arXiv:2602.05724v1 [eess.SP] 5 Feb 2026

基于MMSE估计的双天线中继器互易性校准

Shoma Hara，*IEEE*研究生会员，Takumi Takahashi，*IEEE*会员，Hiroki Iimori，*IEEE*会员，Hideki Ochiai，*IEEE*会士，Erik G. Larsson，*IEEE*会士

摘要——本文提出一种新颖的贝叶斯互易性校准方法，能够在中继器辅助的多输入多输出（MIMO）系统中持续确保上行链路与下行链路的信道互易性。所提算法依据最小均方误差（MMSE）准则构建。其贝叶斯框架纳入了信号模型、噪声及先验分布的完整统计知识，从而实现兼顾低计算复杂度与高校准精度的一致性设计。为进一步提升对校准任务至关重要的相位对齐精度，我们开发了一种冯·米塞斯去噪器，利用目标参数位于复平面圆周上这一事实。仿真结果表明，所提出的MMSE算法相较传统确定性非线性最小二乘（NLS）方法显著提高了估计精度，同时保持了相当的计算复杂度。此外，所提方法展现出极快的收敛速度，因而非常适合实际实现。

索引词—**互易性校准、双天线中继器、MIMO系统、贝叶斯推断、最小均方误差**

I. 引言

大规模多输入多输出（MIMO）系统依靠时分双工（TDD）信道互易性，使基站（BS）能够以可扩展的方式获取信道状态信息（CSI） [1], [2]。在TDD运行模式下，基站根据导频传输估计上行信道，并将其用于下行预编码；同时利用大型天线阵列，对低功率上行信号进行相干合并。然而，尽管物理传播信道具有互易性，实际收发器硬件却会因发射（TX）和接收（RX）射频（RF）链路之间存在失配而表现出非互易性 [3], ，从而使基站观测到的有效上行和下行信道发生失真。因此，若不进行补偿，便无法直接将基于上行链路的信道估计用于下行预编码。这使互易性校准成为实现精确波束成形、充分发挥基于互易性的大规模MIMO系统性能潜力的基本要求 [4]。

基于时分双工的多输入多输出系统中的互易性校准问题在早期研究 [5]。这一研究方向确立了一种基础校准模型，其中观测到的上行链路和下行链路信道因乘性硬件响应而存在差异，并指出，要实现准确的下行波束成形，必须明确补偿这些非互易效应。基于这一表述，大量研究发展了适用于采用共址天线阵列的传统大规模MIMO系统的互易性校准方法。具有代表性的方法包括利用同一天线阵列内天线之间的内部信号交换或互耦的相对校准技术 [6], ，以及基于双向导频传输的空口方案 [7]。在单跳信道和硬件共址的假设下，这些方法能够估计基站天线单元的相对发射和接收响应，其结果存在一个共同的标量模糊；这已足以支持多用户下行预编码，因为该标量不会影响波束成形方向或干扰抑制。

近期的蜂窝网络架构正从传统的共址基站演进为更加分布式的空间部署，以提升覆盖范围、空间复用能力以及小区边缘的用户体验 [8]–[12]。尽管分布式大规模MIMO系统已得到广泛研究，旨在通过在地理位置上分散天线单元来提升空间分集、覆盖均匀性和小区边缘性能，但其实际部署仍面临挑战。分布式MIMO架构通常依赖高容量、低时延的前传或回传连接来协调多个远端射频单元，这会显著增加部署成本和运营复杂性 [13]。

鉴于这些挑战，中继辅助大规模MIMO已成为一种务实的替代方案，能够在无需采用完全分布式部署架构的情况下实现网络致密化 [14]–[23]。通过在小区内部署体积小、成本低的无线中继器，可以增加有效的散射路径，从而改善覆盖范围和信道秩，同时将所有多天线信号处理集中在基站。与分布式接入点不同，中继器无需专用回传、严格的站间同步或对现有网络架构进行重大改动，因此在部署和维护方面都颇具吸引力。围绕网络控制中继器持续开展的标准化工作，进一步凸显了其作为迈向接近分布式大规模MIMO系统性能水平的现实路径的重要性 [24]。

传统互易性校准方法是在以下假设下开发的：基站与用户设备（UE）之间的信道是单跳传播信道，仅受收发器端点处的发射和接收硬件链路影响。在中继器辅助架构中，这一

S. Hara、T. Takahashi 和 H. Ochiai 任职于大阪大学工学研究科，地址：日本吹田市山田丘 2-1，邮编 565-0871（电子邮箱：s-hara@wcs.comm.eng.osaka-u.ac.jp、｛takahashi、ochiai｝@comm.eng.osaka-u.ac.jp）。H. Iimori 任职于爱立信研究院，地址：日本爱立信日本株式会社，横滨 SYMPHOSTAGE WestTower 12F，港未来 5-1-2，横滨，邮编 220-0012（电子邮箱：hiroki.iimori@ericsson.com）。E. G. Larsson 任职于林雪平大学电气工程系（ISY），地址：瑞典林雪平 581 83（电子邮箱：erik.g.larsson@liu.se）。

本研究的会议版计划投稿至 IEEE SPAWC 2026 的一个特别会议。本研究已提交至 IEEE，供其考虑是否发表。版权可能会在不另行通知的情况下转让，届时本版本可能无法再访问。

假设不再成立。中继器会引入额外的发射和接收链路，这些链路具有各自的模拟损伤、放大增益和相位响应，并与基站—用户设备信道 [25] 级联。因此，有效的端到端信道不仅由传播环境和基站/用户设备硬件构成，还包含中继器未知且可能不对称的正向和反向响应。这种级联非互易性从根本上改变了校准问题的结构，使得针对共址阵列和直连链路设计的标准大规模 MIMO 校准技术不再适用。

直到最近，文献中才开始明确讨论中继器引起的非互易性，尽管早期研究例如 [26]，已经指出，由于中继器具有固有的非互易特性，它们在时分双工系统中可能无法有效工作。为解决这一问题，近期提出了一种校准方法，通过耦合最小二乘（LS）表述 [27]，对中继器引起的硬件不对称性进行显式建模，并将其纳入互易性校准问题。所得估计问题采用非线性最小二乘（NLS）方法和交替非线性最小二乘方法求解，以处理多个未知硬件响应之间的耦合。尽管这一表述是迈向中继器感知互易性校准的重要一步，但所提出的解决方案是在确定性最小二乘框架内构建的，并将中继器引起的硬件损伤视为未知但无结构的量。在实际应用中，可能存在有关这些损伤的先验信息。利用此类先验信息有望显著提升校准的鲁棒性和准确性，尤其是在训练开销有限或信噪比（SNR）较低的情况下。这一观察促使本文提出了校准算法。本文的具体贡献如下：

- 我们针对考虑中继器的互易性校准问题，构建了一种贝叶斯表述，明确纳入中继器引起的非互易性和测量噪声的先验统计模型，从而能够以有原则的方式处理硬件引起的性能损伤。
- 基于这一表述，我们提出了一种新颖的基于最小均方误差（MMSE）的校准算法，利用概率数据关联（PDA）进行迭代贝叶斯双线性推断 [28]–[30]，；其中，最小均方误差更新通过冯·米塞斯去噪器实现，未知先验统计量则通过矩估计法（MoM）估计。
- 通过大量仿真，我们证明，所提出的算法在大规模和低信噪比场景下，相较于最先进的确定性非线性最小二乘方法 [27]，显著提升了鲁棒性和准确性，同时保持了与基本非线性最小二乘方法相当的计算复杂度。

符号说明：实数和复数集合分别记为 $\mathbb{R}$ 和 $\mathbb{C}$。向量和矩阵分别用小写字母和大写字母表示。矩阵 $\boldsymbol{A}$ 的第（i，j）个元素记为 $\boldsymbol{A}$（i，j），$\boldsymbol{A}$ 的第i列记为（$\boldsymbol{A}$）$_i$。$a \times a$ 阶单位矩阵记为 $\boldsymbol{I}_a$。对于标量 a_1，...，a_M，diag（a_1，...，a_M）表示以这些元素为主对角线元素的对角矩阵

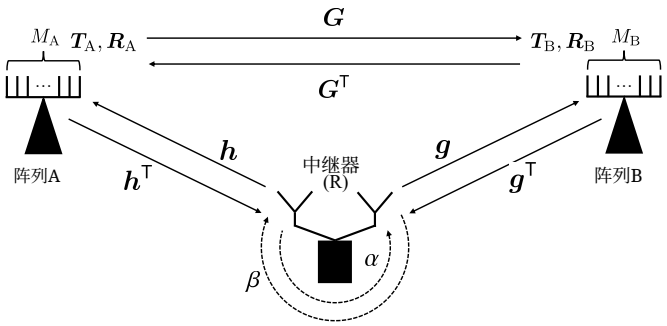


图 1. 两个天线阵列 A 和 B，以及一个中继器（R）。 $\boldsymbol{G}$ 表示 R 关闭时从 A 到 B 的传播信道。无线电信道用实线表示，中继器增益用虚线表示。该图改编自 [27]。

a_1，...，a_M。共轭、转置和共轭转置分别记为（$\cdot$）*、（$\cdot$）$^{\mathsf{T}}$和（$\cdot$）$^{\mathsf{H}}$。 $\boldsymbol{A}$ 的迹和行列式分别记为 tr（$\boldsymbol{A}$）和 det（$\boldsymbol{A}$）。矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的克罗内克积记为 $\boldsymbol{A} \otimes \boldsymbol{B}$。将矩阵 $\boldsymbol{A}$ 的各列依次堆叠，记为 vec（$\boldsymbol{A}$）。符号 $\|\cdot\|_2$ 和 $\|\cdot\|_{\mathrm{F}}$ 分别表示向量的欧几里得范数和矩阵的弗罗贝尼乌斯范数。复数的实部记为 $\Re\{\cdot\}$，虚数单位记为 j $\triangleq \sqrt{-1}$。复数c 的模和相位分别记为 $|c|$ 和 arg（c）。符号$a \sim \mathcal{P}$ 表示随机变量 a 服从概率分布 $\mathcal{P}$。均值为 μ、协方差为 $\boldsymbol{\Lambda}$ 的圆对称复高斯分布记为$\mathcal{CN}$（μ，$\boldsymbol{\Lambda}$）。均值为 μ 、方差为 σ^2 的实高斯分布记为 $\mathcal{N}$（μ，σ^2）。$[a, b$上的均匀分布记为$\mathcal{U}[a, b)$。 x 的概率密度函数（PDF）记为 p_{x}（$\cdot$），给定 y 时x 的条件概率密度函数记为$p_{\mathsf{x}|\mathsf{y}}$（$\cdot|\cdot$）。期望记为 $\mathbb{E}[\cdot]$。最后，符号 $\mathcal{O}$（$\cdot$）表示计算复杂度阶。

II. 系统模型

考虑一个通信场景，其中包含两个天线阵列 A 和 B，以及一个双天线中继器 R，如图 1 [27]所示。阵列 A 和 B 的天线数分别记为 M_{A} 和 M_{B}。实际中，A 和 B 可以分别代表一个基站和一个移动用户终端，也可以代表分布式 MIMO 系统中的两个分布式接入点。

A和B之间的传播信道由直达路径以及经过中继器 R 的间接路径组成。关闭中继器时，A 到 B 的直达信道记为 $\boldsymbol{G} \in \mathbb{C}^{M_{\mathrm{B}} \times M_{\mathrm{A}}}$，根据互易性，B 到 A 的信道记为 $\boldsymbol{G}^{\mathsf{T}}$。我们不对 $\boldsymbol{G}$施加任何结构假设；它可以同时包含视距（LoS）和多径分量。即使 R 未激活，它也可能散射入射的无线电波，从而充当互易无源元件，这种效应已隐含地包含在 $\boldsymbol{G}$ 中。

从 A 到 R 的信道记为 $\boldsymbol{h}^{\mathsf{T}} \in \mathbb{C}^{1 \times M_{\mathrm{A}}}$，从 R 到 A 的反向信道记为 $\boldsymbol{h} \in \mathbb{C}^{M_{\mathrm{A}} \times 1}$。类似地，从 B 到 R 的信道表示为 $\boldsymbol{g}^{\mathsf{T}} \in \mathbb{C}^{1 \times M_{\mathrm{B}}}$，从 R 到 B 的反向信道记为 $\boldsymbol{g} \in \mathbb{C}^{M_{\mathrm{B}} \times 1}$。当中继器工作在其标称

配置下，前向路径和反向路径的复增益分别记为 α 和 β。一般而言，$\alpha \neq \beta$，这意味着该中继器在标称上是非互易的。这种非互易性源于热涨落和硬件非理想性等因素。

此外，阵列 A 和 B 的发射支路与接收支路会因模拟前端的温度漂移和元件长期老化等缺陷而引入硬件导致的非互易性。这种非互易性由四个对角矩阵建模，

$$\boldsymbol{T}_{\mathrm{A}} \triangleq \operatorname{diag}\left(t_{\mathrm{A},1}, t_{\mathrm{A},2}, \ldots, t_{\mathrm{A},M_{\mathrm{A}}}\right), \tag{1a}$$
$$\boldsymbol{R}_{\mathrm{A}} \triangleq \operatorname{diag}\left(r_{\mathrm{A},1}, r_{\mathrm{A},2}, \ldots, r_{\mathrm{A},M_{\mathrm{A}}}\right), \tag{1b}$$
$$\boldsymbol{T}_{\mathrm{B}} \triangleq \operatorname{diag}\left(t_{\mathrm{B},1}, t_{\mathrm{B},2}, \ldots, t_{\mathrm{B},M_{\mathrm{B}}}\right), \tag{1c}$$
$$\boldsymbol{R}_{\mathrm{B}} \triangleq \operatorname{diag}\left(r_{\mathrm{B},1}, r_{\mathrm{B},2}, \ldots, r_{\mathrm{B},M_{\mathrm{B}}}\right), \tag{1d}$$

其对角元素表示各个天线单元的复数互易性系数，其中包括 A 和 B 处振荡器可能发生的漂移。

根据上述定义，从 A 到 B 的有效无噪声信道可表示为

$$\boldsymbol{R}_{\mathrm{B}}\left(\boldsymbol{G}+\alpha \boldsymbol{g}\boldsymbol{h}^{\mathsf{T}}\right)\boldsymbol{T}_{\mathrm{A}}, \tag{2}$$

