功率分配策略下大规模多输入多输出与非正交多址接入的每天线比特效率

Thiago A. Bruza Alves，**Taufik Abrão**隆德里纳州立大学（UEL），电气工程系，巴西隆德里纳-巴拉那州。

arXiv:2301.03101v1 [cs.IT] 8 Jan 2023

摘要——本文针对下行链路（DL）单小区大规模多输入多输出（mMIMO）和非正交多址接入（NOMA）系统开展对比资源分配分析，评估接收端每天线比特频谱效率（SE）和能量效率（EE）。系统中的基站（BS）配备多根（M）天线，而 K 设备各使用单天线；设备负载$\rho = \frac{K}{M}$在 $0 < \rho \leq 2$ 范围内变化，并采用三种不同的功率分配（PA）策略：信道功率分配的逆（PICPA）、改进的注水（Δ-WF）分配方法，以及作为基准方法的等功率分配（EPA）。由于NOMA系统中每个簇的两个设备在功率域上重叠，因此需要对信道矩阵进行变换，以执行mMIMO中采用的迫零（ZF）预编码。因此，在多天线条件下运行的NOMA能够使具有更高阵列增益的一组设备受益，从而胜过mMIMO，并在较高负载范围 $0.6 < \rho < 2.0$内表现出优势。在这种情况下，通过计算SE和能量效率曲线下的面积，分别衡量每天线比特效率和每瓦每天线比特效率，是一种更现实且更有用的指标。数值结果证实，NOMA相较于mMIMO具有明显优势：SE面积指标约为其3倍，能量效率面积指标约为其2倍。

索引词—非正交多址接入（NOMA）；大规模多输入多输出（mMIMO）；能量效率（EE）；频谱效率（SE）。

一、引言

第五代（5G）之后的无线通信系统必须支持超密集连接，并满足差异极大的异构需求。网络面临的挑战依然存在，包括联合提升频谱效率（SE）和能量效率（EE）、改善频谱效率与能量效率之间的权衡，以及保障服务质量（QoS），始终致力于满足接入网络的设备数量不断增长这一需求。在解决这些挑战的各种方案中，大规模多输入多输出（mMIMO）系统是首要方案；它通过在基站（BS）部署大量天线，利用多径传播来提高链路容量 [1], [2]。另一项重要的赋能技术是非正交多址接入（NOMA），它将功率域作为多址接入技术的一种替代方式，有助于缓解频谱耗尽问题，并使每个资源块能够服务多个设备[3]。

尽管许多研究将大规模多输入多输出归类为正交技术，但借助空间分集，可将来自设备的信号分配到同一资源块，因此也可以把它归类为一种非正交技术 [4]。大量文献表明，与正交多址接入（OMA）技术相比，非正交多址接入的频谱效率更优 [5]。此前已有研究探讨了通过结合多输入多输出（天线数量较少 M）和非正交多址接入来提升通信系统性能 [6]–[9]。

文献 [4], [10], [11]提出了对单小区中非正交多址接入和大规模多输入多输出进行比较的研究。文献 [10]提出了通过导频获取信道状态信息（CSI）来应用于NOMA系统的方法。文献 [11], 提出了在大规模多输入多输出方案中应用非正交多址接入，并在所提出的比较中取得了更好的结果。此外，文献 [4] 分析了视距和非视距条件下非正交多址接入与大规模多输入多输出的性能。

经典的大规模多输入多输出系统是指基站配备的天线数量远多于活跃设备数量 M；简而言之，$K\ M \gg K$ 即构成大规模多输入多输出配置。典型的非正交多址接入通过将所选设备的信号叠加，在功率域形成一个簇，并在同一信号上进行复用，再由同一波束成形服务，从而提高频谱效率。不过，非正交多址接入的成功取决于连续干扰消除（SIC）。

功率域非正交多址接入可以成为密集网络中的候选技术 [12]。为了提升性能并尽量减小影响，假设采用理想的串行干扰消除 [13], ，将设备分为两组。配对并形成簇后，每对设备形成一个簇，且两者的信道条件差异较大。信道条件较好的设备可以解码发送给信道条件较差设备的信号。因此，可以通过串行干扰消除消除干扰。此前研究了在配备大量天线的基站中使用非正交多址接入时的频谱效率 [4], [10]；我们在类似的系统配置下，提出将负载提高至基站天线数量的两倍，并分析频谱效率、能效以及频谱效率与能效权衡。

