面向簇式大规模机器类通信的联合活动检测与信道估计

Leatile Marata，IEEE 会员，Onel Luis Alcaraz L´opez，IEEE 会员，Andreas Hauptmann，IEEE 高级会员，Hamza Djelouat，IEEE 学生会员，以及 Hirley Alves，IEEE 会员

arXiv:2305.02935v2 [eess.SP] 3 Oct 2023

摘要——压缩感知多用户检测（CS-MUD）算法在优化大规模机器类通信（mMTC）的免授权（GF）非正交多址（NOMA）中发挥着关键作用。然而，当前的 CS-MUD 算法无法实现高效并行化，随着所部署机器型设备（MTD）数量的增加，联合活动检测与信道估计（JADCE）的实现将带来高昂的计算开销。为解决这一问题，本文提出了新型 JADCE 算法，通过利用导频序列的结构，可针对不同的 MTD 簇并行应用。这两种算法分别为近似误差方法（AEM）-交替方向乘子法（ADMM）和 AEM-稀疏贝叶斯学习（SBL）。以归一化均方误差和漏检概率表示的结果表明，与传统算法相比，这些算法具有相当的性能。然而，AEM-ADMM 和 AEM-SBL 算法的计算复杂度和运行时间均显著降低，从而有助于提升网络可扩展性。

I. 引言

DETECTION、信道估计和数据解码是无线通信网络 [1]–[3]中接收机执行的基本操作。然而，在以往的无线通信系统（即第四代（4G）及更早的系统）中，为这些操作设计的大多数算法都专门针对下行人类通信（HTC）[4], [5]进行了优化。随着通信技术的发展，新的通信标准，即第五代（5G）及更高版本（5GB），原生支持一类称为机器类设备（MTD）的新型设备，这些设备在物联网（IoT）范式下执行各种感知任务 [6]–[8]。值得注意的是，机器型设备受到能量受限的约束，但在某些情况下，它们需要部署在偏远地区，无法方便地进行充电。因此，机器型设备被设计为通过仅在感知到数据后切换至主动传输模式，并在没有数据时保持休眠模式来节省能量。这种间歇性运行模式降低了能耗。

Leatile Marata、Onel Luis Alcaraz L´opez、Hamza Djelouat 和 Hirley Alves 任职于芬兰奥卢大学无线通信中心——无线电技术，邮政编码为 FI-90014。电子邮件：{leatile.marata,onel.alcarazlopez,hamza.djelouat, hirley.alves}@oulu.fi。Andreas Hauptmann 任职于芬兰奥卢大学数学科学研究组（邮政编码：FI-90014），同时任职于英国伦敦大学学院计算机科学系，地址为英国伦敦 WC1E 6BT。电子邮件：andreas.hauptmann@oulu.fi。

本研究得到芬兰科学院资助（资助编号：319485、340171、346208（6G旗舰计划）、338408、353093）。Leatile Marata 的研究部分得到 Riitta ja Jorma J. Takanen 基金会、芬兰技术促进基金会以及博茨瓦纳国际科学技术大学的资助。芬兰技术促进基金会部分资助了 Onel L´opez 的研究。Tauno T¨onning 基金会部分资助了 HamzaDjelouat 的研究。

从而产生间歇性的上行链路流量，这与 HTC 的常规做法不同。由于机器型设备的运行方式与现有 HTC 设备不兼容，直接将现有接收算法应用于大规模机器类通信（mMTC）网络，可能会降低通信性能。幸运的是，采用低复杂度传输方案可以同时解决上述问题，其中免授权（GF）非正交多址接入（NOMA）发挥着关键作用。

免授权非正交多址接入技术已被提出，作为一种低复杂度传输方案，用于机器型设备 [9], [10] 的非协调传输。在这些方案中，活跃设备无需获得基站 (BS) 的许可即可传输数据，从而绕过与握手/调度过程相关的信令开销，进而减少通信开销和接入时延 [9], [10]。然而，缺乏调度以及不可避免地使用非正交导频序列，会导致碰撞和多用户干扰 (MUI) 增加。免授权非正交多址接入使用效率低下，会对前述操作（检测、信道估计和数据解码）造成不利影响，这是免授权非正交多址接入的主要缺点之一。归根结底，免授权非正交多址接入方案的性能取决于能否有效解决碰撞和多用户干扰 [11]。

对高效免授权非正交多址接入的需求，推动了采用压缩感知进行多用户检测，以实现联合活动检测与信道估计和/或无源随机接入 [12]。前者关注用户识别和数据解码，而后者关注对传输数据进行解码，而非识别实际发送数据的机器型设备 [13]。因此，联合活动检测与信道估计适用于状态更新场景，其中包含不同类型的消息，例如，不同的机器型设备分别用于感知环境湿度、酸度和温度。相反，无源随机接入可用于多个机器型设备传输关于同一物理现象的观测值的场景，例如在智能工厂中传输温度测量值，从而获取该现象的平均信息。

总体而言，JADCE 和 URA 都依赖这样一个事实：尽管机器型设备的数量庞大，但在给定的相干间隔（CI）内，同时处于激活状态的机器型设备数量相对较少。例如，预计未来网络每平方公里将容纳多达 1000 万台机器型设备[2]，，而同一时间只有其中很小一部分处于激活状态 [14]。对活跃 MTD 的识别自然可以被表述为一个压缩感知多用户检测问题，其计算复杂度可能很高 [15], [16]。然而，随着物联网部署的不断增加，一些机器型设备

势必会表现出相似的特征和性能要求。这些相似性自然有助于形成簇 [17]–[20]。基本上，可以根据信道统计特性（例如信道协方差矩阵）、性能要求、业务特征或激活概率等因素对机器型设备进行分簇[20]。鉴于这些因素，对机器型设备进行分簇有助于优化资源分配，从而使网络设计更加灵活且具备更强的可扩展性。这可以简化某些压缩感知多用户检测问题，并促进面向大规模机器型设备的免授权非正交多址接入。

目前有一项值得关注的研究，旨在开发可扩展的压缩感知多用户检测算法，以优化免授权非正交多址接入。通常，这些方法采用大规模多输入多输出（mMIMO）技术，从而实现分布式或并行信号处理。为了支持海量机器型设备，大多数研究依赖于基于完全非正交序列设计的导频数据 [21]–[25]。尽管导频序列的非正交性对于服务海量机器型设备至关重要，但也可以设计出能够划分为有限个正交子空间的导频序列，以区分或实现不同的簇。利用这些正交子空间，可以并行运行压缩感知多用户检测算法，同时确保簇间不存在多用户干扰。最终，这减少了算法并行运行期间对信息交换的需求。同样重要的是，这类导频序列必须采用非独立同分布（i.i.d.）序列构造，例如哈达玛、Zadoff Chu 和傅里叶矩阵；这些序列均符合 3GPP Release 17 标准 [26]。正如 Liu 等人 [27], 所指出的，使用独立同分布导频序列并不切实际，而且迄今为止设计的大多数基于这一假设的压缩感知多用户检测算法，在实际场景中都会面临挑战。因此，开发实用方法至关重要，这些方法应能在考虑大规模机器类通信中存在簇的情况下，高效实现压缩感知多用户检测算法。尽管现有研究为并行联合活动检测与信道估计设计奠定了良好基础，但据我们所知，目前还没有针对簇化机器类型通信、纳入正交导频子空间的联合活动检测与信道估计算法。这正是本文的研究目标。所提算法具有若干优势：i）降低多用户干扰；ii）提高资源利用效率；iii）可扩展性；iv）网络设计灵活性；所有这些都有助于实现高效的信号恢复。为提供背景，我们简要回顾相关研究。

A. 相关文献

近年来，基于大规模多输入多输出（mMIMO）的大规模机器类通信（mMTC）已成为一个活跃的研究领域。mMIMO 对 mMTC 至关重要，因为它能够提高频谱效率、数据速率和链路可靠性 [28]。目前，已有一些研究致力于利用 mMIMO 为免授权非正交多址接入（GF-NOMA）开发高效的压缩感知多用户检测算法。例如，He 等人提出了一种基于期望传播（EP）的分布式检测算法 [29] ，以促进其在中央处理单元（CPU）中的集成。该研究还对分布式无小区多输入多输出（CF-MIMO）中的 EP 进行了性能分析。然而，尽管该方案提高了检测性能，但也增加了 CPU 的计算复杂度。同样，Li 等人在 [30] 中提出了一种基于协方差的设备活动检测算法，该算法利用正交导频序列来降低多用户干扰（MUI）。与现有研究相比，他们的结果表明，在低信噪比（SNR）和较短导频长度下，性能得到了提升。另一方面，Ganesan 等人提出了一种用于 CF-MIMO 的基于最大似然（ML）的设备检测算法 [23]。结果表明，与共址多输入多输出相比，采用 CF-MIMO 时性能有所提升。总体而言， [23], [30] 中提出的 ML 算法具有较高的计算复杂度。

为了降低接收机的计算复杂度，Donoho 等人在 [31] 首次提出的近似消息传递（AMP）算法在不同场景下得到了广泛研究。例如，Bai 等人在 [32] 提出了一种基于似然比的分布式近似消息传递算法，并融入了状态演化的结构。然而，需要注意的是，大多数近似消息传递算法只有在导频矩阵的列服从高斯分布且彼此不相关时，才能保证收敛。一般而言，近似消息传递算法的这一特性限制了其实际应用。针对这一挑战，Rangan 等人在 [33] 提出了向量近似消息传递（VAMP），从而保证算法在导频序列列具有更广泛结构的情况下收敛。类似地，Ma 等人在 [34], 提出了正交近似消息传递（OAMP），该算法能够在导频序列由部分正交矩阵生成时实现收敛。因此，VAMP 和 OAMP 都已被广泛应用于联合活动检测与信道估计问题，例如在 [35], 中，Cheng 等人提出使用 OAMP 解决空时相关信道问题。尽管如此，AMP、VAMP 和 OAMP 都是为高维问题设计的，因此应用于低维问题时表现不佳 [10], [35]。

值得注意的是， [23], [28]–[30], [32], [35] 均假设设备已实现同步，这促使 Li 等人 [36] 提出了一种适用于无小区大规模多输入多输出系统的异步设备活动检测方法，并对基站与中央处理器之间的通信进行了优化。根据 [36], 的结果可以明显看出，去中心化算法的瓶颈之一在于不同子处理器之间信令增加所导致的通信开销。类似地，Chen 等人在 [37]中提出了一种适用于无小区大规模多输入多输出的结构化大规模接入方法，采用了逐组算法和 IB-K-means 聚类算法。他们的工作表明，所提出的导频分配策略具有更高的频谱效率，优于传统导频分配。此外，Figueredo 等人在 [38] 中开展了一项可行性研究，通过对机器型设备进行聚类，使其能够共享相同的时频资源块，从而提升系统容量。Iimori 等人在 [39] 中提出了一种双线性消息传递算法，该算法利用超大规模多输入多输出系统子阵列中的稀疏性，高效地检测设备簇。

尽管利用正交导频子空间捕获机器型设备的簇，有望促进 JADCE 算法的高效并行实现，但文献中尚未对此进行探索。受 [40], 中 Marata 等人提出的若干导频设计策略的启发，

为实现不同业务的友好共存，我们提出了新颖的压缩感知多用户检测算法，利用大规模机器类通信场景中机器型设备簇的导频结构。

B. 主要贡献

我们考虑一个为异构机器型设备簇[1]提供服务的大规模多输入多输出网络。通过利用正交导频子空间来刻画机器型设备的异构特性，我们提出了一个联合活动检测与信道估计问题，并采用并行算法求解。请注意，与 [23], [28]–[30], [32], 等研究不同，本研究采用的压缩感知多用户检测算法即使在并行实现中，也会交换部分迭代步骤。我们的主要贡献如下:

- 我们基于导频子空间构建联合活动检测与信道估计问题，即每个簇中机器型设备的大量非正交导频序列均由正交子空间生成。所提出建模方式的主要优势包括：稀疏信号恢复算法的并行实现、网络设计灵活性以及可扩展性。

- 我们提出了数据驱动算法，采用逆问题领域中建立的近似误差方法（AEM），以处理感知矩阵和似然函数中的误差 [41]–[43]。在本文中，AEM用于处理理想测量与各簇中用于执行联合活动检测与信道估计的测量之间的不匹配。首先，我们提出AEM-交替方向乘子法（ADMM），利用学习得到的不匹配统计特性执行迭代软阈值。其次，我们提出近似误差稀疏贝叶斯学习算法，该算法在贝叶斯框架中利用修正似然函数。AEM-ADMM不考虑先验分布，因此适用于参数分布未知的情形。另一方面，AEM-SBL依赖统计分布，因此能够利用更多信息来提升联合活动检测与信道估计的性能。

- 我们将所提出的 JADCE框架与不采用聚类的传统方法进行比较，并对其性能进行数值量化。具体而言，我们表明，所提出的算法在信道估计精度和检测能力方面可媲美经典方法，同时运行时间更短。

C. 组织结构与符号说明

本文其余部分的组织结构如下。第二节介绍系统模型。第三节阐述基于簇的设备活动检测问题。第四节提出该问题的解决方案，第五节则展示结果并进行讨论。最后，第六节总结全文，并探讨未来的一些研究方向。

