arXiv:2602.03283v1 [math.ST] 3 Feb 2026

旋转不变噪声下矩形尖峰矩阵模型的正交近似消息传递算法

Haohua Chen* Songbin Liu† Junjie Ma*

*中国科学院数学与系统科学研究院，北京，中国

†哥伦比亚大学电气工程系，美国纽约 电子邮箱：chenhaohua25@mails.ucas.ac.cn, sl5878@columbia.edu, majunjie@lsec.cc.ac.cn

摘要

我们针对含一般旋转不变（RI）噪声的矩形尖峰矩阵模型中的信号估计，提出了一种正交近似消息传递（OAMP）算法。我们建立了严格的状态演化，能够精确刻画该算法的高维动态。基于这一框架，我们推导出 OAMP 的一个最优变体，可在每次迭代中使预测的均方误差最小化。对于独立同分布高斯噪声这一特殊情形，所提出的 OAMP 算法的不动点与标准 AMP 算法的不动点一致。对于一般 RI 噪声模型，我们猜想，最优 OAMP 算法在一大类迭代方法中具有统计最优性，并能在某些情形下达到贝叶斯最优性能。

索引词——近似消息传递、尖峰模型、旋转不变随机矩阵、状态演化。

1. 引言

我们考虑从矩形尖峰矩阵模型中估计秩一信号 $\boldsymbol{u}_* \in \mathbb{R}^M$, $\boldsymbol{v}_* \in \mathbb{R}^N$:

$$\boldsymbol{Y} = \frac{\theta}{\sqrt{MN}}\boldsymbol{u}_*\boldsymbol{v}_*^\top + \boldsymbol{W} \in \mathbb{R}^{M\times N}, \tag{1}$$

其中 $\boldsymbol{W} \in \mathbb{R}^{M\times N}$ 是一个旋转不变（RI）噪声矩阵，$\theta \geq 0$ 是信噪比（SNR）参数。尖峰模型是分析特征数量与样本数量处于相近量级的数据的有力工具，在金融数据分析[1],和社区检测 [2]等领域有着显著应用。

对于具有独立同分布高斯噪声的式（1）中的尖峰模型，已有公认的相变结果表明，通过主成分分析（PCA）进行信号恢复时存在一个明确的信噪比（SNR）阈值[3, 4]。已知，当且仅当信噪比超过这一临界阈值时，主特征向量才会与底层信号呈非平凡对齐。利用关于信号的先验结构信息（例如稀疏性），可以进一步提升标准主成分分析的性能[5] – [9]。在此背景下，最初为压缩感知 [10] – [14], 开发的近似消息传递（AMP）算法，在广义一阶方法中能够实现可证明的最优性能[15], , 并且在某些情况下达到统计物理学所预测的最小均方误差（MMSE） [16]– [19]。

对于具有旋转不变（RI）噪声的矩形尖峰模型，主成分分析相变也已得到充分理解 [20], , 但在先验分布可用时，其性能通常并非最优。为此，人们提出了适用于 RI 模型的 AMP 类型算法（此处称为 RI-AMP）[21, 22], , 不过其均方误差性能最优性尚未得到确立 [21]。对于对称对应版本（1）， [23] 取得了重大进展。基于 OAMP框架 [12] ，并受到 [24],[23] 的见解启发，研究者证明，在一大类算法中实现最优性，不仅需要利用先验信息进行迭代去噪，还需要对观测矩阵进行预处理 $\boldsymbol{Y}$，而该预处理基于噪声矩阵的谱。这种预处理与协方差矩阵收缩和矩阵去噪领域的经典研究成果自然相联系[25]—— [28]。此外，[23] 中的 OAMP 算法被证明在一大类算法中性能最优，并且在一定条件下与副本预测相符[24, 29]。

在本文中，我们将 OAMP框架 [23] 扩展到矩形尖峰矩阵模型。我们针对所提出的 OAMP 算法建立了一个易于处理的状态演化（SE），并推导出相应的最优去噪器。进一步地，我们证明 OAMP 算法的不动点方程在独立同分布高斯情形 [15], 下可恢复已知结果，从而支持其潜在最优性的猜想。

我们的主要结果的详细分析和完整证明见本文的扩展版本 [30]。

1.1. 记号与假设。

设 $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K)$ 为一组向量，其中每个向量 $\boldsymbol{x}_k = (x_{k,1}, \ldots, x_{k,d})^\top \in \mathbb{R}^d$。该向量组以 2-瓦瑟斯坦距离收敛到随机变量组 $(\mathrm{X}_1, \ldots, \mathrm{X}_K)$，记作 $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K) \xrightarrow{\mathcal{W}_2} (\mathrm{X}_1, \ldots, \mathrm{X}_K)$，其定义为：对于所有阶数为 $h \quad \mathbb{R}^K \to$ 的伪利普 $\mathbb{R} \qquad 2,$ $\qquad\qquad d\to\infty \frac{1}{d}\sum_{i=1}^{d} h\ x_{1\,i}\ x_{2\,i}\ \ldots\ x_{K\,i}$ 希茨测试函数，几乎处处成立极限 $\lim(,,,,)$ $\xrightarrow{a.s.}(\mathrm{X}\,\mathrm{X}\,,,\mathrm{X}\,)$。 $\mathbb{E}[h \quad _1, \ _2 \ \cdots \quad _K\ ]$

