 。我们将 ϕ_ϵ 定义为：$\mathbb{R}\times\mathbb{R}\to\mathbb{R}_{\geq 0}\cup\{\infty\}$，其中 $\epsilon>0$:

$$\phi_\epsilon(x,\tau)\coloneqq\begin{cases}\dfrac{1}{2\tau}\left(\dfrac{|x|}{\epsilon}+1\right)^2+\dfrac{1}{2}\log\tau-\dfrac{1}{2}, & \text{if } \tau\geq 1;\\ \infty, & \text{otherwise.}\end{cases}\tag{4}$$

我们设计该函数，使其在 τ 上的最小值恢复对数惩罚，如引理 1 所证明。

引理 1 （ $\phi_\epsilon(x,\tau)$的最小值）。 对于任意固定的 $x\in\mathbb{R}$和 $\epsilon>0$, $\phi_\epsilon(x,\tau)$ 在 $\tau\in\mathbb{R}$上具有唯一的全局极小化器 $\tau^*\ =(|x|/\epsilon+1)^2$，且最小值为：

$$\min_{\tau\in\mathbb{R}}\phi_\epsilon(x,\tau)=\min_{\tau\geq 1}\phi_\epsilon(x,\tau)=\log\left(\frac{|x|}{\epsilon}+1\right).$$

证明。 令 $z:\ =(|x|/\epsilon+1)^2$。由于 $x\in\mathbb{R}$ 和 $\epsilon>0$，注意到 $z\geq 1$。对于 $\tau\geq 1$，将 $\phi_\epsilon(x,\tau)$视为 τ的函数，其导数为

$$\frac{\partial\phi_\epsilon}{\partial\tau}(x,\tau)=-\frac{z}{2\tau^2}+\frac{1}{2\tau}=\frac{\tau-z}{2\tau^2}.$$

因此， $\phi_\epsilon(x,\tau)$在 $1\leq\tau<z$上严格递减，在 $\tau>z$上严格递增，从而确认 $\tau^*=z$ 是唯一极小值点。代入 $\tau^*=z$ 即得最小值：

$$\phi_\epsilon(x,z)=\frac{z}{2z}+\frac{1}{2}\log z-\frac{1}{2}=\log\left(\frac{|x|}{\epsilon}+1\right).$$

至此，证明完成。 □

2) 对数*LOP-*ℓ_2/ℓ_1： 将此 ϕ_ϵ代入一般框架 (3)，得到对数LOP-ℓ_2/ℓ_1 惩罚项：

$$\Psi_{\alpha,\epsilon}(\boldsymbol{x})\coloneqq\min_{\substack{\boldsymbol{\sigma}\in\mathbb{R}^N\\ \|\boldsymbol{D}\boldsymbol{\sigma}\|_1\leq\alpha}}\sum_{n=1}^{N}\phi_\epsilon(x_n,\sigma_n).$$

该惩罚项能够发现诸如 LOP-ℓ_2/ℓ_1 的块划分，同时利用对数尺度评估这些块的大小，从而直接解决低估问题。

3) 渐近行为： 分区参数 $\alpha\geq 0$在逐元素对数和稀疏性与对数块稀疏惩罚项之间实现无缝插值：

定理1 （对数LOP-ℓ_2/ℓ_1的渐近行为）。 *(i)* 当$\alpha\to\infty$ （最细分区）时，$\Psi_{\alpha,\ \epsilon}$ 可恢复对数和惩罚：

$$\lim_{\alpha\to\infty}\Psi_{\alpha,\epsilon}(\boldsymbol{x})=\Omega_\epsilon(\boldsymbol{x})=\sum_{n=1}^{N}\log\left(\frac{|x_n|}{\epsilon}+1\right).$$

(ii) 当 $\alpha\to 0$ （最粗分区）时，如图 *3b* 所示，它收敛到对数 ℓ_2 惩罚项：

$$\lim_{\alpha\to 0}\Psi_{\alpha,\epsilon}(\boldsymbol{x})=\Psi_{0,\epsilon}(\boldsymbol{x})=N\log\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\frac{|x_n|}{\epsilon}+1\right)^2}.$$

证明。 参见附录 B，与 [35, 附录 C]类似。 □

4) 块分区的有效性： 任何基于 LOP 的惩罚项都必须能够区分非零块和零块，这是一项关键要求，详见 [35, 备注 1]。我们将特定块 $\mathcal{B}$ 上的惩罚项定义为：

$$\Theta_\epsilon(\boldsymbol{x}_\mathcal{B})\coloneqq\min_{\tau\geq 1}\sum_{n\in\mathcal{B}}\phi_\epsilon(x_n,\tau)=|\mathcal{B}|\log\sqrt{\frac{1}{|\mathcal{B}|}\sum_{n\in\mathcal{B}}\left(\frac{|x_n|}{\epsilon}+1\right)^2}.$$

对数LOP-ℓ_2/ℓ_1 惩罚项满足以下性质：

命题1 (块分解性质)。 对于任意可分解为非零子块 $\mathcal{B}'$ *(*$\boldsymbol{x}_{\mathcal{B}'}\neq\boldsymbol{0}$*)* 和零值子块 $\mathcal{B}''$ *(*$\boldsymbol{x}_{\mathcal{B}''}=\boldsymbol{0}$*)* 的块$\mathcal{B}$ ，分离后惩罚项严格减小：

$$\Theta_\epsilon(\boldsymbol{x}_\mathcal{B})>\Theta_\epsilon(\boldsymbol{x}_{\mathcal{B}'})+\Theta_\epsilon(\boldsymbol{x}_{\mathcal{B}''})=\Theta_\epsilon(\boldsymbol{x}_{\mathcal{B}'}),$$

其中 Θ_ϵ 表示一个块上的惩罚值。

证明。 我们分析分离零块前后的惩罚差 Δ 。令 $z\coloneqq\sum_{n\in\mathcal{B}'}(|x_n|/\epsilon+1)^2$。由于 $(|x_n|/\epsilon+1)^2\geq 1$ 对所有 n 均成立，且在 $x_n\neq 0$时严格大于 1，因此 $\boldsymbol{x}_{\mathcal{B}'}\neq\boldsymbol{0}$ 保证了 $z>|\mathcal{B}'|$。我们证明，对于 $z>|\mathcal{B}'|$， $\Delta(z)$ 严格递增，且 $\Delta(|\mathcal{B}'|)=0$。这意味着对于任意 $z>|\mathcal{B}'|$, $\Delta(z)>0$ ，从而证明分离总是会降低惩罚项。证明见附录C。 □

这一性质确保对数LOP-ℓ_2/ℓ_1 惩罚项能够主动推动各分区将非零信号分量与零区域分离，从而确保精确的支撑集恢复。

C. Adaptively Weighted LOP-ℓ_2/ℓ_1 *(AdaLOP-*ℓ_2/ℓ_1*)*

1) 变分形式： 我们通过选择一个非凸变分函数$\phi_{\gamma,\ w}$ 来实现自适应LOP-ℓ_2/ℓ_1惩罚项。受 EMCP [15,16],自适应加权机制的启发，我们将该函数明确定义如下：

$$\phi_{\gamma,w}(x,\tau)\coloneqq\begin{cases}w\phi(x,\tau)-\frac{\phi(x,\tau)^2}{2\gamma}, & \text{if } \phi(x,\tau)\leq\gamma w;\\ \frac{\gamma w^2}{2}, & \text{if } \phi(x,\tau)>\gamma w,\end{cases}\tag{5}$$

其中， $\phi(x,\ \tau)$ 是变分函数 (2)。该函数提供有界惩罚项，并且可以通过涉及自适应权重 ω 的变分形式等价地表示：

$$\phi_{\gamma,w}(x,\tau)=\min_{\omega\in\mathbb{R}_{\geq 0}}\left\{\omega\phi(x,\tau)+\frac{\gamma}{2}(\omega-w)^2\right\}.\tag{6}$$

最小值在 $\omega = \max(0,\ w - \gamma^{-1}\phi(x,\ \tau))$ 处取得。这里，我们采用规则 $0 \cdot \infty = 0$。这确保即使在 $\phi(x,\ \tau) = \infty$ 时，最小化问题仍有良好定义（在 $\omega = 0$ 处取得），并与显式定义 (5) 一致，从而保留第四节所述交替方向乘子法优化所需的下半连续性。

引理2 （$\phi_{\gamma,\ w}(x,\tau)$ 的最小值。 对于任意固定的 $x \in \mathbb{R}$、 $\gamma > 0$ 和 $w > 0$，关于 $\tau \in \mathbb{R}$ 的$\phi_{\gamma,\ w}(x,\tau)$ 的最小值可恢复*MCP*惩罚项 $\Omega_{\gamma,\ w}(x)$ ：

$$\min_{\tau\in\mathbb{R}} \phi_{\gamma,w}(x,\tau) = \min_{\tau\in\mathbb{R}_{\geq 0}} \phi_{\gamma,w}(x,\tau) = \Omega_{\gamma,w}(x).$$

最小值在 $\tau = |x|$ 处取得，即 $\phi(x,\tau) = |x|$。

证明。 参见附录D。 □

*2）AdaLOP-*的定义： 将$\phi_{\gamma,\ w}$ 代入一般框架（3），得到AdaLOP-ℓ_2/ℓ_1 惩罚项:

$$\begin{aligned}\Psi_{\alpha,\gamma,\boldsymbol{w}}(\boldsymbol{x}) &:= \min_{\substack{\boldsymbol{\sigma}\in\mathbb{R}^N \\ \|\boldsymbol{D}\boldsymbol{\sigma}\|_1\leq\alpha}} \sum_{n=1}^{N} \phi_{\gamma,w_n}(x_n,\sigma_n) \\ &= \min_{\substack{\boldsymbol{\sigma}\in\mathbb{R}^N \\ \|\boldsymbol{D}\boldsymbol{\sigma}\|_1\leq\alpha}} \min_{\boldsymbol{\omega}\in\mathbb{R}^N_{\geq 0}} \left\{ \sum_{n=1}^{N} \omega_n\phi(x_n,\sigma_n) + \frac{\gamma}{2}\|\boldsymbol{\omega}-\boldsymbol{w}\|_2^2 \right\}.\end{aligned}$$

该公式包含三重优化：信号估计（$\boldsymbol{x}$）、分区发现（$\boldsymbol{\sigma}$）和加权（$\boldsymbol{\omega}$）。请注意，尽管 $\boldsymbol{w}$ 在此定义中作为固定参数向量出现，但在所提出的算法中会对其进行迭代更新，以自适应地控制正则化强度（参见第四节）。

3) 渐近行为：参数 $\alpha \geq 0$ 控制块划分的粒度，例如 LOP-ℓ_2/ℓ_1。

定理2 （自适应LOP-ℓ_2/ℓ_1的渐近行为）。 *(i)* 当$\alpha \to \infty$， $\Psi_{\alpha,\ \gamma,\ \boldsymbol{w}}(\boldsymbol{x})$ 时，收敛到向量 *EMCP*：

$$\Omega_{\gamma,\boldsymbol{w}}(\boldsymbol{x}) = \min_{\boldsymbol{\omega}\in\mathbb{R}^N_{\geq 0}} \left\{ \|\boldsymbol{x}\|_{\boldsymbol{\omega},1} + \frac{\gamma}{2}\|\boldsymbol{\omega}-\boldsymbol{w}\|_2^2 \right\}.$$