符号说明：粗体小写和大写字母分别表示列向量和矩阵。此外，a_i and

[1]此处，异质性指机器型设备在特征和性能要求方面的差异。

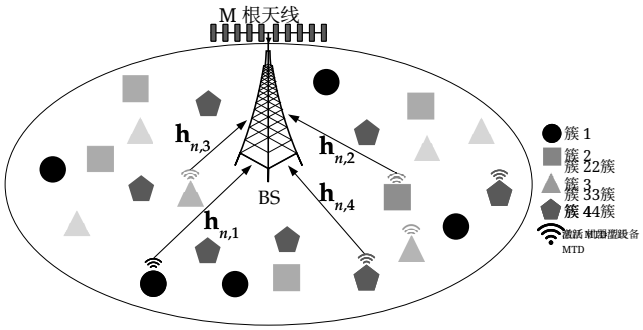

图1：一种大规模机器类通信场景，其中一个配备 M根天线的基站为分成 G 个簇的N 个机器型设备提供服务，其中平均共有 ϵN 个机器型设备处于激活状态。

$a_{i,j}$ 分别是矩阵 A 的第 i列和第 i行、第 j列的元素，而 a_i 是向量 a 的第 i个元素。上标$(\cdot)^*$、$(\cdot)^{\mathrm{T}}$和$(\cdot)^{\mathrm{H}}$ 分别表示共轭、转置和共轭转置运算。对于矩阵和向量，帽符号表示估计值，例如，$\hat{}x$ 是 x的估计值。此外，$\mathbb{C}$ 和 $\mathbb{R}$ 分别指复数域和实数域。我们用 $\mathcal{CN}(\mathrm{a}, \mathrm{B})$表示均值为 a、协方差为 B 的圆对称复高斯分布，而 $\mathbb{E}\{\cdot\}$ 和$\mathbb{V}\{\cdot\}$ 分别表示期望算子和协方差算子。此外， $\mathcal{U}(a,b)$表示界为 a 和 b的均匀分布。$\mathrm{diag}\{a_1, a_2, \cdots, a_n\}$生成一个对角矩阵，其主对角线元素为 $a_1, a_2, \cdots, a_n$。最后， $\|\cdot\|_F$, $\|.\|_p$ 和$\|\cdot\|_{n,p}$ 分别表示弗罗贝尼乌斯范数、 ℓ_p范数和混合 n,p 范数；随机变量的概率分布定义为 $\mathcal{P}(\cdot)$，而$\mathcal{P}(\cdot|\cdot)$表示条件概率。

II. 系统模型

我们考虑图1所示的上行大规模MIMO场景，其中一个配备有一组 M 天线的基站，即$\mathcal{M} = \{1, \cdots, M\}$，为一组 $\mathcal{N} = \{1, \cdots, N\}$静止的机器型设备提供服务。其中，基数为 $K = |\mathcal{K}|$的子集 $\mathcal{K} \subset \mathcal{N}$处于激活状态，且基站对此并不知情。假设机器型设备以概率 $\epsilon \ll 1$处于激活状态，因此，在给定的相干间隔内，网络中活跃设备的平均数量为 ϵN。此外，$\mathcal{N}$ 被划分为一组簇，簇的子索引为 $\mathcal{G} =\{1, \cdots, G\}$，其中 $G \leq N$，且每个机器型设备仅属于一个簇 $\mathcal{C}_g \subseteq \mathcal{N}$， $g \in \mathcal{G}$。第 g个簇的基数定义为 N_g，使得$\sum_{g=1}^{G} N_g = N$。

我们还假设信道为准静态块衰落信道，即在每个由 T 个符号组成的相干间隔（CI）内信道保持不变，而在不同相干间隔之间独立变化。我们假设基站（BS）仅掌握大尺度信道状态信息（CSI），即路径损耗信息，而不掌握瞬时信道状态信息。给定相干间隔内， g簇中的第 n个机器型设备（MTD）与基站之间的上行信道定义为 $\mathrm{h}_{n,g} \sim \mathcal{CN}(\mathrm{0}, \mathrm{Q}_{n,g}) \in \mathbb{C}^{M\times 1}$，其中 $\mathrm{Q}_{n,g}$是信道协方差矩阵。令 $\alpha_{n,g}$ 表示 g簇中设备 n 的激活状态，

$$\alpha_{n,g} = \begin{cases} 1, \text{ if device } n \text{ in the } g\text{-th cluster is active} \\ 0, \text{ otherwise} \end{cases}, \quad (1)$$

因此，整个网络的活动指示器可表示为向量 $\boldsymbol{\alpha}=[\boldsymbol{\alpha}_1^{\mathrm{T}},\boldsymbol{\alpha}_2^{\mathrm{T}},\cdots,\boldsymbol{\alpha}_G^{\mathrm{T}}]^{\mathrm{T}} \in \{0,1\}^{N\times 1}$，其中 $\boldsymbol{\alpha}_g = [\alpha_{g,1},\alpha_{g,2},\cdots,\alpha_{g,N_g}]^{\mathrm{T}} \in \{0,1\}^{N_g\times 1}, \forall g$。每个相干间隔（CI）的第一阶段（定义为前 L 个符号）对应于元数据处理块，即执行联合活动检测与信道估计（JADCE）的阶段，而其余 $T-L$ 个符号用于传输目标消息（有效载荷）。因此，为每个机器型设备（MTD）预先分配一个导频序列 $\mathbf{s}_{n,g} \in \mathbb{C}^{L\times 1}$，基站和机器型设备均知晓该序列，以用于联合活动检测与信道估计（JADCE）[2]。然而，由于机器型设备数量庞大，每个导频序列的长度通常远小于设备总数，即 $L \ll N$。因此，基站必须为所有机器型设备预先分配非正交导频序列，并利用压缩感知多用户检测（CS-MUD），根据训练阶段接收到的信号来求解联合活动检测与信道估计问题，该信号表示为

$$\begin{aligned}\mathbf{Y} &= \sum_{\forall n,g} \alpha_{n,g}\sqrt{p_{n,g}}\mathbf{s}_{n,g}\mathbf{h}_{n,g}^{\mathrm{T}} + \mathbf{W} \\ &= \sum_{g=1}^{G} \mathbf{S}_g\mathbf{X}_g + \mathbf{W} \\ &= \mathbf{S}\mathbf{X} + \mathbf{W}, \end{aligned} \tag{2}$$

其中，$\mathbf{S}_g=[\mathbf{s}_{n,g},\cdots,\mathbf{s}_{N_g,g}]$, $\mathbf{X}_g=[\mathbf{x}1,_g,\cdots,\mathbf{x}_{N_g,g}]^{\mathrm{T}}$ $\forall g$, $\mathbf{S} = [\mathbf{S}_1^{\mathrm{T}},\cdots,\mathbf{S}_G^{\mathrm{T}}]^{\mathrm{T}}$, 和 $\mathbf{X} = [\mathbf{X}_1^{\mathrm{T}},\cdots,\mathbf{X}_G^{\mathrm{T}}]^{\mathrm{T}}$。同时，$\mathbf{W} \in \mathbb{C}^{L\times M}$ 是接收机噪声，其各列为独立同分布的 $\{\mathbf{w}_m\} \sim \mathcal{CN}(0,\sigma^2\mathbf{I}) \in \mathbb{C}^{L\times 1}$；而 $\mathbf{x}_{n,g} = \sqrt{p_{n,g}}\alpha_{n,g}\mathbf{h}_{n,g} \in \mathbb{C}^{M\times 1}$ 和 $p_{n,g}$ 分别表示第 n 个设备在第 g 个簇中的有效行稀疏信道向量和发射功率。

由 (2) 可知，如果 $\{\mathbf{S}_g\}, \forall g$, 由正交子空间构成，则可以通过分别检测各个簇，完全分散化 JADCE 过程，从而降低问题的复杂度。为此，接下来我们将讨论如何利用正交子空间构造 $\mathbf{S}_g, g=1,\cdots,G,$，以实现完全分散式检测和信道估计。

III. 将分散式检测作为压缩感知问题

为了实现去中心化，我们假设与每个簇相关联的导频序列 $\{\mathbf{S}_g \in \mathbb{C}^{L\times N_g}\}$ 严格生成为正交基矩阵的线性组合。请注意，正交基矩阵的列是相互正交（垂直）的向量 [44]。因此，正交基矩阵为处理结构化导频设计提供了良好的基础。用于设计 MTC 导频序列的正交基矩阵包括单位矩阵、哈达玛矩阵和傅里叶矩阵等 [45]。然而，单位矩阵缺乏感知矩阵所需的多样性 [46]。尽管傅里叶矩阵和哈达玛矩阵都可以用作正交基矩阵，但我们采用了哈达玛矩阵，因为它具有良好的计算和存储特性。设 $\mathbf{B} \in \mathbb{C}^{L\times L}$ 为这样一个方阵，其列可以划分为 G 个不同的正交矩阵，即，$\mathbf{B} = [\mathbf{B}_1,\mathbf{B}_2,\cdots,\mathbf{B}_G] \in \mathbb{C}^{L\times L}$, $\left\|\mathbf{B}_i^{\mathrm{H}}\mathbf{B}_j\right\|_F = 0, \forall i \neq j$。因此，第 g 个簇的导频序列 $\mathbf{S}_g$ 由 $\mathbf{B}_g \in \mathbb{C}^{L\times\kappa_g}$ 生成，其中 κ_g 表示用于第 g 个簇的 $\mathbf{B}_g$ 的列数，使得 $\sum_{\forall g}\kappa_g \leq L$。

如第一节所述，机器型设备的数量通常极其庞大，因此 $N_g \geq L$ 是一个合理的假设。因此，生成大型 $\mathbf{B}_g$在计算上难以承受。此外，相干间隔始终有限，因此无法为属于同一簇的机器型设备分配相互正交的导频序列。由于条件 $N_g \geq L$，有必要生成每个 $\mathbf{S}_g$，同时保证每个簇都能实现信号恢复；这是传感（测量）矩阵设计中的一个基本问题 [47]。从压缩感知的角度来看，如果矩阵 $\mathbf{S}_g$ 满足受限等距性质（RIP），则可保证信号恢复，其正式定义如下：

$$(1-\delta_\epsilon)\left\|\mathbf{X}_g\right\|_F^2 \leq \left\|\mathbf{S}_g\mathbf{X}_g\right\|_F^2 \leq (1+\delta_\epsilon)\left\|\mathbf{X}_g\right\|_F^2, \forall \mathbf{X}_g, \tag{3}$$

使得 $\left\|\mathbf{X}_g\right\|_{2,\,0} \leq K_g$，其中 K_g 是给定簇中的活跃设备平均数量，$\delta_\epsilon > 0$ 是受限等距常数 [48]。

备注 1：RIP 可以理解为矩阵 $\mathbf{S}_g$ 将 $\mathbf{X}_g$ 映射到测量空间，同时保持 $\mathbf{X}_g$ 中不同样本之间的间隔的能力。这使得无需歧义即可恢复 $\mathbf{X}_g$ 中的不同样本。

根据备注 1，且不失一般性，注意到，为簇内两个不同的 MTD 创建导频序列，即 $\mathbf{s}_{i,\,g}$ 和 $\mathbf{s}_{j,\,g}$，其中 $i \neq j$，需要最大化两条不同导频序列之间的最小距离，即

$$\begin{aligned} \underset{\mathbf{S}_g}{\text{maximize}} \quad & \min_{1\leq i<j\leq|\mathcal{C}_g|} d\left(\mathbf{B}_g\mathbf{z}_i,\mathbf{B}_g\mathbf{z}_j\right) && \text{(4a)} \\ \text{subject to} \quad & \|\mathbf{z}_i\|_0 = \|\mathbf{z}_j\|_0, && \text{(4b)} \\ & s_{k,g} = \mathbf{B}_g\mathbf{z}_k, && \text{(4c)} \end{aligned}$$

其中 $d(\cdot,\cdot)$ 是两项之间的一般距离度量，而 $\mathbf{z}_k \in \mathbb{C}^{\kappa_g\times 1}$, $k \in \mathcal{C}_g$ 是一个包含随机合并权重且基数经过优化的向量。注意，$\mathbf{z}_k$ 必须是稀疏的，以确保良好的可检测性，因此 $\mathbf{s}_{k,\,g} = \mathbf{B}_g\mathbf{z}_k, k \in \mathcal{C}_g$，而等式约束 (4b) 确保了公平性。

由于 (4) 中的问题具有组合性质，因此是 NP 难问题。为了求得其近似解，我们利用 [40] 的第IV-B节中讨论的过程。这将为每个簇生成 $\{\mathbf{S}_g \in \mathbb{C}^{L\times N_g}\}$，从而得到网络中所有设备的拼接导频矩阵 $\mathbf{S}=[\mathbf{S}_1,\mathbf{S}_2,\cdots,\mathbf{S}_G] \in \mathbb{C}^{L\times N}$[3]。这实质上使不同簇的导频序列彼此正交[4], , 即$|\mathbf{S}_i^{\mathrm{H}}\mathbf{S}_j|_F = 0$, $\forall i \neq j$。为了阐明如何为第 g 个簇生成导频序列，我们在下面给出一个简单示例。假设 B 由顶部的方程 (5) 给出。