对于定义在 $\mathbb{R}$ 上的任意有限测度 $\chi(\lambda)$ 以及任意有界的博雷尔可测函数 $p: \mathbb{R} \mapsto \mathbb{R}$，相对于该测度的内积记作 $\langle p(\lambda)\rangle_\chi$,$\int p(\lambda)d\chi(\lambda)$。有限测度 μ 的斯蒂尔切斯变换定义为

$$\mathcal{S}_\mu(z) \triangleq \int \frac{1}{z-\lambda}\mathrm{d}\mu(\lambda), \quad \forall z \in \mathbb{C}\setminus \mathrm{supp}(\mu).$$

本文始终采用以下假设。

假设1（信号与噪声模型）。

(a) 我们考虑高维极限，其中 M、$N \to \infty$ 的长宽比为 $M/N \to \delta \in (0, 1]$。 (b) 信号及其侧信息分别按 $(\boldsymbol{u}_*, \boldsymbol{a}) \xrightarrow{\mathcal{W}_2} \pi_U$ 和 $(\boldsymbol{v}_*, \boldsymbol{b}) \xrightarrow{\mathcal{W}_2} \pi_V$ 收敛，并被归一化为单位方差，使得 $\|\boldsymbol{u}_*\|^2/M \xrightarrow{a.s.} 1$ 和 $\|\boldsymbol{v}_*\|^2/N \xrightarrow{a.s.} 1$。

(c) 噪声矩阵定义为 $\boldsymbol{W}=\boldsymbol{U}\mathrm{diag}(\boldsymbol{\sigma})\boldsymbol{V}^{\mathsf{T}}\in\mathbb{R}^{M\times N}$，其中 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 是相互独立且服从 Haar 分布的正交矩阵。我们假设，对于某个与维度无关的常数 C，有 $\|\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}\|_{\mathrm{op}}\le C$。矩阵 $\boldsymbol{W}\boldsymbol{W}^{\mathsf{T}}$ 的经验谱分布弱收敛于确定性测度 μ，该测度在紧支撑上具有绝对连续且 H¨older 连续的密度。因此，$\boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}$ 的极限谱测度为 $\tilde{\mu}$，$\delta\mu+(1-\delta)\delta_{\{0\}}$。

2. 谱分析预备知识

我们首先介绍一些谱分析结果，这些结果将在本文的主要结果（第3节）中用到。

定义1（信号–特征空间谱测度）。令 $(\lambda_i(\cdot)\ \boldsymbol{u}_i(\cdot))$ 表示一个矩阵的特征值–特征向量对。

(a) 定义以下加权经验谱分布：

$$\nu_{M,1}\triangleq\frac{1}{M}\sum_{i=1}^{M}\langle\boldsymbol{u}_i(\boldsymbol{Y}\boldsymbol{Y}^{\top}),\boldsymbol{u}_*\rangle^2\,\delta_{\lambda_i(\boldsymbol{Y}\boldsymbol{Y}^{\top})},$$
$$\nu_{N,2}\triangleq\frac{1}{N}\sum_{i=1}^{N}\langle\boldsymbol{u}_i(\boldsymbol{Y}^{\top}\boldsymbol{Y}),\boldsymbol{v}_*\rangle^2\,\delta_{\lambda_i(\boldsymbol{Y}^{\top}\boldsymbol{Y})}.$$

(b) 考虑 $\boldsymbol{Y}$ 的对称扩张：

$$\widehat{\boldsymbol{Y}}\triangleq\begin{bmatrix}\boldsymbol{0}&\boldsymbol{Y}\\\boldsymbol{Y}^{\top}&\boldsymbol{0}\end{bmatrix}\in\mathbb{R}^{L\times L},\qquad L\triangleq M+N.$$

定义 $\widehat{\boldsymbol{u}}_*$、$[\boldsymbol{u}_*^{\top},\boldsymbol{0}^{\top}]^{\top}$ 和 $\widehat{\boldsymbol{v}}_*$、$[\boldsymbol{0}^{\top},\boldsymbol{v}_*^{\top}]^{\top}$。定义相应的经验有符号谱测度：

$$\nu_{L,3}\triangleq\frac{1}{L}\sum_{i=1}^{L}\langle\boldsymbol{u}_i(\widehat{\boldsymbol{Y}}),\widehat{\boldsymbol{u}}_*\rangle\langle\boldsymbol{u}_i(\widehat{\boldsymbol{Y}}),\widehat{\boldsymbol{v}}_*\rangle\,\delta_{\lambda_i(\widehat{\boldsymbol{Y}})}.$$

与上述经验测度的极限相关的三个收缩函数 φ_1、φ_2 和 φ_3，对于我们的 OAMP 算法至关重要。为陈述结果，定义：

$$\mathcal{C}(z)\triangleq z\mathcal{S}_\mu(z)\left(\delta\mathcal{S}_\mu(z)+(1-\delta)/z\right),\quad\forall z\in\mathbb{C}\backslash\mathrm{supp}(\mu),\tag{2}$$