而从 B 到 A 的信道则为

$$\boldsymbol{R}_{\mathrm{A}}\left(\boldsymbol{G}^{\mathsf{T}}+\beta \boldsymbol{h}\boldsymbol{g}^{\mathsf{T}}\right)\boldsymbol{T}_{\mathrm{B}}. \tag{3}$$

在每个表达式中，第一项表示直达传播信道，第二项则表示中继器辅助路径。

为保持一般性，假设阵列A和阵列B均未进行预校准。因此，所有参数

$$\alpha,\ \beta,\ \boldsymbol{G},\ \boldsymbol{g},\ \boldsymbol{h},\ \boldsymbol{R}_{\mathrm{A}},\ \boldsymbol{T}_{\mathrm{A}},\ \boldsymbol{R}_{\mathrm{B}},\ \boldsymbol{T}_{\mathrm{B}}$$

均视为先验未知。

III. 最先进的校准方案

现在，我们介绍文献 [27]中提出的双天线中继器最先进的互易性校准方案。其核心思想是在两种不同的中继器配置下，通过阵列A和阵列B之间的双向导频传输，估计正向和反向中继器增益 α 与 β之间的比值。具体而言，在一个信道相干间隔内进行两组测量：

（i）R以其标称配置进行双向传输；（ii）将R配置为对正向和反向增益均施加π相移后进行双向传输。

该过程得到如下四个信道测量值：

$$\boldsymbol{X}_{\mathrm{AB}}^{0} = \boldsymbol{R}_{\mathrm{B}}\left(\boldsymbol{G}+\alpha \boldsymbol{g}\boldsymbol{h}^{\mathsf{T}}\right)\boldsymbol{T}_{\mathrm{A}}+\boldsymbol{W}_{\mathrm{B}}^{0}, \tag{4a}$$
$$\boldsymbol{X}_{\mathrm{BA}}^{0} = \boldsymbol{R}_{\mathrm{A}}\left(\boldsymbol{G}^{\mathsf{T}}+\beta \boldsymbol{h}\boldsymbol{g}^{\mathsf{T}}\right)\boldsymbol{T}_{\mathrm{B}}+\boldsymbol{W}_{\mathrm{A}}^{0}, \tag{4b}$$
$$\boldsymbol{X}_{\mathrm{AB}}^{1} = \boldsymbol{R}_{\mathrm{B}}\left(\boldsymbol{G}-\alpha \boldsymbol{g}\boldsymbol{h}^{\mathsf{T}}\right)\boldsymbol{T}_{\mathrm{A}}+\boldsymbol{W}_{\mathrm{B}}^{1}, \tag{4c}$$
$$\boldsymbol{X}_{\mathrm{BA}}^{1} = \boldsymbol{R}_{\mathrm{A}}\left(\boldsymbol{G}^{\mathsf{T}}-\beta \boldsymbol{h}\boldsymbol{g}^{\mathsf{T}}\right)\boldsymbol{T}_{\mathrm{B}}+\boldsymbol{W}_{\mathrm{A}}^{1}, \tag{4d}$$

其中，$\boldsymbol{W}^{0}{}_{\mathrm{B}} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{A}}}$、$\boldsymbol{W}_{\mathrm{A}}^{0} \in \mathbb{C}^{M_{\mathrm{A}}\times M_{\mathrm{B}}}$、$\boldsymbol{W}_{\mathrm{B}}^{1} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{A}}}$和 $\boldsymbol{W}_{\mathrm{A}}^{1} \in \mathbb{C}^{M_{\mathrm{A}}\times M_{\mathrm{B}}}$ 表示测量噪声矩阵。

校准的目的是确定比值

$$\gamma \triangleq \frac{\beta}{\alpha}, \tag{5}$$

由于已知 γ，中继器便可配置为使得 $\alpha=\beta$，从而与网络互易。

为解决这一估计问题，将模型重新参数化为

$$\boldsymbol{H} \triangleq \boldsymbol{R}_{\mathrm{B}}\boldsymbol{G}\boldsymbol{T}_{\mathrm{A}} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{A}}}, \tag{6a}$$
$$\boldsymbol{A} \triangleq \boldsymbol{T}_{\mathrm{A}}^{-1}\boldsymbol{R}_{\mathrm{A}} \in \mathbb{C}^{M_{\mathrm{A}}\times M_{\mathrm{A}}}, \tag{6b}$$
$$\boldsymbol{B} \triangleq \boldsymbol{T}_{\mathrm{B}}\boldsymbol{R}_{\mathrm{B}}^{-1} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{B}}}, \tag{6c}$$
$$\boldsymbol{Z} \triangleq \alpha \boldsymbol{R}_{\mathrm{B}}\boldsymbol{g}\boldsymbol{h}^{\mathsf{T}}\boldsymbol{T}_{\mathrm{A}} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{A}}}, \tag{6d}$$

由此得到等价的测量方程：

$$\boldsymbol{X}_{\mathrm{AB}}^{0} = \boldsymbol{H}+\boldsymbol{Z}+\boldsymbol{W}_{\mathrm{B}}^{0}, \tag{7a}$$
$$\boldsymbol{X}_{\mathrm{BA}}^{0} = \boldsymbol{A}(\boldsymbol{H}+\gamma \boldsymbol{Z})^{\mathsf{T}}\boldsymbol{B}+\boldsymbol{W}_{\mathrm{A}}^{0}, \tag{7b}$$
$$\boldsymbol{X}_{\mathrm{AB}}^{1} = \boldsymbol{H}-\boldsymbol{Z}+\boldsymbol{W}_{\mathrm{B}}^{1}, \tag{7c}$$
$$\boldsymbol{X}_{\mathrm{BA}}^{1} = \boldsymbol{A}(\boldsymbol{H}-\gamma \boldsymbol{Z})^{\mathsf{T}}\boldsymbol{B}+\boldsymbol{W}_{\mathrm{A}}^{1}. \tag{7d}$$

经过简单计算，校准任务所需的预处理测量值如下：

$$\boldsymbol{R}_{1} \triangleq \frac{1}{2}\left(\boldsymbol{X}_{\mathrm{AB}}^{0}+\boldsymbol{X}_{\mathrm{AB}}^{1}\right) = \boldsymbol{H}+\boldsymbol{W}_{1}, \tag{8a}$$
$$\boldsymbol{R}_{2} \triangleq \frac{1}{2}\left(\boldsymbol{X}_{\mathrm{AB}}^{0}-\boldsymbol{X}_{\mathrm{AB}}^{1}\right) = \boldsymbol{Z}+\boldsymbol{W}_{2}, \tag{8b}$$
$$\boldsymbol{R}_{3} \triangleq \frac{1}{2}\left(\boldsymbol{X}_{\mathrm{BA}}^{0}+\boldsymbol{X}_{\mathrm{BA}}^{1}\right) = \boldsymbol{A}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{B}+\boldsymbol{W}_{3}, \tag{8c}$$
$$\boldsymbol{R}_{4} \triangleq \frac{1}{2}\left(\boldsymbol{X}_{\mathrm{BA}}^{0}-\boldsymbol{X}_{\mathrm{BA}}^{1}\right) = \gamma \boldsymbol{A}\boldsymbol{Z}^{\mathsf{T}}\boldsymbol{B}+\boldsymbol{W}_{4}, \tag{8d}$$

其中

$$\boldsymbol{W}_{1} \triangleq \frac{1}{2}\left(\boldsymbol{W}_{\mathrm{B}}^{0}+\boldsymbol{W}_{\mathrm{B}}^{1}\right), \quad \boldsymbol{W}_{2} \triangleq \frac{1}{2}\left(\boldsymbol{W}_{\mathrm{B}}^{0}-\boldsymbol{W}_{\mathrm{B}}^{1}\right),$$
$$\boldsymbol{W}_{3} \triangleq \frac{1}{2}\left(\boldsymbol{W}_{\mathrm{A}}^{0}+\boldsymbol{W}_{\mathrm{A}}^{1}\right), \quad \boldsymbol{W}_{4} \triangleq \frac{1}{2}\left(\boldsymbol{W}_{\mathrm{A}}^{0}-\boldsymbol{W}_{\mathrm{A}}^{1}\right).$$

根据式（8），通过最小化以下非线性最小二乘准则 [27], 联合估计式（6）中定义的参数

$$f = \|\boldsymbol{R}_{1}-\boldsymbol{H}\|_{\mathrm{F}}^{2}+\|\boldsymbol{R}_{2}-\boldsymbol{Z}\|_{\mathrm{F}}^{2} + \|\boldsymbol{R}_{3}-\boldsymbol{A}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{B}\|_{\mathrm{F}}^{2}+\|\boldsymbol{R}_{4}-\gamma \boldsymbol{A}\boldsymbol{Z}^{\mathsf{T}}\boldsymbol{B}\|_{\mathrm{F}}^{2}. \tag{9}$$

所得解提供了 γ 的统计一致估计，从而实现中继器的互易性校准。

A. NLS算法

基于 (9)，实际参数估计采用以下逐步算法。

1) $\boldsymbol{H}$ 的估计：根据 (8a) 中的平均测量值，相对于 $\boldsymbol{H}$ 的极小化解为

$$\hat{\boldsymbol{H}} = \boldsymbol{R}_{1}. \tag{10}$$

2) $\boldsymbol{Z}$ 的估计：利用 (8b)，$\boldsymbol{Z}$ 的估计值可得为

$$\hat{\boldsymbol{Z}} = \mathcal{S}\{\boldsymbol{R}_{2}\}, \tag{11}$$

其中，$\mathcal{S}\{\cdot\}$ 表示矩阵的最佳秩一近似，可通过其主奇异向量对获得。

3) 估计 $\boldsymbol{A}$和 $\boldsymbol{B}$：根据(8c)，通过交替投影估计矩阵 $\boldsymbol{A}$和 $\boldsymbol{B}$ 。从初始值 $\boldsymbol{A}=\boldsymbol{I}_{M_{\mathrm{A}}}$ 和 $\boldsymbol{B}=\boldsymbol{I}_{M_{\mathrm{B}}}$开始，每个对角元素

算法1：基于非线性最小二乘法的校准算法

输入： $\boldsymbol{R}_1,\ \boldsymbol{R}_2,\ \boldsymbol{R}_3,\ \boldsymbol{R}_4$

输出： $\hat{\boldsymbol{H}}, \hat{\boldsymbol{Z}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\gamma}$

// 1) $\boldsymbol{H}$的估计

$\hat{\boldsymbol{H}} \leftarrow \boldsymbol{R}_1$

// 2) $\boldsymbol{Z}$的估计

$\hat{\boldsymbol{Z}} \leftarrow \mathcal{S}\{\boldsymbol{R}_2\}$

// 3) $\boldsymbol{A}$ 和 $\boldsymbol{B}$的估计

$\hat{\boldsymbol{A}} \leftarrow \boldsymbol{I}_{M_\mathrm{A}},\ \hat{\boldsymbol{B}} \leftarrow \boldsymbol{I}_{M_\mathrm{B}}$ ▷ 初始化

重复

$\forall i \in \mathcal{A}$: $\hat{\boldsymbol{A}}(i,\ i) \leftarrow \dfrac{(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}})_i^{\mathsf{H}}(\boldsymbol{R}_3^{\mathsf{T}})_i}{\|(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}})_i\|_2^2}$

$\forall j \in \mathcal{B}\ \hat{\boldsymbol{B}}(j,j) \leftarrow \dfrac{(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}})_j^{\mathsf{H}}(\boldsymbol{R}_3)_j}{\|(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}})_j\|_2^2}$

$\hat{\boldsymbol{A}} \leftarrow \hat{\boldsymbol{A}} \cdot \|\hat{\boldsymbol{B}}\|_\mathrm{F}$

$\hat{\boldsymbol{B}} \leftarrow \hat{\boldsymbol{B}} / \|\hat{\boldsymbol{B}}\|_\mathrm{F}$

直到 迭代次数达到 N_Iter

// 4) γ的估计

$\hat{\gamma} \leftarrow \dfrac{\mathrm{tr}\{(\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}})^{\mathsf{H}}\boldsymbol{R}_4\}}{\|\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}}\|_\mathrm{F}^2}$

迭代更新如下

$$\hat{\boldsymbol{A}}(i,i) \ = \ \frac{(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}})_i^{\mathsf{H}}(\boldsymbol{R}_3^{\mathsf{T}})_i}{\|(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}})_i\|_2^2},\ \forall i \in \mathcal{A}, \tag{12a}$$

$$\hat{\boldsymbol{B}}(j,j) \ = \ \frac{(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}})_j^{\mathsf{H}}(\boldsymbol{R}_3)_j}{\|(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}})_j\|_2^2},\ \forall j \in \mathcal{B}, \tag{12b}$$

其中 $\mathcal{A} \triangleq \{1,2,\ldots,M_\mathrm{A}\}$和 $\mathcal{B} \triangleq \{1,2,\ldots,M_\mathrm{B}\}$。每次迭代都会进行归一化，以防止数值不稳定性 [27]。

4) γ的估计：使用 (8d),

$$\boldsymbol{R}_4 = \tfrac{1}{2}\left(\boldsymbol{X}_{\mathrm{BA}}^0 - \boldsymbol{X}_{\mathrm{BA}}^1\right), \tag{13}$$

增益比估计为

$$\hat{\gamma} = \frac{\mathrm{tr}\Big\{(\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}})^{\mathsf{H}}\boldsymbol{R}_4\Big\}}{\|\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}}\|_\mathrm{F}^2}. \tag{14}$$

该NLS算法给出了(9)中原始准则的近似解。上述过程总结于算法1。

B. 交替优化*NLS*算法

由(8)可知， $\boldsymbol{H}$ 的信息不仅包含在$\boldsymbol{R}_1$ 中，也包含在 $\boldsymbol{R}_3$中，而 $\boldsymbol{Z}$ 的信息不仅包含在 $\boldsymbol{R}_2$ 中，也包含在 $\boldsymbol{R}_4$中。因此，通过对所有未知变量进行交替优化，即 $\boldsymbol{H}$、 $\boldsymbol{Z}$、 $\boldsymbol{A}$、 $\boldsymbol{B}$和 $\boldsymbol{\gamma}$，可以进一步提升性能。详情请读者参阅 [27], ；为完整起见，算法2中仅提供伪代码。然而，该方法会带来更高的计算复杂度。

IV. 所提出的MMSE算法

最先进方法基于NLS准则设计，因此无法利用可用的统计信

Algorithm 2 : Alternating-opt. NLS-based algorithm

Input: $\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, \boldsymbol{R}_4$

Output: $\hat{\boldsymbol{H}}, \hat{\boldsymbol{Z}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\gamma}$

/* ———— Initialization ———— */

Obtain initial estimates $\hat{\boldsymbol{H}}, \hat{\boldsymbol{Z}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\gamma}$ using Alg. 1.