[2]鉴于基站先验已知导频序列，因此可以优化其结构，以改进接收算法。为此，我们还在第三节提出了一个组合优化问题，基站可以通过求解该问题来提升检测能力。

[3]为简化表述，假设矩阵拼接按照 $g=1,\cdots,G$ 的顺序进行。4

导频序列可以在较长时间内保持固定，以降低为每个机器型设备指示导频序列索引所带来的计算复杂度。在这种情况下，机器型设备可以在本地存储器中存储导频序列列表，而基站可以使用 $\log_2 N$ 比特指示 N 索引。或者，对于采用 d 个符号的调制方案，基站可以使用 $L\log_2 d$ 比特指示序列。

$$\mathbf{B}=\underbrace{\begin{pmatrix}1+j & 1+j & 1+j\\ 1+j & -1-j & 1+j\\ 1+j & 1+j & -1-j\\ 1+j & -1-j & -1-j\\ 1+j & 1+j & 1+j\\ 1+j & -1-j & 1+j\\ 1+j & 1+j & -1-j\\ 1+j & -1-j & -1-j\end{pmatrix}}_{\mathbf{B}_1}\underbrace{\begin{pmatrix}1+j & 1+j & 1+j\\ -1-j & 1+j & -1-j\\ -1-j & 1+j & 1+j\\ 1+j & 1+j & -1-j\\ 1+j & -1-j & -1-j\\ -1-j & -1-j & 1+j\\ -1-j & -1-j & -1-j\\ 1+j & -1-j & 1+j\end{pmatrix}}_{\mathbf{B}_2}\underbrace{\begin{pmatrix}1+j & 1+j\\ 1+j & 1+j\\ 1+j & -1-j\\ 1+j & -1-j\\ 1+j & -1-j\\ 1+j & -1-j\\ 1+j & 1+j\\ 1+j & -1-j\end{pmatrix}}_{\mathbf{B}_3},\tag{5}$$

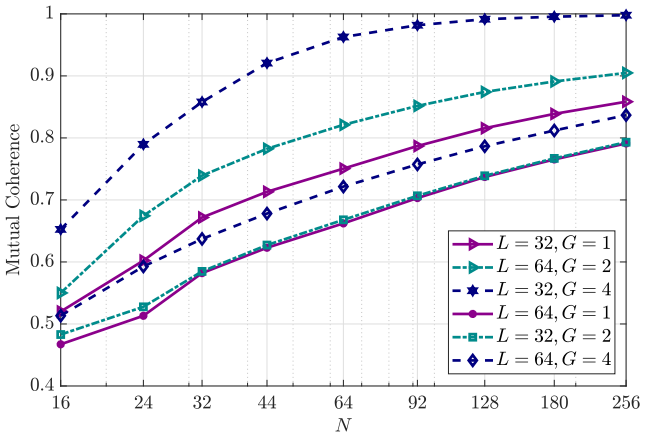

图2：互相干性随 L、 G和 N变化的函数关系。

下一页的内容，因此 $L = 8$、 $\kappa_1 = 3$、 $\kappa_2 = 3$ 和 $\kappa_3 = 2$。现在聚焦于簇 2，其中 S_2 由 $\mathbf{B}_2 \in \mathbb{C}^{L\times 3}$的列的线性组合构成。例如，设 $\mathbf{z}_1 = [1\ 0\ 1]^{\mathsf{T}}$ 和 $\mathbf{z}_2 = [0\ 1\ 1]^{\mathsf{T}}$、 $\bar{\mathsf{s}}_{1,2}$ 和 $\bar{\mathsf{s}}_{2,2}$ 为由以下内容生成的向量

$$\bar{\mathbf{s}}_{1,2}=\mathbf{B}_2\mathbf{z}_1=[2+2j,-2+2j,0,0,0,0,-2-2j,2+2j]^{\mathsf{T}},\tag{6}$$

$$\bar{\mathbf{s}}_{2,2}=\mathbf{B}_2\mathbf{z}_2=[2+2j,0,2+2j,0,-2-2j,0,-2-2j,0]^{\mathsf{T}},\tag{7}$$

这些变量经过归一化，即 $\mathsf{s}_{1,2}=\bar{\mathbf{s}}_{1,2}/\|\bar{\mathbf{s}}_{1,2}\|_2$ 和 $\mathsf{s}_{2,2}=\bar{\mathbf{s}}_{2,2}/\|\bar{\mathbf{s}}_{2,2}\|_2$，以构成 $\mathsf{S}_2 = [\mathsf{s}_{1,2}, \mathsf{s}_{2,2}, \cdots, \mathsf{s}_{N_g,2}]$。请注意，所得矩阵 $\{\mathsf{S}_g\}$ 可能具有高度相关的列，这会限制某些依赖近似消息传递框架的最先进的压缩感知多用户检测算法的适用性 [10], [49]。例如，我们在图2中展示了两列不同列之间的最大相关性，即互相干性随机器型设备数量（N）和簇数 G变化的情况。从结果可以看出，簇数相对较多且 L 相对较小时，互相干性可能增大。因此，对于给定数量的机器型设备，这类结果可以为合理选择 L 和G 提供指导。

A. 将设备活动检测视为压缩感知问题

根据定义（1）以及机器型设备（MTDs）的业务流量通常具有间歇性的事实，$\mathbf{X}$ 的恢复构成了一个压缩感知多用户检测（CS-MUD）问题，可以利用稀疏信号恢复概念来解决。具体而言，基站（BS）必须从压缩后的信号中识别出激活的设备。

在已知 $\mathbf{X}$ 具有行稀疏性的情况下，测量值 $\mathbf{Y}$。因此，我们可以定义一个通用的逆运算 f，其形式为 $\hat{\mathbf{X}} = f(\mathbf{Y}, \mathbf{S})$，用于将测量值映射到有效信道空间。此外，由于子空间之间具有正交性，如果能够精确获知 $\mathbf{Y}^g$，则可以进一步将该逆运算表示为 $\hat{\mathbf{X}}_g = f(\mathbf{Y}_g, \mathbf{S}_g)$。在此可以明显看出，$\hat{\mathbf{X}}$ 可以是诸如 [10], [50]，中所述的集中式问题的解，而 $\mathbf{X}_g$ 则可以为每个簇分别提供一个解。这种逆运算可以通过利用贝叶斯理论和/或松弛的凸优化框架来构建。

在贝叶斯表述中，采用伯努利-高斯混合分布对有效信道 $\mathbf{X}$ 在每个组级别以及组内的稀疏性进行建模，如下所示： [39]

$$\mathcal{P}(\mathbf{X})=\prod_{n,g}\Big((1-\epsilon)\delta(\mathbf{x}_{n,g})+\epsilon\mathcal{CN}(\mathbf{x}_{n,g};\mathbf{0},\tilde{\mathbf{Q}}_{n,g})\Big),\tag{8}$$

其中， $\delta(\cdot)$ 是狄拉克δ函数，以概率 $1-\epsilon$ 强制令 $\mathbf{x}_{n,g} = \mathbf{0}$，而 $\mathbf{Q}_{n,g} = p_{n,g}\tilde{\mathbf{Q}}_{n,g}$ 是有效协方差矩阵。从数学角度来看，准确掌握线性问题 (2) 和 (8)，就能够通过最大后验（MAP）估计计算贝叶斯最优 f，从而恢复 $\mathbf{X}$。因此，最优联合活动检测与信道估计算法通过以下方式选择一对 $\boldsymbol{\alpha}$ 和 $\mathbf{X}$:

$$\hat{\boldsymbol{\alpha}}=\underset{\boldsymbol{\alpha}}{\operatorname{argmax}}\int\mathcal{P}(\mathbf{X},\boldsymbol{\alpha}|\mathbf{Y})d\mathbf{X},\tag{9}$$

$$\hat{\mathbf{X}}=\underset{\mathbf{X}}{\operatorname{argmax}}\int\mathcal{P}(\mathbf{X},\boldsymbol{\alpha}|\mathbf{Y})d\boldsymbol{\alpha}\tag{10}$$

根据联合后验分布

$$\mathcal{P}(\mathbf{X},\boldsymbol{\alpha}|\mathbf{Y})\propto\mathcal{P}(\mathbf{Y}|\mathbf{X},\boldsymbol{\alpha})\mathcal{P}(\mathbf{X}|\boldsymbol{\alpha})\mathcal{P}(\boldsymbol{\alpha}).\tag{11}$$

然而，在接收机中，实际无法实现用于计算 (9) 和 (10) 的检测器。首先，缺乏关于机器型设备激活概率的准确信息，因此难以准确构造先验分布 $\mathcal{P}(\mathbf{X})$。其次，即使能够准确计算先验分布，(11) 中的边缘化仍需要针对大量变量进行维数极高、计算量难以承受的积分/求和，这在大规模机器类通信场景中尤为突出。因此，针对 (9) 和 (10) 的大多数求解方案都采用替代方法。

一种常见的方法是放宽复杂的边缘化—

方法是利用置信传播 [51]近似联合后验分布。这种方法通过和积算法得出高效的解决方案。另一种方法是利用平均场技术近似后验分布，并通过变分消息传递 [52], [53]求解。或者，也可以通过松弛的凸优化问题来求解。

$$\text{最小化}_{\mathbf{X}} \quad \frac{1}{2}\|\mathbf{Y}-\mathbf{SX}\|_F^2+\lambda\|\mathbf{X}\|_{2,1}, \tag{12}$$

其中，$\lambda \in \mathbb{R}^+$ 充当惩罚项，用于权衡测量保真度与混合$\ell_{2,1}$范数所体现的稀疏结构。不过，值得注意的是，对于大规模问题，式（12）的计算负担也可能很重，而且其效果取决于是否恰当选择惩罚项。与其基于贝叶斯的方法相比，这显然不是最优的。然而，它仍能取得合理的结果，并且已有一些研究致力于并行求解该问题，例如 [54], [55]。如前所述，由于这些并行实现依赖于分解目标函数的能力，而非可能将机器型设备的簇彼此隔离，因此无法充分发挥所提出的基于导频的簇模型的优势。最终，这些并行实现需要交换拉格朗日乘子的更新值。随后，我们将介绍受 AEM 启发的稀疏信号恢复方案，该方案在考虑簇存在的情况下执行联合活动检测与信道估计。

IV. 受近似误差法启发的稀疏信号恢复

考虑到问题 (2) 的构建方式使得 $\mathcal{C}_i$ 和 $\mathcal{C}_j$、$i \neq j$ 的导频序列 $\mathbf{S}_i$ 和 $\mathbf{S}_j$ 分别彼此正交，因此，联合活动检测与信道估计（JADCE）被分解为若干个较小的问题，每个簇对应一个问题。例如，第 g 个簇中的信号处理可以从

$$\tilde{\mathbf{X}}_g = \mathbf{S}_g^{\mathrm{H}}\mathbf{Y}. \tag{13}$$

有趣的是，如果设备遵循用户随机接入或具有相关活动，则可以使用 $\|\mathbf{X}_g\|_F \geq \zeta$ 估计其联合活动，其中 ζ 应大于但相对接近 $\frac{1}{\sqrt{L}}\sigma$，以获得良好的性能。同时处理用户随机接入和相关活动的能力，为导频序列分配提供了急需的灵活性。例如，可以灵活地为需要更多资源的设备分配更大的基矩阵 $\mathbf{B}_g$，例如超可靠低时延通信系统中使用的设备，正如 Lopez 等人在 [20] 中所讨论的那样。

需要注意，去相关步骤（13）分离出了不同的˜簇。然而，$\mathbf{X}_g$ 是通过将测量值反向投影到解空间中得到的，因此，在联合活动检测与信道估计（JADCE）中依赖这一统计量可能导致次优性能。这是因为现有的大多数算法都是专门设计用于利用 Y 所捕获的低维测量值 ˆ ˜来计算 X。显然，$\mathbf{X}_g$ 会导致信号中关键结构的丢失，而这些结构正是各个簇在进行联合活动检测与信道估计时所使用的。

备注 2：从式 (13) 可以看出，大多数压缩感知多用户检测算法都包含相关步骤 $\mathbf{S}_j^{\mathrm{H}}\mathbf{Y}$，$j \in \mathcal{G}$，其中五月婷婷

Algorithm 1: CB-SOMP, $\forall g \in \mathcal{G}$

Input: $\mathbf{Y}$, Δ

$\hat{\mathbf{X}}_g^{(0)} = \mathbf{0}$

$\mathbf{R}_g^{(0)} = \mathbf{Y}$, $t = 0$, $\mathcal{H}_g^{(0)} = \emptyset$

repeat

  $\mathbf{D}^{(t)} = \mathbf{S}_g^{\mathrm{H}}\mathbf{R}_g^{(t)}$

  $j^{(t)} = \mathrm{argmax}_j \left\{ \frac{\left\|\mathbf{d}_j^{(t)}\right\|_1}{\|\mathbf{s}_j\|_2} \right\}$

  $\mathcal{H}^{(t)} = \mathcal{H}^{(t)} \cup j^{(t)}$

  $\mathbf{X}^{(t)}_{[\mathcal{H}^{(t)}]} = \mathbf{S}^{\dagger}_{[\mathcal{H}^{(t)}]}\mathbf{Y}$

  $\mathbf{R}^{(t)} = \mathbf{Y} - \mathbf{S}\mathbf{X}^{(t)}$

  $t = t + 1$

until $\frac{\|\mathbf{X}^{(t)}-\mathbf{X}^{(t-1)}\|_F}{\|\mathbf{X}^{(t)}\|_F} < \Delta$;

Output: $\hat{\mathbf{X}}_{\mathbf{g}} = \mathbf{X}^{(t)}$

迭代步骤可以为基于簇的联合活动检测与信道估计提供一种朴素的解决方案。

根据备注 2，我们通过同时正交匹配追踪（SOMP） [56] 示例说明朴素解的实现。为将其与传统的 SOMP 实现区分开来，我们将基于簇的实现称为基于簇的同时正交匹配追踪（CB-SOMP），其流程概述于算法 1。在所有算法中，Δ是停止准则中使用的误差容限。

如前所述，反向投影可能会对接收机的性能产生不利影响。将 $\mathbf{X}_g$ 映射回各个簇的测量空间的一种经典解决方案˜是使用

$$\hat{\mathbf{Y}}_{\mathbf{g}} = (\mathbf{S}_g\mathbf{S}_g^{\mathrm{H}})^{-1}\mathbf{S}_g\tilde{\mathbf{X}}_g. \tag{14}$$