其中 $\mathcal{S}_\mu$ 是噪声矩阵 μ的极限谱的斯蒂尔切斯变换；参见假设 1。函数 φ_i 定义如下：

$$\varphi_1(\lambda)\triangleq\frac{1+\delta\theta^2\pi^2\lambda\left(\mathcal{H}_\mu(\lambda)^2+\mu(\lambda)^2\right)}{\lim_{\epsilon\to0^+}|1-\theta^2\mathcal{C}(\lambda-i\epsilon)|^2},\tag{3a}$$
$$\varphi_3(\lambda)\triangleq\frac{\theta\left(1-\delta+2\delta\pi\lambda\mathcal{H}_\mu(\lambda)\right)}{\lim_{\epsilon\to0^+}|1-\theta^2\mathcal{C}(\lambda-i\epsilon)|^2}\cdot\mathbf{1}_{\{\lambda\neq0\}},\tag{3b}$$
$$\varphi_2(\lambda)\triangleq\begin{cases}\delta\varphi_1(\lambda)+\frac{\theta(1-\delta)}{\lambda}\varphi_3(\lambda), & \text{if }\lambda>0,\\ \frac{\delta}{1-\theta^2(1-\delta)\pi\mathcal{H}(0)}, & \text{if }\lambda=0,\end{cases}\tag{3c}$$

其中，$\mu(\lambda)$ 表示测度 μ 的密度函数，$\mathcal{H}_\mu$ 表示其希尔伯特变换：

$$\mathcal{H}_\mu(x)\triangleq\frac{1}{\pi}\mathrm{P.V.}\int\frac{1}{x-\lambda}\mu(\lambda)\,\mathrm{d}\lambda.\tag{4}$$

通过应用索霍茨基–普莱梅利公式 [31]，可以利用 $\mu(\lambda)$ 及其希尔伯特变换显式地表示函数 φ_i：

$$\begin{aligned}&\lim_{\epsilon\to0^+}|1-\theta^2\mathcal{C}(\lambda-i\epsilon)|^2\\&=\left\{1-\theta^2\left[(1-\delta)\pi\mathcal{H}_\mu(\lambda)-\delta\pi^2\lambda(\mu^2(\lambda)-\mathcal{H}_\mu^2(\lambda))\right]\right\}^2\\&\quad+\left\{\pi\theta^2\mu(\lambda)\left[1-\delta+2\delta\pi\lambda\mathcal{H}_\mu(\lambda)\right]\right\}^2.\end{aligned}\tag{5}$$

下面的引理刻画了定义 1中谱测度的弱极限。关于对称情形的类似结论见 [23, 引理 1]。

引理 1。满足以下结论：

(a) 定义 $\nu_{M,1}$、$\nu_{N,2}$、$\nu_{L,3}$ 中的测度在几乎处处意义下弱收敛至确定性的紧支撑测度 ν_1、ν_2 和 ν_3。此外，ν_1 和 ν_2 是 $\mathbb{R}_+$上的概率测度，而 ν_3 是 $\mathbb{R}$上的符号测度。

(b) 对于 $z\in\mathbb{C}\setminus\mathbb{R}$，$\nu_1$、$\nu_2$ 和 ν_3 的斯蒂尔切斯变换为：

$$\mathcal{S}_{\nu_1}(z)=\frac{\mathcal{S}_\mu(z)}{1-\theta^2\mathcal{C}(z)},\quad\mathcal{S}_{\nu_2}(z)=\frac{\delta\mathcal{S}_\mu(z)+\frac{1-\delta}{z}}{1-\theta^2\mathcal{C}(z)},\tag{6}$$
$$\mathcal{S}_{\nu_3}(z)=\frac{\sqrt{\delta}}{1+\delta}\cdot\frac{\theta\mathcal{C}(z^2)}{1-\theta^2\mathcal{C}(z^2)}.\tag{7}$$

(c) 令 $\nu_i=\nu_i^{\parallel}+\nu_i^{\perp}$ 为 ν_i 的勒贝格分解，其中包含绝对连续部分 $\nu_i^{\parallel}$ 和奇异部分 $\nu_i^{\perp}$，对于 $i\in\{1,2,3\}$。绝对连续部分的密度为：

$$\frac{d\nu_i^{\parallel}}{d\lambda}=\mu(\lambda)\varphi_i(\lambda),\quad\textit{for } i\in\{1,2\},\tag{8}$$
$$\frac{d\nu_3^{\parallel}}{d\sigma}=\frac{\sqrt{\delta}}{1+\delta}\mathrm{sign}(\sigma)\cdot\mu(\sigma^2)\varphi_3(\sigma^2).\tag{9}$$

(d) 这些测度在根 o 处有点质量 f

$$1-\theta^2\mathcal{C}(\lambda)=0,\quad\textit{for }\lambda\in\mathbb{R}\backslash\mathrm{supp}(\mu).\tag{10}$$

令 $\lambda_*>0$ 为这样的非负根。相应的质量为：

$$\nu_1(\{\lambda_*\})=\frac{\mathcal{S}_\mu(\lambda_*)}{-\theta^2\mathcal{C}'(\lambda_*)},\quad\nu_2(\{\lambda_*\})=\frac{\delta\mathcal{S}_\mu(\lambda_*)+\frac{1-\delta}{\lambda_*}}{-\theta^2\mathcal{C}'(\lambda_*)},$$
$$\nu_3(\{\pm\sigma_*\})=\mp\frac{\sqrt{\delta}}{1+\delta}\frac{1}{2\theta^3\sigma_*\mathcal{C}'(\sigma_*^2)},\quad\forall\sigma_*=\sqrt{\lambda_*}>0.$$