能效指标是一种常用的性能指标，用于分析功耗与数据速率之间的平衡。能效是有效传输的数据速率与传输过程中消耗的总功率之比，其中包括瞬时分量和静态分量。借助能效指标，可以评估系统利用有限能源资源传输数据的效率，并对这一比值进行优化。它还可以展示在追求设备间公平性的情况下能耗的变化趋势。

迫零（ZF）是一种简单且广受欢迎的干扰抑制波束成形替代方案。在完美信道状态信息条件下，它在实际的不完美信道状态信息情况下也能取得令人满意的效果。本研究采用完美信道状态信息，因此

本研究部分由巴西国家科学技术发展委员会（CNPq）通过资助项目310681/2019-7提供支持，部分由CAPES-巴西-财务代码001以及隆德里纳州立大学-巴拉那州政府（UEL）提供支持。

T. A. Bruza Alves和Taufik Abrão任职于隆德里纳州立大学（UEL）电气工程系，地址：巴西隆德里纳-巴拉那州。电子邮箱：thiagobruza@hotmail.com、taufik@uel.br

需要导频。采用配备大量天线的 NOMA系统时，需要部署一个定义明确的等效信道以抑制干扰；根据 NOMA原理，由于利用了 NOMA中的功率域，等效信道矩阵会小于原始信道矩阵。

已有多种传输拓扑研究了大规模多输入多输出中的能效问题，确定了天线数量、小区内设备数量以及最大能效 [2], [14]。 [15], 对非正交多址接入系统进行了能效分析，并证明了其相较于传统正交多址（OMA）系统的优越性。近期研究致力于提升NOMA性能，例如， [16], 定义了最小配对距离，并将其与OMA进行比较； [17] 则比较了OMA与配备大规模多输入多输出-NOMA的无蜂窝系统。 [18]提出了太赫兹（THz）-NOMA-多输入多输出（MIMO）中的能效分析。然而，活跃设备的数量仍远少于基站中的天线数量； [19] 是一篇关于功率域NOMA的综述，明确指出了能效分析以及多天线NOMA与大规模多输入多输出之间比较方面的空白。

近期研究提出将NOMA与其他技术结合，以构建更有效的传输方案；例如，在 [20] 中，研究者在双层网络中联合考虑NOMA和大规模多输入多输出，以适应海量流量。此外，在 [21], 中，作者将NOMA应用于分布式天线系统（DAS），旨在获得优于单独采用传统NOMA或DAS技术的性能。 [22] 则深入综述了功率域NOMA变体的最新进展，并系统梳理了基于NOMA的应用中若干尚未解决的问题和研究挑战。NOMA系统存在一些缺点，例如硬件复杂度（包括SIC）、信道反馈、接收机设计，以及精细的功率和导频分配策略 [12], [19], [23]。

本研究旨在揭示在基站天线数量庞大的情况下，与NOMA方案相比，采用大规模多输入多输出方案的优势，并考察不同设备负载的影响，即移动设备数量与基站天线数量之比 $\rho = \frac{K}{M}$，同时改变功率分配策略。此外，我们采用了 [2] 中的系统功耗实际模型，并根据本研究需求进行了调整，以便对系统资源分配进行适当分析。

贡献： 本研究的贡献有四点。**a)** 对 mMIMO 系统与 NOMA 系统的频谱效率（SE）性能进行了全面的比较分析，在特定的三种不同功率分配方法下改变系统负载，并利用系统 SE（$\mathcal{S}^{\text{SYST}}$）曲线下的面积，作为有效、实用且公平的性能与效率指标；**b)** 基于详细的能耗模型开展能量效率（EE）分析，其中包含与天线数量和设备数量分别相关的固定和可变电路功耗项，并在现实运行场景下对 NOMA 系统与 mMIMO 系统进行了全面的比较分析，同时利用系统 EE（$\mathcal{E}^{\text{SYST}}$）曲线下的面积；**c)** 在广泛的设备负载范围内分析频谱效率-能效权衡，并验证设备间的公平性；**d)** 最后，在温和条件下，我们提供证据表明，NOMA 能够服务比 mMIMO 系统更多的设备。

本文其余部分安排如下。第二节介绍本研究采用的非正交多址接入和大规模多输入多输出系统模型。第三节提出适用于非正交多址接入和大规模多输入多输出系统的能效-频谱效率公式。第四节分析数值结果。第五节总结全文。