同时，类似于式（11），所得的简化 MAP 问题具有联合后验分布。

$$\mathcal{P}(\mathbf{X}_g, \boldsymbol{\alpha}_g|\hat{\mathbf{Y}}_g) \propto \mathcal{P}(\hat{\mathbf{Y}}_g|\mathbf{X}_g, \boldsymbol{\alpha}_g)\mathcal{P}(\mathbf{X}_g|\boldsymbol{\alpha}_g)\mathcal{P}(\boldsymbol{\alpha}_g). \tag{15}$$

需要注意的是，所得的 MAP 问题比不采用簇时处理的问题维度更低。尽管如此，一个主要缺点是，式 (4) 中的矩阵 $\mathbf{S}_g$ 秩非常低，因此其伪逆会导致测量失配，即 $\mathbf{Y}_g \neq \mathbf{Y}_g$，即使在无噪声条件下也是如此。受逆问题中的近似误差模型方法和 [42], 中贝叶斯近似误差模型的启发，我们通过施加数据驱动模型校正来处理这种失配。请注意，这种数据驱动校正基于统计学习，旨在估计失配误差的均值和协方差，这与机器学习技术不同。下一小节将简要介绍受近似误差模型启发的解决方案的背景。

A. AEM 概述

本小节简要介绍用于构建所提出的 JADCE 算法的 AEM [42], [43], [57],。ˆ AEM 用于处理伪逆运算 (14) 所导致的 $\mathbf{Y}_g$ 与 $\mathbf{Y}_g$ 之间的数值误差，尤其是在后文讨论的 $\mathbf{S}_g$ 为低秩矩阵时。

ˆ 因此，Y_g 不能可靠地测量与第 g 个簇对应的接收信号。然而，JADCE 的目标是利用一组测量值，并依赖它们理想的线性关系来恢复/估计隐变量 X_g 。该线性关系由测量矩阵 S_g 定义，假设其已预先充分知晓，并包含 X_g 如何映射为 Y_g 所需的全部信息。在无噪声场景中，这种关系可建模为

$$\mathbf{Y}_g = \mathbf{S}_g \mathbf{X}_g. \tag{16}$$

ˆ 为此，仅依靠 Y_g 恢复隐变量 X_g ，相当于使用了错误的线性模型，因此也等同于使用了错误的测量矩阵。

˜设 S^g 为未知且不准确的测量矩阵，则产生 Y_g 的关系可建模为

$$\hat{\mathbf{Y}}_g = \tilde{\mathbf{S}}_g \mathbf{X}_g, \tag{17}$$

并引入系统性模型误差

$$\Delta \mathbf{Y}_g = \mathbf{Y}_g - \hat{\mathbf{Y}}_g. \tag{18}$$

未能考虑此误差会对 JADCE 性能产生负面影响。幸运的是，可以借助逆问题领域的 AEM 概念来校正这一差异 [42], [57]。具体而言，通过使用经校正的测量矩阵 S_g求解˜X_g ，从而为压缩感知多用户检测提供一个通用框架，该框架适用于感知矩阵存在误差的情况[58]。本质上，我们可以将 S_g 近似为 $\mathrm{C}_g\mathrm{S}_g \approx \mathrm{S}_g$，其中 C_g 表示与模型误差 (18) 相关的校正项，并且可以通过统计训练获得。由于机器型设备被视为静止不动，因此可以离线获取每个 C_g 。此外，该训练过程不会增加 JADCE 解决方案的计算复杂度。具体而言，C_g 根据模型误差的协方差矩阵计算，即 $\mathbf{\Phi}_g$，如第IV-B节开头所述。首先，可以在每个簇中获取模型误差，并使用 τ 个训练样本进行考虑。

$$\mathbf{E}_g(i) = \mathbf{S}_g \mathbf{X}_g(i) - \hat{\mathbf{Y}}_g(i), i = 1, \cdots, \tau. \tag{19}$$

这里，$\mathrm{E}_g(i) = [\mathrm{e}_1(i) \cdots, \mathrm{e}_M(i)]$, 其中，$\tilde{\mathrm{e}}_m(i) = \mathrm{S}_g \mathrm{x}_m(i) - \mathrm{S}_g\mathrm{x}_m(i)$，表示第 m 根天线在第 i 个训练样本上的失配误差。通过使用 (19) 进行训练，第 m 根天线的平均失配误差为 $\boldsymbol{\mu}_m = \sum_{i=1}^{\tau} \mathrm{e}_m(i)/\tau$。随后，簇中第 m 根天线上的失配误差样本误差协方差矩阵为

$$\mathbf{\Omega}_m = \frac{1}{\tau - 1} \sum_{i=1}^{\tau} \mathbf{e}_m(i) \mathbf{e}_m(i)^{\mathrm{H}} - \boldsymbol{\mu}_m \boldsymbol{\mu}_m^{\mathrm{H}}, \tag{20}$$

当$\tau \to \infty$时，该矩阵收敛于总体误差协方差矩阵。

注 3：鉴于机器型设备处于静止状态，并依据大数定律，第 g 个簇的失配误差平均协方差为 $\mathbf{\Phi}_g = \mathbb{E}\{\Omega_m\}$。

根据上述说明，接下来我们将 C_g 引入优化问题（12），随后提出受 AEM 启发、针对每个簇的数据驱动型 JADCE 解决方案。

B. 通过 AEM-ADMM 求解

第一个受 AEM 启发的解决方案通过依赖迭代软阈值算法（ISTA）的 ADMM框架推导得出。根据对 AEM 的讨论，我们注意到，每个簇的噪声测量值与其伪逆的结果相关，具体为

$$\mathbf{Y}_g = \mathbf{S}_g \mathbf{X}_g + \mathbf{E}_g + \mathbf{W} = \hat{\mathbf{Y}}_g + \mathbf{E}_g \ , \forall g, \tag{21}$$

其中，E_g 是 (19) 中的模型误差，W 来自 (2)。据此，将噪声和模型误差造成的总误差定义为 $\mathbf{\Xi}_g = [\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_M]$, 使得 $\mathbf{\Xi}_g = \mathrm{E}_g + \mathrm{W}$。依据大数定律，我们将 $\mathbf{\Xi}_g$ 的每一列近似为一个高斯变量，即 $\boldsymbol{\xi}_m = \mathrm{e}_1 + \mathrm{w}_i \sim \mathcal{CN}(\boldsymbol{\psi}_m, \Omega_m)$, 其中 $\boldsymbol{\psi}_m$ 通过训练获得，与 $\boldsymbol{\mu}_m$类似。鉴于其高斯特性，$\boldsymbol{\xi}_m$ 具有精度矩阵，其乔列斯基分解为 $\mathrm{C}_g^{\mathrm{T}}\mathrm{C}_g = \mathbf{\Phi}_g^{-1}$，从而便于构建指数似然函数 [42]

$$\mathcal{P}(\mathbf{Y}_g | \mathbf{X}_g) \propto \exp\left(-\frac{1}{2} \left\| \mathbf{C}_g \left(\mathbf{S}_g \mathbf{X}_g - \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g \right) \right\|_F^2 \right), \tag{22}$$

其中 $\mathbf{\Psi}_g = [\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_M]$。因此，JADCE解为

$$\begin{aligned} \hat{\mathbf{X}}_g &= \underset{\mathbf{X}_g}{\operatorname{argmax}} \mathcal{P}(\mathbf{Y}_g | \mathbf{X}_g) \mathcal{P}(\mathbf{X}_g) \\ &\overset{(a)}{=} \underset{\mathbf{X}_g}{\operatorname{argmax}} \ln \mathcal{P}(\mathbf{Y}_g | \mathbf{X}_g) + \ln \mathcal{P}(\mathbf{X}_g) \\ &\overset{(b)}{=} \underset{\mathbf{X}_g}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{C}_g \left(\mathbf{S}_g \mathbf{X}_g - \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g \right) \right\|_F^2 + \lambda \| \mathbf{X}_g \|_{2,1}, \end{aligned} \tag{23}$$

其中，(a)源于采用对数形式，而(b)利用了(22)，同时将 $\lambda\|\mathrm{X}_g\|_{2,\ 1}$ 作为–ln$\mathcal{P}$(X_g)的近似值。请注意，λ 在测量保真度与解的稀疏性之间进行权衡。

为了通过ADMM框架 [54], [55],实现(23)，我们针对每个簇 $g \in \mathcal{G}$ 将(23)重新表述如下

$$\begin{aligned} \underset{\mathbf{X}_g, \mathbf{Z}_g}{\text{minimize}} \quad & \frac{1}{2} \left\| \mathbf{C}_g \left(\mathbf{S}_g \mathbf{Z}_g - \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g \right) \right\|_F^2 \\ & + \lambda \| \mathbf{X}_g \|_{2,1} + \frac{\rho}{2} \| \mathbf{X}_g - \mathbf{Z}_g \|_F^2 \quad \text{(24a)} \\ \text{subject to} \quad & \mathbf{X}_g = \mathbf{Z}_g, \quad \text{(24b)} \end{aligned}$$

其中，ρ为ADMM步长，$\mathrm{Z}_g = [\mathrm{z}_{n,\ g}, \cdots, \mathrm{z}_{N_g,\ g}]^{\mathrm{T}}$为辅助变量，通过Moreau-Yosida正则化 [54]实现对X_g 估计值的闭式更新。根据公式(24)，增广拉格朗日函数表示为 [54]

$$\begin{aligned} \mathcal{L}(\mathbf{X}_g, \mathbf{\Theta}_g) &\overset{(a)}{=} \frac{1}{2} \left\| \mathbf{C}_g \left(\mathbf{S}_g \mathbf{Z}_g - \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g \right) \right\|_F^2 + \lambda \| \mathbf{X}_g \|_{2,1} \\ &\quad + \mathbf{\Theta}_g^{\mathrm{T}} (\mathbf{X}_g - \mathbf{Z}_g) + \frac{\rho}{2} \| \mathbf{X}_g - \mathbf{Z}_g \|_F^2 \\ &\overset{(b)}{=} \frac{1}{2} \left\| \mathbf{C}_g \left(\mathbf{S}_g \mathbf{Z}_g - \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g \right) \right\|_F^2 + \lambda \| \mathbf{X}_g \|_{2,1} \\ &\quad + \left\| \mathbf{X}_g - \mathbf{Z}_g + \frac{\mathbf{\Theta}_g}{\rho} \right\|_F^2 - \frac{\| \mathbf{\Theta}_g \|_F^2}{2\rho}, \end{aligned} \tag{25}$$

其中，$\mathbf{\Theta}_g = [\boldsymbol{\theta}_{1,g}, \cdots, \boldsymbol{\theta}_{N_g,g}] \in \mathbb{C}^{M \times N_g}$是增广拉格朗日函数的对偶矩阵。($a$) 中 $\mathcal{L}$(X_g, $\mathbf{\Theta}_g$) 的表达式是标准的增广拉格朗日函数，而 (b) 是缩放拉格朗日形式 [54]。通过更新变量集，使用 AEM-ADMM 求解 JADCE

Algorithm 2: AEM-ADMM, $\forall g \in \mathcal{G}$

Input: $\hat{\mathbf{Y}}_g$, Δ, $\mathbf{C}_g$, $\mathbf{\Psi}_g$

Initialisation: $\mathbf{X}_g^{(0)} = \mathbf{0}$, $\mathbf{Z}_g^{(0)} = \mathbf{0}$, $\mathbf{\Theta}_g^{(0)} = \mathbf{0}$, $t = 0$

repeat

  Update $\mathbf{Z}_g^{(t+1)}$ using (30)

  Update $\mathbf{X}_g^{(t+1)}$ using (32)

  $\mathbf{\Theta}_g^{(t+1)} = \mathbf{\Theta}_g^{(t)} + \rho\left(\mathbf{X}_g^{(t+1)} - \mathbf{Z}_g^{(t+1)}\right)$

  $t = t + 1$

until $\frac{\left\|\mathbf{Z}_g^{(t+1)} - \mathbf{Z}_g^{(t)}\right\|_F}{\left\|\mathbf{Z}_g^{(t)}\right\|_F} < \Delta$;

Output: $\hat{\mathbf{X}}_g = \mathbf{Z}_g$

$\{\mathbf{X}_g, \mathbf{Z}_g, \mathbf{\Theta}_g\}$ 以交替方式进行。具体而言，在第 $(t+1)$ 次迭代中，各变量对应的子问题分别为

$$\begin{aligned}\mathbf{Z}_g^{(t+1)} =& \min_{\mathbf{Z}_g} \mathcal{L}\left(\mathbf{X}_g^{(t)}, \mathbf{Z}_g, \mathbf{\Theta}_g^{(t)}\right) \\ =& \min_{\mathbf{Z}_g} \frac{1}{2}\left\|\mathbf{C}_g\left(\mathbf{S}_g\mathbf{Z}_g - \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g\right)\right\|_F^2 \\ &+ \left\|\mathbf{X}_g^{(t)} - \mathbf{Z}_g + \frac{\mathbf{\Theta}_g^{(t)}}{\rho}\right\|_F^2, \end{aligned} \tag{26}$$

$$\begin{aligned}\mathbf{X}_g^{(t+1)} =& \min_{\mathbf{X}_g} \mathcal{L}\left(\mathbf{X}_g, \mathbf{Z}_g^{(t+1)}, \mathbf{\Theta}_g^{(t)}\right) \\ =& \min_{\mathbf{X}_g} \|\mathbf{X}_g\|_{2,1} + \frac{\rho}{2}\|\mathbf{X}_g - \mathbf{Z}_g^{(t+1)} + \frac{1}{\rho}\mathbf{\Theta}_g^{(t)}\|_F^2, \end{aligned} \tag{27}$$

$$\mathbf{\Theta}_g^{(t+1)} = \mathbf{\Theta}_g^{(t)} + \rho\left(\mathbf{X}_g^{(t+1)} - \mathbf{Z}_g^{(t+1)}\right), \tag{28}$$