此外，对于 $\delta<1$，ν_2 在原点处有一个点质量：

$$\nu_2(\{0\})=\frac{1-\delta}{1-\theta^2(1-\delta)\pi\mathcal{H}(0)}.\tag{11}$$

3. ORTHOGONALAMP算法

本节介绍主要结果。

3.1. OAMP 算法

我们为矩形模型（1）提出了一类正交近似消息传递（OAMP）算法。这类算法扩展了针对对称尖峰矩阵模型提出的 OAMP 算法 [23]。

定义2（OAMP算法）。OAMP 算法通过以下更新规则生成迭代量序列 $\{\boldsymbol{u}_t\}$ 和 $\{\boldsymbol{v}_t\}$：

$$\begin{aligned}\boldsymbol{u}_t &= F_t(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}})f_t(\boldsymbol{u}_{\le t-1};\boldsymbol{a}) + \widetilde{F}_t(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}})\boldsymbol{Y}g_t(\boldsymbol{v}_{\le t-1};\boldsymbol{b}), \quad (12)\\ \boldsymbol{v}_t &= G_t(\boldsymbol{Y}^{\mathsf{T}}\boldsymbol{Y})g_t(\boldsymbol{v}_{\le t-1};\boldsymbol{b}) + \widetilde{G}_t(\boldsymbol{Y}^{\mathsf{T}}\boldsymbol{Y})\boldsymbol{Y}^{\mathsf{T}}f_t(\boldsymbol{u}_{\le t-1};\boldsymbol{a}),\end{aligned}$$

其中，f_t, g_t 是逐元素迭代去噪器，依赖于过去迭代量的历史记录 $\boldsymbol{u}_{\le t-1}, (\boldsymbol{u}_1,\dots,\boldsymbol{u}_{t-1})$，$\boldsymbol{v}_{\le t-1}, (\boldsymbol{v}_1,\dots,\boldsymbol{v}_{t-1})$ 以及侧信息 $\boldsymbol{a}, \boldsymbol{b}$。矩阵去噪器 F_t, F_t, G_t, G_t 是作用于特征值、且不改变特征向量的标量函数。例如，设 $\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}}$ 的特征分解为 $\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}} = \boldsymbol{U}\mathrm{diag}(\lambda_i)\boldsymbol{U}^{\mathsf{T}}$，则 $\tilde{F}_t$ 的作用定义为 $\boldsymbol{U}\mathrm{diag}(\tilde{F}_t(\lambda_i))\boldsymbol{U}^{\mathsf{T}}$。其他矩阵去噪器的定义类似。第 t 次迭代的最终估计值由逐元素后处理函数 $\phi_{u,t}$,$\phi_{v,t}$生成：

$$\hat{\boldsymbol{u}}_t = \phi_{u,t}(\boldsymbol{u}_{\le t};\boldsymbol{a}), \quad \hat{\boldsymbol{v}}_t = \phi_{v,t}(\boldsymbol{v}_{\le t};\boldsymbol{b}). \tag{13}$$

关键的是，矩阵去噪器 F_t 和 G_t 必须满足以下意义上的迹为零：

$$\begin{aligned}\lim_{M\to\infty}\frac{1}{M}\mathrm{Tr}(F_t(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}})) &\xrightarrow{a.s.} \langle F_t(\lambda)\rangle_{\mu} = 0, \quad (14)\\ \lim_{N\to\infty}\frac{1}{N}\mathrm{Tr}(G_t(\boldsymbol{Y}^{\mathsf{T}}\boldsymbol{Y})) &\xrightarrow{a.s.} \langle G_t(\lambda)\rangle_{\tilde{\mu}} = 0,\end{aligned}$$

其中，内积分别是相对于极限谱测度 μ 和 $\tilde{\mu}$ 定义的。迭代去噪器必须满足散度为零：

$$\begin{aligned}\mathbb{E}[\partial_s f_t(\mathsf{U}_{\le t-1};\mathsf{A})] = 0, \quad \forall s\in\{1,\dots,t-1\}, \quad (15)\\ \mathbb{E}[\partial_s g_t(\mathsf{V}_{\le t-1};\mathsf{B})] = 0, \quad \forall s\in\{1,\cdots t-1\}, \quad (16)\end{aligned}$$

其中，$(\mathsf{U}_{\le t-1},\mathsf{V}_{\le t-1})$ 是将在下一节中定义的状态演化随机变量。请注意，状态演化随机变量和迭代去噪器是以递归方式定义的；相关讨论见 [23, 备注 2] 。

3.2. 状态演化

OAMP 的状态演化（SE）理论预测，迭代量与信号的经验联合分布将（在 2-瓦瑟斯坦距离意义下）收敛到一列标量随机变量；这些变量通过高斯信道演化，并由真实信号生成的相互独立的 $\mathsf{Z}_{u,t}$、$\mathsf{Z}_{v,t} \sim \mathcal{N}(0,1)$ 驱动：

$$\begin{aligned}(\mathsf{U}_*,\mathsf{A}) &\sim \pi_U, \quad \mathsf{U}_t = \mu_{u,t}\mathsf{U}_* + \sigma_{u,t}\mathsf{Z}_{u,t}, \quad (17)\\ (\mathsf{V}_*,\mathsf{B}) &\sim \pi_V, \quad \mathsf{V}_t = \mu_{v,t}\mathsf{V}_* + \sigma_{v,t}\mathsf{Z}_{v,t}, \quad (18)\end{aligned}$$