/* ———— Alternating optimization ———— */

repeat

// 1) Update of $\boldsymbol{H}$

$\boldsymbol{r} \leftarrow \left[\mathrm{vec}(\boldsymbol{R}_1);\ \mathrm{vec}(\boldsymbol{R}_3^{\mathsf{T}})\right]$, $\boldsymbol{\Xi} \leftarrow \left[\boldsymbol{I}_{M_\mathrm{A}} \otimes \boldsymbol{I}_{M_\mathrm{B}};\ \hat{\boldsymbol{A}} \otimes \hat{\boldsymbol{B}}\right]$

$\mathrm{vec}(\hat{\boldsymbol{H}}) \leftarrow (\boldsymbol{\Xi}^{\mathsf{H}}\boldsymbol{\Xi})^{-1}\boldsymbol{\Xi}^{\mathsf{H}}r$

// 2) Update of $\boldsymbol{A}$ and $\boldsymbol{B}$

repeat

$\forall i \in \mathcal{A} : \hat{\boldsymbol{A}}(i,i) \leftarrow \dfrac{(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}})_i^{\mathsf{H}}(\boldsymbol{R}_3^{\mathsf{T}})_i + \hat{\gamma}^*(\hat{\boldsymbol{B}}\hat{\boldsymbol{Z}})_i^{\mathsf{H}}(\boldsymbol{R}_4^{\mathsf{T}})_i}{\|(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}})_i\|_2^2 + |\hat{\gamma}|^2\|(\hat{\boldsymbol{B}}\hat{\boldsymbol{Z}})_i\|_2^2}$

$\forall j \in \mathcal{B} : \hat{\boldsymbol{B}}(j,j) \leftarrow \dfrac{(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}})_j^{\mathsf{H}}(\boldsymbol{R}_3)_j + \hat{\gamma}^*(\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}})_j^{\mathsf{H}}(\boldsymbol{R}_4)_j}{\|(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}})_j\|_2^2 + |\hat{\gamma}|^2\|(\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}})_j\|_2^2}$

$\hat{\boldsymbol{A}} \leftarrow \hat{\boldsymbol{A}} \cdot \|\hat{\boldsymbol{B}}\|_\mathrm{F}$

$\hat{\boldsymbol{B}} \leftarrow \hat{\boldsymbol{B}} / \|\hat{\boldsymbol{B}}\|_\mathrm{F}$

until Num. of iterations reaches N_Iter.

// 3) Update of $\boldsymbol{Z}$

$\hat{\boldsymbol{Z}} \leftarrow \mathcal{S}\left\{\boldsymbol{R}_2 + \dfrac{\hat{\gamma}^*\hat{\boldsymbol{B}}^{-1}\boldsymbol{R}_4^{\mathsf{T}}\hat{\boldsymbol{A}}^{-1}}{1+|\hat{\gamma}|^2}\right\}$

// 4) Update of γ

$\hat{\gamma} \leftarrow \dfrac{\mathrm{tr}\Big\{(\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}\hat{\boldsymbol{B}})^{\mathsf{H}}\boldsymbol{R}_4\Big\}}{\|\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}\hat{\boldsymbol{B}}\|_\mathrm{F}^2}$

until the objective function f no longer decreases.

息，例如信噪比以及未知变量的先验分布。因此，将适当的数学模型作为未知变量的先验信息，并依据MMSE准则设计算法，有望显著提升性能。受这一观察的启发，本节提出一种贝叶斯推断算法，该算法显式纳入先验信息并进行MMSE估计。

A. 未知变量的先验统计建模

在无线通信系统中，未知变量的统计特性和数学模型通常可以根据先验观测或估计推断得出，尤其是当这些变量源自硬件引起的效应或长期统计行为时。本小节将明确介绍所提出的方法中使用的统计量和数学模型。

1) 复值互易系数：广泛用于表征 MIMO 阵列中硬件损伤的一种成熟概率模型，是乘性随机损伤模型 [31], [32]。在该模型中，补偿后的残余损伤表示为乘性幅度误差和相位误差。在这一假设下，式 (1) 中引入的复值互易系数建模为

$$t_{\mathrm{A},i} = (1+\epsilon_{\mathrm{TA},i})\, e^{\mathrm{j}\theta_{\mathrm{TA},i}}, \quad i \in \mathcal{A}, \tag{15a}$$

$$r_{\mathrm{A},i} = (1+\epsilon_{\mathrm{RA},i})\, e^{\mathrm{j}\theta_{\mathrm{RA},i}}, \quad i \in \mathcal{A}, \tag{15b}$$

$$t_{\mathrm{B},j} = (1+\epsilon_{\mathrm{TB},j})\, e^{\mathrm{j}\theta_{\mathrm{TB},j}}, \quad j \in \mathcal{B}, \tag{15c}$$
$$r_{\mathrm{B},j} = (1+\epsilon_{\mathrm{RB},j})\, e^{\mathrm{j}\theta_{\mathrm{RB},j}}, \quad j \in \mathcal{B}, \tag{15d}$$

其中，$\epsilon_{\mathrm{TA},\,i}$，$\epsilon_{\mathrm{RA},\,i}$，$\epsilon_{\mathrm{TB},\,j}$和$\epsilon_{\mathrm{RB},\,j}$ 分别表示幅度误差，而$\theta_{\mathrm{TA},\,i}$，$\theta_{\mathrm{RA},\,i}$，$\theta_{\mathrm{TB},\,j}$和$\theta_{\mathrm{RB},\,j}$ 分别表示相位误差。如第二节所述，各个天线单元的复值互易系数的主要来源，是由 A 和 B 采用的不同振荡器引起的相位漂移和相位误差。因此，在将补偿后的残余幅度误差建模为服从 $\mathcal{N}(0,\ \sigma_\epsilon^2)$ [31], 的实值高斯随机变量时，我们假设其方差足够小，即，$\sigma_\epsilon \ll 1$。

2）测量噪声：观测噪声除了可能包含源自热噪声的高斯噪声外，还可能包含空间有色干扰。此外，根据采样方案的不同，还可能产生时间相关性。因此，本文为了适当考虑此类时空相关性，将式（4）中的观测噪声矩阵建模为圆对称多元复高斯随机矩阵。假设空间（列方向）相关性和时间（行方向）相关性是可分离的，并且可以采用克罗内克表示，则式（4）中观测噪声矩阵的概率密度函数表示为

$$p_{\mathbf{W}_{\mathrm{A}}^{\mathrm{x}}}\left(\boldsymbol{W}_{\mathrm{A}}^{\mathrm{x}}\right) = \frac{\exp\left[-\mathrm{tr}\left(\bar{\boldsymbol{\Omega}}_{\mathrm{A}}^{-1}\boldsymbol{W}_{\mathrm{A}}^{\mathrm{x}}\bar{\boldsymbol{\Psi}}_{\mathrm{A}}^{-1}\left(\boldsymbol{W}_{\mathrm{A}}^{\mathrm{x}}\right)^{\mathrm{H}}\right)\right]}{\pi^{M_{\mathrm{A}}M_{\mathrm{B}}}\det\left(\bar{\boldsymbol{\Omega}}_{\mathrm{A}}\right)^{M_{\mathrm{B}}}\det\left(\bar{\boldsymbol{\Psi}}_{\mathrm{A}}\right)^{M_{\mathrm{A}}}}, \tag{16a}$$

$$p_{\mathbf{W}_{\mathrm{B}}^{\mathrm{x}}}\left(\boldsymbol{W}_{\mathrm{B}}^{\mathrm{x}}\right) = \frac{\exp\left[-\mathrm{tr}\left(\bar{\boldsymbol{\Omega}}_{\mathrm{B}}^{-1}\boldsymbol{W}_{\mathrm{B}}^{\mathrm{x}}\bar{\boldsymbol{\Psi}}_{\mathrm{B}}^{-1}\left(\boldsymbol{W}_{\mathrm{B}}^{\mathrm{x}}\right)^{\mathrm{H}}\right)\right]}{\pi^{M_{\mathrm{B}}M_{\mathrm{A}}}\det\left(\bar{\boldsymbol{\Omega}}_{\mathrm{B}}\right)^{M_{\mathrm{A}}}\det\left(\bar{\boldsymbol{\Psi}}_{\mathrm{B}}\right)^{M_{\mathrm{B}}}}, \tag{16b}$$

其中 $\mathrm{x} \in \{0,1\}$。A和B处的空间相关矩阵分别表示为 $\bar{\boldsymbol{\Omega}}_{\mathrm{A}} \in \mathbb{C}^{M_{\mathrm{A}}\times M_{\mathrm{A}}}$ 和 $\bar{\boldsymbol{\Omega}}_{\mathrm{B}} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{B}}}$，而时间相关矩阵分别表示为 $\bar{\boldsymbol{\Psi}}_{\mathrm{A}} \in \mathbb{C}^{M_{\mathrm{B}}\times M_{\mathrm{B}}}$ 和 $\bar{\boldsymbol{\Psi}}_{\mathrm{B}} \in \mathbb{C}^{M_{\mathrm{A}}\times M_{\mathrm{A}}}$。为简化记号，我们进一步定义 $\boldsymbol{\Omega}_{\mathrm{A}} \triangleq \bar{\boldsymbol{\Omega}}_{\mathrm{A}}/2$、$\boldsymbol{\Omega}_{\mathrm{B}} \triangleq \bar{\boldsymbol{\Omega}}_{\mathrm{B}}/2$、$\boldsymbol{\Sigma}_{\mathrm{A}} \triangleq \bar{\boldsymbol{\Sigma}}_{\mathrm{A}}/2$和 $\boldsymbol{\Sigma}_{\mathrm{B}} \triangleq \bar{\boldsymbol{\Sigma}}_{\mathrm{B}}/2$，分别表示式（8）中预处理测量噪声的协方差矩阵。

下文假设这些协方差矩阵已知（或可先验准确估计），并将其用于所提出的方法。然而，如果无法获得此类信息，则可仅利用可用信息构建协方差矩阵，并据此执行该方法。需要注意的是，所提出的方法不要求完整掌握协方差矩阵。即使可以获得完整协方差矩阵，也可以仅使用其对角元素（即方差矩阵）来实现所提出的方法，从而降低计算复杂度。可由系统设计者根据实际约束决定所提出方法的具体实现方式。

B. 概述

为概述所提出的算法，我们在图2中给出了与互易校准任务对应的FG。在该图中，方形节点（■）表示与观测相关联的因子节点（FNs），其中f_1、f_2、f_3和f_4 分别对应于（8a）、（8b）、（8c）和（8d）。圆形节点（◯）表示与未知变量对应的变量节点（VNs）。

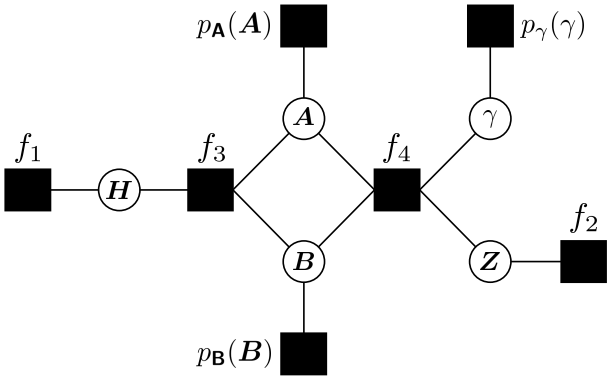


图 2. 为所提出的 MMSE 算法设计的因子图示意图。

当FG中存在环路时，估计过程对于确定算法的准确性和稳定性起着至关重要的作用。所提出的算法遵循与基本非线性最小二乘算法类似的估计过程，在该过程中，未知变量的估计值及其相关的均方误差（MSEs）（如果可用）会依次在FG中传播。最后，通过整合从所有观测沿FG传播的消息，以最小均方误差（MMSE）意义对 γ 进行估计。

相较于 NLS 算法，性能得以提升的关键因素有三点：i）在贝叶斯框架中明确纳入测量噪声的统计特性；ii）将式 (15) 中针对互易系数（即，$\boldsymbol{A}$ 和 $\boldsymbol{B}$）引入的先验统计建模用作先验概率分布；以及 iii）通过矩估计法（MoM）估计 γ 的先验分布，然后基于所得先验进行 MMSE 估计。下文将详细介绍所提出的算法。

*C. MMSE*算法

1) $\boldsymbol{H}$的估计：根据（6a），由于 $\boldsymbol{H}$ 包含从 A 到 B 的直达信道 $\boldsymbol{G}$，因此实际上很难为 $\boldsymbol{H}$假设统计先验模型。因此，与 NLS算法一样，我们使用（10）作为 $\boldsymbol{H}$ 的点估计，并传播表示估计误差信息的相关协方差矩阵，即，$\boldsymbol{\Omega}_{\mathrm{B}}$ 和 $\boldsymbol{\Psi}_{\mathrm{B}}$。

2) $\boldsymbol{Z}$的估计：类似地，根据（6d），由于 $\boldsymbol{Z}$ 包含经由 R 的间接信道 $\boldsymbol{g}$ 和 $\boldsymbol{h}$，因此很难为 $\boldsymbol{Z}$假设统计先验模型。因此，与 NLS算法一样，我们传播（11）作为 $\boldsymbol{Z}$的点估计。由于无法对秩一近似的估计误差进行解析评估，因此仅传播该估计值。