(ii) 当 $\alpha \to 0$时，它收敛到如图*3c*所示的全局自适应加权 ℓ_2惩罚项：

$$\Psi_{0,\gamma,\boldsymbol{w}}(\boldsymbol{x}) = \min_{\boldsymbol{\omega}\in\mathbb{R}^N_{\geq 0}} \left\{ \|\boldsymbol{\omega}\|_1^{\frac{1}{2}} \|\boldsymbol{x}\|_{\boldsymbol{\omega},2} + \frac{\gamma}{2}\|\boldsymbol{\omega}-\boldsymbol{w}\|_2^2 \right\}. \quad (7)$$

其中$\|\boldsymbol{x}\|_{\boldsymbol{\omega},\ 2} = \sqrt{\sum_{n=1}^{N} \omega_n |x_n|^2}$是加权 ℓ_2 惩罚项。

Proof. See Appendix E analogous to [35, App. C]. □

该定理证实，自适应LOP-ℓ_2/ℓ_1 惩罚项是从 MCP 到 LOP 型惩罚的严格推广。

注1 (γ的参数选择)。 参数 γ 不仅控制惩罚曲率，还充当权重更新的逆步长（见 *(17)*）。我们的实验表明，当稀疏性参数 w 较大时，采用较小的 γ （较大的步长）效果更好。这会加快已识别非零块上的权重衰减，从而提高准确恢复的效果。

IV. 优化算法

A. 所提方法的实现

我们使用交替方向乘子法 [47]求解 LogLOP-ℓ_2/ℓ_1和自适应LOP-ℓ_2/ℓ_1 惩罚项的优化问题。该框架有效地将非凸正则项与数据保真项解耦，从而能够进行高效的迭代更新。

对于对数LOP-ℓ_2/ℓ_1，该问题表述为:

$$\underset{\boldsymbol{x}\in\mathbb{R}^N}{\text{minimize}}\, f(\boldsymbol{A}\boldsymbol{x}) + \lambda\Psi_{\alpha,\epsilon}(\boldsymbol{L}\boldsymbol{x}). \quad (8)$$

类似地，对于自适应LOP-ℓ_2/ℓ_1:

$$\underset{\boldsymbol{x}\in\mathbb{R}^N}{\text{minimize}}\, f(\boldsymbol{A}\boldsymbol{x}) + \Psi_{\alpha,\gamma,\boldsymbol{w}}(\boldsymbol{L}\boldsymbol{x}). \quad (9)$$

请注意，在自适应LOP-ℓ_2/ℓ_1中，遵循 [15],中的方法，权重向量 $\boldsymbol{w}$ 也会被迭代优化，以主动减轻估计偏差。在下文中，为简便起见，我们假设函数 f 是正常的、下半连续的、强制的且凸的。算法 1 和算法 2 给出了详细算法。我们在下文描述每个变量的具体更新规则。附录 F 和 G 提供了完整推导。此处，$\mu_1, \ldots, \mu_4$ 表示增广拉格朗日约束的惩罚参数， $\rho \geq 1$ 表示其更新率。对于对数LOP-ℓ_2/ℓ_1，我们建议将 $\mu_4 = \lambda/54 + \delta$ 设为一个较小的常数 $\delta > 0$ ，以保证ξ更新的严格凸性（见附录 F）。

1) $\boldsymbol{x}$-更新（线性系统）： 两种方法采用相同的 $\boldsymbol{x}$-更新，通过最小化二次惩罚项来求解。对 $\boldsymbol{x}$ 求最小值可得以下线性系统:

$$\left(\mu_1\boldsymbol{A}^\top\boldsymbol{A} + \mu_2\boldsymbol{L}^\top\boldsymbol{L}\right)\boldsymbol{x} = \mu_1\boldsymbol{A}^\top\left(\boldsymbol{u} + \frac{\boldsymbol{r}_1}{\mu_1}\right) + \mu_2\boldsymbol{L}^\top\left(\boldsymbol{v} + \frac{\boldsymbol{r}_2}{\mu_2}\right). \quad (10)$$

该系统为半正定系统，可使用共轭梯度法（CG）或基于快速傅里叶变换的求解器[1]高效求解。

2) $\boldsymbol{u}$-更新（ f的近端算子）： 此步骤应用数据保真项的近端算子:

$$\boldsymbol{u} \leftarrow \operatorname{prox}_{\mu_1^{-1}f}\left(\boldsymbol{A}\boldsymbol{x} - \mu_1^{-1}\boldsymbol{r}_1\right). \quad (11)$$

对于平方误差 $f(\boldsymbol{u}) = \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{u}\|_2^2$，$\operatorname{prox}_{\beta f}(\boldsymbol{u}) = (\beta\boldsymbol{y} + \boldsymbol{u})/(\beta+1)$。对于绝对误差 $f(\boldsymbol{u}) = \|\boldsymbol{y}-\boldsymbol{u}\|_1$，$\operatorname{prox}_{\beta f}(\boldsymbol{u}) = (u_j + \beta(y_j - u_j)/\max\{|y_j - u_j|, \beta\})_{j=1}^{J}$。

3) $\boldsymbol{\sigma}$-更新（线性系统）： 此更新将辅助变量 $\boldsymbol{\eta}$ 和 $\boldsymbol{\xi}$ 联系起来，并最小化二次惩罚项:

$$\left(\mu_3\boldsymbol{D}^\top\boldsymbol{D} + \mu_4\boldsymbol{I}\right)\boldsymbol{\sigma} = \mu_3\boldsymbol{D}^\top\left(\boldsymbol{\eta} + \mu_3^{-1}\boldsymbol{r}_3\right) + \mu_4\left(\boldsymbol{\xi} + \mu_4^{-1}\boldsymbol{r}_4\right). \quad (12)$$

4) $\boldsymbol{\eta}$-更新（投影到 ℓ_1-球）： 变量 $\boldsymbol{\eta}$负责处理关于 $\boldsymbol{\sigma}$的全变差约束。更新操作是投影到 ℓ_1-球 $B_1^\alpha := \{\boldsymbol{z} \mid \|\boldsymbol{z}\|_1 \leq \alpha\}$:

$$\boldsymbol{\eta} \leftarrow P_{B_1^\alpha}\left(\boldsymbol{D}\boldsymbol{\sigma} - \mu_3^{-1}\boldsymbol{r}_3\right). \quad (13)$$

投影 $P_{B_1^\alpha}(\boldsymbol{z})$在 $\|\boldsymbol{z}\|_1 \leq \alpha$时计算为 $\boldsymbol{z}$ ，否则计算为$(\max\{|z_k| - \theta, 0\} \cdot \operatorname{sign}(z_k))_{k=1}^{K}$，其中 θ 是一个阈值，可在期望线性时间内找到 [48]。

[1]如果矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{L}$ 可通过快速傅里叶变换对角化，例如因周期边界条件而形成的循环矩阵，则可以使用基于FFT的求解器。

5) $\boldsymbol{v}$ 和 $\boldsymbol{\xi}$ 更新： $\boldsymbol{v}$ 和 $\boldsymbol{\xi}$ 的更新在这两种方法之间有所不同。

对数*LOP-ℓ_2/ℓ_1*： $\boldsymbol{v}$ 更新是弹性网络 [40]的近端算子：

$$\boldsymbol{v} \leftarrow \operatorname{prox}_{\lambda_1\|\cdot\|_1+\lambda_2\|\cdot\|_2^2}\left(\boldsymbol{L}\boldsymbol{x}-\mu_2^{-1}\boldsymbol{r}_2\right), \tag{14}$$

其中 $\lambda_1=\lambda/(\mu_2\xi_n\epsilon)$、 $\lambda_2=\lambda/(2\mu_2\xi_n\epsilon^2)$，并且

$$\operatorname{prox}_{\lambda_1\|\cdot\|_1+\lambda_2\|\cdot\|_2^2}(\boldsymbol{v})=\frac{1}{1+2\lambda_2}\operatorname{sign}(\boldsymbol{v})\odot\max\{0,|\boldsymbol{v}|-\lambda_1\}.$$

$\boldsymbol{\xi}$ 最小化问题具有严格凸性，并在充分条件 $\mu_4>\lambda/54$ 下存在唯一解（见附录 F）。对于每个分量，我们通过二分法求解该问题，以找到最优的 $\xi_n\geq 1$:

$$\xi_n \leftarrow \operatorname*{argmin}_{\xi\geq 1}\frac{\lambda}{2\xi}\left(\frac{|v_n|}{\epsilon}+1\right)^2+\frac{\lambda}{2}\log\xi+\frac{\mu_4}{2}\left(\xi-\sigma_n+\frac{r_{4,n}}{\mu_4}\right)^2. \tag{15}$$

二分区间的合适初始点分别为 $\xi_n=\max\{1,\min\{a,b\}\}$和 $\xi_n=\max\{a,b\}$，其中 $a=(|v_n|/\epsilon+1)^2$ 且 $b=\sigma_n-r_{4,n}/\mu_4$。严格凸性和区间的详细说明见附录 F。

自适应*LOP-ℓ_2/ℓ_1*: </style}变量 $\boldsymbol{v}$ 和 $\boldsymbol{\xi}$ 通过加权变分函数的近端算子联合更新：

$$(\boldsymbol{v},\boldsymbol{\xi})\leftarrow\operatorname{prox}_{(\mu_2^{-1}\boldsymbol{\omega})\odot\phi}\left(\boldsymbol{L}\boldsymbol{x}-\mu_2^{-1}\boldsymbol{r}_2,\boldsymbol{\sigma}-\mu_4^{-1}\boldsymbol{r}_4\right). \tag{16}$$

该算子存在闭式解 [49]:

$$\operatorname{prox}_{\omega\phi}(\boldsymbol{v},\boldsymbol{\xi})=\begin{cases}(0,0), & \text{if } 2\omega\xi+|v|^2\leq\omega^2;\\ (0,\xi-\omega/2), & \text{if } v=0 \text{ and } 2\xi>\omega;\\ \left(v-s\omega\operatorname{sign}(v),\xi+\frac{s^2-1}{2}\omega\right), & \text{otherwise,}\end{cases}$$

其中， $s>0$ 是 $s^3+(2\xi/\omega+1)s-2|v|/\omega=0$的唯一正根，可通过卡尔丹公式 [35, 38, 50]求解。此外，对于任意 $n\in[N]$, 权重 $\boldsymbol{\omega}$ 按如下方式更新：

$$\omega_n\leftarrow\max\left\{0,w_n-\gamma^{-1}\phi(v_n,\xi_n)\right\}. \tag{17}$$

6）对偶变量更新： 拉格朗日乘子通过标准上升法进行更新：

$$\begin{cases}\boldsymbol{r}_1\leftarrow\boldsymbol{r}_1+\mu_1(\boldsymbol{u}-\boldsymbol{A}\boldsymbol{x}),\\ \boldsymbol{r}_2\leftarrow\boldsymbol{r}_2+\mu_2(\boldsymbol{v}-\boldsymbol{L}\boldsymbol{x}),\\ \boldsymbol{r}_3\leftarrow\boldsymbol{r}_3+\mu_3(\boldsymbol{\eta}-\boldsymbol{D}\boldsymbol{\sigma}),\\ \boldsymbol{r}_4\leftarrow\boldsymbol{r}_4+\mu_4(\boldsymbol{\xi}-\boldsymbol{\sigma}).\end{cases} \tag{18}$$

注2 （算法 1 和 2 的收敛）。由于我们的算法采用交替最小化，增广拉格朗日函数的值（见附录 *F* 和 *G* 中的式（*21*）、（*22*））单调递减。该值也以下界零为界，因此收敛到一个非负值。然而，证明序列收敛颇具挑战。尽管我们的目标函数满足 *Kurdyka-Łojasiewicz*（*KL*）性质（见附录 *H* 至 *K*），现有依赖于该性质的分析 *[51, 52]* 假设线性约束是满射的

Algorithm 1 ADMM algorithm for LogLOP-ℓ_2/ℓ_1 (8)