其中，每次迭代中去噪器的输出分别建模为 $\mathsf{F}_t = f_t(\mathsf{U}_{\le t-1};\mathsf{A})$ 和 $\mathsf{G}_t = g_t(\mathsf{V}_{\le t-1};\mathsf{B})$。这些信道的参数会在每次迭代 $t\in\mathbb{N}$ 中递归更新：

$$\begin{aligned}\mu_{u,t} &= \alpha_t\langle F_t(\lambda)\rangle_{\nu_1} + \beta_t(1+\delta^{-1})\langle\sigma\widetilde{F}_t(\sigma^2)\rangle_{\nu_3}, \quad (19)\\ \mu_{v,t} &= \beta_t\langle G_t(\lambda)\rangle_{\nu_2} + \alpha_t(1+\delta)\langle\sigma\widetilde{G}_t(\sigma^2)\rangle_{\nu_3},\\ \sigma^2_{u,t} &= \alpha_t^2\langle F_t^2(\lambda)\rangle_{\nu_1} + \beta_t^2\delta^{-1}\langle\lambda\widetilde{F}_t^2(\lambda)\rangle_{\nu_2} + \sigma^2_{f,t}\langle F_t^2(\lambda)\rangle_{\mu}\\ &\quad + 2\alpha_t\beta_t(1+\delta^{-1})\langle\sigma F_t(\sigma^2)\widetilde{F}_t(\sigma^2)\rangle_{\nu_3} - \mu^2_{u,t}\\ &\quad + \sigma^2_{g,t}\delta^{-1}\langle\lambda\widetilde{F}_t^2(\lambda)\rangle_{\tilde{\mu}},\\ \sigma^2_{v,t} &= \beta_t^2\langle G_t^2(\lambda)\rangle_{\nu_2} + \alpha_t^2\delta\langle\lambda\widetilde{G}_t^2(\lambda)\rangle_{\nu_1} + \sigma^2_{f,t}\delta\langle\lambda\widetilde{G}_t^2(\lambda)\rangle_{\mu}\\ &\quad + 2\alpha_t\beta_t(1+\delta)\langle\sigma G_t(\sigma^2)\widetilde{G}_t(\sigma^2)\rangle_{\nu_3} - \mu^2_{v,t}\\ &\quad + \sigma^2_{g,t}\langle G_t^2(\lambda)\rangle_{\tilde{\mu}},\end{aligned}$$

其中，信号对齐度 α_t、β_t 以及去噪器输出的残差方差 $\sigma^2_{f,t}$、$\sigma^2_{g,t}$ 定义如下：

$$\begin{aligned}\alpha_t &\triangleq \mathbb{E}[\mathsf{U}_*\mathsf{F}_t], \qquad \beta_t \triangleq \mathbb{E}[\mathsf{V}_*\mathsf{G}_t], \quad (20)\\ \sigma^2_{f,t} &\triangleq \mathbb{E}[\mathsf{F}_t^2] - \alpha_t^2, \qquad \sigma^2_{g,t} \triangleq \mathbb{E}[\mathsf{G}_t^2] - \beta_t^2. \quad (21)\end{aligned}$$

在上述表达式中，内积是相对于极限谱测度 ν_1，ν_2，ν_3定义的，其刻画见引理 1。

下面的定理刻画了 OAMP 算法的高维动态；其证明见本文完整版的 [30, 附录B和C]。

定理 1（状态演化）。在假设 *1*下，定义 *2* 中的 OAMP 算法的迭代量（其初始化迭代量与 $\boldsymbol{W}$相互独立）在 $\mathcal{W}_2$ 距离下，对于任意有限的 $t\in\mathbb{N}$，均收敛于（17）-（18）中的状态演化随机变量：

$$\begin{aligned}(\boldsymbol{u}_*,\boldsymbol{u}_1,\dots,\boldsymbol{u}_t;\boldsymbol{a}) &\xrightarrow{\mathcal{W}_2} (\mathsf{U}_*,\mathsf{U}_1,\dots,\mathsf{U}_t;\mathsf{A}), \quad (22)\\ (\boldsymbol{v}_*,\boldsymbol{v}_1,\dots,\boldsymbol{v}_t;\boldsymbol{b}) &\xrightarrow{\mathcal{W}_2} (\mathsf{V}_*,\mathsf{V}_1,\dots,\mathsf{V}_t;\mathsf{B}), \quad (23)\end{aligned}$$

其中，状态演化随机变量 $(\mathsf{U}_t、\mathsf{V}_t)_{t\ge1}$ 见（17）-（21）中的描述。

3.3. OptimalOAMP算法

可以直接根据状态演化推导出最优 OAMP 算法。详细推导见 [30, 附录D]；此处仅给出最终形式。

所得的最优 OAMP 算法如下：

$$\begin{aligned}\boldsymbol{u}_t^* &= \frac{1}{\sqrt{w_{1,t}}}\big[\boldsymbol{F}_t^*\bar{\phi}(\boldsymbol{u}_{t-1}^*|w_{1,t-1}) + \widetilde{\boldsymbol{F}}_t^*\boldsymbol{Y}\bar{\phi}(\boldsymbol{v}_{t-1}^*|w_{2,t-1})\big],\\ \boldsymbol{v}_t^* &= \frac{1}{\sqrt{w_{2,t}}}\big[\boldsymbol{G}_t^*\bar{\phi}(\boldsymbol{v}_{t-1}^*|w_{2,t-1}) + \widetilde{\boldsymbol{G}}_t^*\boldsymbol{Y}^{\mathsf{T}}\bar{\phi}(\boldsymbol{u}_{t-1}^*|w_{1,t-1})\big].\end{aligned}$$