3) 对 $\boldsymbol{A}$ 与 $\boldsymbol{B}$的估计：基于(8c)，我们对 $\boldsymbol{A}$ 和 $\boldsymbol{B}$进行MMSE估计。为利用 $\boldsymbol{H}$ 的点估计及其相关误差协方差矩阵，并结合(15)中引入的先验统计模型，我们在贝叶斯双线性推断框架 [28]–[30], [33],[34]内采用基于概率数据关联的迭代交替估计方法。为便于理解消息更新规则，图3详细展示了由对应于(8c)的因子节点 f_3 以及与 $\boldsymbol{A}$和 $\boldsymbol{B}$ 中未知变量相关联的变量节点所组成的因子图。如图所示，以 f_3 为中心的因子图相对稀疏，并且由于不存在短

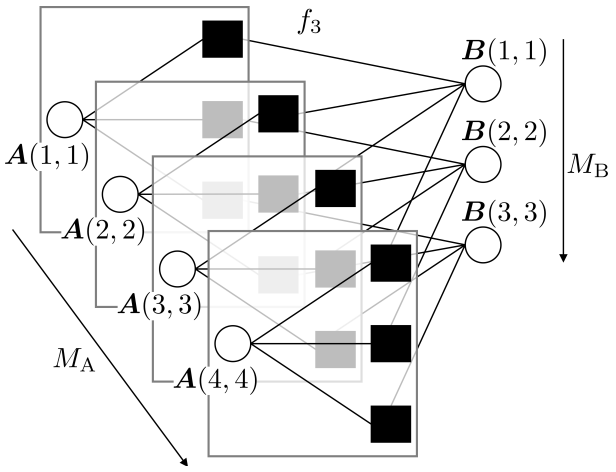


图 3. 用于对 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 进行双线性推断的三部因子图示意图，其中 $(M_{\mathrm{A}},\ M_{\mathrm{B}}) = (4,\ 3)$。

环路，则可预期以相对较少的迭代次数实现高精度收敛。

$\hat{}\hat{}$ 将 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的估计值分别记为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，则它们针对所有 $i \in \mathcal{A}$ 和 $j \in \mathcal{B}$ 的均方误差为

$$\hat{v}_{\mathrm{A},i} \triangleq \mathbb{E}\left[|\tilde{\boldsymbol{A}}(i,i)|^2\right],\ \tilde{\boldsymbol{A}}(i,i) \triangleq \boldsymbol{A}(i,i) - \hat{\boldsymbol{A}}(i,i), \tag{17a}$$

$$\hat{v}_{\mathrm{B},j} \triangleq \mathbb{E}\left[|\tilde{\boldsymbol{B}}(j,j)|^2\right],\ \tilde{\boldsymbol{B}}(j,j) \triangleq \boldsymbol{B}(j,j) - \hat{\boldsymbol{B}}(j,j), \tag{17b}$$

$\tilde{}$ $\tilde{}$其中，$\boldsymbol{A}$和 $\boldsymbol{B}$ 表示相应的估计误差矩阵。在第一次迭代时，所有 $i \in \mathcal{A}$ 和 $j \in \mathcal{B}$对应的估计值及其均方误差 $\hat{}\hat{}$ 初始化为 $\boldsymbol{A} = \boldsymbol{I}_{M_{\mathrm{A}}}$、$\boldsymbol{B} = \boldsymbol{I}_{M_{\mathrm{B}}}$和 $\hat{v}_{\mathrm{A},\ i} = \hat{v}_{\mathrm{B},\ j} = 1$。

首先，为了估计 $\boldsymbol{A}$的第 $(i,\ i)$ 个元素，我们关注式 (8c) 中观测的第 i 行，由此得到以下观测方程：

$$\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i = (\boldsymbol{B}\boldsymbol{H})_i\,\boldsymbol{A}(i,i) + \left(\boldsymbol{W}_3^{\mathsf{T}}\right)_i \in \mathbb{C}^{M_{\mathrm{B}}\times 1}. \tag{18}$$

$\hat{}\tilde{}$ $\hat{}\tilde{}$将 $\boldsymbol{H} = \hat{\boldsymbol{H}} + \tilde{\boldsymbol{H}}$和 $\boldsymbol{B} = \hat{\boldsymbol{B}} + \tilde{\boldsymbol{B}}$ 代入 (18)，得到

$$\begin{aligned}\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i = &\left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i)\\ &+ \left(\tilde{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i) + \left(\hat{\boldsymbol{B}}\tilde{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i)\\ &+ \left(\tilde{\boldsymbol{B}}\tilde{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i) + \left(\boldsymbol{W}_3^{\mathsf{T}}\right)_i.\end{aligned} \tag{19}$$

根据中心极限定理（CLT），将 (19) 第二行和第三行中的各项合并近似为一个多元复高斯随机向量。这种近似称为向量高斯近似（VGA） [35]。因此，给定 $\boldsymbol{A}(i,\ i)$ 时，(19) 的条件概率密度函数可表示为

$$\begin{aligned}&p_{\left(\mathbf{R}_3^{\mathsf{T}}\right)_i|\mathbf{A}(i,i)}\left(\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i|\boldsymbol{A}(i,i)\right)\\ &\propto \exp\Big[-\left(\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i - \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i)\right)^{\mathsf{H}} \boldsymbol{V}_{\mathrm{A},i}^{-1}\\ &\qquad\times\left(\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i - \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i)\right)\Big],\end{aligned} \tag{20}$$

其中

$$\begin{aligned}\boldsymbol{V}_{\mathrm{A},i} \triangleq\ &\mathbb{E}\Big[\left(\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i - \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i)\right)\\ &\times\left(\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i - \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i \boldsymbol{A}(i,i)\right)^{\mathsf{H}}\Big]\\ =\ &\boldsymbol{\Psi}_{\mathrm{A}} + \hat{\boldsymbol{B}}\boldsymbol{\Omega}_{\mathrm{B}}\hat{\boldsymbol{B}}^{\mathsf{H}} + \mathrm{diag}\left(v_{\mathrm{A},i1},\ldots,v_{\mathrm{A},iM_{\mathrm{B}}}\right),\end{aligned} \tag{21}$$

with

$$v_{\mathrm{A},ij} \triangleq \left(|\hat{\boldsymbol{H}}(j,i)|^2 + \boldsymbol{\Omega}_{\mathrm{B}}(j,j)\right)\hat{v}_{\mathrm{B},j}. \tag{22}$$

接下来，通过展开并重新整理式 (20) 中的指数项，并对 $\boldsymbol{A}(i,\ i)$ 配方，我们得到

$$\begin{aligned}&p_{\left(\mathbf{R}_3^{\mathsf{T}}\right)_i|\mathbf{A}(i,i)}\left(\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i|\boldsymbol{A}(i,i)\right)\\ &\propto \exp\left[-\frac{|\bar{\boldsymbol{A}}(i,i) - \boldsymbol{A}(i,i)|^2}{\bar{v}_{\mathrm{A},i}}\right],\end{aligned} \tag{23}$$

with

$$\bar{\boldsymbol{A}}(i,i) = \frac{1}{\psi_{\mathrm{A},i}}\left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i^{\mathsf{H}} \boldsymbol{V}_{\mathrm{A},i}^{-1}\left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i, \tag{24a}$$

$$\bar{v}_{\mathrm{A},i} = \frac{1}{\psi_{\mathrm{A},i}}, \tag{24b}$$

其中

$$\psi_{\mathrm{A},i} \triangleq \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i^{\mathsf{H}} \boldsymbol{V}_{\mathrm{A},i}^{-1}\left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i. \tag{25}$$

¯由 (23) 可知，$\boldsymbol{A}(i,\ i)$ 可视为对 $\boldsymbol{A}(i.i)$ 的加性白高斯噪声（AWGN）观测。

最后，基于 (15) 中的先验统计模型，我们考虑 $\boldsymbol{A}(i,\ i)$ 的先验分布。将 (1a) 和 (1b) 代入 (6b)，可得

$$\begin{aligned}\boldsymbol{A}(i,i) = \frac{r_{\mathrm{A},i}}{t_{\mathrm{A},i}} &= \frac{(1+\epsilon_{\mathrm{RA},i})\,e^{\mathrm{j}\theta_{\mathrm{RA},i}}}{(1+\epsilon_{\mathrm{TA},i})\,e^{\mathrm{j}\theta_{\mathrm{TA},i}}} = \left[1 + \frac{\epsilon_i}{1+\epsilon_{\mathrm{TA},i}}\right]e^{\mathrm{j}\theta_i}\\ &\approx e^{\mathrm{j}\theta_i} + (1-\epsilon_{\mathrm{TA},i})\,\epsilon_i e^{\mathrm{j}\theta_i} \approx e^{\mathrm{j}\theta_i} + \epsilon_i e^{\mathrm{j}\theta_i}\end{aligned} \tag{26}$$

其中 $\epsilon_i \triangleq \epsilon_{\mathrm{RA},\ i} - \epsilon_{\mathrm{TA},\ i}$ 和 $\theta_i \triangleq \theta_{\mathrm{RA},\ i} - \theta_{\mathrm{TA},\ i}$。上述近似是在假设 $\sigma_\epsilon \ll 1$ 的条件下得到的，其中第一个近似采用一阶展开 $1/(1+\epsilon_{\mathrm{TA},\ i}) \approx 1 - \epsilon_{\mathrm{TA},\ i}$，第二个近似则忽略了二阶项 $\epsilon_{\mathrm{TA},\ i}\epsilon_i$。由于第二项 $\epsilon_i e^{\mathrm{j}\theta_i}$ 的均值为零，且总方差为 $2\sigma_\epsilon^2$，因此以极高概率而言，与第一项相比，其贡献足够小；故本文忽略第二项。在这种情况下，$\boldsymbol{A}(i,\ i)$ 的先验分布可基于指数族近似，采用冯·米塞斯分布建模。

引理 *1*：当 $\boldsymbol{A}(i,\ i)$ 位于复单位圆上时，可以使用冯·米塞斯去噪器计算 $\boldsymbol{A}(i,\ i)$ 的最小均方误差估计，如下所示：

$$\hat{\boldsymbol{A}}(i,i) = \eta\left(\bar{\boldsymbol{A}}(i,i);\bar{v}_{\mathrm{A},i}\right) = \frac{I_1(|\zeta_{\mathrm{A}}|)}{I_0(|\zeta_{\mathrm{A}}|)}e^{\mathrm{j}\cdot\arg(\zeta_{\mathrm{A}})}, \tag{27}$$

其中，$I_n(\cdot)$ 表示 n 阶第一类修正贝塞尔函数，且

$$\zeta_{\mathrm{A}} = \frac{2\bar{\boldsymbol{A}}(i,i)}{\bar{v}_{\mathrm{A},i}}. \tag{28}$$

相应的后验均方误差为

$$\hat{v}_{\mathrm{A},i} = \bar{v}_{\mathrm{A},i}\cdot\frac{\partial\eta(\bar{\boldsymbol{A}}(i,i);\bar{v}_{\mathrm{A},i})}{\partial\bar{\boldsymbol{A}}(i,i)} = 1 - \left(\frac{I_1(|\zeta_{\mathrm{A}}|)}{I_0(|\zeta_{\mathrm{A}}|)}\right)^2. \tag{29}$$

证明：该结果可通过在附录A中将冯·米塞斯去噪器分别代入 $\beta = 0$ 和 $r = 1$ 得出。 ■

在实际应用中，（未缠绕的）相位误差方差通常难以准确估计，因此无法可靠地确定冯·米塞斯分布对应的集中参数 β 。因此，本文在去噪过程中对相位采用非信息先验，令 $\beta = 0$，这对应于圆周均匀去噪器。

ˆ 这就完成了 $\boldsymbol{A}$的最小均方误差更新。

$\boldsymbol{B}(j,\ j)$的估计与 $\boldsymbol{A}(i,\ i)$的估计完全采用相同的方法。首先，为了估计 $\boldsymbol{B}$的$(j,\ j)$项，我们关注式（8c）中观测的第 j列，从而得到如下观测方程：

$$(\boldsymbol{R}_3)_j = \left(\boldsymbol{A}\boldsymbol{H}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j) + (\boldsymbol{W}_3)_j \in \mathbb{C}^{M_{\mathrm{A}}\times 1}. \tag{30}$$

ˆ˜ ˆ˜将 $\boldsymbol{H} = \hat{\boldsymbol{H}} + \tilde{\boldsymbol{H}}$和 $\boldsymbol{A} = \hat{\boldsymbol{A}} + \tilde{\boldsymbol{A}}$ 代入（30）可得

$$\begin{aligned}(\boldsymbol{R}_3)_j = &\left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j) \\ &+ \left(\tilde{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j) + \left(\hat{\boldsymbol{A}}\tilde{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j) \\ &+ \left(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j) + (\boldsymbol{W}_3)_j .\end{aligned} \tag{31}$$

根据中心极限定理，（31）第二行和第三行中的各项可以整体近似为一个多元复高斯随机向量。因此，给定 $\boldsymbol{B}(j,\ j)$时，（31）的条件概率密度函数可表示为

$$\begin{aligned}&p_{(\mathbf{R}_3)_j|\mathbf{B}(j,j)}\left((\boldsymbol{R}_3)_j \,|\boldsymbol{B}(j,j)\right) \\ &\propto \exp\Big[-\left((\boldsymbol{R}_3)_j - \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j)\right)^{\mathsf{H}} \boldsymbol{V}_{\mathrm{B},j}^{-1} \\ &\qquad \times \left((\boldsymbol{R}_3)_j - \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j)\right)\Big],\end{aligned} \tag{32}$$

其中

$$\begin{aligned}\boldsymbol{V}_{\mathrm{B},j} &\triangleq \mathbb{E}\Big[\left((\boldsymbol{R}_3)_j - \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j)\right) \\ &\qquad \times \left((\boldsymbol{R}_3)_j - \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j \boldsymbol{B}(j,j)\right)^{\mathsf{H}}\Big] \\ &= \boldsymbol{\Omega}_{\mathrm{A}} + \hat{\boldsymbol{A}}\boldsymbol{\Psi}_{\mathrm{B}}\hat{\boldsymbol{A}}^{\mathsf{H}} + \mathrm{diag}\left(v_{\mathrm{B},j1},\ldots,v_{\mathrm{B},jM_{\mathrm{A}}}\right),\end{aligned} \tag{33}$$

with

$$v_{\mathrm{B},ji} \triangleq \left(|\hat{\boldsymbol{H}}(j,i)|^2 + \boldsymbol{\Psi}_{\mathrm{B}}(i,i)\right)\hat{v}_{\mathrm{A},i}. \tag{34}$$