Input: $\boldsymbol{x}^0,\boldsymbol{\sigma}^0,\boldsymbol{u}^0,\boldsymbol{v}^0,\boldsymbol{\eta}^0,\boldsymbol{\xi}^0,\boldsymbol{r}_1^0,\boldsymbol{r}_2^0,\boldsymbol{r}_3^0,\boldsymbol{r}_4^0,\epsilon>0,\lambda\geq 0,\alpha\geq 0,\mu_1>0,\mu_2>0,\mu_3>0,\mu_4>0,\rho\geq 1$

for $k\rightarrow 0,1,\ldots,k_{\max}$ **do**
- Updating $\boldsymbol{x}^{k+1}$ via (10)
- $\boldsymbol{u}^{k+1}\leftarrow\operatorname{prox}_{\mu_1^{-1}f}\left(\boldsymbol{A}\boldsymbol{x}^{k+1}-\mu_1^{-1}\boldsymbol{r}_1^k\right)$ (11)
- $\boldsymbol{v}^{k+1}\leftarrow\operatorname{prox}_{\lambda_1\|\cdot\|_1+\lambda_2\|\cdot\|_2^2}\left(\boldsymbol{L}\boldsymbol{x}^{k+1}-\mu_2^{-1}\boldsymbol{r}_2^k\right)$ (14)
- Updating $\boldsymbol{\sigma}^{k+1}$ via (12)
- $\boldsymbol{\eta}^{k+1}\leftarrow P_{B_1^\alpha}\left(\boldsymbol{D}\boldsymbol{\sigma}^{k+1}-\mu_3^{-1}\boldsymbol{r}_3^k\right)$ (13)
- Updating $\boldsymbol{\xi}^{k+1}$ via (15)
- Updating $(\boldsymbol{r}_1^{k+1},\boldsymbol{r}_2^{k+1},\boldsymbol{r}_3^{k+1},\boldsymbol{r}_4^{k+1})$ via (18)
- $(\mu_1^{k+1},\mu_2^{k+1},\mu_3^{k+1},\mu_4^{k+1})\leftarrow\rho\cdot(\mu_1^k,\mu_2^k,\mu_3^k,\mu_4^k)$

Output: $\boldsymbol{x}^{k_{\max}+1}\in\mathbb{R}^N,\boldsymbol{\sigma}^{k_{\max}+1}\in\mathbb{R}^K$

算法2 用于 AdaLOP-ℓ_2/ℓ_1 (9) 的 ADMM 算法

Input: $\boldsymbol{x}^0,\boldsymbol{\sigma}^0,\boldsymbol{u}^0,\boldsymbol{v}^0,\boldsymbol{\eta}^0,\boldsymbol{\xi}^0,\boldsymbol{\omega}^0,\boldsymbol{w}^0,\boldsymbol{r}_1^0,\boldsymbol{r}_2^0,\boldsymbol{r}_3^0,\boldsymbol{r}_4^0,\gamma>1,\alpha\geq 0,\mu_1>0,\mu_2>0,\mu_3>0,\mu_4=\mu_2,\rho\geq 1$

for $k\rightarrow 0,1,\ldots,k_{\max}$ **do**
- $\omega_n^{k+1}\leftarrow\max\left\{0,w_n^k-\gamma^{-1}\phi(v_n^k,\xi_n^k)\right\}$ (17)
- $\boldsymbol{w}^{k+1}\leftarrow\boldsymbol{\omega}^{k+1}$
- Updating $\boldsymbol{x}^{k+1}$ via (10)
- $\boldsymbol{u}^{k+1}\leftarrow\operatorname{prox}_{\mu_1^{-1}f}\left(\boldsymbol{A}\boldsymbol{x}^{k+1}-\mu_1^{-1}\boldsymbol{r}_1^k\right)$ (11)
- Updating $\boldsymbol{\sigma}^{k+1}$ via (12)
- $\boldsymbol{\eta}^{k+1}\leftarrow P_{B_1^\alpha}\left(\boldsymbol{D}\boldsymbol{\sigma}^{k+1}-\mu_3^{-1}\boldsymbol{r}_3^k\right)$ (13)
- Updating $(\boldsymbol{v}^{k+1},\boldsymbol{\xi}^{k+1})$ via (16)
- Updating $(\boldsymbol{r}_1^{k+1},\boldsymbol{r}_2^{k+1},\boldsymbol{r}_3^{k+1},\boldsymbol{r}_4^{k+1})$ via (18)
- $(\mu_1^{k+1},\mu_2^{k+1},\mu_3^{k+1},\mu_4^{k+1})\leftarrow\rho\cdot(\mu_1^k,\mu_2^k,\mu_3^k,\mu_4^k)$

Output: $\boldsymbol{x}^{k_{\max}+1}\in\mathbb{R}^N,\boldsymbol{\sigma}^{k_{\max}+1}\in\mathbb{R}^K$

正如近期一篇综述 *[53]*中所述。在我们的公式中，线性约束表示为

$$\begin{pmatrix}-\boldsymbol{I}&\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\\ \boldsymbol{O}&-\boldsymbol{I}&\boldsymbol{O}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}&-\boldsymbol{I}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}&-\boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{v}\\ \boldsymbol{\eta}\\ \boldsymbol{\xi}\end{pmatrix}+\underbrace{\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\ \boldsymbol{L}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{D}\\ \boldsymbol{O}&\boldsymbol{I}\end{pmatrix}}_{=:\boldsymbol{C}}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{\sigma}\end{pmatrix}=\boldsymbol{0},$$

其中 $\boldsymbol{C}$ 通常不是满射的，从而违反了该假设。替代惩罚项方法 *[54]* 在理论上可避免这一问题，但会导致病态子问题。尽管存在这些理论缺口，算法 *1* 和算法 *2* 仍表现出稳定的经验收敛，这得益于子问题的凸性。

B. 计算复杂度分析

我们分析所提算法对数LOP-ℓ_2/ℓ_1 （算法1）和自适应LOP-ℓ_2/ℓ_1（算法2）的计算复杂度，重点关注其每次迭代成本。该复杂度受目标信号 $\boldsymbol{x}$、观测 $\boldsymbol{y}$和变换信号 $\boldsymbol{L}\boldsymbol{x}$ 的维度，以及应用算子 $\boldsymbol{A}$和 $\boldsymbol{L}$的相关成本的影响。

令 N 为目标信号维度， J 为观测维度， K 为变换信号维度， C_A 和C_L 分别为应用算子 $\boldsymbol{A}$， $\boldsymbol{A}^\top$ 和 $\boldsymbol{L}$, $\boldsymbol{L}^\top$ 的计算开销， N_{cg} 为共轭梯度法迭代次数（由条件数和容差 [55]决定），并且

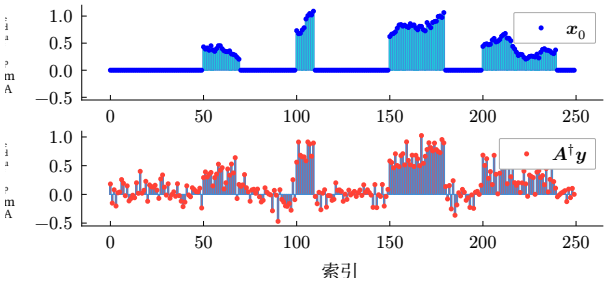


图 4：压缩感知的实验设置。（上）具有四个非零块的真实块稀疏信号 $\boldsymbol{x}_0 \in \mathbb{R}^{250}$ 。（下）初始伪逆恢复 $\boldsymbol{A}^\dagger \boldsymbol{y}$ （信噪比：7.6 dB）。

N_{bisect} 为二分迭代次数（由初始区间和容差决定）。

对于对数LOP-ℓ_2/ℓ_1，每次迭代的复杂度为：

$$\mathcal{O}(N_{\text{cg}}(C_A + C_L) + N_{\text{bisect}} \cdot K)$$

对于自适应LOP-ℓ_2/ℓ_1，每次迭代的复杂度为：

$$\mathcal{O}(N_{\text{cg}}(C_A + C_L) + K)$$

当 $\boldsymbol{A}$ 和 $\boldsymbol{L}$是稀疏算子（如差分算子）时， C_A， $C_L \approx \mathcal{O}(N)$。如果它们可由傅里叶变换对角化，则可直接通过$\mathcal{O}(N \log N)$次运算求解该线性系统，而无需使用共轭梯度法。

V. 数值实验

我们通过数值实验评估所提出的 LogLOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1 惩罚项。我们使用合成数据进行可控分析，并在角功率谱（APS）估计和纳米孔电流去噪等实际应用中，展示它们在偏差缓解和块划分恢复方面的有效性。

A. 压缩感知中的合成数据示例

我们首先在合成压缩感知任务中评估所提出的方法，重点关注偏差缓解和真实支撑集的恢复。

1) 实验设置：我们考虑线性模型 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\epsilon}$，从含噪测量 $\boldsymbol{y} \in \mathbb{R}^J$中恢复块稀疏信号 $\boldsymbol{x} \in \mathbb{R}^N$。设定 $N = 250$ 和 $J = 200$。感知矩阵 $\boldsymbol{A}$的元素独立同分布，服从 $\mathcal{N}(0, 1)$。真实值 $\boldsymbol{x}_0$ （图 4）包含四个具有块内变化的非零块。添加高斯噪声 $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$，使输入信噪比达到 40 dB，其中 SNR $:=20 \log_{10}(\|\boldsymbol{x}_0\|_2 / \|\boldsymbol{x} - \boldsymbol{x}_0\|_2)$。

2) 性能评估：我们将 LogLOP-ℓ_2/ℓ_1和自适应LOP-ℓ_2/ℓ_1 与五种基线方法进行比较： ℓ_1 范数、凸 LOP-ℓ_2/ℓ_1 [35]、非凸 GME-LOP-ℓ_2/ℓ_1 [38]、PC-SBL [28], 和 TV-SBL [33]。对于 PC-SBL 和 TV-SBL，我们均采用 PC-SBL 求解器估计的 λ。所有方法均求解正则化问题 $\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mathcal{R}(\boldsymbol{x})$，其中，对于 ℓ_1、LOP、GME-LOP 和 LogLOP-ℓ_2/ℓ_1， $\mathcal{R}(\boldsymbol{x}) = \lambda\Psi(\boldsymbol{x})$；而对于自适应LOP-$\ell_2/\ell_1$，在初始权重 $\boldsymbol{w} = \lambda\boldsymbol{1}$下， $\mathcal{R}(\boldsymbol{x}) = \Psi_{\alpha,\ \gamma,\ \boldsymbol{w}(\boldsymbol{x})}$。对于基于LOP的方法，块划分参数固定为 $\alpha = 40.0$ ，该值对于 LOP-而言近乎最优

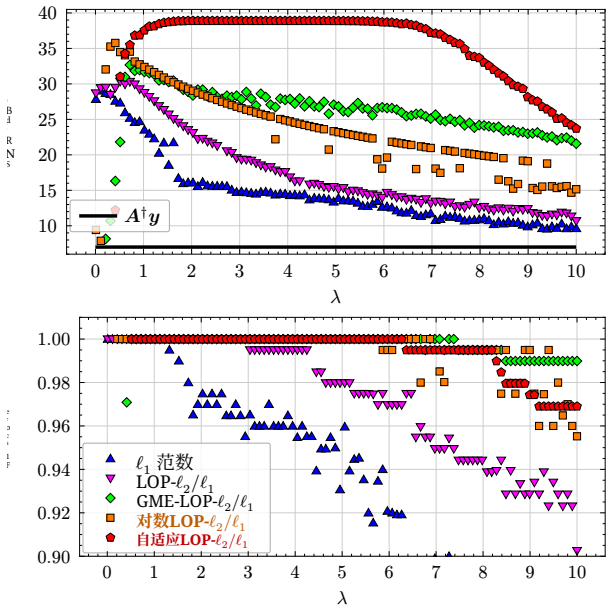


图5：定量评估结果。（上）信噪比与 λ的关系。（下）F1 分数与λ的关系。自适应LOP-ℓ_2/ℓ_1 实现了最高信噪比，并在广泛的 λ 范围内保持稳健性。所提出的方法保持了近乎完美的支撑集恢复

（F1 ≈ 1.0）显著优于凸基线方法。

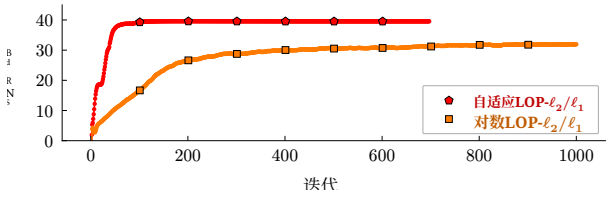


图6：固定 $\lambda = 1.0$、 $\alpha = 4$时信噪比的收敛情况 0.0.