更新矩阵为 $\boldsymbol{F}_t^*$、$F_t^*(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}})$ 以及 $\widetilde{\boldsymbol{F}}_t^*$、$\tilde{F}_t^*(\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{T}})$，而 $\boldsymbol{G}_t^*$ 和 $\widetilde{\boldsymbol{G}}_t^*$ 应用于 $\boldsymbol{Y}^{\mathsf{T}}\boldsymbol{Y}$ 时的定义类似。最终估计值如下：

$$\hat{\boldsymbol{u}}_t^* \triangleq \phi(\boldsymbol{u}_t^*|w_{1,t}), \quad \hat{\boldsymbol{v}}_t^* \triangleq \phi(\boldsymbol{v}_t^*|w_{2,t}). \tag{24}$$

算法的组成部分定义如下：

1. 迭代去噪器：对于标量高斯信道 $\mathsf{X} = \sqrt{\omega}\mathsf{X}_* + \sqrt{1-\omega}\mathsf{Z}$，信号 X_* 在侧信息 C 已知时的 MMSE 估计器为 $\phi(x;\ c|\omega) \quad \mathbb{E}[\mathsf{X}_*|\mathsf{X}=x,\ \mathsf{C}=c]$。根据 散度为零条件（15），DMMSE 估计器 ϕ由下式给出 [23]:

$$\bar{\phi}(x;c|\omega) \triangleq \frac{\phi(x;c|\omega) - \frac{\mathbb{E}[Z\phi(X;C|\omega)]}{\sqrt{1-\omega}}x}{1 - \frac{\sqrt{\omega}}{\sqrt{1-\omega}}\mathbb{E}[Z\phi(X;C|\omega)]}, \quad \forall \omega \in (0,1). \tag{25}$$

2. 矩阵去噪器：无迹去噪器为：

$$F_t^*(\lambda) \triangleq (1+\rho_{1,t}^{-1}) \cdot \left[1 - \langle P_t^*(\lambda)\rangle_\mu^{-1} \cdot P_t^*(\lambda)\right],$$
$$\widetilde{F}_t^*(\lambda) \triangleq (1+\rho_{2,t}^{-1}) \cdot \left[\langle P_t^*(\lambda)\rangle_\mu^{-1} \cdot \widetilde{P}_t^*(\lambda)\right], \tag{26a}$$
$$G_t^*(\lambda) \triangleq (1+\rho_{2,t}^{-1}) \cdot \left[1 - \langle Q_t^*(\lambda)\rangle_{\tilde{\mu}}^{-1} \cdot Q_t^*(\lambda)\right],$$
$$\widetilde{G}_t^*(\lambda) \triangleq (1+\rho_{1,t}^{-1}) \cdot \left[\langle Q_t^*(\lambda)\rangle_{\tilde{\mu}}^{-1} \cdot \widetilde{Q}_t^*(\lambda)\right]. \tag{26b}$$

其中

$$P_t^*(\lambda;\rho_{1,t},\rho_{2,t}) \triangleq \frac{\lambda(\rho_{2,t}\varphi_2(\lambda)+\delta)}{D_t(\lambda)}, \tag{26c}$$
$$\widetilde{P}_t^*(\lambda;\rho_{1,t},\rho_{2,t}) \triangleq \frac{\sqrt{\delta}\rho_{2,t}\varphi_3(\lambda)}{D_t(\lambda)}, \tag{26d}$$
$$Q_t^*(\lambda;\rho_{1,t},\rho_{2,t}) \triangleq \frac{\delta\lambda(\rho_{1,t}\varphi_1(\lambda)+1)}{D_t(\lambda)}, \tag{26e}$$
$$\widetilde{Q}_t^*(\lambda;\rho_{1,t},\rho_{2,t}) \triangleq \frac{\sqrt{\delta}\rho_{1,t}\varphi_3(\lambda)}{D_t(\lambda)}, \tag{26f}$$

and

$$D_t(\lambda) \triangleq (\rho_{1,t}\varphi_1(\lambda)+1)(\rho_{2,t}\varphi_2(\lambda)+\delta)\lambda - \rho_{1,t}\rho_{2,t}\varphi_3^2(\lambda).$$

The shrinkage functions $\varphi_1, \varphi_2, \varphi_3$ are defined in (3).

3.4. 最优 OAMP 的状态演化

最优 OAMP 算法的均方误差性能可简化为以下形式。

命题 1. 对于最优 OAMP 算法，$\|\hat{\boldsymbol{u}}_t - \boldsymbol{u}_*\|^2/M \xrightarrow{\text{a.s.}}$ 最小均方误差$_{\mathsf{U}}(w_{1,t})$ 和 $\|\hat{\boldsymbol{v}}_t - \boldsymbol{v}_*\|^2/N \xrightarrow{\text{a.s.}}$ 最小均方误差$_{\mathsf{V}}(w_{2,t})$，其中

$$\rho_{1,t} = \frac{1}{\mathrm{mmse}_{\mathsf{U}}(w_{1,t-1})} - \frac{1}{1-w_{1,t-1}},$$
$$\rho_{2,t} = \frac{1}{\mathrm{mmse}_{\mathsf{V}}(w_{2,t-1})} - \frac{1}{1-w_{2,t-1}},$$
$$w_{1,t} = 1 - \frac{1-\langle P_t^*(\lambda)\rangle_\mu}{\langle P_t^*(\lambda)\rangle_\mu} \cdot \frac{1}{\rho_{1,t}},$$
$$w_{2,t} = 1 - \frac{1-\langle Q_t^*(\lambda)\rangle_{\tilde{\mu}}}{\langle Q_t^*(\lambda)\rangle_{\tilde{\mu}}} \cdot \frac{1}{\rho_{2,t}}, \tag{27}$$