符号说明。在本文中，粗体小写和大写字符分别表示向量和矩阵。算子 $(x)^+ = \max(0, x)$。算子 $[\cdot]^{\mathsf{T}}$、 $\mathbb{E}[\cdot]$ 和 $|\cdot|$ 分别表示转置、期望和基数。随机向量 $\mathbf{x} \sim \mathcal{CN}\{0, \mathbf{I}_m\}$ 服从均值为 0、协方差矩阵为 $\mathbf{I}_m$的圆对称高斯分布。$\mathbf{I}_m$ 是 $m \times m$ 单位矩阵。

II. 系统模型

考虑一个采用时分双工（TDD）的多用户单小区下行链路（DL）传输系统，其中有K 个配备单天线的活跃设备与一台基站通信，该基站在非视距（NLOS）条件下配备 M 根发射天线。集合 $\mathcal{K}$ 由 K 个设备构成，这些设备随机分布在半径为 $d_{\max}$的圆盘内；圆盘区域由两个子圆盘组成，且每个子区域内的设备数量相同；这两个子集分别记为 $\mathcal{K}_H$ 和 $\mathcal{K}_L$。在第一个子集中， $\mathcal{K}_H$ 表示信道系数较高的设备索引，并按降序排列；另一个子集 $\mathcal{K}_L$ 由信道系数较低的设备构成，并按升序排列；索引$k \in \mathcal{K}_H$ 和 $k \in \mathcal{K}_L$ 满足：

$$\begin{aligned} &\mathcal{K} = \mathcal{K}_H \cup \mathcal{K}_L, \quad \text{where} \\ &\mathcal{K}_H = \{1, ..., K/2\} \text{ and } \mathcal{K}_L = \{K/2+1, ..., K\}. \end{aligned} \tag{1}$$

设备 k 的信道向量模型可描述如下：

$$\mathbf{h}_k = \sqrt{\beta_k}\mathbf{h}'_k, \quad k = 1, ..., K, \tag{2}$$

其中， β_k 为大尺度衰落系数，且满足

$$\beta_j > \beta_i, \quad \forall j \in \mathcal{K}_H, \ \forall i \in \mathcal{K}_L. \tag{3}$$

此处， [dB] 中的路径损耗模型定义如下：

$$\beta_k = \beta_0 + 10 \cdot \xi \cdot \log_{10}(d_k), \tag{4}$$

其中， d_k 为用户 k 到基站的距离， ξ 为路径损耗系数， β_0 为参考距离处的衰减。

在每个相干区间内，式（2）中设备 k 的 $\mathbf{h}'_k$ 是从相互独立的瑞利衰落分布 $\mathbf{h}'_k \sim \mathcal{CN}(0, \mathbf{I}_M)$, $k = 1, ..., K$. 中独立抽取的小尺度衰落随机实现。发射信号 $\mathbf{x}_k \in \mathbb{C}^M$ 是设备 k 的波束成形数据信号：

$$\mathbf{x}_k = \mathbf{g}_k \sqrt{p_k} s_k, \tag{5}$$

其中，$\mathbf{g}_k$ 是归一化波束成形向量， p_k 是归一化发射功率， $s_k \sim \mathcal{CN}(0, 1)$是设备 k的数据信号，周期为 T_{s}。第 k 个设备接收到的信号：

$$\begin{aligned} y_k &= \sum_{k'=1}^{K} \mathbf{h}_k^{\mathsf{T}} \mathbf{x}_{k'} + n_k, \\ &= \sqrt{\beta_k}\mathbf{h'}_k^{\mathsf{T}} \sum_{k'=1}^{K} \mathbf{g}_{k'} \sqrt{p_{k'}} s_{k'} + n_k, \\ &= \sqrt{\beta_k p_k}\mathbf{h'}_k^{\mathsf{T}} \mathbf{g}_k s_k + \sqrt{\beta_k}\mathbf{h'}_k^{\mathsf{T}} \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \mathbf{g}_{k'} \sqrt{p_{k'}} s_{k'} + n_k, \end{aligned} \tag{6}$$

其中 $n_k \sim \mathcal{CN}(0, 1)$ 为加性噪声。请注意，该建模适用于非正交多址接入和 mMIMO 系统，但波束成形的选择方式有所不同，相关内容将在后续章节中讨论。

A. 先前操作

由于基站需要预先了解与信道及小区内分布的设备相关的重要信息，包括设备位置、所需速率和信道系数，因此所需的先验信息可能因所考虑的多址方案 [12] 而异。

mMIMO 和 NOMA 系统的方案是在确保满足设备需求的前提下实施的，这一初始步骤已成功完成。第三节 E 小节简要讨论了在 mMIMO 和 NOMA 系统中执行不同功率分配（PA）流程所需的初步信息。