选择的值。对于所有方法，我们均使用从独立同分布标准正态分布中抽取的随机值作为初始点，以减轻初始化偏差。

结果： 图5展示了基于LOP的方法在一次试验中的定量结果。在较宽的 λ范围内，AdaLOP-ℓ_2/ℓ_1 始终取得最高的信噪比，并在支撑集恢复中保持近乎完美的F1分数。这凸显了所提出的非凸形式相较于凸方法快速退化所具有的鲁棒性。此外，图6证实，所提出的方法从随机初始点出发能够稳定收敛至更优解（对100次独立试验取平均，$\lambda = 1.0$）。

从定性角度看，图7展示了偏差缓解能力。尽管凸方法存在幅度收缩，AdaLOP-ℓ_2/ℓ_1 仍能高精度地恢复信号幅值（38.91 dB），与真实值高度吻合。

图8进一步分析了鲁棒性，其中结果为500次独立试验的平均值。在不同观测维度 J 和噪声水平 σ下，AdaLOP-ℓ_2/ℓ_1 始终优于基线方法，证实了其在各种条件下的可靠性。在此次评估中，每种方法的参数 λ 均调至其最佳值。

图 9 考察了 AdaLOP-ℓ_2/ℓ_1 对 γ 的敏感性（对 20 次试验取平均）。我们观察到，最优的 γ 随着 λ 的增大而减小。由于 γ^{-1} 充当

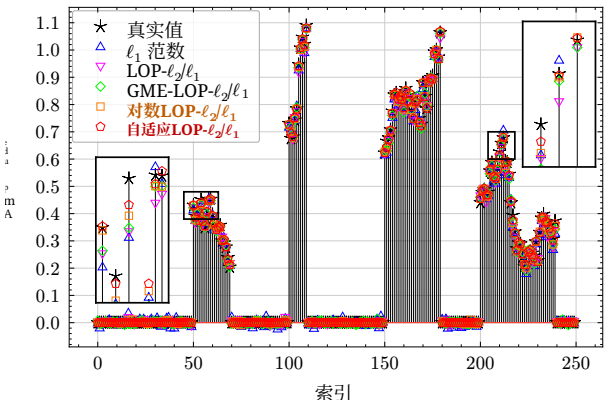


图7：单次试验中恢复信号的定性比较。凸方法（ℓ_1、LOP-ℓ_2/ℓ_1）会产生幅度收缩。相比之下，AdaLOP-ℓ_2/ℓ_1 能够实现高度准确的恢复，有效减轻偏差。

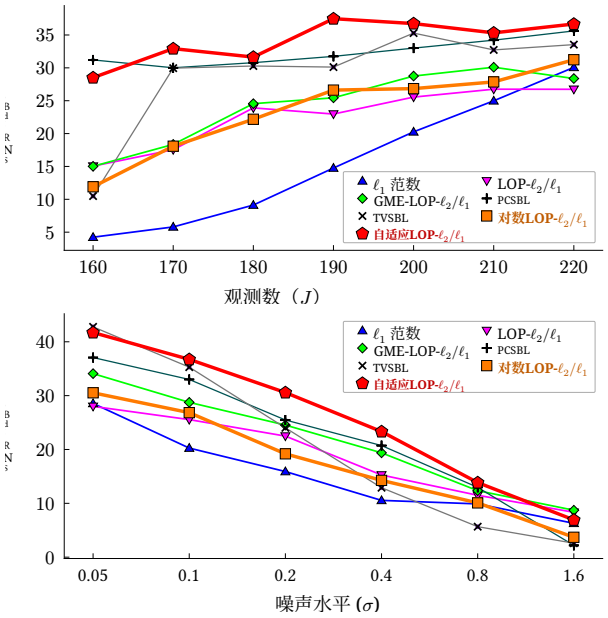


图8：鲁棒性分析（对500次独立试验取平均值）。（上）固定 $\sigma = 0.1$ 时，信噪比与观测数 J 的关系。（下）固定 $J = 200$时，信噪比与噪声水平 σ 的关系。在不同条件下，AdaLOP-ℓ_2/ℓ_1 始终优于基线方法。

式 (17) 中更新 $\boldsymbol{w}$ 的学习率，这意味着更强的正则化需要更大的步长才能有效恢复。

B. 角功率谱估计的应用

我们将所提出的方法应用于MIMO通信系统中的角功率谱（APS）估计问题。由于天线数量相对于角分辨率有限，该问题通常是不适定的。在实际应用中，这一限制至关重要 [56], ，因此需要有效的正则化。

1）问题表述：考虑一个具有 M 个定向天线的均匀线性阵列（ULA），其遵循 3GPP 文档 [57] 和 [56, 示例 1]。接收信号协方差矩阵 $\boldsymbol{R} \in \mathbb{C}^{M\times M}$ 与定义在角度网格 $\{\theta_n\}_{n=1}^N$ 上的角功率谱（APS） $\boldsymbol{x} \in \mathbb{R}^N_{\geq 0}$ 的关系为： $\boldsymbol{R} = \sum_{n=1}^N x_n \boldsymbol{a}(\theta_n)\boldsymbol{a}(\theta_n)^{\mathsf{H}}$,

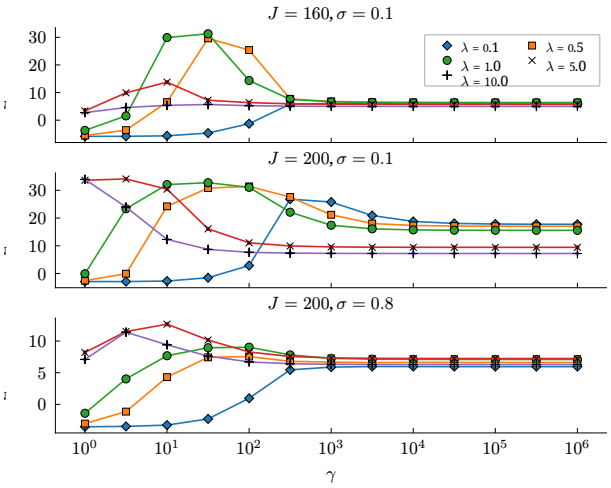


图 9：AdaLOP-ℓ_2/ℓ_1 对 $\gamma > 1$ 的参数敏感性。随着 λ 增大，为获得更高的信噪比，较小的 γ 更为合适。

其中， $\boldsymbol{a}(\theta)$ 是阵列响应向量， $\mathbb{C}$ 表示复数集合。给定由 T 个快拍计算得到的样本协方差矩阵 $\hat{\boldsymbol{R}}$ ，目标是估计 $\boldsymbol{x}$。我们将其表述为一个正则化优化问题。遵循 [56],, 我们引入基于过去观测值的数据驱动先验项。目标函数为：

$$\min_{\boldsymbol{x}\in\mathbb{R}^N_{\geq 0}} \frac{1}{2}\|\mathcal{A}(\boldsymbol{x}) - \hat{\boldsymbol{r}}\|_2^2 + \frac{\mu}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_{\boldsymbol{P}}^2 + \mathcal{R}(\boldsymbol{x}), \tag{19}$$

其中， $\mathcal{A}(\boldsymbol{x})$ 是从角功率谱到向量化协方差的线性算子， $\hat{\boldsymbol{r}}$ 是从 $\hat{\boldsymbol{R}}$ 提取的观测向量， $\frac{\mu}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_{\boldsymbol{P}}^2$ 是数据驱动项，其均值为 $\bar{\boldsymbol{x}}$ ，精度矩阵为 $\boldsymbol{P}$ ，两者均由数据集推导得出。此处，$\|\boldsymbol{z}\|_{\boldsymbol{P}}^2 := \boldsymbol{z}^\top \boldsymbol{P}\boldsymbol{z}$ 表示马氏距离平方。对于 LOP、GME-LOP 和对数 LOP-ℓ_2/ℓ_1，正则化项定义为 $\mathcal{R}(\boldsymbol{x}) = \lambda\Psi(\boldsymbol{x})$；对于具有初始权重 $\boldsymbol{w} = \lambda\mathbf{1}$ 的自适应LOP-ℓ_2/ℓ_1 ，正则化项定义为 $\mathcal{R}(\boldsymbol{x}) = \Psi_{\alpha,\ \gamma,\ \boldsymbol{w}(\boldsymbol{x})}$。

2) 实验设置：实验设置遵循 [56]。我们模拟了一个场景，其中 $N = 100$ 个角度网格点在 $[-\pi/2, \pi/2]$ 中均匀分布。天线数量 M 从 4 到 32 变化。真实角功率谱由 $Q \in \{1, 2\}$ 个高斯形状块组成，其位置 $\hat{}$ 和幅值随机选取。样本协方差 $\boldsymbol{R}$ 根据 $T = 1000$ 个快拍估计，输入信噪比为 30 dB。我们生成了一个包含 $L = 1000$ 个历史角功率谱样本的数据集，以计算先验统计量 $\bar{\boldsymbol{x}}$ 和 $\boldsymbol{P}$。

我们将所提出的对数LOP和自适应LOP-ℓ_2/ℓ_1 的估计精度与若干基线方法进行比较。所有块稀疏正则化方法，包括基于SBL的方法（PC-SBL [28] 和TV-SBL [33]），均纳入了数据驱动先验项$\frac{\mu}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_{\boldsymbol{P}}^2$。对比方法包括：混合方法 [58]（对应于在(19)中设定 $\lambda = 0$ ），以及现有的LOP和GME-LOP-ℓ_2/ℓ_1。我们评估估计角功率谱的归一化均方误差（NMSE），并对100次独立试验取平均。各方法的超参数均通过网格搜索进行优化： $\mu \in \{10^i\}_{i=-2}^1$、 $\lambda \in \{10^i\}_{i=-4}^0$、 $\alpha \in \{10^i\}_{i=1}^4$和 $\gamma \in \{1, 10, 100\}$ 分别用于LOP、GME-LOP、对数LOP、自适应LOP-ℓ_2/ℓ_1。对于PC-SBL和TV-SBL，我们均使用了由 λ 估