其中 最小均方误差$\sqrt{w}$ $_*$+$\sqrt{1-w}$ P_t^* $P^* \cdot \rho_{1\,t}$ ρ_2 () $\mathbb{E}\left[(\mathsf{X}_* - \mathbb{E}[\mathsf{X}_*|\mathsf{X}])^2\right]$ 对于信道 $\mathsf{X}\mathsf{X}\mathsf{Z}$。这里，我们使用了简写 $\triangleq(;,,,t)$ 在 (26c) 中，并类似地用于 Q_t^* 在 (26e)。

现在，我们将结果专门应用于噪声矩阵元素独立同分布且服从高斯分布的情形。在这种情况下，最优 OAMP 算法的不动点方程与标准 AMP 算法的相同，[15], 只需进行重新参数化。

命题 2. 假设 $\boldsymbol{W}$ 具有独立同分布的 $\mathcal{N}(0,1/N)$ 元素。那么，(27) 中的不动点方程可简化为：

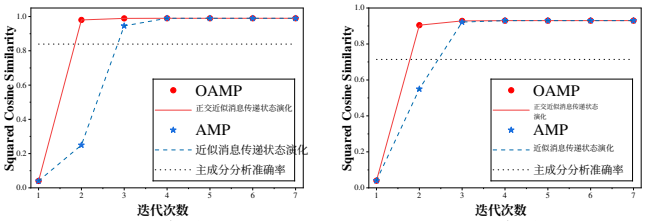
$$\mathrm{mmse}_{\mathrm{U}}(w_1) = 1 - \frac{1}{\theta^2}\frac{w_2}{1-w_2},$$
$$\mathrm{mmse}_{\mathrm{V}}(w_2) = 1 - \frac{\delta}{\theta^2}\frac{w_1}{1-w_1}.$$

命题1 和2 的证明见 [30]。

4. 仿真结果

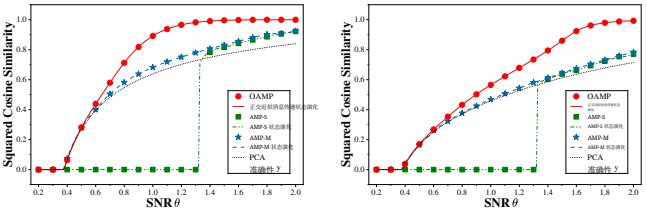
4.1. 独立同分布高斯噪声

我们先考察独立同分布高斯噪声下的模拟。真实信号 $\boldsymbol{u}_*$ 和 $\boldsymbol{v}_*$ 的先验分布均为独立同分布的拉德马赫分布。图1 展示了主成分分析、近似消息传递和所提出的 OAMP 算法在平方余弦相似度方面的表现。结果表明，OAMP 的实测性能与其状态演化预测高度吻合。此外，所提出的 OAMP 算法收敛至与近似消息传递相同的不动点，这与命题2一致。



(a) $\boldsymbol{u}$-信道　(b) $\boldsymbol{v}$-信道

图 1：独立同分布高斯噪声模型的仿真结果。经验结果（标记）是对 50 次运行取平均得到的，其中 $M = 4000, N = 8000$，且 $\theta = 2$。对于两个信号，算法均以 0.2 的先验余弦相似度初始化。



(a) $\boldsymbol{u}$-信道　(b) $\boldsymbol{v}$-信道

图 2：非高斯噪声情形的仿真结果。的特征值 $\boldsymbol{W}\boldsymbol{W}^\top$ 服从经缩放/平移的 Beta(1 5, 1 5 . .) 分布，其支撑集为 $[1,3]$。信号分量服从拉德马赫先验。其他设置与图 1相同。

4.2. 非高斯噪声

对于旋转不变噪声模型，我们将所提出的 OAMP 与 RI-AMP 进行比较 [21]。我们评估了 RI-AMP 的两种变体：单次迭代版本（AMP-S）和多次迭代版本（AMP-M）。在 AMP-S 中，去噪器是应用于当前迭代的 MMSE函数，而在 AMP-M 中，去噪器在应用 MMSE 去噪器之前，利用协方差最优地组合所有过去的信号加噪声观测（见[21, 备注 3。3] 详情）。图2 表明，最优 OAMP 算法始终优于主成分分析以及两种 RI-AMP 变体。