因此，通过在保持其他变量不变的情况下，最小化关于 $\mathbf{Z}_g$ 的式 (26) 来更新 $\mathbf{Z}_g^{(t+1)}$。也就是说，对 $\mathbf{Z}_g$ 求导，将其设为零，并求解 $\mathbf{Z}_g$，如下所示

$$\begin{aligned}&\frac{\partial \mathcal{L}\left(\mathbf{X}_g^{(t)}, \mathbf{Z}_g, \mathbf{\Theta}_g^{(t)}\right)}{\partial \mathbf{Z}_g} \\ &= \mathbf{S}_g^{\mathrm{H}}\mathbf{C}_g^{\mathrm{H}}\mathbf{C}_g\mathbf{Z}_g + \rho\mathbf{Z} - \rho\mathbf{X}_g^{(t)} + \mathbf{\Theta}_g^{(t)} \\ &\quad - \mathbf{S}_g^{\mathrm{H}}\mathbf{C}_g^{\mathrm{H}}\mathbf{C}_g\hat{\mathbf{Y}}_g + \mathbf{S}_g^{\mathrm{H}}\mathbf{C}_g^{\mathrm{H}}\mathbf{C}_g\mathbf{\Psi}_g = \mathbf{0}, \end{aligned} \tag{29}$$

由此得到

$$\begin{aligned}\mathbf{Z}_g^{(t+1)} =& \left(\mathbf{S}_g^{\mathrm{H}}\mathbf{C}_g^{\mathrm{H}}\mathbf{C}_g\mathbf{S}_g + \rho\mathbf{I}_{N_g}\right)^{-1} \times \\ &\left(\rho\mathbf{X}^{(t)} - \mathbf{\Theta}_g^{(t)} + \mathbf{S}_g^{\mathrm{H}}\mathbf{C}_g^{\mathrm{H}}\mathbf{C}_g\hat{\mathbf{Y}}_g - \mathbf{S}_g^{\mathrm{H}}\mathbf{C}_g^{\mathrm{H}}\mathbf{C}_g\mathbf{\Psi}_g\right). \end{aligned} \tag{30}$$

类似地，计算 $\mathbf{X}_g^{(t+1)}$ 时，需要对 $\mathbf{X}_g$ 进行最小化。注意到式 (27) 是混合范数 $\|\mathbf{X}_g\|_{2,1}$ 的 Moreau包络，我们可以根据式 (30) 得到的结果更新 $\mathbf{X}_g$。因此，为了保证 JADCE 解的稀疏性，$\mathbf{X}_g$ 各行的 ℓ_2 范数必须是稀疏的。我们将 $\mathbf{X}_g$ 各行的范数记为 $\bar{\mathbf{x}}_g \in \mathbb{R}^{N_g \times 1}$，据此通过求解以下问题来更新 $\mathbf{X}_g$

$$\mathbf{X}_g^{(t+1)} = \min_{\mathbf{X}_g} \|\bar{\mathbf{x}}_g\|_1 + \frac{\rho}{2}\left\|\mathbf{X}_g - \mathbf{\Pi}_g^{(t+1)}\right\|_F^2, \tag{31}$$

其中 $\mathbf{\Pi}_g^{(t+1)} = \mathbf{Z}_g^{(t+1)} + \frac{1}{\rho}\mathbf{\Theta}_g^{(t)}$，并使用近端算子法 [54], 可得到 N_g 解耦解 [59]

$$\begin{aligned}\mathbf{x}_{n,g}^{(t+1)} &= \mathrm{prox}_{\lambda, \|\bar{\mathbf{x}}_g\|_1}(\mathbf{\Pi}_g^{(t+1)}, \rho) \\ &= \boldsymbol{\pi}_{n,g}^{(t+1)} \frac{\max\{\|\boldsymbol{\pi}_{n,g}^{(t+1)}\| - \rho^{-1}, 0\}}{\|\boldsymbol{\pi}_{n,g}^{(t+1)}\|}, n = 1, \cdots, N_g. \end{aligned} \tag{32}$$

算法2总结了 AEM-ADMM。请注意，AEM-ADMM 继承了经典ADMM [54] 的性质，因此收敛速度较慢，难以达到最佳精度；尽管如此，其大多数结果对于 JADCE框架而言是可以接受的。接下来，我们介绍 AEM-SBL，它是贝叶斯解，因此利用了观测值和先验的统计分布。

C. 通过 AEM-SBL 求解

第 IV-B 节提出的 AEM-ADMM 依赖于促进稀疏性的惩罚项，而非显式先验分布 $\mathcal{P}(\mathbf{X}_g)$，因此可能远非最优。为此，我们引入 AEM-SBL，这是一种受 AEM 启发的另一种 JADCE 解决方案，依赖于 SBL 框架 [22], [60]。凭借其贝叶斯特性，AEM-SBL 利用 $\mathcal{P}(\mathbf{X}_g)$，这是相较于 AEM-ADMM 的明显优势。AEM-SBL 使用联合后验分布

$$\begin{aligned}\mathcal{P}(\mathbf{X}_g, \mathbf{\Gamma}_g | \mathbf{Y}_g) &= \frac{\mathcal{P}(\mathbf{Y}_g | \mathbf{X}_g, \mathbf{\Gamma}_g)\mathcal{P}(\mathbf{X}_g | \mathbf{\Gamma}_g)\mathcal{P}(\mathbf{\Gamma}_g)}{\mathcal{P}(\mathbf{Y}_g)} \\ &= \frac{\mathcal{P}(\hat{\mathbf{Y}}_g + \mathbf{\Psi}_g | \mathbf{X}_g, \mathbf{\Gamma}_g)\mathcal{P}(\mathbf{X}_g | \mathbf{\Gamma}_g)\mathcal{P}(\mathbf{\Gamma}_g)}{\mathcal{P}(\mathbf{Y}_g)}, \end{aligned} \tag{33}$$

其中，第二步由修正似然（22）得出；由于 AEM，$\mathcal{P}(\boldsymbol{y}_m | \mathbf{x}_m) \approx \mathcal{CN}(\mathbf{S}_m\mathbf{x}_m, \mathbf{\Phi}_g)$；而

$$\mathbf{\Gamma}_g = \quad \{\boldsymbol{\gamma}_g\} \qquad \boldsymbol{\gamma}_g = \{\gamma_{1,g}, \cdots, \gamma_{N_g,g}\} \in \mathrm{diag},$$

$\mathbb{R}_+^{N_g \times 1}$

其中 是每个簇中的稀疏性促进超参数。注意，这些超参数的对角精度矩阵 $\mathbf{\Gamma}_g$ 在 $\mathbf{X}_g$ 的每一^行中形成固定的稀疏模式，从而可以将每个簇中的联合后验分布分解为

$$\begin{aligned}\mathcal{P}(\mathbf{X}_g, \mathbf{\Gamma}_g | \mathbf{Y}_g) \propto& \prod_{m=1}^{M} \mathcal{P}\left(\hat{\mathbf{y}}_m + \boldsymbol{\psi}_m | \mathbf{S}_g\mathbf{x}_m\right) \times \\ &\prod_{n=1}^{N_g} \mathcal{P}\left(\mathbf{x}_{n,g} | \gamma_{n,g}\right) \prod_{n=1}^{N_g} \mathcal{P}\left(\gamma_{n,g}\right). \end{aligned} \tag{34}$$

这样便可独立更新 $\gamma_{n,\,g}$ 和 $\mathbf{x}_{n,\,g}$，如下文所示。

遵循传统的SBL框架，AEM-SBL使用（34），分别通过AEM校正期望（E）步和最大化（M）步求得 $\boldsymbol{\gamma}_g$ 和 $\mathbf{X}_g$ [60],[61]。为此，我们将每个簇处的AEM-SBL代价定义为 $\boldsymbol{\gamma}_g$ 的函数，即

$$\begin{aligned}\boldsymbol{\gamma}_g &= \underset{\boldsymbol{\gamma}_g}{\mathrm{argmax}} \ln \mathcal{P}(\boldsymbol{\gamma}_g | \hat{\mathbf{Y}}_g + \mathbf{\Psi}_g) \\ &\propto \underset{\boldsymbol{\gamma}_g}{\mathrm{argmax}} \ln \mathcal{P}(\mathbf{Y}_g | \boldsymbol{\gamma}_g)\mathcal{P}(\boldsymbol{\gamma}_g). \end{aligned} \tag{35}$$

注意，问题（35）需要对 $\mathcal{P}(\mathbf{X}_g, \mathbf{Y}_g, \boldsymbol{\gamma}_g)$ 关于 $\mathbf{X}_g$ 进行边缘化，这可以通过期望最大化（EM）迭代求解。因此，在第 $(t+1)$ 次迭代中，修正后的 E 步使用完整联合分布的对数似然，并基于后验分布

算法3：AEM-SBL， $\forall g \in \mathcal{G}$

输入：$\hat{\mathbf{Y}}_g$，Δ，$\mathbf{C}_g$

初始化：$\mathbf{X}^0=\mathbf{0}$，$\mathbf{X}^0=\mathbf{0}$，$\mathbf{\Gamma}^{(0)}=\mathbf{1}_{N_g\times 1}$，$t=0$

重复

  使用（38）更新 $\mathbf{X}_g^{(t+1)}$

  使用（42）更新 $\mathbf{\Gamma}^{(t+1)}$

  $t=t+1$

直到 $\frac{\left\|\mathbf{X}_g^{(t+1)}-\mathbf{X}_g^{(t)}\right\|_F}{\left\|\mathbf{X}_g^{(t+1)}\right\|_F}<\Delta$;

输出：$\hat{\mathbf{X}}_g=\mathbf{X}_g^{(t)}$

分布计算，该后验分布以$\boldsymbol{\gamma}_g$的先前估计值为参数，即 $\mathcal{P}(\mathbf{x}_m|\mathbf{y}_m,\boldsymbol{\gamma}_g^{(t)})$。将此期望定义为

$$\begin{aligned}J(\boldsymbol{\gamma}_g,\hat{\boldsymbol{\gamma}}_g^{(t)})&=\mathbb{E}_{\mathcal{P}(\mathbf{x}_m|\mathbf{y}_m,\hat{\boldsymbol{\gamma}}_g^{(t)})}\ln\mathcal{P}(\mathbf{x}_m,\mathbf{y}_m|\boldsymbol{\gamma}_g)\\&=\mathbb{E}_{\mathcal{P}(\mathbf{x}_m|\mathbf{y}_m,\boldsymbol{\gamma}_g^{(t)})}[\ln\mathcal{P}(\mathbf{x}_m|\boldsymbol{\gamma}_g)\\&\quad+\ln\mathcal{P}(\mathbf{y}_m|\mathbf{x}_m)+\ln\mathcal{P}(\boldsymbol{\gamma}_g)],\forall m\in\mathcal{M}.\end{aligned}\tag{36}$$

(34)中给出的分解使得完整期望可表示为

$$J(\boldsymbol{\gamma}_g,\hat{\boldsymbol{\gamma}}_g)\propto\sum_{n=1}^{N_g}\ln\gamma_{n,g}-\mathbb{E}_{\mathcal{P}(\mathbf{x}_g|\mathbf{y}_g,\boldsymbol{\gamma}_g^{(t)})}[\|\mathbf{x}_{n,g}\|^2],\tag{37}$$

其中，(36)中的后验分布 $\mathcal{P}(\mathbf{x}_m|\mathbf{y}_m,\hat{\boldsymbol{\gamma}}_g^{(t)})$ 在 AEM框架中由校正后的均值和协方差参数化，分别为

$$\begin{aligned}\hat{\mathbf{x}}_m&=\mathbf{\Sigma}_g\mathbf{S}_g^{\mathrm{H}}\mathbf{\Phi}_g^{-1}\left(\hat{\mathbf{y}}_m+\boldsymbol{\psi}_m\right),\forall m\in\mathcal{M},\\\mathbf{\Sigma}_g&=\left(\mathbf{S}_g^{\mathrm{H}}\mathbf{\Phi}_g^{-1}\mathbf{S}_g+\mathbf{\Gamma}_g^{(t)}\right)^{-1}.\end{aligned}\tag{38}$$

注意，$\mathbf{\Sigma}_g$ 的每个对角元素（记为$\{\nu_{n,g}\}$）对于 $\mathbf{x}_{n,g}$ 的所有元素均相同。此外，(37)中的第二项可通过 [62] 化简

$$\begin{aligned}\mathbb{E}_{\mathcal{P}(\mathbf{x}_g|\mathbf{y}_g,\hat{\boldsymbol{\gamma}}_g^{(t)})}[\|\mathbf{x}_{n,g}\|^2]&=\mathbb{E}_{\mathcal{P}(\mathbf{x}_g|\mathbf{y}_g,\boldsymbol{\gamma}_g^{(t)})}[\|\mathbf{x}_{n,g}\|]^2\\&\quad+\mathbb{V}_{\mathcal{P}(\mathbf{x}_g|\mathbf{y}_g,\boldsymbol{\gamma}_g^{(t)})}[\mathbf{x}_{n,g}],\end{aligned}\tag{39}$$

因此，对于第 n 个簇中的第 g 个设备

$$J(\gamma_{n,g},\hat{\gamma}_{n,g})\propto\ln\gamma_{n,g}-\gamma_{n,g}(\|\hat{\mathbf{x}}_{n,g}\|^2+\nu_{n,g}).\tag{40}$$