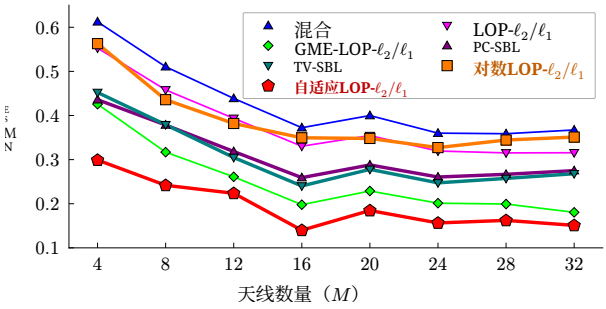


图10：APS估计的归一化均方误差随天线数量的变化（对100次独立试验取平均）。自适应LOP-ℓ_2/ℓ_1 的归一化均方误差最低。

计在PC-SBL求解器内部得到的结果，这是基于SBL的方法的标准无参数方法。

3）性能评估：如图10所示，自适应LOP-ℓ_2/ℓ_1 在所有 M下均始终取得最低的归一化均方误差，显著优于包括最先进的GME-LOP-ℓ_2/ℓ_1 和基于SBL的方法在内的所有其他方法。当天线数量 M 较少时（例如 $M=4$，8），性能差距最为明显，此时问题的病态性较为严重。这表明，即使观测值有限，所提出的方法也能有效利用块稀疏先验来消除角度模糊性。

C. 在纳米孔*Cu*去噪中的应用

我们将所提出的方法应用于纳米孔电流去噪，这是基于纳米孔的 DNA/RNA 测序中的关键步骤。纳米孔电流的特点是由不同核苷酸对应的分段常数区间组成，但由于膜电容和放大器滤波，它们通常会出现瞬态斜率 [59, 60]。尽管实际纳米孔传感器的噪声十分复杂，且包含 $1/f$噪声，本研究通过采用混合泊松-高斯（MPG）分布来建模噪声环境，该分布能够刻画信号相关散粒噪声和热噪声 [59]。

为应对这些特性，我们采用移位 I-散度（SID）作为适用于 MPG 噪声的凸数据保真项 [61]。我们将优化问题表述为：

$$\min_{\boldsymbol{x}\in\mathbb{R}^N_{\geq 0}} \sum_{n=1}^{N} \left\{x_n+\nu^2-y_n\log\left(x_n+\nu^2\right)\right\}+\lambda\Psi(\boldsymbol{D}\boldsymbol{x}),$$

其中，$\boldsymbol{y}\in\mathbb{R}^N_{>0}$ 为观测信号，ν^2 表示热噪声方差，$\boldsymbol{D}$ 为一阶差分算子。正则化项 $\Psi(\boldsymbol{D}\boldsymbol{x})$促进梯度域中的块稀疏性，从而有效刻画信号的分段线性结构（斜率），同时避免引入标准全变分正则化中常见的阶梯伪影。我们所提出框架的一项优势是能够兼容一般的数据保真项。移位 I-散度的近端算子具有闭式解 [39, 引理 2.5 (ii) 和例 4.9]，使得能够通过所提出的基于ADMM的算法（算法 1 和算法 2）高效地进行优化。

1）实验设置：我们生成了长度为 $N=1024$ 的合成纳米孔信号，用于模拟随机阻塞事件，其深度在 $[30,100]$ 皮安范围内均匀分布，且

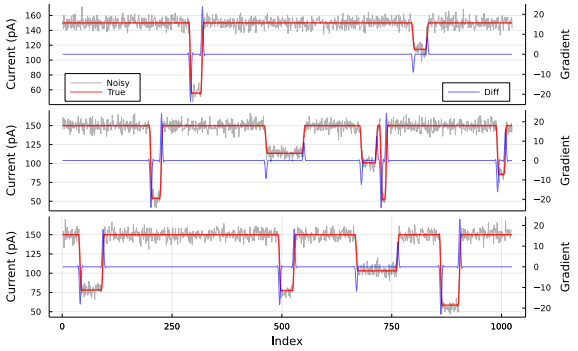


图11：合成纳米孔电流信号示例。这些信号呈现分段常数结构，并带有瞬态斜率（即块稀疏梯度），且受到泊松-高斯混合噪声的污染。

表I：在500次独立试验中，合成纳米孔信号的去噪性能（信噪比（dB））。

方法	Mean	中位数	标准差
LOP-ℓ_2/ℓ_1 (L2)	28.41	28.39	0.40
GME-LOP-ℓ_2/ℓ_1 (L2)	27.89	27.89	0.32
LOP-ℓ_2/ℓ_1 (SID)	32.77	32.24	2.55
对数LOP-ℓ_2/ℓ_1 (SID)	32.02	31.76	1.16
自适应LOP-ℓ_2/ℓ_1 (SID)	**33.41**	**33.17**	1.88

驻留时间服从伽马分布 $\Gamma(2,\ 20)$，基线为 150 皮安。我们使用 4 阶贝塞尔滤波器对信号进行滤波，时间常数为 $\tau_{\text{cap}}=4$，以模拟传感器电容。加入了尺度为 $\alpha_{\text{shot}}=20$ 的 MPG 噪声和 $\sigma_{\text{thermal}}=5$ 的热噪声。生成的纳米孔信号示例见图11。我们将所提出的对数LOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1 （使用 SID）与 LOP-ℓ_2/ℓ_1 （使用 L2 和 SID）以及 GME-LOP-ℓ_2/ℓ_1 （仅使用 L2）进行了比较。我们进行了 500 次独立试验。

2) 结果：表I从信噪比角度总结了去噪性能。使用SID的方法显著优于使用标准平方误差损失（L2）的方法，证实了适当噪声建模的重要性。在基于SID的方法中，所提出的自适应LOP-ℓ_2/ℓ_1取得了最高的平均信噪比，为 33.41 dB。对数LOP-ℓ_2/ℓ_1同样表现出稳健的性能，达到 32.02 dB。这些结果表明，所提出的非凸惩罚能够在存在MPG噪声的情况下有效恢复底层信号。

VI. 结论

我们提出了对数LOP-ℓ_2/ℓ_1 和自适应LOP-ℓ_2/ℓ_1这两种用于块稀疏恢复的非凸正则化框架，它们无需预先知道块划分。这些方法通过非凸变分函数逼近 ℓ_0 伪范数，以减轻低估偏差，同时能够灵活处理高斯假设之外的多种噪声模型。我们证明，在合成与真实世界的信号处理任务中，它们的性能均优于最先进的基准方法。尽管由于非凸性，理论上的全局收敛仍是一个开放问题，但我们基于交替方向乘子法的算法展现出稳定的经验收敛。一个关键局限在于依赖经验性的超参数选择。开发自动化策略以最优确定块划分粒度等参数，仍是未来研究的重要方向。

参考文献

[1] R. Tibshirani，“通过套索法进行回归收缩与选择”，英国皇家统计学会杂志：B辑（统计方法学）58.1（1996），第267–288页，JSTOR: 2346178。[2] S. S. Chen、D. L. Donoho与M. A. Saunders，“基追踪原子分解”，《SIAM科学计算杂志》20.1（1998年1月），第33–61页。[3] Z. Zhang、Y. Xu、J. Yang、X. Li与D. Zhang，“稀疏表示综述：算法与应用”，IEEE Access 3（2015），第490–530页。[4] F. Wen、L. Chu、P. Liu与R. C. Qiu，“信号处理、统计学与机器学习中基于非凸正则化的稀疏与低秩恢复综述”，IEEE Access 6（2018），第69883–69906页。[5] John Wright与Yi Ma，《基于低维模型的高维数据分析：原理、计算与应用》，第1版，剑桥大学出版社，2022年1月。[6] D. He、H. Cao、S. Wang与X. Chen，“时间重分配同步挤压变换：算法及其在机械信号处理中的应用”，《机械系统与信号处理》117（2019年2月），第255–279页。[7] A. Sandryhaila与J. M. F. Moura，“图上的离散信号处理：图傅里叶变换”，载于IEEE国际声学、语音与信号处理会议（ICASSP）论文集，电气与电子工程师协会，2013年5月，第6167–6170页。[8] P. Rodr´ıguez，“针对受不同噪声模型污染图像的全变分正则化算法：综述”，《电气与计算机工程杂志》2013.1（2013），第217021页。[9] H. Wang、J. Peng、W. Qin、J. Wang与D. Meng，“融合低秩性与平滑性的有保证张量恢复”，IEEE模式分析与机器智能汇刊45.9（2023年9月），第10990–11007页。[10] J. Xue、Y. Zhao、S. Huang、W. Liao、J. C.-W. Chan与S. G. Kong，“用于低秩张量补全的基于多层稀疏性的张量分解”，IEEE神经网络与学习系统汇刊33.11（2022年11月），第6916–6930页。[11] C.-H. Zhang，“极小极大凹惩罚下近乎无偏的变量选择”，《统计年刊》38.2（2010年4月），第894–942页。[12] P.-L. Loh与M. J. Wainwright，“无需非相干性的支撑集恢复：支持非凸正则化的一个案例”，《统计年刊》45.6（2017年12月），第2455–2482页。[13] E. J. Cand`es、M. B. Wakin与S. P. Boyd，“通过重加权L1最小化增强稀疏性”，《傅里叶分析与应用杂志》14.5（2008年12月），第877–905页。[14] I. Selesnick，“通过凸分析进行稀疏正则化”，IEEE信号处理汇刊65.17（2017年9月），第4481–4494页。[15] P. K. Pokala、R. V. Hemadri与C. S. Seelamantula，“用于精确低秩加稀疏矩阵分解的迭代重加权极小极大凹惩罚最小化”，IEEE模式分析与机器智能汇刊44.12（2022年2月），第8992–9010页。[16] H. Zhang、H. Fan、Y. Li、X. Liu、C. Liu与X. Zhu，“基于新型非凸函数极小极大对数凹惩罚函数的张量恢复”，IEEE图像处理汇刊32（2023），第3413–3428页。[17] W. Zhu、Z. Huang、J. Chen与Z. Peng，“用于欠定线性方程组稀疏解的迭代加权阈值同伦法”，《中国科学：数学》64.3（2021年3月），第639–664页。[18] Q. Zhao、Y. Lin、F. Wang与D. Meng，“面向基于加权核范数的矩阵/张量补全的自适应加权函数”，《机器学习与控制论国际期刊》15.2（2024年2月），第697–718页。[19] T. Sasaki、Y. Bandoh与M. Kitahara，“基于反向有序加权L1范数的稀疏正则化及其在边缘保持平滑中的应用”，载于IEEE国际声学、语音与信号处理会议（ICASSP）论文集，2024年4月，第9531–9535页。[20] M. Yuan与Y. Lin，“分组变量回归中的模型选择与估计”，英国皇家统计学会杂志：B辑（统计方法学）68.1（2006），第49–67页。[21] L. Jacob、G. Obozinski与J.-P. Vert，“重叠组套索与图套索”，载于国际机器学习会议（ICML）论文集，2009。