致谢

本工作得到国家重点研发计划（项目编号：2024YFA1014200）和国家自然科学基金（项目编号：62571526）的资助。

5. 参考文献

[1] J.-P. Bouchaud 和 M. Potters，“随机矩阵理论的金融应用：简要综述”，载于《牛津随机矩阵理论手册》。牛津大学出版社，2015 年。[2] E. Abbe、A. S. Bandeira 和 G. Hall，“随机块模型中的精确恢复”，《IEEE 信息论汇刊》，第 62 卷，第 1 期，第 471–487 页，2016 年。[3] D. F´eral 和 S. P´ech´e，“大型 Wigner 矩阵秩一形变的最大特征值”，《数学物理通讯》，第 272 卷，第 185–228 页，2007 年。[4] Z. Bai 和 J. W. Silverstein，《大维随机矩阵的谱分析》，施普林格，纽约，第 2 版，2010 年。[5] H. Zou、T. Hastie 和 R. Tibshirani，“稀疏主成分分析”，《计算与图形统计学杂志》，第 15 卷，第 2 期，第 265–286 页，2006 年。[6] Y. Deshpande 和 A. Montanari，“信息论最优稀疏主成分分析”，载于 2014 年 IEEE 信息论国际研讨会。2014 年，第 2197–2201 页，IEEE。[7] T. Lesieur、F. Krzakala 和 L. Zdeborov´a，“稀疏主成分分析中的相变”，载于 2015 年 IEEE 信息论国际研讨会。2015 年，第 1635–1639 页，IEEE。[8] A. Montanari 和 E. Richard，“非负主成分分析：消息传递算法与精确渐近分析”，《IEEE 信息论汇刊》，第 62 卷，第 1458–1484 页，2016 年。[9] T. Lesieur、F. Krzakala 和 L. Zdeborov´a，“约束低秩矩阵估计：相变、近似消息传递及其应用”，《统计力学杂志：理论与实验》，第 2017 卷，第 073403 页，2017 年。[10] D. L. Donoho、A. Maleki 和 A. Montanari，“压缩感知的消息传递算法”，《美国国家科学院院刊》，第 106 卷，第 45 期，第 18914–18919 页，2009 年。[11] M. Bayati 和 A. Montanari，“稠密图上的消息传递动态及其在压缩感知中的应用”，《IEEE 信息论汇刊》，第 57 卷，第 2 期，第 764–785 页，2011 年。[12] J. Ma 和 Li Ping，“正交AMP”，《IEEE Access》，第 5 卷，第 2020–2033 页，2017 年。[13] S. Rangan、P. Schniter 和 A. K. Fletcher，“向量近似消息传递”，《IEEE 信息论汇刊》，第 65 卷，第 10 期，第 6664–6684 页，2019 年。[14] K. Takeuchi，“基于期望传播的酉不变测量信号恢复的严格动态分析”，《IEEE 信息论汇刊》，第 66 卷，第 368–386 页，2020 年。[15] A. Montanari 和 R. Venkataramanan，“通过近似消息传递估计低秩矩阵”，《统计学年刊》，第 49 卷，第 1 期，第 321–345 页，2021 年。[16] J. Barbier、M. Dia、N. Macris、F. Krzakala、T. Lesieur 和 L. Zdeborov´a，“对称秩一矩阵估计的互信息：副本公式的证明”，载于第 30 届神经信息处理系统国际会议论文集

, 2016, 第 424–432 页。

[17]L. Miolane，《低秩矩阵估计的基本极限：非对称情形》，arXiv 预印本 arXiv:1702.00473，2018 年。[18] A. E. Alaoui 和 F. Krzakala，《尖峰 Wigner 模型中的估计：副本公式的简短证明》，2018 年 IEEE 信息论国际研讨会，第 1874–1878 页，2018 年。[19] M. Lelarge 和 L. Miolane，《对称低秩矩阵估计的基本极限》，《概率论及其相关领域》，第 173 卷，第 3 期，第 859–929 页，2019 年。[20] F. Benaych-Georges 和 R. R. Nadakuditi，《大型矩形随机矩阵低秩扰动的奇异值和奇异向量》，《多元分析杂志》，第 111 卷，第 120–135 页，2012 年。[21] Z. Fan，《旋转不变矩阵的近似消息传递算法》，《统计学年刊》，第 50 卷，第 1 期，第 197–224 页，2022 年。[22] X. Zhong 、 T. Wang 和 Z. Fan，《正交不变系综的近似消息传递：多元非线性与谱初始化》，《信息与推断：IMA 杂志》，第 13 卷，第 3 期，文章号 iaae024，2024 年 9 月。[23] R. Dudeja、S. Liu 和 J. Ma，《旋转不变噪声尖峰矩阵模型中近似消息传递算法的最优性》，arXiv 预印本 arXiv:2405.18081，2024 年。[24] J. Barbier、F. Camilli、M. Mondelli 和 M. S´aenz，《结构化主成分分析的基本极限及其实现方法》，《美国国家科学院院刊》，第 120 卷，文章号 e2302028120，2023 年。[25] O. Ledoit 和 M. Wolf，《高维协方差矩阵的非线性收缩估计》，《统计学年刊》，第 40 卷，第 1024–1060 页，2012 年。[26] R. R. Nadakuditi，《OptShrink：通过最优数据驱动的奇异值收缩改进低秩信号矩阵去噪的算法》，《IEEE 信息论汇刊》，第 60 卷，第 5 期，第 3002–3018 页，2014 年。[27] J. Bun、R. Allez、J.-P. Bouchaud 和 M. Potters，《一般噪声矩阵的旋转不变估计量》，《IEEE 信息论汇刊》，第 62 卷，7490，2016 12，第 7475 页，期号、页码信息有误。[28] F. Pourkamali 和 N. Macris，《高秩矩阵估计的矩形旋转不变估计量》，《信息与推断：IMA 杂志》，第 14 卷，015，2025 年第 3 期，页码信息为 iaaf。[29] J. Barbier、F. Camilli、Y. Xu 和 M. Mondelli，《结构化噪声尖峰矩阵模型的信息极限与 Thouless-Anderson-Palmer 方程》，《物理评论研究》，第 7 卷，第 1 期，文章号 013081，2025 年。[30] H. Chen、S. Liu 和 J. Ma，《具有最优谱初始化的矩形尖峰矩阵模型正交近似消息传递》，arXiv 预印本 arXiv:2512 19334，2025 年。[31] P. Blanchard 和 E. Br¨uning，《数学物理方法》，《数学物理进展》。Birkh¨auser，第二版，2015 年。