接下来，通过展开并重新整理(32)中的指数项，并对 $\boldsymbol{B}(j,j)$使用配方法，我们得到

$$\begin{aligned}&p_{(\mathbf{R}_3)_j|\mathbf{B}(j,j)}\left((\boldsymbol{R}_3)_j \,|\boldsymbol{B}(j,j)\right) \\ &\propto \exp\left[-\frac{|\bar{\boldsymbol{B}}(j,j) - \boldsymbol{B}(j,j)|^2}{\bar{v}_{\mathrm{B},j}}\right],\end{aligned} \tag{35}$$

with

$$\bar{\boldsymbol{B}}(j,j) = \frac{1}{\psi_{\mathrm{B},j}}\left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j^{\mathsf{H}} \boldsymbol{V}_{\mathrm{B},j}^{-1}(\boldsymbol{R}_3)_j , \tag{36a}$$

$$\bar{v}_{\mathrm{B},j} = \frac{1}{\psi_{\mathrm{B},j}}, \tag{36b}$$

其中

$$\psi_{\mathrm{B},j} \triangleq \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j^{\mathsf{H}} \boldsymbol{V}_{\mathrm{B},j}^{-1}\left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j . \tag{37}$$

最后，对于所有 $j \in \mathcal{B}$，$\boldsymbol{B}(j,\ j)$ 的先验分布也可以用冯·米塞斯分布建模，方法与式 (26) 完全相同。应用冯·米塞斯去噪器，并使用 $\beta = 0$ 和 $r = 1$，即可得到相应的最小均方误差估计及其均方误差，即，

$$\hat{\boldsymbol{B}}(j,j) = \eta\left(\bar{\boldsymbol{B}}(j,j);\bar{v}_{\mathrm{B},j}\right) = \frac{I_1(|\zeta_{\mathrm{B}}|)}{I_0(|\zeta_{\mathrm{B}}|)} e^{\mathrm{j}\cdot\arg(\zeta_{\mathrm{B}})}, \tag{38a}$$

$$\hat{v}_{\mathrm{B},j} = \bar{v}_{\mathrm{B},j}\cdot\frac{\partial\eta\left(\bar{\boldsymbol{B}}(j,j);\bar{v}_{\mathrm{B},j}\right)}{\partial\bar{\boldsymbol{B}}(j,j)} = 1 - \left(\frac{I_1(|\zeta_{\mathrm{B}}|)}{I_0(|\zeta_{\mathrm{B}}|)}\right)^2, \tag{38b}$$

其中

$$\zeta_{\mathrm{B}} = \frac{2\bar{\boldsymbol{B}}(j,j)}{\bar{v}_{\mathrm{B},j}}. \tag{39}$$

4) 对 γ的估计：将前述步骤中得到的所有估计值代入(8d)，从而得到

$$\begin{aligned}\boldsymbol{R}_4 &= \gamma\left(\hat{\boldsymbol{A}} + \tilde{\boldsymbol{A}}\right)\hat{\boldsymbol{Z}}^{\mathsf{T}}\left(\hat{\boldsymbol{B}} + \tilde{\boldsymbol{B}}\right) + \boldsymbol{W}_4 \\ &= \gamma\underbrace{\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}}}_{\triangleq\hat{\boldsymbol{D}}} + \gamma\underbrace{\left(\tilde{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}} + \hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\tilde{\boldsymbol{B}} + \tilde{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\tilde{\boldsymbol{B}}\right)}_{\triangleq\tilde{\boldsymbol{D}}} + \boldsymbol{W}_4 \\ &= \gamma\hat{\boldsymbol{D}} + \gamma\tilde{\boldsymbol{D}} + \boldsymbol{W}_4,\end{aligned} \tag{40}$$

ˆ其中我们假设 $\boldsymbol{Z} \approx \boldsymbol{Z}$。对于 γ的MMSE估计，将 (40) 其改写为如下向量化形式：

$$\mathrm{vec}\left(\boldsymbol{R}_4\right) = \underbrace{\mathrm{vec}\left(\hat{\boldsymbol{D}}\right)}_{\triangleq\hat{\boldsymbol{d}}}\gamma + \underbrace{\mathrm{vec}\left(\tilde{\boldsymbol{D}}\right)}_{\triangleq\tilde{\boldsymbol{d}}}\gamma + \underbrace{\mathrm{vec}\left(\boldsymbol{W}_4\right)}_{\triangleq\boldsymbol{w}_4}. \tag{41}$$

根据中心极限定理，(41)中的第二项和第三项可共同近似为一个多元复高斯随机向量。因此，在给定 $\boldsymbol{R}_4$的条件下，(41)的条件概率密度函数可表示为

$$\begin{aligned}&p_{\mathrm{vec}(\mathbf{R}_4)|\gamma}\left(\mathrm{vec}\left(\boldsymbol{R}_4\right)|\gamma\right) \\ &\propto \exp\Big[-\left(\mathrm{vec}\left(\boldsymbol{R}_4\right) - \hat{\boldsymbol{d}}\gamma\right)^{\mathsf{H}} \boldsymbol{V}_\gamma^{-1}\left(\mathrm{vec}\left(\boldsymbol{R}_4\right) - \hat{\boldsymbol{d}}\gamma\right)\Big],\end{aligned} \tag{42}$$

其中

$$\begin{aligned}\boldsymbol{V}_\gamma &\triangleq \mathbb{E}\Big[\left(\mathrm{vec}\left(\boldsymbol{R}_4\right) - \hat{\boldsymbol{d}}\gamma\right)\left(\mathrm{vec}\left(\boldsymbol{R}_4\right) - \hat{\boldsymbol{d}}\gamma\right)^{\mathsf{H}}\Big] \\ &= \boldsymbol{\Sigma}_{\mathrm{A}} + \boldsymbol{V}_\gamma^{\mathrm{D}},\end{aligned} \tag{43}$$

其中 $\boldsymbol{\Sigma}_{\mathrm{A}} \triangleq \boldsymbol{\Psi}_{\mathrm{A}} \otimes \boldsymbol{\Omega}_{\mathrm{A}}$，$\boldsymbol{V}_{\mathrm{D}_\gamma} \triangleq \mathrm{diag}\left(v_{\gamma,\ 11},\ \ldots,\ v_{\gamma,\ M_{\mathrm{A}}M_{\mathrm{B}}}\right)$,

$$\begin{aligned}v_{\gamma,ij} &\triangleq \phi_\gamma|\hat{\boldsymbol{Z}}(j,i)|^2 \\ &\quad\times\left(\hat{v}_{\mathrm{A},i}|\hat{\boldsymbol{B}}(j,j)|^2 + |\hat{\boldsymbol{A}}(i,i)|^2\hat{v}_{\mathrm{B},j} + \hat{v}_{\mathrm{A},i}\hat{v}_{\mathrm{B},j}\right),\end{aligned} \tag{44}$$

以及 $\phi_\gamma \triangleq \mathbb{E}\left[|\gamma|^2\right]$。

通过展开并重新整理(42)中的指数项，并对 γ使用配方法，我们得到

$$p_{\mathrm{vec}(\mathbf{R}_4)|\gamma}\left(\mathrm{vec}\left(\boldsymbol{R}_4\right)|\gamma\right) \propto \exp\left[-\frac{|\bar{\gamma} - \gamma|^2}{\bar{v}_\gamma}\right], \tag{45}$$

with

$$\bar{\gamma} = \frac{\hat{\boldsymbol{d}}^{\mathsf{H}}\boldsymbol{V}_\gamma^{-1}\mathrm{vec}\left(\boldsymbol{R}_4\right)}{\psi_\gamma}, \quad \bar{v}_\gamma = \frac{1}{\psi_\gamma}, \tag{46}$$

其中

$$\psi_\gamma \triangleq \hat{\boldsymbol{d}}^{\mathsf{H}}\boldsymbol{V}_\gamma^{-1}\hat{\boldsymbol{d}}. \tag{47}$$

由于（44）中使用的 ϕ_γ 通常无法获得，因此需要进一步讨论其处理方式。理想情况下，在系统设置期间可以获得真实的长期统计量 $\phi_\gamma \triangleq \mathbb{E}\left[|\gamma|^2\right]$ ，此时可直接使用该统计量。当 ϕ_γ 未知时，可以考虑几种替代方案。最简单的方法是让系统以 $\phi_\gamma = 1$ 运行，同时注意到校准的最终目标是实现 $\alpha = \beta$。这相当于仅根据信噪比来确定估计值的可靠性。本文采用一种更具自适应性的替代方法，通过下文所述的 MoM，对 $|\gamma|^2$ 进行瞬时估计，并以此近似真实的长期统计量 ϕ_γ。

引理*2*：基于式（8d）中的观测，可通过矩估计法（MoM）获得 $|\gamma|^2$ 的一致估计量，如下所示：

$$|\check{\gamma}|^2 = \frac{|q|^2 - u}{u^2 + s}, \tag{48}$$

其中

$$q \triangleq \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \mathrm{vec}\left(\boldsymbol{R}_4\right), \tag{49a}$$
$$u \triangleq \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}}, \tag{49b}$$
$$s \triangleq \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \boldsymbol{V}_\gamma^{\mathrm{D}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}}. \tag{49c}$$

证明：根据（49a）中的定义以及（41）中的向量化观测模型，我们有

$$q = \underbrace{\hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}}}_{=u} \gamma + \underbrace{\hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \tilde{\boldsymbol{d}}}_{\triangleq \kappa_1} \gamma + \underbrace{\hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \boldsymbol{w}_4}_{\triangleq \kappa_2}. \tag{50}$$

由于 $\mathbb{E}[\boldsymbol{d}] = \mathbb{E}[\boldsymbol{w}_4] = \boldsymbol{0}$，因此可得 $\mathbb{E}[\kappa_1] = \mathbb{E}[\kappa_2] = 0$。此外，

$$\mathbb{E}\left[|\kappa_1|^2\right] = \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \boldsymbol{V}_\gamma^{\mathrm{D}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}} = s, \tag{51a}$$
$$\mathbb{E}\left[|\kappa_2|^2\right] = \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}} = u. \tag{51b}$$

对（50）两边平方并取期望，可得

$$\mathbb{E}\left[|q|^2\right] = \left(u^2 + s\right)|\gamma|^2 + u, \tag{52}$$

其中，我们利用了 $\boldsymbol{d}$ 与 $\boldsymbol{w}_4$的独立性。令理论矩 $\mathbb{E}\left[|q|^2\right]$ 等于其单样本对应值$|q|^2$，并求解 $|\gamma|^2$，即可得到（48）。 ■

最后，由于对 $|\gamma|$ 取平方根即可得到其估计值，因此我们可以将 γ的先验分布建模为复平面上半径为 $|\check{\gamma}|$的圆周均匀分布。采用 $\beta = 0$ 和$r = |\check{\gamma}|$的冯·米塞斯去噪器，即可得到相应的最小均方误差估计及其对应的均方误差，即，

$$\hat{\gamma} = \eta\left(\bar{\gamma}; \bar{v}_\gamma\right) = |\check{\gamma}| \frac{I_1(|\zeta_\gamma|)}{I_0(|\zeta_\gamma|)} e^{\mathrm{j}\cdot\arg(\zeta_\gamma)}, \tag{53a}$$

$$\hat{v}_\gamma = \bar{v}_\gamma \cdot \frac{\partial \eta\left(\bar{\gamma}; \bar{v}_\gamma\right)}{\partial \bar{\gamma}} = |\check{\gamma}|^2 \left(1 - \left(\frac{I_1(|\zeta_\gamma|)}{I_0(|\zeta_\gamma|)}\right)\right)^2, \tag{53b}$$

其中

$$\zeta_\gamma = \frac{2|\check{\gamma}|}{\bar{v}_\gamma} \bar{\gamma}. \tag{54}$$

通过上述推导，所提出的算法得以明确，并总结于算法3中。

Algorithm 3 : MMSE-based calibration algorithm

Input: $\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, \boldsymbol{R}_4, (\boldsymbol{\Omega}_{\mathrm{A}}, \boldsymbol{\Omega}_{\mathrm{B}}, \boldsymbol{\Psi}_{\mathrm{A}}, \boldsymbol{\Psi}_{\mathrm{B}}$ if available)
Output: $\hat{\boldsymbol{H}}, \hat{\boldsymbol{Z}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\gamma}$