[22] Y. C. Eldar、P. Kuppinger 和 H. Bolcskei，“块稀疏信号：不确定性关系与高效恢复”，IEEE信号处理汇刊 58.6（2010），第 3042–3054 页。[23] J. Huang、P. Breheny 和 S. Ma，“高维模型中组选择的精选综述”，Statistical Science 27.4（2012 年 11 月），第 481–499 页。[24] Y. Hu、C. Li、K. Meng、J. Qin 和 X. Yang，“通过 Lp,q 正则化进行组稀疏优化”，机器学习研究杂志 18.30（2017），第 1–52 页。[25] Z. Li、Y. Li、B. Tan、S. Ding 和 S. Xie，“采用组对数正则项的结构化稀疏编码及关键帧提取”，IEEE/CAA自动化学报 9.10（2022 年 10 月），第 1818–1830 页。[26] K. Zhan 和 A. Wan，“Lp/Lq 最小化方法与近似块k稀疏信号的稳定恢复”，IEEE信号处理汇刊（2025），第 1–16 页。[27] G. Obozinski 、 L. Jacob 和 J.-P. Vert，“ 重叠组 Lasso：潜在组Lasso 方法”，2011，arXiv：1110.0413。[28] J. Fang、Y. Shen、H. Li 和 P. Wang，“用于块稀疏信号恢复的模式耦合稀疏贝叶斯学习”，IEEE信号处理汇刊 63.2（2015），第 360–372 页。[29] J. Fang、L. Zhang 和 H. Li，“通过广义近似消息传递进行二维模式耦合稀疏贝叶斯学习”，IEEE图像处理汇刊 25.6（2016 年 6 月），第 2920–2930 页。[30] L. Wang、L. Zhao、S. Rahardja 和 G. Bi，“扩展块稀疏贝叶斯学习的替代方法及其与模式耦合稀疏贝叶斯学习的关系”，IEEE信号处理汇刊 66.10（2018 年 5 月），第 2759–2771 页。[31] Q. He、J. Fang、Z. Chen 和 S. Li，“一种用于未知块划分下块稀疏信号恢复的迭代重加权方法”，载于 IEEE国际声学、语音与信号处理会议（ICASSP）论文集，2016 年 3 月，第 4488–4492 页。[32] Y. Feng 和 I. Selesnick，“潜在融合套索”，载于 IEEE国际声学、语音与信号处理会议（ICASSP）论文集，2020 年 5 月，第 5969–5973 页。[33] A. Sant、M. Leinonen 和 B. D. Rao，“通过广义全变差正则化稀疏贝叶斯学习进行块稀疏信号恢复”，IEEE信号处理汇刊 70（2022），第 1056–1071 页。[34] Y. Zhang、Z. Zhu 和 Y. Xia，“块稀疏贝叶斯学习：一种多样化方案”，载于《神经信息处理系统进展》第 37 卷，Curran Associates, Inc.，2024 年，第 129988–130017 页。[35] H. Kuroda 和 D. Kitahara，“ 采用最优块划分的块稀疏恢复 ”，IEEE 信号处理汇刊 70（2022），第 1506–1520 页。[36] H. Kuroda 和 D. Kitahara，“通过凸优化进行图结构稀疏正则化”，载于 IEEE国际声学、语音与信号处理会议（ICASSP）论文集，电气与电子工程师协会，2022 年 5 月，第 5538–5542 页。[37] T. Furuhashi、H. Hontani 和 T. Yokota，“任意线性变换下的自适应块稀疏正则化”，载于欧洲信号处理会议（EUSIPCO）论文集，2024 年 8 月，第 2437–2441 页。[38] H. Kuroda，“一种用于增强最小化诱导惩罚项的凸—非凸框架”，《富兰克林研究所杂志》362.15（2025 年 10 月），第 107969 页。[39] C. Chaux、P. L. Combettes、J.-C. Pesquet 和 V. R. Wajs，“基于框架的逆问题的变分形式”，Inverse Problems 23.4（2007 年 6 月），第 1495 页。[40] G. Chierchia、E. Chouzenoux、P. L. Combettes 和 J.-C. Pesquet，“邻近算子库”，https://proximity-operator.net/，2016，（访问于 2025 年 12 月 14 日）。[41] J. Fan 和 R. Li，“通过非凹惩罚似然进行变量选择及其预言机性质”，美国统计学会杂志 96.456（2001 年 12 月），第 1348–1360 页。[42] E. Strekalovskiy 和 D. Cremers，“分段光滑 Mumford-Shah 泛函的实时最小化”，载于欧洲计算机视觉会议（ECCV）论文集，Springer International Publishing，2014 年，第 127–141 页。[43] A. Lanza、S. Morigi、I. W. Selesnick 和 F. Sgallari，“用于凸图像处理的稀疏性诱导非凸非可分正则化”，SIAM影像科学杂志 12.2（2019 年 1 月），第 1099–1134 页。[44] J. Abe、M. Yamagishi 和 I. Yamada，“满足整体凸性条件的线性参与广义 Moreau 增强模型及其近端分裂算法”

在整体凸性条件下”，*Inverse Problems* 36.3（2020 年 3 月），第 035012 页。[45] A. H. Al-Shabili、Y. Feng 和 I. Selesnick，“通过平滑锐化稀疏正则项”，*IEEE* 开放获取信号处理期刊 2（2021），第 396–409 页。[46] Alfred Auslender 和 Marc Teboulle，优化与变分不等式中的渐近锥与函数，《数学施普林格专著》，施普林格出版社，2003。[47] S. Boyd，“通过交替方向乘子法进行分布式优化与统计学习”，《机器学习基础与趋势》3.1（2010），第 1–122 页。[48] L. Condat，“快速投影到单纯形和 L1 球”，《数学规划》158.1（2016 年 7 月），第 575–585 页。[49] P. L. Combettes 和 C. L. M¨uller，“通过近端分解进行透视最大似然型估计”，《电子统计学期刊》14.1（2020 年 1 月），第 207–238 页。[50] H. H. Bauschke、M. K. Lal 和 X. Wang，“实系数三次多项式的实根与优化”，2023，arXiv：2302.10731。[51] Y. Wang、W. Yin 和 J. Zeng，“非凸非光滑优化中 ADMM 的全局收敛”，《科学计算杂志》78.1（2019 年 1 月），第 29–63 页。[52] R. I. Bot¸ 和 D.-K. Nguyen，“非凸情形下的近端交替方向乘子法：收敛分析与收敛速率”，《运筹学数学》45.2（2020 年 5 月），第 682–712 页。[53] D.-R. Han，“交替方向乘子法近期若干进展综述”，《中国运筹学会学报》10.1（2022 年 3 月），第 1–52 页。[54] H. Attouch、J. Bolte、P. Redont 和 A. Soubeyran，“非凸问题的近端交替最小化与投影方法：基于 Kurdyka-Łojasiewicz 不等式的方法”，《运筹学数学》35.2（2010 年 5 月），第 438–457 页。[55] J. R. Shewchuk，共轭梯度法简介：不再令人痛苦，技术报告，卡内基梅隆大学，1994 年 2 月。[56] H. Kuroda、R. L. G. Cavalcante 和 M. Yukawa，“潜在最优分区 L2/L1 惩罚项的理论验证及其在角功率谱估计中的应用”，2025，arXiv：2509.13745。[57] 3第三代合作伙伴计划（3GPP），*0.5* 至 *100 GHz* 频段的信道模型研究，技术报告 TR 38.901 V14.0.0，第三代合作伙伴计划（3GPP），2017 年 5 月。[58] R. L. G. Cavalcante 和 S. Sta´nczak，“用于角功率谱估计的数据与模型混合驱动算法”，载于 *IEEE* 全球通信大会（*GLOBECOM*）论文集，2020 年 2 月，第 1–6 页。[59] S. Liang、F. Xiang、Z. Tang、R. Nouri、X. He、M. Dong 和 W. Guan，“纳米孔传感器中的噪声：来源、模型、降低方法与基准测试”，《纳米技术与精密工程》3.1（2020 年 3 月），第 9–17 页。[60] C. Wen、D. Dematties 和 S.-L. Zhang，“纳米孔传感器信号处理算法指南”，《*ACS* 传感器》6.10（2021 年 10 月），第 3536–3555 页。[61] A. Chakrabarti 和 T. Zickler，“采用信号相关相机噪声的图像复原”，2012，arXiv：1204.2994。[62] S. P. Boyd 和 L. Vandenberghe，凸优化，第 29 版，剑桥大学出版社，2023。[63] N. Bourbaki，一般拓扑学，《数学基础》第 3 卷，施普林格，1989。

附录 A
一般 LOP 惩罚项的极小值点存在性

我们验证一般 LOP 型惩罚 (3) 中联合变量 $(\boldsymbol{x},\boldsymbol{\sigma})$ 的极小值点的存在性。考虑函数

$$H(\boldsymbol{x},\boldsymbol{\sigma}) = h(\boldsymbol{x},\boldsymbol{\sigma}) + \iota_{\tilde{C}}(\boldsymbol{x},\boldsymbol{\sigma}) \qquad h(\boldsymbol{x},\boldsymbol{\sigma}) = \sum_{n=1}^{N}\varphi(x_n,\sigma_n)。$$

此处，$\varphi(\cdot,\cdot):(\cdot,\cdot)$ 由一个满足渐近水平稳定（als）性质 [46,定义 3.3.2] 的正常下半连续函数构成。约束集为 $\tilde{C} := \mathbb{R}^N \times C$，其中 $C = \{\boldsymbol{\sigma} \in \mathbb{R}^N \mid \|\boldsymbol{D}\boldsymbol{\sigma}\|_1 \le \alpha\}$ 是一个多面体集。

约束集性质： 线约束集 [46, 定义 2.3.2 和命题 2.3.1] 因为它是一个多面体集 [62, 第 2.2 节。4]。

变分函数的性质： 现在我们验证可分和 h 的 ALS 性质，假设每个 φ 都满足该性质。考虑一个序列 $\{\boldsymbol{z}^k\}$，满足 $\|\boldsymbol{z}^k\| \to \infty$ 且 $\boldsymbol{z}^k/\|\boldsymbol{z}^k\| \to \bar{\boldsymbol{z}} \in \mathbb{R}^{2N}$。将 $\boldsymbol{z}^k$ 分解为 N 个对 $\{(z_n^k)\}_{n=1}^N$，其中 z_n 由 $(x_n,\ \sigma_n)$ 构成。对于发散分量，φ 的 ALS 性质确保对于任意 $\rho > 0$ 以及充分大的 k，有 $\varphi(z_n^k - \rho\bar{z}_n) \le \varphi(z_n^k)$。对于有界分量，$\bar{z}_n = \lim_{k\to\infty} z_n^k/\|\boldsymbol{z}^k\| = 0$，从而得到 $\varphi(z_n^k - \rho\bar{z}_n) = \varphi(z_n^k)$。

将这些不等式相加，可得

$$h(\boldsymbol{z}^k - \rho\bar{\boldsymbol{z}}) = \sum_{n=1}^{N}\varphi(z_n^k - \rho\bar{z}_n) \le \sum_{n=1}^{N}\varphi(z_n^k) = h(\boldsymbol{z}^k).$$

因此，如果 $\boldsymbol{z}^k$ 属于 h的某个水平集，那么 $\boldsymbol{z}^k - \rho\bar{\boldsymbol{z}}$ 也属于同一个水平集，从而验证了 h的 ALS 性质。

因此，根据 [46, 命题 3.3.3(c)]，和 $H = h + \iota_{\tilde{C}}$ 也是 ALS。由于 H 下有界，[46, 推论 3.42] 保证 H 存在极小值点。该结果可推广到正则化问题 (1)，其中数据保真项 f 是适定的、下半连续的、强制的且为凸函数。该推广之所以成立，是因为 $f \circ \boldsymbol{A}$ 是弱强制的（ [46, 例 3.21],），并且通过结合弱强制性与 ALS 的定义（ [46, 定义 3.2.1 和定义 3.32] ）以及渐近函数的和法则（ [46,命题 2.6.1(b)]），和 $f \circ \boldsymbol{A} + H$仍具有 ALS 性质。

对于特定惩罚项，对数LOP-ℓ_2/ℓ_1 满足 ALS 性质，因为其变分函数 ϕ_ϵ （见式（4））天然具有强制性。自适应 LOP-ℓ_2/ℓ_1 也满足 ALS 性质。对于$\phi_{\gamma,\ w} = g\circ\phi$，ALS 性质可直接由 ϕ的几何性质推出。由于 ϕ 是凸且具有强制性的 [35, 引理1], ，其水平集有界，且沿任意径向方向单调递增。这确保对于任意一个在方向 $\bar{\boldsymbol{z}}$ 上趋于无穷且 k充分大的序列 $\{\boldsymbol{z}^k\}$，都有 $\phi(\boldsymbol{z}^k - \rho\bar{\boldsymbol{z}}) \le \phi(\boldsymbol{z}^k)$。$g$ 的非递减性使得这一不等式对$\phi_{\gamma,\ w}$仍然成立，从而验证了 ALS 性质。