B. 信道状态信息的导频开销

图 1 比较了所研究的 mMIMO 和 NOMA 系统在一个信道相干间隔内的导频-数据传输结构。注意，T_{s} 是传输一个数据信号符号（Data）所需的时间。此外，假设信道相干时间间隔 T 是数据信号符号周期 T $=\iota\cdot T_{\mathrm{s}}$ 的整数倍。训练阶段分配给每个导频的功率足够。相比之下，两个系统中的导频数量以及用于数据传输的相干间隔专用部分均假设相同。

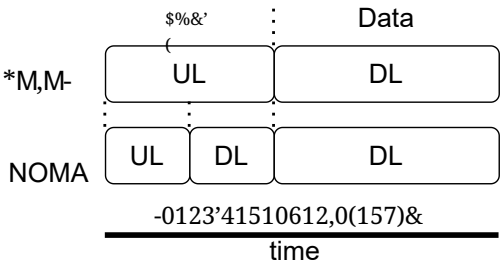


图1：相干时间间隔结构：TDD非视距设置下，大规模多输入多输出和NOMA方案的训练与数据传输结构。

请注意，在NOMA传输中，导频传输步骤分为两部分，一部分用于上行传输导频，另一部分用于接收下行导频确认；这是因为，为了执行串行干扰消除，小区中心设备需要了解由波束成形建立的有效信道。此外，波束成形向量是基于小区中心设备生成的，因此小区边缘设备可达到的速率有限。另一方面，大规模多输入多输出方案的一项显著优势是无需下行导频，因为波束成形所形成的有效信道具有很强的可预测性，即由于信道硬化效应，其增益和相位几乎是确定的 [4]。

假设1：在NOMA系统中，分配给每个下行导频的功率足以到达目标设备。

C. NOMA和大规模多输入多输出系统的波束成形

在大规模多输入多输出系统中，每个设备都由一个波束成形向量提供服务。零迫技术是大规模多输入多输出系统中常用的干扰抑制波束成形方案，因为它为每个设备单独进行波束成形，从而消除所有用户间干扰；而有利传播则有助于在大规模多输入多输出配置中实现此类干扰抑制。此外，要在 NOMA系统中进行零迫预编码，必须理解 NOMA 用户配对的概念。

用户配对：NOMA系统固有地支持多种用户聚类方式：先进行用户排序，再将用户分类为中心用户子集和边缘用户子集，然后进行聚类。由于我们知道，NOMA 的频谱效率与用户之间路径损耗的差值成正比，因此一种自然的选择是尽可能配对路径损耗差异较大的用户 [6]:

$$\Delta\beta_k = \beta_k - \beta_{K+1-k}, \tag{7}$$

形成簇 k，用于 $k=1$，...，$K/2$。配对完成后，需要谨慎选择波束成形向量。因此，在NOMA中，我们假设配对用户使用相同的波束成形向量，即对所有 $k=1$，...，$K/2$，均有 $\mathbf{g}_k=\mathbf{g}_{K+1-k}$。

假设2：在用户配对过程中，我们假设配对用户与基站对齐，从而相同的波束成形可以同时服务所有配对用户。因此，假定每对设备都与基站空间对齐，并使用例如[23], [24], 中所述的定位工具，在NOMA系统中还应预先假设一个用户配对步骤。

假设3：在NOMA系统中，波束成形可同时服务多个对齐的设备；具体而言，本文根据用户配对步骤，允许每个簇中有两个对齐的设备，同时在所采用的完美信道状态信息条件下消除簇间干扰（有利传播）。

在本文中，我们采用由下列向量定义的线性零迫预编码:

$$\mathbf{g}_k = \mathbf{h'}_k(\mathbf{h'}_k^{\mathsf{T}}\mathbf{h'}_k)^{-1}, \tag{8}$$

并满足 $\mathbf{h'}_i^{\mathsf{T}}\,\mathbf{g}_k = 0, \forall\, i \neq k$，即属于不同簇的用户之间具有有利传播效应。

III. 非正交多址接入和 mMIMO 系统中的频谱效率-能效

我们讨论非正交多址接入和 mMIMO 系统中的频谱效率与能效配置。NOMA系统的运行需要对设备进行配对，使同一簇中设备的信道系数适当地不同，从而能够利用功率域。如前所述，通过波束成形进行干扰消除的过程存在一些问题，我们将在下文中加以说明。