// 1) Estimation of $\boldsymbol{H}$
$\hat{\boldsymbol{H}} \leftarrow \boldsymbol{R}_1$
// 2) Estimation of $\boldsymbol{Z}$
$\hat{\boldsymbol{Z}} \leftarrow \mathcal{S}\{\boldsymbol{R}_2\}$
// 3) Estimation of $\boldsymbol{A}$ and $\boldsymbol{B}$
$\hat{\boldsymbol{A}} \leftarrow \boldsymbol{I}_{M_{\mathrm{A}}}$, $\hat{\boldsymbol{B}} \leftarrow \boldsymbol{I}_{M_{\mathrm{B}}}$ ▷ Init初始化
$\forall i \in \mathcal{A}, \forall j \in \mathcal{B} : \hat{v}_{\mathrm{A},i} \leftarrow 1, \hat{v}_{\mathrm{B},j} \leftarrow 1$
repeat
$\forall i \in \mathcal{A}$: Compute $\boldsymbol{V}_{\mathrm{A},i}$ using (21).
$\forall i \in \mathcal{A} : \psi_{\mathrm{A},i} \leftarrow \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i^{\mathsf{H}} \boldsymbol{V}_{\mathrm{A},i}^{-1} \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i$
$\forall i \in \mathcal{A} : \bar{\boldsymbol{A}}(i,i) \leftarrow \frac{1}{\psi_{\mathrm{A},i}} \left(\hat{\boldsymbol{B}}\hat{\boldsymbol{H}}\right)_i^{\mathsf{H}} \boldsymbol{V}_{\mathrm{A},i}^{-1} \left(\boldsymbol{R}_3^{\mathsf{T}}\right)_i$
$\forall i \in \mathcal{A} : \bar{v}_{\mathrm{A},i} \leftarrow \frac{1}{\psi_{\mathrm{A},i}}$
$\forall i \in \mathcal{A} : \zeta_{\mathrm{A}} \leftarrow \frac{2}{\bar{v}_{\mathrm{A},i}} \bar{\boldsymbol{A}}(i,i)$
$\forall i \in \mathcal{A} : \hat{\boldsymbol{A}}(i,i) \leftarrow \frac{I_1(|\zeta_{\mathrm{A}}|)}{I_0(|\zeta_{\mathrm{A}}|)} e^{\mathrm{j}\cdot\arg(\zeta_{\mathrm{A}})}$
$\forall i \in \mathcal{A} : \hat{v}_{\mathrm{A},i} \leftarrow \left(1 - \left(\frac{I_1(|\zeta_{\mathrm{A}}|)}{I_0(|\zeta_{\mathrm{A}}|)}\right)^2\right)$
$\forall j \in \mathcal{B}$: Compute $\boldsymbol{V}_{\mathrm{B},j}$ using (33).
$\forall j \in \mathcal{B} : \psi_{\mathrm{B},j} \leftarrow \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j^{\mathsf{H}} \boldsymbol{V}_{\mathrm{B},j}^{-1} \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j$
$\forall j \in \mathcal{B} : \bar{\boldsymbol{B}}(j,j) \leftarrow \frac{1}{\psi_{\mathrm{B},j}} \left(\hat{\boldsymbol{A}}\hat{\boldsymbol{H}}^{\mathsf{T}}\right)_j^{\mathsf{H}} \boldsymbol{V}_{\mathrm{B},j}^{-1} \left(\boldsymbol{R}_3\right)_j$
$\forall j \in \mathcal{B} : \bar{v}_{\mathrm{B},j} \leftarrow \frac{1}{\psi_{\mathrm{B},j}}$
$\forall j \in \mathcal{B} : \zeta_{\mathrm{B}} \leftarrow \frac{2}{\bar{v}_{\mathrm{B},j}} \bar{\boldsymbol{B}}(j,j)$
$\forall j \in \mathcal{B} : \hat{\boldsymbol{B}}(j,j) \leftarrow \frac{I_1(|\zeta_{\mathrm{B}}|)}{I_0(|\zeta_{\mathrm{B}}|)} e^{\mathrm{j}\cdot\arg(\zeta_{\mathrm{B}})}$
$\forall j \in \mathcal{B} : \hat{v}_{\mathrm{B},j} \leftarrow \left(1 - \left(\frac{I_1(|\zeta_{\mathrm{B}}|)}{I_0(|\zeta_{\mathrm{B}}|)}\right)^2\right)$
until Num. of iterations reaches N_{Iter}
// 4) Estimation of γ
Compute $\boldsymbol{V}_\gamma$ and $\boldsymbol{V}_\gamma^{\mathrm{D}}$ using (43).
$\hat{\boldsymbol{d}} \leftarrow \mathrm{vec}\left(\hat{\boldsymbol{A}}\hat{\boldsymbol{Z}}^{\mathsf{T}}\hat{\boldsymbol{B}}\right)$, $q \leftarrow \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \mathrm{vec}\left(\boldsymbol{R}_4\right)$
$u \leftarrow \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}}$, $s \leftarrow \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \boldsymbol{V}_\gamma^{\mathrm{D}} \boldsymbol{\Sigma}_{\mathrm{A}}^{-1} \hat{\boldsymbol{d}}$.
$|\check{\gamma}|^2 \leftarrow \frac{|q|^2 - u}{u^2 + s}$
$\psi_\gamma \leftarrow \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{V}_\gamma^{-1} \hat{\boldsymbol{d}}$
$\bar{\gamma} \leftarrow \frac{1}{\psi_\gamma} \hat{\boldsymbol{d}}^{\mathsf{H}} \boldsymbol{V}_\gamma^{-1} \mathrm{vec}\left(\boldsymbol{R}_4\right)$
$\zeta_\gamma = \frac{2|\check{\gamma}|}{\bar{v}_\gamma} \bar{\gamma}$
$\hat{\gamma} \leftarrow |\check{\gamma}| \frac{I_1(|\zeta_\gamma|)}{I_0(|\zeta_\gamma|)} e^{\mathrm{j}\cdot\arg(\zeta_\gamma)}$

D. 复杂度分析

最后，我们将所提出的 MMSE 算法（算法3）与基本非线性最小二乘（算法1）和交替优化非线性最小二乘（算法2）进行比较，评估其计算复杂度。对于所有方法，估计 $\boldsymbol{Z}$ 都需要通过算子 $\mathcal{S}\{\cdot\}$ 计算 $\boldsymbol{R}_2$ 的最佳秩一近似；最坏情况下，这通常需要进行完整的奇异值分解（SVD）。因此，计算复杂度为

$$C_{\mathcal{S}} = \mathcal{O}\left(M_{\mathrm{A}} M_{\mathrm{B}} \min\{M_{\mathrm{A}}, M_{\mathrm{B}}\}\right). \tag{55}$$

然而，在实际应用中，可以采用 Krylov 子空间方法（如幂法或 Lanczos 算法）迭代计算最大奇异值及其对应的奇异向量，从而将计算开销降低至 $\mathcal{O}(\tau M_{\mathrm{A}} M_{\mathrm{B}})$，其中 τ 表示迭代次数，在实践中通常是一个较小的常数。

1) 非线性最小二乘（算法*1*）：对于基本非线性最小二乘算法，主要计算开销来自用于估计 $\boldsymbol{Z}$ 的秩一近似，以及用于估计 $\boldsymbol{A}$和 $\boldsymbol{B}$的 N_{Iter} 交替更新。由于 $\boldsymbol{A}$和 $\boldsymbol{B}$是对角矩阵，其更新可简化为内积计算和逐元素运算；因此，每次迭代更新 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的计算复杂度为 $\mathcal{O}(M_{\mathrm{A}} M_{\mathrm{B}})$量级。此外，$\gamma$ 的最终估计也仅使用逐元素运算和迹计算完成，所需复杂度为 $\mathcal{O}(M_{\mathrm{A}} M_{\mathrm{B}})$。因此，总体计算复杂度为

$$C_{\mathrm{NLS}} = \mathcal{O}\left(C_{\mathcal{S}} + N_{\mathrm{Iter}} M_{\mathrm{A}} M_{\mathrm{B}}\right). \tag{56}$$

2) 交替优化非线性最小二乘（算法*2*）：交替优化非线性最小二乘首先使用基本非线性最小二乘获得初始估计，然后通过带有外层迭代的交替优化提高估计精度。在每次外层迭代中，秩一近似在更新 $\boldsymbol{Z}$时执行一次，而更新 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的内层迭代重复 N_{Iter} 次。与算法1相比，新引入的 $\boldsymbol{H}$ 更新基于最小二乘准则表示为矩阵求逆。然而，由于 $\boldsymbol{A}$和 $\boldsymbol{B}$ 是对角矩阵，该操作可以逐元素实现，从而将计算复杂度降低至 $\mathcal{O}(M_{\mathrm{A}} M_{\mathrm{B}})$量级。计入使用基本非线性最小二乘进行初始估计的开销后，总体计算复杂度为

$$C_{\mathrm{AO-NLS}} = \mathcal{O}\left((N_{\mathrm{opt}} + 1)(C_{\mathcal{S}} + N_{\mathrm{Iter}} M_{\mathrm{A}} M_{\mathrm{B}})\right), \tag{57}$$

其中，N_{opt} 表示收敛所需的外层迭代次数，即代价函数的下降幅度足够小时为止。

3) MMSE（算法*3*）：在算法3中，决定计算复杂度的主要因素是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$的估计，以及 γ的估计。需要注意的是，基于冯·米塞斯去噪器的去噪运算均以标量方式执行，并可通过查找修正贝塞尔函数表高效实现。如上所述，所提出的MMSE算法经过推广，可利用观测噪声协方差矩阵的信息，其计算复杂度在很大程度上取决于估计过程中利用了多少此类信息。下面评估两种情况：i）使用式（16）中的完整协方差矩阵；ii）仅使用其对角元素（即方差）。

i) 完整协方差矩阵： 使用完整协方差矩阵时，$\boldsymbol{A}$ 的估计涉及对所有 $i \in \mathcal{A}$ 的 $\boldsymbol{V}_{\mathrm{A}}$矩阵求逆，$i$ 其计算复杂度为每次更新 $\mathcal{O}\left(M_{\mathrm{A}} M_{\mathrm{B}}^3\right)$ 量级。同样，$\hat{}\boldsymbol{B}$的更新需要 $\mathcal{O}\left({M^3}_{\mathrm{A}} M_{\mathrm{B}}\right)$次运算。这些交替更新重复进行 N_{Iter} 次。此外，γ 的估计涉及一个矩阵求逆，$\boldsymbol{V}_\gamma \in \mathbb{C}^{M_{\mathrm{A}} M_{\mathrm{B}} \times M_{\mathrm{A}} M_{\mathrm{B}}}$，其计算复杂度为 $\mathcal{O}\left(M_{\mathrm{A}}^3 M_{\mathrm{B}}^3\right)$ 量级。因此，该算法的总体计算复杂度为

$$C_{\mathrm{MMSE}} = \mathcal{O}\Big(C_{\mathcal{S}} + N_{\mathrm{Iter}}\left(M_{\mathrm{A}} M_{\mathrm{B}}^3 + M_{\mathrm{A}}^3 M_{\mathrm{B}}\right) + M_{\mathrm{A}}^3 M_{\mathrm{B}}^3\Big). \tag{58}$$

由于协方差矩阵表示长期统计特性，通常随时间缓慢变化。因此，与 $\boldsymbol{\Sigma}_{\mathrm{A}}$ 相关的矩阵求逆可视为离线操作，并从在线复杂度评估中排除。在这种情况下，γ 的估计主要由矩阵–向量乘法决定，整体计算复杂度可降至

$$C_{\mathrm{MMSE}} = \mathcal{O}\Big(C_{\mathcal{S}} + N_{\mathrm{Iter}}\left(M_{\mathrm{A}} M_{\mathrm{B}}^3 + M_{\mathrm{A}}^3 M_{\mathrm{B}}\right) + M_{\mathrm{A}}^2 M_{\mathrm{B}}^2\Big). \tag{59}$$

值得注意的是，利用完整协方差信息提高估计精度，必然要以增加计算复杂度为代价。

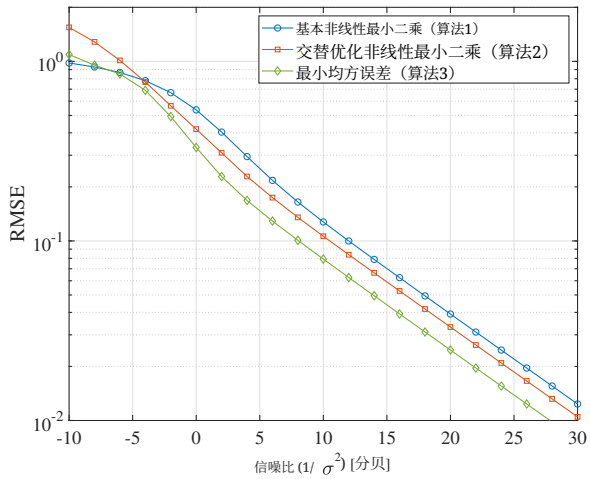
ii) 对角方差矩阵： 当仅使用协方差矩阵的对角元素时，$\boldsymbol{V}_{\mathrm{A},i}$、$\boldsymbol{V}_{\mathrm{B},j}$和 $\boldsymbol{V}_\gamma$ 的矩阵求逆可简化为逐元素取倒数运算。因此，估计 $\boldsymbol{A}$ 和 $\boldsymbol{B}$以及估计γ的每次迭代复杂度均降至 $\mathcal{O}(M_{\mathrm{A}} M_{\mathrm{B}})$，与基于NLS的算法相当。因此，在这种情况下，计算复杂度为

$$C_{\mathrm{MMSE}} = \mathcal{O}\left(C_{\mathcal{S}} + N_{\mathrm{Iter}} M_{\mathrm{A}} M_{\mathrm{B}}\right). \tag{60}$$

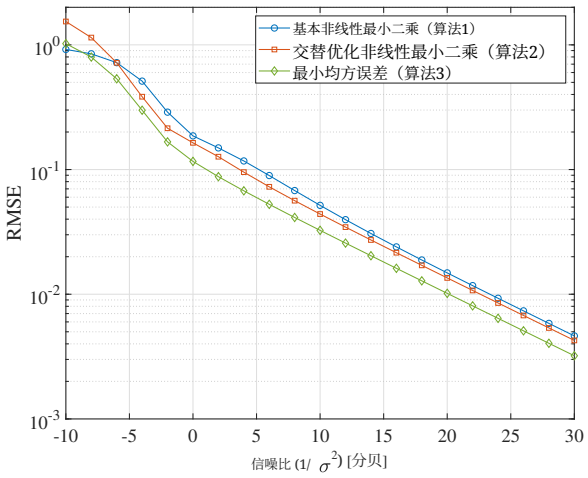
换言之，当观测噪声的先验知识仅限于方差矩阵时，所提出的MMSE算法的计算复杂度与基本非线性最小二乘算法处于同一数量级。