附录 B 定理1的证明

我们分析$\Psi_{\alpha,\ \epsilon}(\boldsymbol{x})$在 α极限下的行为。当 $\alpha \to \infty$时，约束 $\|\boldsymbol{D}\boldsymbol{\sigma}\|_1 \le \alpha$ 变为非活动约束。根据引理 1，当$\sigma_n = (|x_n|/\epsilon + 1)^2$，时，项$\sum_{n=1}^N \phi_\epsilon(x_n,\ \sigma_n)$取得最小值，这意味着 $\lim_{\alpha\to\infty}\Psi_{\alpha,\ \epsilon}(\boldsymbol{x}) = \Omega_\epsilon(\boldsymbol{x})$。

反之，当 $\alpha \to 0$ 时，该约束强制满足 $\|\boldsymbol{D}\boldsymbol{\sigma}\|_1 = 0$，这意味着对于所有 n，存在某个 $\tau \ge 1$，使得 $\sigma_n = \tau$。因此：

$$\begin{aligned}
\Psi_{0,\epsilon}(\boldsymbol{x}) &= \min_{\tau\ge1}\sum_{n=1}^{N}\left\{\frac{1}{2\tau}\left(\frac{|x_n|}{\epsilon}+1\right)^2 + \frac{1}{2}\log\tau - \frac{1}{2}\right\} \\
&= \min_{\tau\ge1}\left\{\frac{1}{2\tau}\sum_{n=1}^{N}\left(\frac{|x_n|}{\epsilon}+1\right)^2 + \frac{N}{2}\log\tau - \frac{N}{2}\right\} \\
&= \frac{N}{2} + N\log\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\frac{|x_n|}{\epsilon}+1\right)^2} - \frac{N}{2} \\
&= N\log\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\frac{|x_n|}{\epsilon}+1\right)^2}.
\end{aligned}$$

从第二行到第三行的过渡，是通过代入最优的 $\tau = N^{-1}\sum_{n=1}^N(|x_n|/\epsilon+1)^2$, 得到的，其可类似于引理 1 推导得出。

附录C 命题1的证明

我们证明，惩罚差 $\Delta(z)$ 对于$z > L$和 $M > L$为正，其中 $M = |\mathcal{B}|$、 $L = |\mathcal{B}'|$和 $z := \sum_{n\in\mathcal{B}'}(|x_n|/\epsilon+1)^2$。惩罚差为：

$$\Delta(z) \coloneqq M\log\left(\frac{z+M-L}{M}\right) - L\log\left(\frac{z}{L}\right).$$

注意， $\Delta(L) = 0$。关于 z 的导数为：

$$\Delta'(z) = \frac{M}{z+M-L} - \frac{L}{z} = \frac{(M-L)(z-L)}{z(z+M-L)}.$$

对于 $z > L$ 和 $M > L$，有 $M-L>0$、 $z-L>0$，且$z(z+M-L) > 0$。因此， $\Delta'(z) > 0$。因为当 $z > L$时， $\Delta(L) = 0$ 且$\Delta(z)$严格递增，所以 $\Delta(z) > 0$。这证实，将零块 $\mathcal{B}''$ 与块$\mathcal{B}$ 分离会严格降低惩罚项。

附录D 引理2的证明

我们证明不等式$\phi_{\gamma,w}(x,\sigma) \geq \Omega_{\gamma,w}(x)$ 对任意$w,\gamma > 0$成立。MCP函数 $g \coloneqq \Omega_{\gamma,w}$ 扩展到扩展实数轴 $[0,\infty]$ 上，定义为

$$g(z) = \begin{cases} wz - \frac{z^2}{2\gamma}, & \text{if } 0 \leq z \leq \gamma w; \\ \frac{w^2\gamma}{2}, & \text{if } z \geq \gamma w. \end{cases} \tag{20}$$

注意，对所有 $x \in \mathbb{R}$， $g(|x|) = \Omega_{\gamma,w}(x)$。 g 在 $[0,\infty]$ 上的连续性和单调性确保 $\phi_{\gamma,w} = g\circ\phi$ 定义良好，并满足所需的不等式。

() 的单调性。对于 $0 \leq z < \gamma w$， $g'(z) = w - z/\gamma > 0$；对于 $z \geq \gamma w$， $g(z)$ 保持为常数 $w^2\gamma/2$，这意味着对于 $z > \gamma w$， $g'(z) = 0$。在 $z = \gamma w$ 处连续性仍然成立，因为 $\lim_{z\uparrow\gamma w} g(z) = \frac{w^2\gamma}{2} = g(\gamma w)$。因此， $g(z)$ 在扩展实数轴 $[0,\infty]$上是非递减的。

由于 $\phi(x,\sigma) \geq |x|$，且 $g(z)$ 是非递减的：

$$g(\phi(x,\sigma)) \geq g(|x|) \implies \phi_{\gamma,w}(x,\sigma) \geq \Omega_{\gamma,w}(x).$$

当 $\tau = |x|$ [35, Lemma 1]时等号成立。

附录E 定理2的证明

我们分析 $\Psi_{\alpha,\gamma,\boldsymbol{w}}(\boldsymbol{x})$ 的极限。当 $\alpha\to\infty$时，忽略 TV 约束 $\|\boldsymbol{D\sigma}\|_1 \leq \alpha$。最小化 $\sum_{n=1}^N\{\omega_n\phi(x_n,\sigma_n)\}$ 得到 $\sigma_n = |x_n|$，因此 $\Psi_{\alpha,\gamma,\boldsymbol{w}}(\boldsymbol{x}) \to \Omega_{\gamma,\boldsymbol{w}}(\boldsymbol{x})$。

当 $\alpha\to 0$ 时，该约束意味着对所有 n 都有 $\sigma_n = \tau$。因此：

$$\Psi_{0,\gamma,\boldsymbol{w}}(\boldsymbol{x}) = \min_{\substack{\tau\in\mathbb{R}_{\geq 0}\\ \boldsymbol{\omega}\in\mathbb{R}_{\geq 0}}} \sum_{n=1}^N\left\{\omega_n\left(\frac{|x_n|^2}{2\tau}+\frac{\tau}{2}\right)\right\} + \frac{\gamma}{2}\|\boldsymbol{\omega}-\boldsymbol{w}\|_2^2$$

$$= \min_{\substack{\tau\in\mathbb{R}_{\geq 0}\\ \boldsymbol{\omega}\in\mathbb{R}_{\geq 0}}} \frac{\|\boldsymbol{x}\|_{\boldsymbol{\omega},2}^2}{2\tau} + \frac{\tau}{2}\|\boldsymbol{\omega}\|_1 + \frac{\gamma}{2}\|\boldsymbol{\omega}-\boldsymbol{w}\|_2^2.$$

对 τ 最小化得到$\tau = \|\boldsymbol{x}\|_{\boldsymbol{\omega},2}/\|\boldsymbol{\omega}\|_1^{1/2}$ （类似于 [35, 引理 1] 的证明）。将其代回即得(7)。

附录 F 算法1的推导

我们推导由对数LOP-ℓ_2/ℓ_1 惩罚项正则化的问题的交替方向乘子法更新：

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} f(\boldsymbol{Ax}) + \lambda\Psi_{\alpha,\epsilon}(\boldsymbol{Lx}).$$

引入辅助变量 $\boldsymbol{u} = \boldsymbol{Ax}$、 $\boldsymbol{v} = \boldsymbol{Lx}$、 $\boldsymbol{\eta} = \boldsymbol{D\sigma}$ 和 $\boldsymbol{\xi} = \boldsymbol{\sigma}$，我们将问题改写为：

$$\min_{\substack{\boldsymbol{x}\in\mathbb{R}^N,\boldsymbol{\sigma}\in\mathbb{R}^K,\boldsymbol{u}\in\mathbb{R}^J,\\ \boldsymbol{v}\in\mathbb{R}^K,\boldsymbol{\eta}\in\mathbb{R}^{K-1},\boldsymbol{\xi}\in\mathbb{R}^K}} f(\boldsymbol{u}) + \lambda\sum_{n=1}^N\phi_\epsilon(v_n,\xi_n) + \iota_{B_1^\alpha}(\boldsymbol{\eta}),$$
$$\text{s.t. } \boldsymbol{u} = \boldsymbol{Ax}, \boldsymbol{v} = \boldsymbol{Lx}, \boldsymbol{\eta} = \boldsymbol{D\sigma}, \boldsymbol{\xi} = \boldsymbol{\sigma}.$$

这里， $\iota_C(\boldsymbol{x})$ 是集合 C 的指示函数；当 $\boldsymbol{x}\in C$ 时取 0，否则取 ∞。增广拉格朗日函数为：

$$\begin{aligned}\mathcal{L} = {} & f(\boldsymbol{u}) + \lambda\sum_{n=1}^N\phi_\epsilon(v_n,\xi_n) + \iota_{B_1^\alpha}(\boldsymbol{\eta}) \\ & + \boldsymbol{r}_1^\top(\boldsymbol{u}-\boldsymbol{Ax}) + \frac{\mu_1}{2}\|\boldsymbol{u}-\boldsymbol{Ax}\|_2^2 + \boldsymbol{r}_2^\top(\boldsymbol{v}-\boldsymbol{Lx}) \\ & + \frac{\mu_2}{2}\|\boldsymbol{v}-\boldsymbol{Lx}\|_2^2 + \boldsymbol{r}_3^\top(\boldsymbol{\eta}-\boldsymbol{D\sigma}) + \frac{\mu_3}{2}\|\boldsymbol{\eta}-\boldsymbol{D\sigma}\|_2^2 \\ & + \boldsymbol{r}_4^\top(\boldsymbol{\xi}-\boldsymbol{\sigma}) + \frac{\mu_4}{2}\|\boldsymbol{\xi}-\boldsymbol{\sigma}\|_2^2,\end{aligned} \tag{21}$$

其中 $\mu_i > 0$ 为交替方向乘子法参数， $\boldsymbol{r}_i$ 为拉格朗日乘子。由于交替方向乘子法的更新规则源自交替最小化，且各子问题均为凸问题，因此当$\mu_4 \geq \lambda/54$ 时，算法1会使 $\mathcal{L}$单调递减。

1) 更新 $\boldsymbol{x}$、 $\boldsymbol{u}$、 $\boldsymbol{v}$、 $\boldsymbol{\sigma}$、 $\boldsymbol{\eta}$：分别对各变量最小化 $\mathcal{L}$，即可得到式（10）至式（14）中的更新式。

2) $\boldsymbol{\xi}$更新：如式（15）所示，该子问题可分解为逐元素最小化问题。为确保严格凸性和解的唯一性，我们要求 $\mu_4 > \lambda/54$。该条件由关于 ξ_n的二阶导数下界推导得出：

$$\frac{\partial^2\mathcal{L}}{\partial\xi_n^2} = \lambda\left(\frac{a}{\xi_n^3} - \frac{1}{2\xi_n^2}\right) + \mu_4 \geq \mu_4 - \frac{\lambda}{54a^2} \geq \mu_4 - \frac{\lambda}{54}.$$

其中 $a = (|v_n|/\epsilon+1)^2$。我们使用二分法求解 ξ_n。一阶导数决定二分区间：

$$\frac{\partial\mathcal{L}}{\partial\xi_n} = \frac{\lambda}{2\xi_n^2}(\xi_n - a) + \mu_4(\xi_n - b),$$