A. 零迫下非正交多址接入的数据速率

设备按照式 (1) 所述分为两组，并根据其大尺度衰落系数由式 (9) 表示，再配对组成簇，如图 2 所示。簇 k 由小区中心集合 $\mathcal{K}_H$ 中的一个设备和小区边缘集合 $\mathcal{K}_L$中的一个设备构成。因此，设备被分为两个子集:

$$\begin{aligned}\mathcal{K}_H &= \{\beta_1 > \beta_2 > ... > \beta_{k/2}\}, \quad \text{(center devices set)} \\ \mathcal{K}_L &= \{\beta_K < \beta_{k-1} < ... < \beta_{k/2+1}\}. \ \text{(edge devices set)}\end{aligned} \tag{9}$$

式（9）采用的用户配对方式与 [6], 提出的方式相同，即为尚未配对的设备选择信道系数差异尽可能大的组合。

假设4：本文假设串行干扰消除（SIC）是完美的，并且每个簇只执行一个完美的串行干扰消除阶段，因为每个簇只允许接入2个设备。

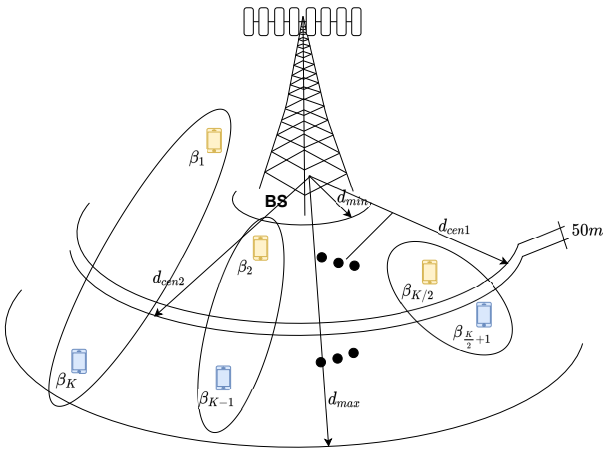


图2：NOMA系统中配对形成的系统模型。大规模多输入多输出系统和NOMA系统在基站部署相同数量的海量天线， M。

簇 k 中设备的瞬时信干噪比（SINR）定义如下：

$$\mathsf{SINR}_k = \frac{\beta_k p_k |\mathbf{h'}_k^{\mathsf{T}} \mathbf{g}_k|^2}{\beta_k \sum_{k' \neq k}^{K} p_{k'} |\mathbf{h'}_k^{\mathsf{T}} \mathbf{g}_{k'}|^2 + 1}. \tag{10}$$

在每个簇中，小区边缘设备将干扰视为噪声并解码其数据信号符号，而小区中心设备可以解码小区边缘设备的数据信号符号并执行串行干扰消除（SIC），从而在假设3下有效消除小区边缘设备造成的干扰。

为了执行串行干扰消除，小区中心设备必须能够解码发给小区边缘设备的数据信号，即设备k处小区边缘设备的遍历信干噪比 sinr_{K+1-k}，定义为 $\mathsf{sinr}_{k,\ K+1-k}$，必须大于或等于第 k 个小区中心设备的遍历信干噪比。因此，给定上行链路（大规模多输入多输出和NOMA）及下行链路（NOMA）的导频开销，并假设所有接收机均具有完美的信道状态信息，同时满足假设4，必须满足以下条件 [4], [10]:

$$\mathbb{E}[\mathsf{SINR}_{k,K+1-k}] \geq \mathbb{E}[\mathsf{SINR}_k], \tag{11}$$

其中

$$\mathsf{SINR}_{k,K+1-k} = \frac{\beta_k p_{K+1-k} |\mathbf{h'}_k^{\mathsf{T}} \mathbf{g}_k|^2}{\beta_k \sum_{k' \neq K+1-k}^{K} p_{k'} |\mathbf{h'}_k^{\mathsf{T}} \mathbf{g}_{k'}|^2 + 1}. \tag{12}$$

在此，必须通过适当选择发射功率来满足式（11）中的条件。

在假设1–4下，簇 k 中设备的可达遍历速率，即 $\mathcal{K}_H$ 子集中设备 k 和 $\mathcal{K}_L$ 子集中设备 $K+1-k$ 的可达遍历速率，由用户中心设备的遍历速率贡献给出：

$$\mathsf{R}_k^{\mathrm{NOMA}} = \tau \mathbb{E}\left[\log_2\left(1 + \mathsf{SINR}_k\right)\right], \qquad \forall k \in \mathcal{K}_H \tag{13}$$