在 M步中，通过求解(35)为每个机器型设备计算超参数，如下所示

$$\frac{\partial J\left(\gamma_{n,g},\gamma_{n,g}^{(t)}\right)}{\partial\gamma_{n,g}}=\frac{1}{\gamma_{n,g}}-(\|\hat{\mathbf{x}}_{n,g}\|^2+\nu_{n,g})=0,\tag{41}$$

由此得到

$$\gamma_{n,g}^{(t+1)}=\frac{1}{\|\hat{\mathbf{x}}_{n,g}\|^2+\nu_{n,g}},\qquad\forall g,n.\tag{42}$$

算法3给出了 AEM-SBL 的概要。

D. 复杂度分析

算法的计算复杂度以大-$\mathcal{O}$ 符号表示，如表I所示，其中考虑了相关的数学运算，例如矩阵乘法和求逆。鉴于 AEM-ADMM 和 AEM-SBL 对预处理步骤 (13) 和 (14) 的依赖性，我们分析了它们的计算复杂度，分别为 $\mathcal{O}(\frac{NLM}{G})$ 和

表I：JADCE 算法的计算复杂度

算法	每次迭代的运算次数	复杂度
ADMM	N^2M+NM^2+MLN	$\mathcal{O}(N^2M+NM^2)$
AEM-ADMM	$\frac{N^2M}{G^2}+\frac{NM^2}{G}+MLN$	$\mathcal{O}(\frac{N^2M}{G^2}+\frac{NM^2}{G})$
SBL	N^2L+N^2+NM	$\mathcal{O}(N^2L)$
AEM-SBL	$\frac{N^2L}{G}+\frac{N^2}{G}+NM$	$\mathcal{O}(\frac{N^2L}{G})$
SOMP	$(2L+1)MN+L(M^2+M+1)$	$\mathcal{O}(LMN)$
CB-SOMP	$(2L+1)MN+L(M^2+M+1)$	$\mathcal{O}(LMN)$

分别为 $\mathcal{O}\left(\frac{L^2N}{G}+L^3+\frac{NLM}{G}\right)$。预处理过程中计算量最大的操作是 $(\mathbf{S}_g\mathbf{S}_g^{\mathrm{H}})^{-1}\mathbf{S}_g$，其计算复杂度为 $\mathcal{O}\left(\frac{L^2N}{G}+L^3\right)$。值得庆幸的是，该操作可以预先计算，并且在导频分配发生变化之前始终有效，例如设备数量变化时。因此，预处理步骤的复杂度可降至 $\mathcal{O}(\frac{NLM}{G})$。总而言之，注意到 $\mathcal{O}(\frac{NLM}{G})$ 会随着簇数量的增加而降低，因此，与传统实现相比，AEM-ADMM 和 AEM-SBL 的计算复杂度更低。由此可知，所提出的逐簇 AEM 算法运行时间必然更短，正如下一节将要展示的那样。

V. 数值结果与讨论

本节将所提出的 AEM-ADMM 和 AEM-SBL 与其他 JADCE 方法进行比较，并展示相关结果。数值结果分别通过归一化均方误差（NMSE）、平均漏检概率（PMD）和算法运行时间，评估信道估计精度、检测能力和可扩展性。需要说明的是，我们只展示簇级性能指标，因为这些指标可以轻松扩展，用于评估 ADMM、SOMP 和 SBL 等集中式算法的性能。

信道估计精度采用如下定义的 NMSE 进行评估：

$$\mathrm{NMSE}_i=\mathbb{E}_{i,g}\left(\frac{\|\mathbf{x}_{i,g}-\hat{\mathbf{x}}_{i,g}\|^2}{\|\mathbf{x}_{i,g}\|^2}\right),i\in\mathcal{K}_g.\tag{43}$$

对于 PMD，我们首先使用以下方法估计活动向量 $\hat{\boldsymbol{\alpha}}_g\in\{0,1\}^{N_g\times 1}$

$$\hat{\alpha}_{n,g}=\begin{cases}1,\ \text{if}\ \|\hat{\mathbf{x}}_{n,g}\|_2\geq\zeta\\0,\ \text{otherwise}\end{cases},\tag{44}$$

其中，ζ 是根据预设的虚警概率（PFA）目标值设定的阈值（见表 II）。由此可使用以下公式计算 PMD：

$$\mathrm{PMD}=\mathbb{E}\left(\frac{\sum_{n=1}^{N_g}\max(0,\alpha_{n,g}-\hat{\alpha}_{n,g})}{|\mathcal{K}_g|}\right).\tag{45}$$

运行时间性能以秒为单位进行评估，因此与迭代次数以及表 I 中所述的计算复杂度分析相关。

A. 仿真设置

我们考虑一个半径为 250 m 的单小区大规模 MIMO 上行网络，其中基站为随机分布在小区内的 1000 个机器型设备提供服务。

表 II：仿真参数

参数	数值
小区半径	250 米
机器型设备数量 (N)	1000
簇的数量(G)	4
带宽	20 兆赫
噪声功率（σ^2）	2×10^{-13} W
相干间隔（T）	300
导频序列长度（L）	64
基站天线数量（M）	32
激活概率（ϵ）	0.01
平均信噪比	10 dB
误差容限（Δ）	10^{-4}
目标虚警概率	10^{-3}

我们假设这些机器型设备均匀划分为 $G = 4$ 个簇。为简单起见，我们采用对数距离路径损耗模型，使得$\beta_{n,g} = -130 - 37.6\log_{10} d_{n,g}$，其中 $d_{n,g}$ 表示第 n 个簇中的第 g 个机器型设备与基站之间的距离。我们考虑两种不同的设置：i）非相关信道，此时 $\mathbf{Q}_{n,g} = \beta_{n,g}\mathbf{I}$；ii）空间相关信道，此时矩阵 $\mathbf{Q}_{n,g}$ 采用 [45]中所述的近似高斯局部散射模型生成，天线间距为半波长。矩阵的第 k 行第 m 列元素计算如下：

$$[\mathbf{Q}_{n,g}]_{k,m} = \frac{\beta_{n,g}}{L_p}\sum_{i=1}^{L_p}\exp(\pi j(k-m)\sin\phi_{n,g}^{(i)})\times \exp(-\frac{1}{2}\sigma^2_{\phi_{n,g}}\pi(k-m)\cos\phi_{n,g}^{(i)}), \quad (46)$$

其中，L_p 是多径分量的数量。此外，$\phi_n^{(i)}, g \sim \mathcal{U}(\bar{\phi}_{n,g} - \frac{2\pi}{9}, \bar{\phi}_{n,g} + \frac{2\pi}{9})$ 是第 i 个多径簇的标称角度，该簇分布在方位角$\bar{\phi}_{n,g}$ 表示第 g 个簇中第 n 个机器型设备相对于基站天线阵列波束轴的角度。多径簇内路径的角度标准差记为 $\sigma_{\phi_{n,g}}$，而导频序列的基矩阵 B 采用复阿达玛矩阵。除非另有说明，仿真均采用表 II 中给出的参数。图中结果为对 10^3 次蒙特卡洛仿真取平均所得。

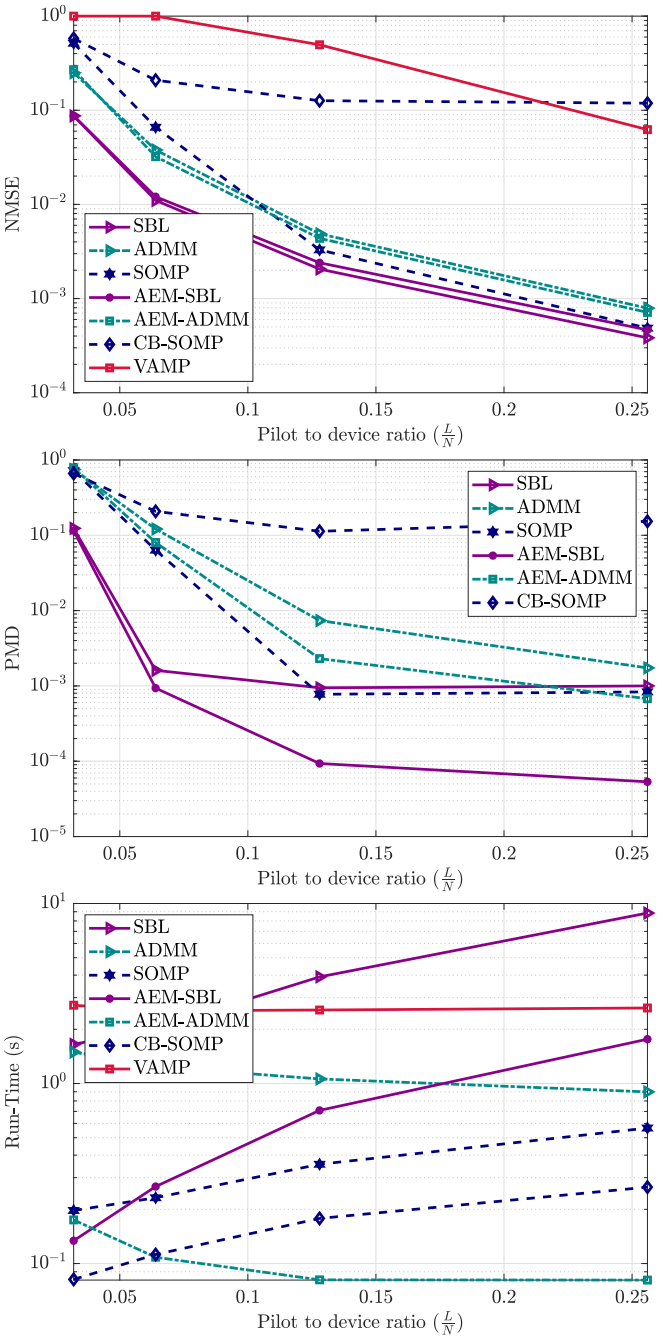

图 3：性能指标随导频长度的变化：(a) 归一化均方误差（顶部）、(b) PMD（中部）和 (c) 运行时间（底部）。

B. 关于导频序列长度

在图 3 中，我们评估导频长度对 JADCE 算法性能的影响。具体而言，图 3(a) 展示了归一化均方误差随欠采样比的变化情况；欠采样比衡量导频长度与网络中设备数量之比，即 $\frac{L}{N}$ 。总体而言，结果表明，信道估计性能会随着导频长度与设备数量之比的增大而提升。这主要是因为较大的 $\frac{L}{N}$ 使基站能够在训练阶段分配更多资源，从而提高信道状态信息的准确性。此外，较高的 $\frac{L}{N}$ 比值意味着导频序列更长，有助于提高不同机器型设备导频序列之间的正交性，从而减少导频碰撞

网络中的信道估计质量更高。因此，随着 $\frac{L}{N}$ 增大，JADCE 算法将受益于矩阵 $\mathbf{S}_g$的RIP得到改善。值得注意的是，AEM-SBL和AEM-ADMM在求解规模更小的问题时，性能与ADMM和SBL相近。因此，AEM-ADMM和AEM-SBL都是在大规模机器类通信的实际条件下进行信道估计的高效方法，尤其是在导频长度远小于网络中的设备数量时，即 $L \ll N$。从相同的结果可以看出，CB-SOMP

（算法 1）的信道估计能力不如经典的SOMP。CB-SOMP表现不佳的原因在于它依赖于完全失配的模型，这凸显了考虑此类失配的重要性，我们的受AEM启发的算法（AEM-ADMM和AEM-SBL）也证明了这一点。相反，尽管SOMP和ADMM未利用聚类结构，但在导频长度较大时性能良好，而这对于大规模机器类通信应用并不实际。从结果可以明显看出，SBL和AEM-SBL算法的性能均显著优于VAMP。尽管在感知矩阵接近高斯矩阵时，VAMP效率很高，但由于感知矩阵的结构偏离高斯假设 [10]，它在这一场景中受到限制。此外，如I-A节所述，对于诸如当前考虑的短导频长度等问题维度，VAMP和其他AMP相关算法并非最优，而短导频长度适用于大规模机器类通信。

图3(b)展示了 PMD 如何受 $\frac{L}{N}$ 影响。一般而言，PMD 会随着导频与设备比的增加而降低。结合图3(a)的结果可知，这是因为不同 MTD 的导频碰撞减少了。值得注意的是，在这一指标上，AEM-SBL 的表现优于传统 SBL，因为它受益于簇间多用户干扰的减少，能够更有效地在恢复矩阵 $\mathbf{X}_g$ 时施加行稀疏性。此外，AEM-ADMM 在这一指标上的表现优于 ADMM，因为它能有效处理多用户干扰，从而为簇化大规模机器类通信提供一种低复杂度的联合活动检测与信道估计方案。为证实这一点，我们在图3(c)中分析了运行时间随 L 的变化。从图中可以看出，与经典方法相比，AEM-ADMM 和 AEM-SBL 的可扩展性更强。例如，随着 L 增加，AEM-SBL 的运行速度比 SBL 快五倍。尽管表 I 显示，AEM-ADMM 和 ADMM 的运行时间对 L 不敏感，但值得一提的是，随着 L 增加，前者的运行速度几乎比后者快十倍，这归功于其更快的收敛速度。尽管 SOMP 的运行时间更短，但如图3(a)和图3(b)所示，CB-SOMP 通常会产生较差的结果。有趣的是，尽管 VAMP 在当前设置下表现不佳，但其可扩展性很强，这与 [10], [31], [33], [35] 中的结果一致。