V. 性能评估

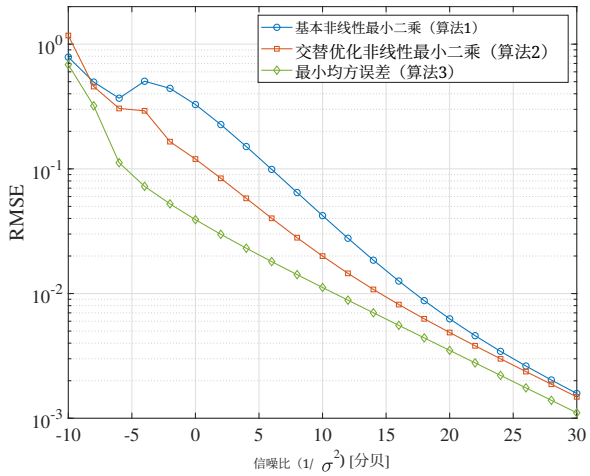
为验证所提方法的有效性，我们进行了计算机仿真。作为示例，采用 [27], 表 I 中列出的系统参数。向量 $\boldsymbol{h}$和 $\boldsymbol{g}$ 分别选取为 $M_{\mathrm{A}} \times M_{\mathrm{A}}$和 $M_{\mathrm{B}} \times M_{\mathrm{B}}$ 离散傅里叶变换（DFT）矩阵中的随机列，用于建模 A 与 R 之间以及 R 与 B 之间的视距传播。A 和 B 之间不经过 R 的直达信道 $\boldsymbol{G}$ 由独立同分布（i.i.d.）的 $\mathcal{CN}(0,\ 1)$元素组成，用于建模瑞利衰落。(1) 中 $\boldsymbol{T}_{\mathrm{A}}$、$\boldsymbol{R}_{\mathrm{A}}$、$\boldsymbol{T}_{\mathrm{B}}$和 $\boldsymbol{R}_{\mathrm{B}}$ 的对角元素具有单位幅度，相位在$[-\pi, \pi]$ [36]上均匀分布。如第二节所述，这一设置对应于相位主导非互易性场景，此时幅度失配可以忽略，非互易性主要由相位漂移/相位误差引起。(4) 中噪声矩阵 ${\boldsymbol{W}^0}_{\mathrm{B}}$、${\boldsymbol{W}^0}_{\mathrm{A}}$、${\boldsymbol{W}^1}_{\mathrm{B}}$和 $\boldsymbol{W}_{\mathrm{A}}^1$的元素独立地取自 $\mathcal{CN}(0,\ \sigma^2)$，其中噪声方差倒数 $1/\sigma^2$表示当 A 的单根天线以单位功率发射且中继器 R 不工作时，B 的任意一根天线处测得的信噪比。这相当于将观测噪声建模为加性高斯白噪声，这与在 (16) 中设置$\bar{\boldsymbol{\Omega}}_{\mathrm{A}} = \bar{\boldsymbol{\Psi}}_{\mathrm{B}} = \sigma^2 \boldsymbol{I}_{M_{\mathrm{A}}}$ 和 $\bar{\boldsymbol{\Omega}}_{\mathrm{B}} = \bar{\boldsymbol{\Psi}}_{\mathrm{A}} = \sigma^2 \boldsymbol{I}_{M_{\mathrm{B}}}$ 等价。



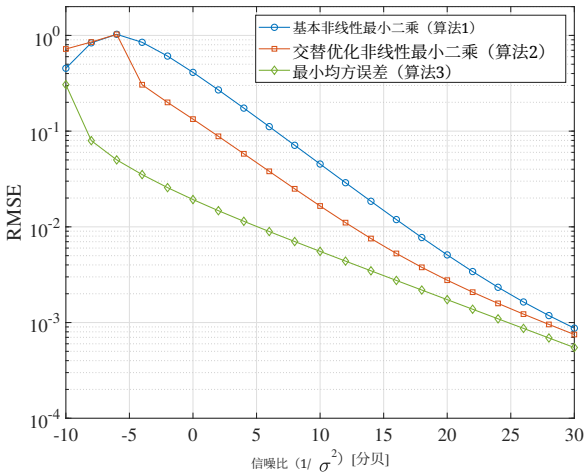
(a) $(M_A, M_B) = (4,3)$.



(b) $(M_A, M_B) = (8,8)$.



(c) $(M_A, M_B) = (32, 16)$。



(d) $(M_A, M_B) = (64,32)$.

图4. 不同天线配置下$\hat{\gamma}$ 的均方根误差（RMSE）。

表 I 仿真参数

参数	符号	值
中继器的正向增益	$\lvert\alpha\rvert^2$ [分贝]	10
中继器的反向增益	$\lvert\beta\rvert^2$ [分贝]	10
$\boldsymbol{A}$估计的迭代次数 $\boldsymbol{B}$，	N_{Iter}	100
仿真实验次数	-	10^5

在上述设置下，值得注意的是，与基于 NLS 的方法相比，所提 MMSE算法利用的额外先验信息很少，且其计算复杂度处于同一数量级，如 (60) 所示。

*A. RMSE*性能

图4展示了 NLS（算法1）、交替优化非线性最小二乘（算法2）和所提出的 MMSE（算法3）中，$\hat{\gamma}$ 随信噪比变化的均方根误差。为评估发射和接收阵列尺寸的影响，我们考虑了四种天线配置，即 $(M_A, M_B) = (4, 3)$、$(8, 8)$、$(32, 16)$ 和 $(64, 32)$。这些结果清晰地凸显了所提出的 MMSE 算法与现有基于 NLS 的方法之间的根本差异。

我们首先关注图4(a)和图4(b)所示的阵列规模相对较小时的结果。在中高信噪比范围（SNR > 5 dB）内，所有方法均表现稳定。对于非线性最小二乘和基于最小均方误差的方法，均方根误差随信噪比近似线性下降（例如，信噪比提高 20 dB，均方根误差约降低至原来的十分之一）。这一现象可以用观测噪声服从高斯分布来解释：对数似然函数的曲率，即费 Fisher信息，与信噪比成正比。在此范围内，真实参数附近的局部线性化是有效的，且雅可比矩阵满秩，因此非线性最小二乘估计器和最小均方误差估计器的误差缩放规律一致，这与估计理论相符。尽管如此，由于最小均方误差估计器利用了先验统计建模和噪声方差，因此在常数因子方面始终表现更优。具体而言，所提出的最小均方误差方法相较于基本非线性最小二乘和交替优化非线性最小二乘，性能分别提升约 4 dB 和 2 dB。

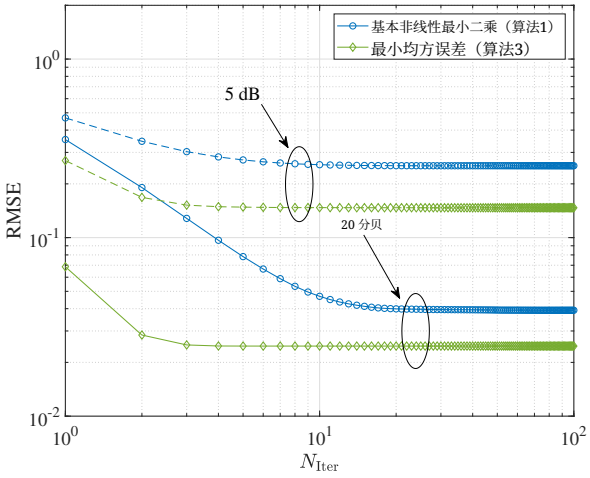
还值得注意的是，交替优化非线性最小二乘会产生与 N_{opt} 成正比的额外计算开销，而所提出的最小均方误差方法在实现这些性能提升的同时，计算复杂度与基本非线性最小二乘相当。

接下来，我们考虑图4(c)和图4(d)所示相对较大阵列规模下的结果，其中非线性最小二乘与最小均方误差之间的差异变得更加显著。随着阵列维度增加，本质上求解几何优化问题的非线性最小二乘估计器对高维效应越来越敏感。具体而言，雅可比矩阵的奇异值分布趋于恶化，且对数似然函数的曲率在许多方向上变得很小。因此，非线性最小二乘估计器的误差协方差矩阵变得病态，所得均方根误差不再与信噪比成比例变化，即使在高信噪比区域也是如此。由于校准任务通常在低信噪比条件下进行，非线性最小二乘估计器的这种行为在实际系统中是不可取的。

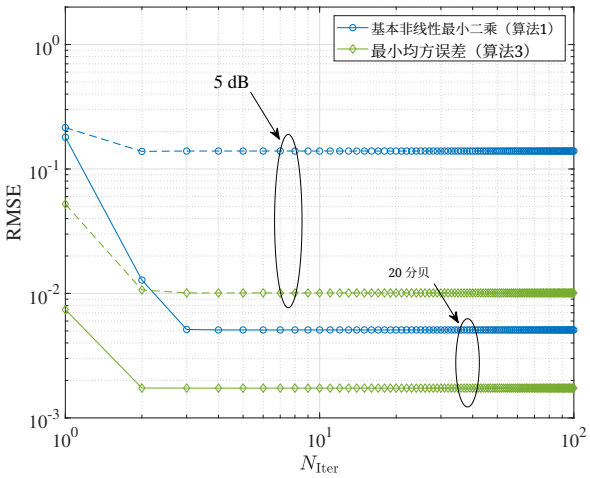
相比之下，对于最小均方误差估计器，先验分布在海森矩阵中引入了正则化项，从而使估计问题保持良态。因此，即使在高维和低信噪比区域，均方根误差也表现出稳定的行为，并保持接近理想1/信噪比缩放。此外，随着维度增加，后验分布由于测度集中现象而趋于集中，从而产生自平均效应，减轻局部非线性与噪声起伏的影响。该效应进一步稳定了最小均方误差估计器，并提高了基于中心极限定理的高斯近似的精度，而这些近似在所提出的算法中被广泛采用。因此，所提出的最小均方误差方法即使在低信噪比[1]下也遵循理论 1/信噪比缩放。特别地，对于图4(d)中最大的阵列配置，在均方根误差 $= 10^{-1}$处，所提出的最小均方误差方法相对于基本非线性最小二乘和交替优化非线性最小二乘分别获得约14分贝和10分贝的性能增益，表明大规模阵列对所提出的最小均方误差方法尤其有利。

B. $\boldsymbol{A}$与 $\boldsymbol{B}$ 估计的收敛行为

为进一步阐明所提出算法的行为，
接下来，我们通过关注 $\boldsymbol{A}$ 和 $\boldsymbol{B}$的双线性推断，研究它们的迭代收敛情况。图 5 展示了 $\hat{}\gamma$ 的均方根误差随迭代次数 N_{Iter}的变化。对于图 5(a) 所示的相对较小的阵列规模，基本非线性最小二乘的估计性能随迭代次数增加而单调提升，最终收敛至稳定解。然而，即使在高信噪比区域（信噪比 $= 20$分贝），也需要约 20 次迭代才能收敛。对于图 5(b) 所示的大规模阵列配置，非线性最小二乘的收敛行为在低信噪比区域显著恶化。在这种情况下，增加迭代次数并不能带来显著的性能



(a) $(M_{\text{A}},\ M_{\text{B}}) = (4,\ 3)$。



(b) $(M_{\text{A}},\ M_{\text{B}}) = (64,\ 32)$。

图5. 基于非线性最小二乘和最小均方误差的互易性校准算法的迭代收敛行为。

提升，估计误差往往会停滞在相对较高的水平。这一现象可归因于双线性最小二乘问题固有的非凸性，以及前文讨论的高维几何效应。

相比之下，所提出的最小均方误差方法呈现出截然不同的收敛行为。由于双线性推断步骤被构建为一系列贝叶斯更新，每次迭代都会通过纳入额外的统计信息来有效地优化后验分布。因此，估计性能随着迭代次数的增加而稳步提升，并且收敛迅速；在所有考察的场景中，通常只需 4 次迭代即可收敛。特别是，对于图 5(b) 所示的大规模阵列，由于高维度带来的自平均效应，收敛速度更快，这与上一小节讨论的结果一致。

VI. 结论

本文提出了一种用于中继器辅助 MIMO 的新型贝叶斯互易性校准框架。通过明确纳入

[1]尽管这些数值结果与理论标度律一致，但要针对当前包含潜变量的双线性推断问题推导严格的闭式克拉美´一罗下界（克拉美-罗下界）并非易事，因此留待未来工作。

互易系数和测量噪声的先验统计模型，所提出的方法为校准问题建立了严谨的贝叶斯表述，从而形成了稳健的估计框架。为有效利用互易系数的相位结构，本文开发了一种冯·米塞斯去噪器。此外，还提出了一种基于矩估计的方法，以自适应地估计所需的二阶统计量，避免依赖无法获得的长期先验信息。仿真结果表明，在计算复杂度相当的情况下，所提出的 MMSE 算法在估计精度方面显著优于现有的确定性基于 NLS 的方法。此外，所提出的方法收敛快速且稳定，并且对阵列规模和信噪比均具有鲁棒性。总之，该方法为中继器辅助 MIMO 系统中的互易性校准提供了一种高效且切实可行的解决方案。

附录A
冯·米塞斯先验分布的去噪器设计

考虑以下加性高斯白噪声观测模型：

$$y = x + w, \quad w \sim \mathcal{CN}(0, v), \tag{61}$$

其中， $x \triangleq re^{\mathrm{j}\theta}$ 是复平面上半径为 r 的圆上的一点，其相位 θ 服从冯·米塞斯分布

$$p_\theta(\theta) = \frac{1}{2\pi I_0(\beta)} \exp(\beta \cos(\theta - \mu)), \tag{62}$$

其中， $I_n(\cdot)$表示阶数为 n 的第一类修正贝塞尔函数，参数 μ 和 β 分别表示该分布的位置和集中度。这些参数类似于高斯分布的均值和精度。

在观测 y 以及(62)中的先验分布已知的假设下，我们基于(61)中的模型推导 x 的最小均方误差估计器为

$$\hat{x}_{\mathrm{MMSE}} = \mathbb{E}[x|y] = r \cdot \mathbb{E}\left[e^{\mathrm{j}\theta}|y\right]. \tag{63}$$

首先，给定观测 y 的似然函数由下式给出

$$\begin{aligned} p_{\mathrm{y}|\theta}(y|\theta) &= \frac{1}{\pi v} \exp\left[-\frac{|y - re^{\mathrm{j}\theta}|^2}{v}\right] \\ &\propto \exp\left[\frac{2r}{v} \Re\left\{y e^{-\mathrm{j}\theta}\right\}\right]. \end{aligned} \tag{64}$$

接下来，利用贝叶斯法则，后验分布可表示为

$$\begin{aligned} p_{\theta|\mathrm{y}}(\theta|y) &\propto p_{\mathrm{y}|\theta}(y|\theta) p_\theta(\theta) \\ &\propto \exp\left[\frac{2r}{v} \Re\left\{y e^{-\mathrm{j}\theta}\right\} + \beta \cos(\theta - \mu)\right] \\ &= \exp\left[\Re\left\{\left(\frac{2r}{v} y + \beta e^{\mathrm{j}\mu}\right) e^{-\mathrm{j}\theta}\right\}\right] \\ &= \exp\left[\Re\left\{\zeta(y, r, v, \mu, \beta)\, e^{-\mathrm{j}\theta}\right\}\right], \end{aligned} \tag{65}$$

其中

$$\zeta(y, r, v, \mu, \beta) \triangleq \frac{2r}{v} y + \beta e^{\mathrm{j}\mu}. \tag{66}$$