单位为 [比特/秒/赫兹], , 用户边缘设备的速率为:

$$\mathsf{R}_{K+1-k}^{\mathrm{NOMA}} = \tau \mathbb{E}\left[\log_2\left(1 + \mathsf{SINR}_{K+1-k}\right)\right], \qquad \forall k \in \mathcal{K}_L \tag{14}$$

其中 $\tau = \left(1 - \frac{K \cdot T_{\mathsf{P}}}{\mathsf{T}}\right)$ 表示每个信道相干间隔（T）中用于数据传输的部分。

假设信道状态信息完美，采用零迫预编码消除簇间干扰，并使用随机矩阵理论结果 [25], , 代入式 (10)、(13) 和 (14)，可得第 k 个簇的非正交多址接入可达速率：

$$\mathsf{R}_{\mathsf{cl}\text{-}k}^{\mathrm{NOMA}} = \tau \mathbb{E}\left[\log_2\left(1 + \bar{M}\beta_k p_k\right)\right] + \tau \mathbb{E}\left[\log_2\left(1 + \frac{\beta_{K+1-k} p_{K+1-k}}{\beta_{K+1-k} p_k + 1}\right),\right] \tag{15}$$

$\forall k \in \mathcal{K}$ 且 $\bar{M} = M + 1 - K/2$。因此，NOMA系统可以运行至 $K < 2M - 1$。本节中这些表达式的详细推导见 [4] 和 [10]。

B. 采用零迫的 mMIMO 数据速率

在采用零迫预编码的 mMIMO 系统中，设备 k 的遍历可达数据速率为：

$$\mathsf{R}_k^{\text{M-MIMO}} = \tau \mathbb{E}\left[\log_2\left(1 + \mathsf{SINR}_k\right)\right], \quad [\text{bits/s/Hz}] \tag{16}$$

其中，SINR_k 定义见式 (10)。因此，上述 mMIMO 可达数据速率公式可写为：

$$\mathsf{R}_k^{\text{M-MIMO}} = \tau \mathbb{E}\left[\log_2\left(1 + (M - K) p_k \beta_k\right)\right], \quad [\text{bits/s/Hz}], \tag{17}$$

其中，$(M - K)$ 通过随机矩阵理论得出，表示接收信号 [25] 的相干阵列增益。在采用线性预编码和合并器时，当 $K < M$ 时，mMIMO 系统可保持稳定运行。最后，系统平均和速率（avg-sum-rate）定义如下：

$$\mathcal{R}^{\text{M-MIMO}} = \sum_{k=1}^{K} \mathsf{R}_k^{\text{M-MIMO}} \qquad \text{and} \qquad \mathcal{R}^{\mathrm{NOMA}} = \sum_{k=1}^{\mathcal{K}_H} \mathsf{R}_{\mathsf{cl}\text{-}k}^{\mathrm{NOMA}}.$$

文献中已对 mMIMO 系统方程进行了深入研究，相关内容可参见 [25] 和 [26]。

C. 能量效率

能效指标是系统为传输和接收/解码此类信息所消耗的总能量与接收到的有效信息比特数之比。系统数据速率决定了目的端接收到的有效信息比特数。通常会忽略发射端和接收端处理信号所需的功耗；在这种情况下，功耗仅按辐射发射功率成比例计算。基站中的天线数量增加以及5G系统中的设备数量增多，可能导致能效目标难以实现。一般而言，平均能效可表示为：

$$\mathsf{EE} = \frac{\sum_{k=1}^{K} R_k}{P_{\mathrm{TOT}}}, \qquad [\text{bits/Joule/Hz}], \tag{18}$$

其中， P_{TOT} 是整个通信系统的总功耗。该功耗应考虑发射功耗，例如射频功率放大器效率低下、基带信号处理和冷却等。因此，需要采用更现实、更详尽的能耗模型。

基于 [2], , 本研究采用的功耗模型考虑两种功耗项：a）固定项；b）随天线数量 M 和设备数量K变化的项。可变项的产生与收发机链路、编码/解码、信道估计和预编码有关。设基站的计算效率为 L 每焦耳运算次数。具体如下：