其中$b = \sigma_n - \mu_4^{-1}r_{4,n}$。将 $\xi_n = a$和 $\xi_n = b$代入导数后，可得 $\frac{\partial\mathcal{L}}{\partial\xi_n}|_{\xi_n=a} = \mu_4(a-b)$和$\frac{\partial\mathcal{L}}{\partial\xi_n}|_{\xi_n=b} = -\frac{\lambda}{2b^2}(a-b)$，二者符号相反。因此，解位于 $[\min(a,b), \max(a,b)]$内。考虑约束 $\xi_n \geq 1$ 以及 $a \geq 1$这一事实，最优解由 $\xi_n = \max\{1, \min\{a,b\}\}$ 和 $\xi_n = \max\{a,b\}$夹定。

3) 对偶变量（$\boldsymbol{r}_i$）$_{i=1}^4$ 的更新见式（18）。

附录 G 算法2的推导

我们推导采用自适应LOP-ℓ_2/ℓ_1 惩罚项正则化的问题的交替方向乘子法更新公式：

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} f(\boldsymbol{Ax}) + \Psi_{\alpha,\gamma,\boldsymbol{w}}(\boldsymbol{Lx}).$$

增广拉格朗日函数 $\mathcal{L}$ 为:

$$\begin{aligned}\mathcal{L} = & f(\boldsymbol{u}) + \sum_{n=1}^{N} \omega_n \phi(v_n, \xi_n) + \frac{\gamma}{2}\|\boldsymbol{\omega} - \boldsymbol{w}\|_2^2 + \iota_{B_1^\alpha}(\boldsymbol{\eta}) \\ & + \boldsymbol{r}_1^\top(\boldsymbol{u} - \boldsymbol{A}\boldsymbol{x}) + \frac{\mu_1}{2}\|\boldsymbol{u} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \boldsymbol{r}_2^\top(\boldsymbol{v} - \boldsymbol{L}\boldsymbol{x}) \\ & + \frac{\mu_2}{2}\|\boldsymbol{v} - \boldsymbol{L}\boldsymbol{x}\|_2^2 + \boldsymbol{r}_3^\top(\boldsymbol{\eta} - \boldsymbol{D}\boldsymbol{\sigma}) + \frac{\mu_3}{2}\|\boldsymbol{\eta} - \boldsymbol{D}\boldsymbol{\sigma}\|_2^2 \\ & + \boldsymbol{r}_4^\top(\boldsymbol{\xi} - \boldsymbol{\sigma}) + \frac{\mu_4}{2}\|\boldsymbol{\xi} - \boldsymbol{\sigma}\|_2^2.\end{aligned} \tag{22}$$

算法2使 $\mathcal{L}$单调递减。

1) $\boldsymbol{x}$、$\boldsymbol{u}$、$\boldsymbol{\sigma}$、$\boldsymbol{\eta}$、$(\boldsymbol{r}_i)_{i=1}^4$的更新：这些更新与(10)至(13)中描述的对数LOP-ℓ_2/ℓ_1 更新相同。

2) 联合($\boldsymbol{v}$、$\boldsymbol{\xi}$)更新：令 $\mu_4 = \mu_2$，该更新即为(16)中描述的 ϕ 的近端算子。

3) $\boldsymbol{\omega}$更新：对 $\omega_n \geq 0$ 进行最小化可得(17)所示的解。虽然理论上需要采用规则 $0\cdot(-\infty) = -\infty$来处理 $\phi(v_n、\ \xi_n) = \infty$，但在实践中，使用可行点初始化可确保每次迭代时 $\phi(v_n、\ \tau_n) < \infty$ 。

4) $\boldsymbol{w}$更新：当 $\boldsymbol{w} \leftarrow \boldsymbol{\omega}$时，取得最小值。

附录 H
$\phi(x,\ \tau)$ 的 Kurdyka-Łojasiewicz 性质

我们证明，在（2）中给出的 $\phi:\ \mathbb{R}\times\mathbb{R}\to\mathbb{R}_{\geq 0}\cup\{+\infty\}$ 满足 Kurdyka-Łojasiewicz（KL）性质。若一个函数是适当的、下半连续的，且其图是半代数集（或在某个 o-极小结构中可定义），则该函数具有 KL 性质。由于 dom(ϕ) $\neq\emptyset$ 且 $\phi(x,\ \tau)\geq 0$，因此它是适当的。 ϕ 的下半连续性证明见 [35，附录 B]。

ϕ 的图为 graph(ϕ) $=\{(x,\ \tau,\ y)\in\mathbb{R}^3 \mid y = \phi(x,\ \tau)$，并且 $\phi(x,\ \tau) < \infty\}$。

- **情形 1：** $\tau > 0$。条件 $y = \frac{x^2}{2\tau} + \frac{\tau}{2}$ 等价于 $2\tau y - x^2 - \tau^2 = 0$。集合 $G_1 = \{(x,\ \tau,\ y)\in\mathbb{R}^3 \mid 2\tau y - x^2 - \tau^2 = 0 \wedge \tau > 0\}$ 是半代数的，因为它由多项式等式和不等式定义。

- **情况2：** $x = 0$ 且 $\tau = 0$。集合 $G_2 = \{(x,\ \tau,\ y)\in\mathbb{R}^3 \mid x = 0 \wedge \tau = 0 \wedge y = 0\}$ 是半代数集。

由于 graph(ϕ) $= G_1 \cup G_2$ 是有限个半代数集的并，因此它是半代数集。由于 $\phi(x,\ \tau)$ 适当、下半连续，且其图是半代数集，因此具有 KL性质。

附录I
$\phi_\epsilon(x,\ \tau)$ 的 Kurdyka-Łojasiewicz性质

我们验证函数 $\phi_\epsilon(x,\ \tau)$ 的 KL性质。

适当性： dom(ϕ_ϵ) $=\{(x,\ \tau) \mid \tau\geq 1\}$ 非空。对于 $\tau\geq 1$，$\phi_\epsilon(x,\ \tau)\geq$对数 $(|x|/\epsilon+1)\geq 0$。因此，ϕ_ϵ是适当的。

下半连续性：我们证明，对于任意 $a\in\mathbb{R}$，水平集 $L_a = \{(x,\tau)\in\mathbb{R}^2 \mid \phi_\epsilon(x,\tau)\leq a\}$ 均为闭集。由于 dom(ϕ_ϵ) $=\{(x,\tau) \mid \tau\geq 1\}$，具有 $\tau < 1$ 的点满足$\phi_\epsilon(x,\tau) = \infty$，因此不属于 L_a。于是，$L_a\subseteq\{(x,\tau) \mid \tau\geq 1\}$。在集合 $\{(x,\tau) \mid \tau\geq 1\}$上，$\phi_\epsilon(x,\tau)$ 是连续的。因此，L_a 是闭集 $\{(x,\tau) \mid \tau\geq 1\}$ 与连续函数下闭集 $(-\infty, a]$ 的原像的交集。因此，L_a 是闭集，这意味着ϕ_ϵ 是下半连续的。

可定义性： 对于 $\tau\geq 1$，$\phi_\epsilon(x,\tau)$，涉及 $|x|$，$1/\tau$ 和对数 τ，它们在对数 - 指数结构 $\mathbb{R}_{\text{an, exp}}$ [54] 中是可定义的。因此，限制在 $\tau\geq 1$ 上的图在 $\mathbb{R}_{\text{an, exp}}$ 中是可定义的。由于 $\phi_\epsilon(x,\tau)$ 是适当的、下半连续的，并且在 $\mathbb{R}_{\text{an, exp}}$ 中可定义，因此它满足 KL 性质。

附录J
$\psi(x,\ \tau,\ \omega,\ w)$ 的 Kurdyka-Łojasiewicz 性质

我们证明了出现在自适应LOP-ℓ_2/ℓ_1最小化问题中的联合函数$\psi(x,\ \tau,\ \omega,\ w) := \omega\phi(x,\ \tau) + \frac{\gamma}{2}(\omega - w)^2$，具有 KL 性质。变分函数 $\phi_{\gamma,\ w}(x,\ \tau)$ 满足 KL 性质（见附录 K）。然而，算法2针对涉及联合变量 $(x,\ \tau,\ \omega,\ w)$ 的增广拉格朗日函数进行运算。因此，我们验证 ψ 关于联合变量具有 KL 性质。

适当性： ψ 的有效定义域为 dom(ψ) =dom(ϕ)$\times\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}$。由于 dom($\phi$) 非空，dom($\psi$) 也非空。此外，由于 $\phi(x,\ \tau)\geq 0$，$\omega\geq 0$, and$\gamma > 0$，该函数以下界 0 为界。因此，ψ 是适当的。

下半连续性（lsc）： 函数 $\phi(x,\ \tau)$ 是下半连续的，如 [35, 附录 B]所示。变量 ω 非负，且映射 $(x,\ \tau,\ \omega,\ w)\mapsto\omega$ 是连续的。非负下半连续函数的乘积在有定义时为下半连续 [63]，因此根据规则 $0\cdot\infty = 0$，项 $\omega\phi(x,\ \tau)$ 是下半连续的。二次项 $\frac{\gamma}{2}(\omega - w)^2$ 显然连续。下半连续函数之和仍为下半连续 [63]，因此 ψ 是下半连续的。

可定义性： $\phi(x,\tau)$的图是半代数的，如附录 H 所证明。诸如 ω 和$(\omega - w)^2$，之类的多项式函数是半代数的。由于半代数函数类在有限和与有限积下封闭 [54, 第4节3]，$\psi(x,\tau,\omega,w)$是半代数的。因此，$\psi(x,\tau,\omega,w)$在o-极小结构中是可定义的。

由于是正常的、下半连续的且可定义的，联合变分函数 $\psi(x,\tau,\omega,w)$满足KL性质。

附录 K
$\phi_{\gamma,w}(x,\tau)$的KURDYKA-ŁOJASIEWICZ性质

我们验证(6)中定义的复合函数$\phi_{\gamma,w}(x,\tau) = g(\phi(x,\tau))$的KL性质。注意，$g(z)$是非递减的分段多项式函数(20)，而$\phi(x,\tau)$是KL函数（见附录 H）。

适当性： 由于 g 定义在扩展实数轴$[0,\infty]$上（见附录D），因此复合函数 $\phi_{\gamma,\ w} = g\circ\phi$ 对所有 $(x,\ \tau)$ 都有定义。鉴于 $0\leq\phi_{\gamma,\ w}(x,\ \tau)\leq\frac{\gamma w^2}{2}$，可知$\phi_{\gamma,\ w}$ 是适当的。

下半连续性： 我们证明，对于任意 $a\in\mathbb{R}$，水平集 $L_a = \{(x,\ \tau) \mid \phi_{\gamma,\ w}(x,\ \tau)\leq a\}$ 都是闭集。

1) 如果 $a\geq\frac{\gamma w^2}{2}$，集合 L_a 是 $\mathbb{R}^2$，因此是闭集。
2) 如果 $a < \frac{\gamma w^2}{2}$，条件 $\phi_{\gamma,\ w}(x,\ \tau)\leq a$意味着$g(\phi(x,\ \tau))\leq a$。在此范围内，$g(z)$ 严格递增。因此，这等价于 $\phi(x,\ \tau)\leq g^{-1}(a)$。由于 $\phi(x,\ \tau)$ 是下半连续的，其水平集是闭集。

因此，$\phi_{\gamma,\ w}$ 是下半连续的。

可定义性： 自代数函数的复合仍是自代数函数 [54]。因此，$\phi_{\gamma,\ w} = g\circ\phi$ 是自代数函数，因而是可定义的 [54, 第 4 节。3]。

因此，$\phi_{\gamma,\ w}$ 满足 KL 性质。