图4中，我们分析了采用 AEM-SBL 时，所提出的导频序列对联合活动检测与信道估计性能的影响，并同时考虑相关和非相关的瑞利衰落信道。总体而言，随着导频长度增加，性能会有所提升，但在相关信道模型下最终会趋于饱和。值得注意的是，尽管高斯导频和伯努利矩阵都不是正交基，但随着 L 增加，它们仍能显著提升联合活动检测与信道估计的性能。尽管并非正交基，较长的高斯导频和伯努利感知导频序列能够更准确地近似 Y，从而提升 AEM-SBL 的检测能力。与 [10], 的结果一致，伯努利导频序列在促进联合活动检测与信道估计方面的能力更强。然而，对于设备数量庞大的场景，采用极低的导频与设备比更为实际，尤其是机器型设备等传输短数据包的场景

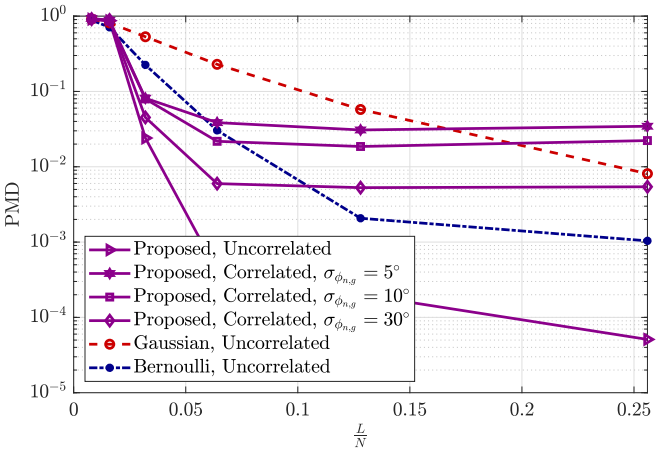

图4：采用 AEM-SBL 时，非相关信道和空间相关信道的性能随导频长度的变化情况$\sigma_{\phi_{n,g}} = \{5°, 10°, 30°\}$

。鉴于此，在实际应用中，所提出的基于簇的导频优于其他导频序列。例如，可以观察到，采用所提出的导频序列时，PMD 可低于 0.001；而当导频与设备比低于 0.1 时，基于伯努利的导频序列和高斯导频的 PMD 均大于 0.01。总体而言，AEM-SBL 在相关信道和非相关信道下呈现相同的性能趋势，唯一的区别是，在强空间相关信道下性能通常较差，例如$\sigma_{\phi_{n,\ g}} = 5°$的情况。不过，这是预期现象，因为相关信道会减弱信道硬化 [45]。

C. 关于平均信噪比

在图5中，我们展示了性能随平均信噪比变化的结果。具体而言，图5(a)展示了归一化均方误差（NMSE）方面的性能。如图所示，当信号功率高于噪声功率时，即平均信噪比提高时，所有 JADCE 算法的性能都会改善。尽管如此，CB-SOMP 由于依赖失配模型，性能并未改善。鉴于此，信噪比越高，失配程度越大，因此其性能逊于其他算法。对于 AEM-SBL 和 AEM-ADMM，随着功率水平提高，AEM 能够更准确地校正失配，因此它们在高信噪比下表现更优。另一方面，图5(b)展示了非相关瑞利衰落信道下 AEM-SBL 和 SBL 在 PMD 方面的性能。此外，我们还评估了不同空间相关程度下 AEM-SBL 的性能。总体而言，随着信噪比提高，两种算法的检测能力均有所增强，表现为 PMD 较低。从图中可见，尽管图5(a)显示 SBL 的 NMSE 略优于 AEM-SBL，但 AEM-SBL 的检测能力优于 SBL。不过，需要注意的是，PMD 的性能取决于阈值设置；下文中该阈值设置为使虚警概率 PFA 达到 $= 10^{-3}$ 。因此，我们

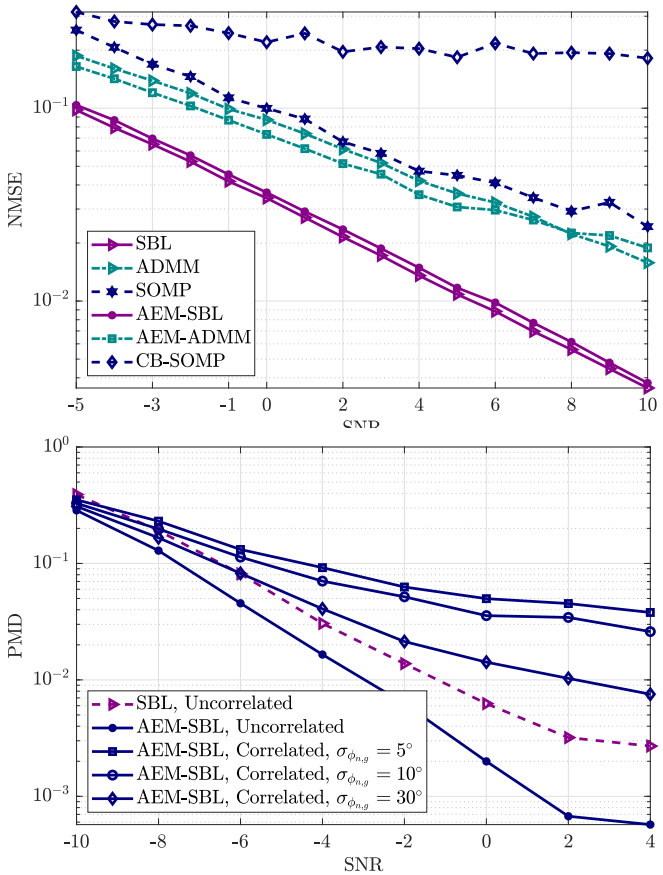

图5：使用 AEM-SBL 时，未相关和空间相关信道下的性能：a) 归一化均方误差（上），b) PMD（下）随信噪比的变化。$\sigma_{\phi_{n,\ g}} \in \{5°, 10°, 30°\}$

不能仅凭这一阈值断言 AEM-SBL 在算法上更优。尽管如此，AEM-SBL 和 SBL 的性能相当，而 AEM-SBL 的运行时间更短。然而，随着信道相关性增强，即 $\sigma_{\phi_{n,\ g}}$ 减小，AEM-SBL 的性能会因信道硬化减弱而下降。

D. 关于基站天线数量

图6展示了性能随基站天线数量变化的结果。图6(a)的整体趋势表明，随着基站天线数量增加，所有算法的信道估计精度均有所提升。这主要是因为增加 M 提供了额外的结构，可在信号恢复过程中加以利用 [10], [50]。然而，如$M = 20$之后性能没有显著提升所示，性能提升最终趋于饱和，这与 [10]的结果一致。我们还注意到，AEM-SBL 和 SBL 均优于其他算法，这是因为它们具有贝叶斯特性。同样，在图6(b)中，PMD方面的性能随着天线数量增加而提升，这是由于天线数量较多时 MMV 问题具有更高的分辨率 [63]。不过，值得注意的是，增加M 会对运行时间和可扩展性产生显著影响

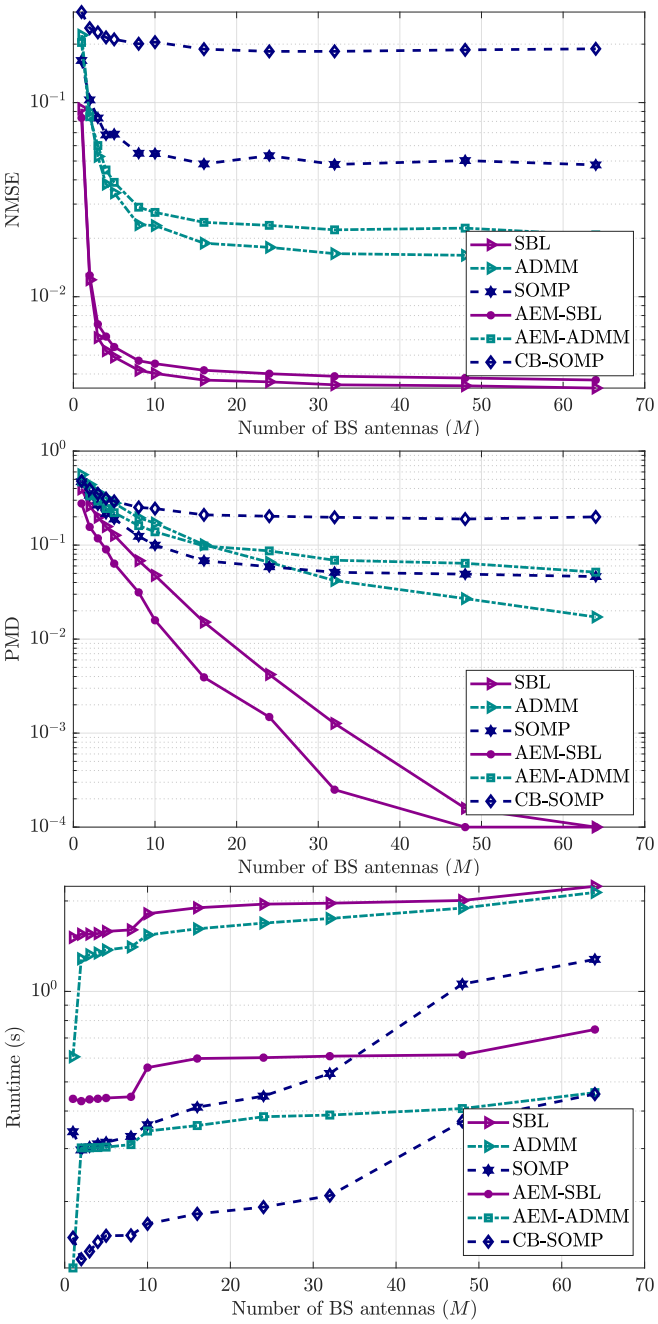

图 6：性能随天线数量变化的情况：a) 归一化均方误差（上），b) PMD（中），以及 c) 运行时间（下）。

如图 6(c) 所示，随着天线数量的增加，算法的运行时间显著增长。这是由于矩阵维度不断增大所致。尽管如此，显而易见，AEM-SBL 和 AEM-ADMM 的运行速度均快于各自的传统方法，因此有助于提高网络的可扩展性。这是所提出算法的一项主要优势，因为它们充分利用了感知矩阵的结构。

E. 关于激活概率

在图 7 中，我们通过评估

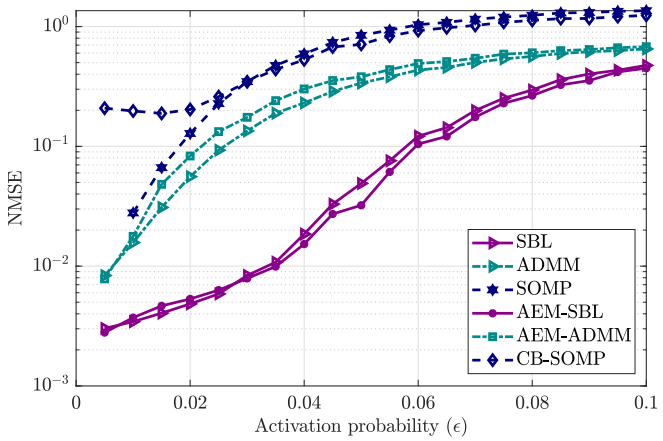

图 7：NMSE 随 ϵ 的变化关系。

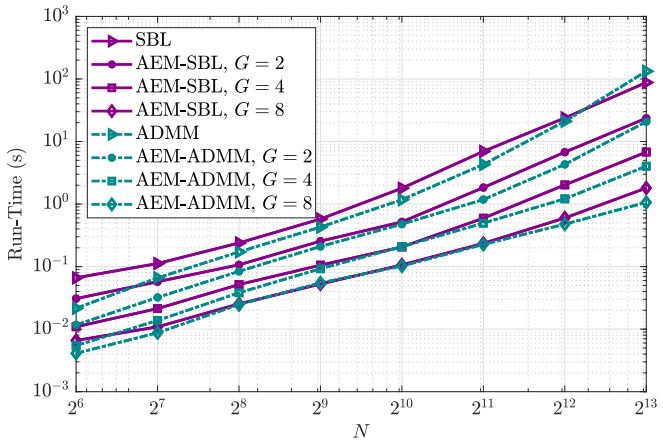

图 8：性能随机器型设备数量和簇大小的变化。

归一化均方误差（NMSE），分析了激活概率对各算法联合活动检测与信道估计性能的影响。可以观察到，随着稀疏度降低，即同时激活的设备增多（ ϵ 较高），所有算法的性能都会变差。这是因为稀疏性降低时，每个簇中的多用户干扰都会增加。值得注意的是，基于 AEM 的算法能够有效降低多用户干扰，同时处理维度更低的联合活动检测与信道估计问题，因此展现出与其他算法相当的效率。

F. 关于簇大小

图 8 展示了运行时间随机器型设备数量和簇数量变化的情况。结果表明，运行时间随设备数量的增加而增加。例如，使用 $N = 64$ 时，SBL 的运行时间不到 0.1 秒，而 $N = 8192$ 大约需要 100 秒。这与表 I 中的复杂度分析相符，并进一步验证了该分析。然而，可以观察到，随着簇数量增加，AEM-SBL 和 AEM-ADMM 的运行时间均短于集中式 SBL 和 ADMM。这些结果证明了所提出的联合活动检测与信道估计方案的可扩展性。