射频功率： P_{RF} 是向达到信干噪比目标的活跃设备传输信号所消耗的功率， $0 < \varpi \leq 1$ 是功率放大器的效率。

固定功耗： P_{FIXED} 是基站消耗的功率，与小区内发射天线和设备的数量无关，由两项构成： P_0 包括回传基础设施、控制信令和基带处理器的功耗； P_{SYN} 为所有基站共用的单振荡器。

$$P_{\text{FIXED}} = P_0 + P_{\text{SYN}}$$

仅取决于 K： P_{K} 由对设备信息符号进行编码与调制所消耗的功率（以 P_{COD} 表示）、基站解码 K 个信息符号序列所消耗的功率（定义为 P_{DEC}），以及接收功率（以 P_{RX} 表示）构成；该项还需乘以 K 。此外，零迫预编码开销 [27] 的一部分仅取决于 K^3。

$$P_{\mathsf{K}} = K(P_{\text{COD}} + P_{\text{DEC}} + P_{\mathsf{RX}}) + K^3\frac{2}{3LT}$$

仅取决于 M。项 P_{M} 表示发射功率（P_{TX}），因此

$$P_{\mathsf{M}} = MP_{\mathsf{TX}}$$

取决于 K 和 M：项 P_{KM} 是零迫预编码的开销（源于基于 LU 分解的矩阵求逆） [27], ，该开销取决于设备数量、天线数量以及信息符号向量。

$$P_{\mathsf{KM}} = MK\frac{3+T}{TL} + MK^2\frac{2}{TL}$$

将各部分相加，可得系统的总功耗：

$$P_{\text{TOT}} = \frac{P_{\text{RF}}}{\varpi} + P_{\text{FIXED}} + P_{\mathsf{K}} + P_{\mathsf{M}} + P_{\mathsf{KM}} \quad [\text{W}]. \tag{19}$$

D. 功率分配策略

接下来，我们介绍三种广为人知且常用的功率分配策略。鉴于非正交多址接入的固有特性，我们对经典注水（WF）算法进行了修改，使其能够应用于 NOMA系统。这些修改，即 Δ-WF，可确保不会丢弃任何配对设备，同时不破坏设备配对。为保证每对设备之间具有一定程度的功率差异，NOMA系统中的功率分配 Δ- WF过程分为两步：a）首先，为各簇分配功率；b）再为配对设备分配功率。由此，我们可以分析系统的行为并比较其结果。

请注意，mMIMO 和 NOMA 系统在基站部署了相同数量的大量天线， M。因此，得益于大规模多输入多输出配置固有的信道硬化效应 [1], [25] ，配备线性零迫预编码的 M 天线上的小尺度衰落会消失，预编码向量如式（ 8 ）所示。因此，对于基于大量天线的系统，可以仅将路径损耗系数 β_k 视为功率分配策略中的主要参数。

1) 等功率分配（EPA）：等功率分配（EPA）是一种简单、朴素的功率分配策略，为所有设备分配相同的功率。在大规模多输入多输出中，无论设备与基站的距离远近，所有设备都分配相同的发射功率。在非正交多址接入中，功率分配分为两个步骤。第一步，在各对设备之间平均分配功率，然后再为每个设备平均分配功率。应用于大规模多输入多输出的等功率分配策略可定义为：

$$p_k = \frac{P_{\text{RF}}}{K} \quad [\text{W}], \quad \forall \text{k} \in \mathcal{K}. \tag{20}$$

对于应用于 NOMA 的 EPA 流程，它由两个步骤组成：第一步，可以简单地将每个簇的参考功率定义为：

$$p_{\text{ref}}^{\mathsf{cl}} = \frac{2 \cdot P_{\text{RF}}}{K} \quad [\text{W}], \quad \forall\, k \in \mathcal{K}_H. \tag{21}$$

第二步，同一簇内各设备之间的功率分配定义为：

$$p_k^{\mathsf{cl}\text{-}k} = p_{K+1-k}^{\mathsf{cl}\text{-}k} = \frac{p_{\text{ref}}^{\mathsf{cl}}}{2}. \tag{22}$$

2）比例信道反演功率分配（PICPA）：这是本研究采用的另一种功率分配技术。与对所有设备分配相同功率的 EPA 技术不同，PICPA 会为信道条件最差的设备分配更多功率，从而提高设备间的公平性。这种功率分配以牺牲平均和速率为代价，换取所有用户之间的公平性。

应用于大规模多输入多输出的信道功率倒数比例分配（PICPA）策略可定义为：

$$p_k = P_{\text{RF}}\frac{\beta_k^{-1}}{\sum_{k=1}^{K}\beta_k^{-1}} \quad [\text{W}], \quad \forall k \in \mathcal{K}, \tag{23}$$

而应用于非正交多址接入的 PICPA 流程分为两个步骤；第一步，将功率平均分配给 $K/2$个簇：

$$p_{\mathsf{ref}}^{\mathsf{cl}} = \frac{2 \cdot P_{\text{RF}}}{K} \quad [\text{W}], \quad \forall k \in \mathcal{K}_H, \tag{24}$$