根据（65）和（66），后验分布仍可表示为冯·米塞斯分布：

$$p_{\theta|\mathrm{y}}(\theta|y) = \frac{1}{2\pi I_0(|\zeta|)} \exp\left[|\zeta| \cos(\theta - \arg(\zeta))\right]. \tag{67}$$

最后，将(67)代入(63)可得

$$\begin{aligned} \hat{x}_{\mathrm{MMSE}} &= r \cdot \mathbb{E}\left[e^{\mathrm{j}\theta}|y\right] \\ &= r \cdot \int_0^{2\pi} e^{\mathrm{j}\theta} \cdot p_{\theta|\mathrm{y}}(\theta|y) d\theta \\ &= \frac{r e^{\mathrm{j}\cdot\arg(\zeta)}}{I_0(|\zeta|)} \cdot \underbrace{\frac{1}{2\pi} \int_0^{2\pi} e^{\mathrm{j}\theta} \exp\left[|\zeta| \cos(\theta)\right] d\theta}_{=I_1(|\zeta|)} \\ &= r \frac{I_1(|\zeta|)}{I_0(|\zeta|)} e^{\mathrm{j}\cdot\arg(\zeta)} \triangleq \eta(y; v), \end{aligned} \tag{68}$$

其中，函数 η 表示将(61)中的最小均方误差解表示为 y 和 v的函数，这种表述通常被称为贝叶斯最优去噪器（*BOD*）。涉及修正贝塞尔函数的数学推导见附录B。

对于BOD，已知最小均方误差解与其均方误差之间成立如下恒等式：

$$v \cdot \frac{\partial \eta(y; v)}{\partial y} = \mathbb{E}\left[|x - \eta(y; v)|^2 |y\right]. \tag{69}$$

因此，后验方差可表示为

$$\mathbb{E}\left[|x - \hat{x}_{\mathrm{MMSE}}|^2 |y\right] = r^2 \left(1 - \left(\frac{I_1(|\zeta|)}{I_0(|\zeta|)}\right)^2\right). \tag{70}$$

附录B 修正贝塞尔函数的推导

为完整起见，我们推导修正贝塞尔函数 $I_n(\kappa)$ 的解析形式。从指数项的展开式出发，可得

$$\int_0^{2\pi} e^{\mathrm{j}n\psi} e^{\kappa \cos\psi} d\psi = \sum_{m=0}^{\infty} \frac{\kappa^m}{m!} \int_0^{2\pi} e^{\mathrm{j}n\psi} \cos^m\psi \, d\psi. \tag{71}$$

根据欧拉公式，$\cos^m\psi$ 可表示为复指数之和：

$$\cos^m\psi = \frac{1}{2^m}(e^{\mathrm{j}\psi} + e^{-\mathrm{j}\psi})^m = \frac{1}{2^m} \sum_{k=0}^{m} \binom{m}{k} e^{\mathrm{j}(2k-m)\psi}. \tag{72}$$

将其代入积分可得

$$\int_0^{2\pi} e^{\mathrm{j}n\psi} \cos^m\psi \, d\psi = \frac{1}{2^m} \sum_{k=0}^{m} \binom{m}{k} \int_0^{2\pi} e^{\mathrm{j}(n+2k-m)\psi} d\psi. \tag{73}$$

积分

$$\int_0^{2\pi} e^{\mathrm{j}q\psi} d\psi = \begin{cases} 2\pi & q = 0, \\ 0 & q \neq 0, \end{cases} \tag{74}$$

这意味着只有满足 $n+2k-m=0$ 的项才对积分有贡献。该条件可参数化为 $m =$

$n+2l$，对于 m的所有其他取值，积分均为零。因此，对于 $m=n+2l$，我们得到

$$\int_0^{2\pi} e^{\mathrm{j}n\psi}\cos^m\psi d\psi = 2\pi\cdot\frac{(n+2l)!}{2^{n+2l}\cdot l!\cdot(n+l)!}. \tag{75}$$

将该结果代回式（71），得到

$$\begin{aligned}\int_0^{2\pi} e^{\mathrm{j}n\psi}e^{\kappa\cos\psi}d\psi &= 2\pi\sum_{l=0}^{\infty}\frac{1}{l!}\cdot\frac{1}{(n+l)!}\cdot\left(\frac{\kappa}{2}\right)^{n+2l}\\ &= 2\pi\left(\frac{\kappa}{2}\right)^n\sum_{l=0}^{\infty}\frac{1}{l!\cdot(n+l)!}\left(\frac{\kappa}{2}\right)^{2l}\\ &= 2\pi I_n(\kappa). \end{aligned}\tag{76}$$

采用标准定义，我们由此得到众所周知的级数展开。

参考文献

[1] T. L. Marzetta，“天线数量无限的非合作蜂窝无线通信”，*IEEE* 无线通信汇刊，第 9 卷，第 11 期，第 3590–3600 页，2010年。[2] E. G. Larsson、O. Edfors、F. Tufvesson 和 T. L. Marzetta，“面向下一代无线系统的大规模MIMO”，*IEEE* 通信杂志，第 52 卷，第 2 期，第 186–195 页，2014年。[3] A. Bourdoux、B. Come 和 N. Khaled，“OFDM/SDMA 系统中的非互易收发器：影响与缓解”，载于*IEEE* 无线电与无线会议论文集（*RAWCON*），美国马萨诸塞州波士顿，2003年。[4] E. Bj¨ornson、J. Hoydis、M. Kountouris 和 M. Debbah，“采用非理想硬件的大规模MIMO系统：能效、估计与容量极限”，*IEEE* 信息论汇刊，2014年。[5] F. Kaltenberger、D. Gesbert 和 R. Knopp，“MIMO/时分双工系统中的相对信道互易性校准”，载于未来网络与移动峰会论文集，意大利佛罗伦萨，2010年6月。[6] J. Vieira、F. Rusek、O. Edfors、S. Malkowsky、L. Liu 和 F. Tufvesson，“大规模MIMO的互易性校准：方案、建模与验证”，*IEEE* 无线通信汇刊，2017年。[7] X. Jiang、A. Decurninge、K. Gopala、F. Kaltenberger、M. Guillaud 和 D. Slock，“时分双工大规模MIMO系统空口互易性校准框架”，*IEEE* 无线通信汇刊，第 5975–5990 页，2018年9月。[8] H. Q. Ngo、A. Ashikhmin、H. Yang、E. G. Larsson 和 T. L. Marzetta，“无蜂窝大规模MIMO与小型蜂窝的比较”，*IEEE* 无线通信汇刊，第 16 卷，第 3 期，第 1834–1850 页，2017年。[9] E. Bj¨ornson 和 L. Sanguinetti，“通过最小均方误差处理和集中式实现，使无蜂窝大规模MIMO更具竞争力”，*IEEE* 无线通信汇刊，第 19 卷，第 1 期，第 77–90 页，2020年。[10] A. Papazafeiropoulos、P. Kourtessis、M. D. Renzo、S. Chatzinotas 和 J. M. Senior，“无蜂窝大规模MIMO系统的性能分析：随机几何方法”，*IEEE* 车辆技术汇刊，第 69 卷，第 4 期，第 3523–3537 页，2020年。[11] H. Iimori、T. Takahashi、K. Ishibashi、G. T. F. de Abreu 和 W. Yu，“通过双线性推断实现低相干导频无蜂窝MIMO的免授权接入”，*IEEE* 无线通信汇刊，第 20 卷，第 11 期，第 7694–7710 页，2021年11月。[12] T. Takahashi、H. Iimori、K. Ando、K. Ishibashi、S. Ibi 和 G. T. F. de Abreu，“通过双线性推断为采用低分辨率模数转换器的无蜂窝大规模MIMO设计贝叶斯接收机”，*IEEE* 无线通信汇刊，第 22 卷，第 7 期，第 4756–4772 页，2023年。[13] H. A. Ammar、R. Adve、S. Shahbazpanahi、G. Boudreau 和 K. V. Srinivas，“用户中心式无蜂窝大规模MIMO网络：机遇、挑战与解决方案综述”，*IEEE* 通信调查与教程，第 24 卷，第 1 期，第 611–652 页，2022年。[14] S. Willhammar、H. Iimori、J. Vieira、L. Sundstr¨om、F. Tufvesson 和 E. G. Larsson，“借助中继器辅助蜂窝大规模MIMO实现分布式MIMO性能”，*IEEE* 通信杂志，第 63 卷，第 3 期，第 114–119 页，2025年3月。[15] T.-H. Vu、N. H. Tu 和 V. N. Q. Bao，“基于可重构中继器的 RSMA 系统研究”，*IEEE* 无线通信快报，第 14 卷，第 5 期，第 1371–1375 页，2025年5月。[16] P. J opanya 和 D. P. M. Osorio，“通过集群中继器辅助MIMO通感一体化实现无人机检测”，*arXiv* 预印本 *arXiv:2509.19119*，2025年9月。 [在线]。网址：https://arxiv.org/abs/2509.19119

[17] H. Iimori、E. Kurihara、T. Yoshida、J. Vieira 和 S. Malomsoky，“通过极大化最小化实现中继器辅助MIMO系统的放大策略”，载于 IEEE 全球通信会议（GLOBECOM）论文集，马来西亚吉隆坡，2023年12月。[18] M. Andersson、A. Chowdhury 和 E. G. Larsson，“中继辅助大规模MIMO是否兼容动态 TDD？ ” arXiv 预印本 arXiv:2510.20998，2026年1月。 [在线]。网址：https://arxiv.org/abs/2510 20998 。O. A. Topal、O. T. Demir、E. Bj¨ornson 和 C. Cavdar，“面向中继辅助大规模MIMO的公平且高能效激活控制机制”，arXiv 预印本 arXiv:2504.03428，2025年4月。 [在线]。网址：https://arxiv.org/abs/2504 03428 。[19] A. Chowdhury 和 E. G. Larsson，“双天线中继器辅助双基地 MIMO 感知与通信一体化的性能研究”，arXiv 预印本 arXiv:2511.17980，2026年1月。 [在线]。网址：https://arxiv.org/abs/2511 17980 。J. Bai、A. Chowdhury、A. Hansson 和 E. G. Larsson，“中继器集群辅助蜂窝系统：交互稳定性与性能分析”，IEEE 无线通信汇刊，第 25 卷，第 10 018–10 034 页，2026年1月。[20] A.-T. Le、T.-H. Vu、N. H. Tu、T. N. Nguyen、L.-T. Tu 和 M. Voznak，“物联网中的有源可重构中继器辅助 NOMA 网络：可靠性、安全性与隐蔽性”，IEEE 物联网汇刊，第 12 卷，第 7 期，第 8759–8772 页，2025年4月。[21] E. G. Larsson 和 J. Bai，“交互式无线中继器的稳定性分析”，载于 IEEE 无线通信信号处理进展国际研讨会（SPAWC）论文集，意大利卢卡，2024年9月。[24] F. I. G. Carvalho、R. V. de O. Paiva、T. F. Maciel、V. F. Monteiro、F. R. M. Lima、D. C. Moreira、D. A. Sousa、B. Makki、M. ° Astr¨om 和 L. Bao，“网络控制中继器简介”，IEEE 通信标准杂志，2025年8月，提前在线发表。[25] R. Nie、L. Chen、N. Zhao、Y. Chen、F. R. Yu 和 G. Wei，“存在互易性失配的中继系统：影响分析与校准”，IEEE 通信汇刊，第 68 卷，第 7 期，第 4035–4049 页，2020年。[26] Y. Ma、D. Zhu、B. Li 和 P. Liang，“中继器增强型大规模MIMO系统中的信道估计误差与波束成形性能”，载于 IEEE 第 26 届年度国际个人、室内与移动无线电通信研讨会（PIMRC）论文集，中国香港，2015年8月。[27] E. G. Larsson、J. Vieira 和 P. Frenger，“双天线中继器的互易性校准”，IEEE 无线通信快报，第 13 卷，第 6 期，第 1606–1610 页，2024年。[28] K. Ito、T. Takahashi、S. Ibi 和 S. Sampei，“采用非正交导频的相关大规模 MIMO 的贝叶斯联合信道与数据估计”，载于 IEEE 国际通信会议（ICC）论文集，加拿大蒙特利尔，2021年6月。[29] T. Takahashi、H. Iimori、K. Ishibashi、S. Ibi 和 G. T. F. de Abreu，“毫米波 MIMO 系统中用于联合信道跟踪与数据检测的贝叶斯双线性推断”，IEEE 无线通信汇刊，第 23 卷，第 9 期，第 11 136–11 153 页，2024年。[30] K. R. R. Ranasinghe、H. Seok Rou、G. Thadeu Freitas de Abreu、T. Takahashi 和 K. Ito，“双重色散信道中 AFDM 系统的联合信道、数据与雷达参数估计”，IEEE 无线通信汇刊，第 24 卷，第 2 期，第 1602–1619 页，2025年。[31] U. Gustavsson、C. Sanch´ez-Perez、T. Eriksson、F. Athley、G. Durisi、P. Landin、K. Hausmair、C. Fager 和 L. Svensson，“硬件损伤对大规模MIMO的影响”，载于 IEEE 全球通信会议研讨会（GLOBECOM WS）论文集，2014年。[32] F. Athley、G. Durisi 和 U. Gustavsson，“采用硬件损伤和不同信道模型的大规模MIMO分析”，载于欧洲天线与传播会议（EuCAP）论文集，葡萄牙里斯本，2015年。[33] K. Ito、T. Takahashi、S. Ibi 和 S. Sampei，“用于采用非正交导频的大规模 MIMO 检测的双线性高斯置信传播”，IEEE 通信汇刊，第 72 卷，第 2 期，第 1045–1061 页，2024年。[34] H. Iimori、T. Takahashi、K. Ishibashi、G. T. F. de Abreu、D. Gonz´alez G. 和 O. Gonsa，“超大规模MIMO 系统的联合活动与信道估计”，IEEE 无线通信汇刊，第 21 卷，第 9 期，第 7253–7270 页，2022年。[35] A. Chockalingam 和 B. S. Rajan，《大规模MIMO 系统》。剑桥大学出版社，2014年。[36] J. Vieira、F. Rusek 和 F. Tufvesson，“基于天线耦合的大规模MIMO互易性校准方法”，载于 IEEE 全球通信会议（GLOBECOM）论文集，美国奥斯汀，2014年12月。