VI. 结论与未来工作展望

本研究提出了一种新颖的基于簇的设备活动检测和信道估计框架，该框架依赖于基于正交导频子空间优化免授权非正交多址接入。通过利用非独立同分布的导频序列，所提出的基于导频的聚类方法促进了高效的设备活动检测，并增强了网络灵活性和可扩展性，因此具有实际应用价值。我们借鉴了逆问题领域的概念，并提出了新颖的数据驱动的联合活动检测与信道估计方案：i）AEM-ADMM，针对没有精确先验的场景采用迭代软阈值处理；ii）AEM-SBL，专为能够构建先验分布的情形设计。与经典算法相比，所提出的算法在保持相近性能的同时，运行时间更短。此外，AEM通过考虑感知矩阵的缺陷，为机器类型通信中的设备活动检测带来了全新视角。我们的方案为5GB 蜂窝网络中的接收机提供了及时的解决方案。

作为未来研究的一个潜在方向，本文提出的 AEM 方法可以通过在每次迭代中引入自适应校正参数来进一步改进，从而提升其性能。此外，可以将平坦衰落信道假设放宽为频率选择性衰落，以适用于基于正交频分复用（OFDM）的大规模机器类通信（mMTC）。此外，本文的研究成果还可以扩展至无小区 MIMO 通信系统。

参考文献

[1] J. A. Zhang 等人，“联合通信与雷达感知的信号处理技术概述”，IEEE 信号处理选题杂志，2021年。 [2] M. Renfors、M. Juntti 和 M. Valkama，“无线收发器的信号处理”，载于《信号处理系统手册》。施普林格，2019年，第 251–310 页。 [3] M. A. Albreem、M. Juntti 和 S. Shahabuddin，“大规模 MIMO 检测技术：综述”，IEEE 通信调查与教程，第 21 卷，第 4 期，第 3109–3132 页，2019年。 [4] Q. H. Spencer 等人，“多用户 MIMO 下行链路简介”，IEEE《通信杂志》，第 42 卷，第 10 期，第 60–67 页，2004 年。 [5] R. W. Heath 等人，“ 毫米波 MIMO 系统的信号处理技术概述”，IEEE 信号处理选题杂志，第 10 卷，第 3 期，第 436–453 页，2016年。 [6] O. L´opez 等人，“ URLLC 统计工具与方法——教程”，arXiv 预印本 arXiv:2212.03292，2022年。 [7] H. Djelouat 等人，“海量机器类型通信中基于近似消息传递的空间相关信道用户活动检测与信道估计”，载于 2021 年第 55 届阿西洛马会议：信号、系统与计算机。IEEE，2021年，第 1200–1204 页。 [8] L. Marata 等人，“ 异构网络中的联合信道估计与设备活动检测 ”，载于第 29 届欧洲信号处理会议（EUSIPCO）。IEEE，2021年，第 836–840 页。 [9] M. B. Shahab 等人，“面向物联网的免授权非正交多址接入：综述”，IEEE 通信调查与教程，第 22 卷，第 3 期，第 1805–1838 页，2020年。 [10] K. Senel 和 E. G. Larsson，“支持免授权海量机器类型通信的海量 MIMO：一种压缩感知方法”，IEEE 通信汇刊，第 66 卷，第 12 期，第 6164–6175 页，2018年。 [11] T. Li 等人，“面向 MIMO 海量无源随机接入的联合设备检测、信道估计与数据解码及冲突消解”，IEEE《通信选题领域期刊》，第 40 卷，第 5 期，第 1535–1555 页，2022年。 [12] A. Fengler 等人，“采用海量 MIMO 接收机的非贝叶斯活动检测、大尺度衰落系数估计与无源随机接入”，IEEE 信息论汇刊，第 67 卷，第 5 期，第 2925–2951 页，2021年。 [13] X. Shao 等人，“协作式活动检测：有源与无源海量随机接入范式”，IEEE 信号处理汇刊，第 68 卷，第 6578–6593 页，2020年。 [14] X. Chen 等人，“面向第五代移动通信及后续发展的海量接入”，IEEE《通信选题领域期刊》，第 39 卷，第 3 期，第 615–637 页，2020年。

[15] B. Li、J. Zheng 和 Y. Gao，“基于压缩感知的免授权 NOMA 多用户检测及动态用户活动”，IEEE 通信快报，第 26 卷，第 1 期，第 143–147 页，2021年。[16] L. Liu 等人，“面向免授权海量连接的稀疏信号处理：物联网随机接入协议的未来范式”，IEEE 信号处理杂志，第 35 卷，第 99 页，2018 5，88第。 – 页。[17] R. B. Di Renna 等人，“大规模机器类通信的检测技术：挑战与解决方案”，IEEE Access，第 8 卷，第 180 928–180 954 页，2020年。[18] Y. Zhu 等人，“频率选择性衰落信道上的基于正交频分复用的海量免授权传输”，IEEE 通信汇刊，2022年。[19] W. Jiang、Y. Jia 和 Y. Cui，“基于正交频分复用的海量免授权接入的统计设备活动检测”，IEEE 无线通信汇刊，2022年。[20] O. L. L´opez 等人，“通过高效的受限信道状态信息方案实现 6G 时代的超低时延、低能耗与海量连接”，IEEE《通信杂志》，第 58 卷，第 11 期，第 56–61 页，2020年。[21] H. Q. Ngo 等人，“无蜂窝大规模 MIMO：为每个人提供同样优质的服务”，载于 IEEE 第 16 届无线通信信号处理进展国际研讨会（SPAWC）。IEEE，2015年，第 201–205 页。[22] A. Rajoriya、S. Rukhsana 和 R. Budhiraja，“大规模机器类通信中的集中式与分布式激活用户检测和信道估计”，IEEE 通信汇刊，第 70 卷，第 3 期，第 1759–1776 页，2022年。[23] U. K. Ganesan、E. Bj¨ornson 和 E. G. Larsson，“无蜂窝大规模 MIMO 免授权随机接入的基于簇的活动检测算法”，IEEE 通信汇刊，第 69 卷，第 11 期，第 7520–7530 页，2021年。[24] S. Kim 、 J. Choi 和 J. Park，“ 采用低分辨率 ADC 的短数据包物联网通信下行链路 NOMA”，IEEE 物联网期刊，2022年。[25] Y. Abdi 和 T. Ristaniemi，“分布式检测中线性化置信传播的优化”，IEEE 通信汇刊，第 68 卷，第 2 期，第 959–973 页，2019年。[26] 3第 3 代合作伙伴计划（3GPP），“第五代移动通信；新空口；物理信道和调制（版本 17.5.0，第 17 版）”，ETSI，技术报告 3GPP TS 38.211，2023年。 [在线]。可访问：https://www.etsi.org/deliver/etsi ts/138200 138299/138211/17.05.00 60/ts_138211v170500p.pdf

[27] S. Liu、H. Zhang 和 Q. Zou，《免授权大规模机器类型通信上行链路的去中心化信道估计》，IEEE 通信汇刊，第 70 卷，第 2 期，第 967–979 页，2021年。[28] S. Elhoushy、M. Ibrahim 和 W. Hamouda，《无蜂窝大规模 MIMO：综述》，IEEE 通信调查与教程，第 24 卷，第 1 期，第 492–523 页，2021年。[29] H. He 等人，《无蜂窝大规模 MIMO 的分布式期望传播检测》，载于 IEEE全球通信大会（GLOBECOM）。IEEE，2021年，第 01–06 页。[30] J. Li 等人，《采用正交导频序列的无蜂窝分布式大规模MIMO系统中基于协方差的活动检测》，IEEE 车辆技术汇刊，2022年。[31] D. L. Donoho、A. Maleki 和 A. Montanari，《用于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 18 914–18 919 页，2009年。[32] J. Bai 和 E. G. Larsson，《分布式 MIMO 中的活动检测：通过似然比融合实现分布式AMP》，IEEE 无线通信快报，第 11 卷，第 10 期，第 2200–2204 页，2022年。[33] S. Rangan、P. Schniter 和 A. K. Fletcher，《矢量近似消息传递》，IEEE 信息论汇刊，第 65 卷，第 10 期，第 6664–6684 页，2019年。[34] J. Ma 和 L. Ping，《正交AMP》，IEEE Access，第 5 卷，第 2020–2033 页，2017年。[35] Y. Cheng、L. Liu 和 L. Ping，《具有空间和时间相关信道中的大规模接入正交 AMP 》，IEEE《通信选题领域期刊》，2020年。[36] Y. Li 等人，《无蜂窝大规模 MIMO 的异步活动检测：从集中式算法到分布式算法》，IEEE 无线通信汇刊，2022年。[37] S. Chen 等人，《可扩展无蜂窝大规模MIMO系统的结构化大规模接入》，IEEE《通信选题领域期刊》，第 39 卷，第 4 期，第 1086–1100 页，2020年。[38] F. A. De Figueiredo 等人，《大规模 MIMO 系统在机器类型通信中的应用》，IEEE Access，第 7 卷，第 2589–2611 页，2018年。[39] H. Iimori 等人，《超大规模MIMO系统的联合活动检测与信道估计》，IEEE 无线通信汇刊，2022年。

[40] L. Marata 等人，《异构网络的相干与非相干联合检测和解码技术》，IEEE 无线通信汇刊，2022年。[41] M. Mozumder 等人，《漫射光学断层成像的基于模型的迭代学习方法》，IEEE 医学成像汇刊，第 41 卷，第 5 期，第 1289–1299 页，2021年。[42] S. Lunz 等人，《逆问题中的学习型算子校正》，SIAM 成像科学杂志，第 14 卷，第 1 期，第 92–127 页，2021年。[43] J. Kaipio 和 E. Somersalo，《统计与计算逆问题》。Springer Science & Business Media，2006年，第 160 卷。[44] G. Strang 等人，《线性代数导论》。Wellesley-Cambridge Press，马萨诸塞州韦尔斯利，1993年，第 3 卷。[45] E. Bj¨ornson 等人，《大规模 MIMO 网络：频谱、能源与硬件效率》，信号处理基础与趋势，第 11 卷，第 3-4 期，第 154–655 页，2017年。[46] J. W. Choi 等人，《无线通信中的压缩感知：实用技巧与窍门》，IEEE 通信调查与教程，第 19 卷，第 3 期，第 1527–1550 页，2017年。[47] J. Zhang 等人，《物理启发的压缩感知：超越深度展开》，IEEE 信号处理杂志，第 40 卷，第 1 期，第 58–72 页，2023年。[48] Y. C. Eldar 和 G. Kutyniok，《压缩感知：理论与应用》。剑桥大学出版社，2012年。[49] Z. Chen、F. Sohrabi 和 W. Yu，《面向海量连接的稀疏活动检测》，IEEE 信号处理汇刊，第 66 卷，第 7 期，第 1890–1904 页，2018年。[50] Z. Wei、D. W. K. Ng 和 J. Yuan，《采用 MRC-SIC 接收机的上行链路 MIMO-NOMA 联合导频与负载功率控制》，IEEE 通信快报，第 22 卷，第 4 期，第 692–695 页，2018年。[51] R. B. Di Renna 和 R. C. de Lamare，《基于动态消息调度策略的 mMTC 联合信道估计、活动检测与数据解码》，IEEE 通信汇刊，第 70 卷，第 4 期，第 2464–2479 页，2022年。[52] L. Liu 等人，《对 GAMP 和 AMP 的新见解》，IEEE 车辆技术汇刊，第 68 卷，第 8 期，第 8264–8269 页，2019年。[53] D. Zhang 等人，《在约束 Bethe 自由能最小化框架下统一消息传递算法》，IEEE 无线通信汇刊，第 20 卷，第 7 期，第 4144–4158 页，2021年。[54] S. Boyd、N. Parikh 和 E. Chu，《通过交替方向乘子法实现分布式优化与统计学习》。Now Publishers Inc，2011年。[55] H. Djelouat、M. Leinonen 和 M. Juntti，《用于大规模 MTC 用户活动检测与信道估计的空间相关性感知压缩感知》，IEEE 无线通信汇刊，2022年。[56] T. T. Cai 和 L. Wang，《含噪稀疏信号恢复的正交匹配追踪》，IEEE 信息论汇刊，第 57 卷，第 7 期，第 4680–4688 页，2011年。[57] S. Arridge 等人，《光扩散断层成像应用中的近似误差与模型降阶》，逆问题，第 22 卷，第 1 期，第 175 页，2006年。[58] A. Wiesel、Y. C. Eldar 和 A. Yeredor，《高斯模型不确定性下的线性回归：算法与界》，IEEE 信号处理汇刊，第 56 卷，第 6 期，第 2194–2205 页，2008年。[59] T. Goldstein、C. Studer 和 R. Baraniuk，《前向-后向分裂实用指南及其 FASTA 实现》，arXiv 预印本 arXiv:1411.3406，2014年。[60] Z. Zhang 和 B. D. Rao，《使用稀疏贝叶斯学习恢复时间相关源向量的稀疏信号》，IEEE 信号处理选题杂志，第 5 卷，第 5 期，第 912–926 页，2011年。[61] A. Sant、M. Leinonen 和 B. D. Rao，《通过一般全变分正则化稀疏贝叶斯学习实现块稀疏信号恢复》，IEEE 信号处理汇刊，第 70 卷，第 1056–1071 页，2022年。[62] M. Al-Shoukairi、P. Schniter 和 B. D. Rao，《基于 GAMP 的低复杂度稀疏贝叶斯学习算法》，IEEE 信号处理汇刊，第 66 卷，第 2 期，第 294–308 页，2017年。[63] B. Zheng 等人，《spark 级稀疏度信号与 MMV 尾部的联合稀疏恢复$\ell_{2,1}$ 最小化》，IEEE 信号处理快报，第 28 卷，第 1130–1134 页，2021年。