之后，将每个簇中设备的功率定义为：大尺度衰落 β_k较小的设备分配更多功率：

$$p_k^{\mathsf{cl}\text{-}k} = p_{\mathsf{ref}}^{\mathsf{cl}}\frac{\beta_{K+1-k}}{\beta_k - \beta_{K+1-k}}, \quad \text{and} \quad p_{K+1-k}^{\mathsf{cl}\text{-}k} = p_{\mathsf{ref}}^{\mathsf{cl}} - p_k^{\mathsf{cl}\text{-}k}, \tag{25}$$

其中， $p_k^{\mathsf{cl}\text{-}k}$是分配给第 cl-$k$ 个簇中设备 k 的功率，而 $p_{K+1-k}^{\mathsf{cl}\text{-}k}$ 是分配给同样属于第 - 个簇的设备（ $K+1-k$ k ）的功率。

3）经典注水（WF）算法：在大规模多输入多输出系统中应用注水（WF）算法，可得到最优（最大）系统和速率解。然而，由于信道条件恶化，部分设备会被剥夺服务。大规模多输入多输出的 WF 功率分配策略如下：

$$\mu = \frac{1}{|\mathcal{K}|}\left(P_{\text{RF}} + \sum_{k=1,k\in\mathcal{K}}^{|\mathcal{K}|}\frac{1}{\beta_k}\right), \tag{26}$$

$$P_{\text{RF}} = \sum_{k=1,k\in\mathcal{K}}^{|\mathcal{K}|} p_k\,, \quad \text{where} \quad p_k = \left(\mu - \frac{1}{\beta_k}\right)^+, \forall k \in \mathcal{K}$$

$$\text{and} \quad \mathbf{p} = [p_1, p_2, ..., p_{|\mathcal{K}|}],$$

使用算子 $(z)^+ = \max(0, z)$。请注意，总可用功率的约束数值设为 P_{RF} [W]。算法1描述了经典WF程序。

另一方面，在 NOMA系统中直接应用注水算法会损害配对形成，即 $\mathcal{K}_L$ 集合中的设备实际上会被排除在服务之外，从而破坏配对。为便于比较，我们提出对经典WF进行如下修改。

算法1： 面向大规模多输入多输出的经典注水（注水）

输入： $\mathcal{K}$，$\mathrm{P_{RF}}$，$K=|\mathcal{K}|$

NP $\neq\oslash$;
当 *(NP* $\neq\oslash$*)* **时执行**
  求解式（26）$\rightarrow$ **p**;
  NP $\leftarrow$ 识别**p**中的零位置;
  $\mathcal{K}/\{k\}_{\mathrm{NP}}\leftarrow$ 排除标记为 NP 的**p** 设备
结束
输出： $\mathbf{p}=[p_1,p_2,\ldots,p_{|\mathcal{K}|}]$

4) 用于NOMA的 Δ-WF：在NOMA中以应用于大规模多输入多输出的相同方式应用经典WF，会导致某些已形成的配对被破坏，因为应用注水算法后，部分设备会从系统中被排除，使同一簇内设备之间的NOMA功率差（Δ）消失。因此，我们建议相应地修改经典WF程序，使其适用于NOMA。 Δ-WF算法的步骤如下。

在非正交多址接入中，功率分配分为两个步骤。第一步是在簇之间进行分配。因此，为避免已形成的配对被拆散，我们提出根据配对设备的大尺度衰落差异进行注水，如式（7）所定义： $\Delta\beta_k=(\beta_k-\beta_{K+1-k})$ 在每个 $\mathcal{K}_L$ 和 $\mathcal{K}_H$ 子集中，见式（9）。在该过程的第二步中，假设采用完美连续干扰消除（SIC），在组内的设备之间分配功率；要实现这一点，必须满足式（11）的条件。非正交多址接入中改进的Δ-注水功率分配策略所采用的新水位定义如下

$$\mu=\frac{2}{\mathcal{K}_H}\left(P_{\mathrm{RF}}+\sum_{k=1,k\in\mathcal{K}_H}^{|\mathcal{K}_H|}\frac{1}{\Delta\beta_k}\right), \tag{27}$$

$$P_{\mathrm{RF}}=\sum_{k=1,k\in\mathcal{K}_H}^{|\mathcal{K}_H|}p_{\text{cl-}k},\qquad \forall k\in\mathcal{K}_H$$

$$\text{where}\quad p_{\text{cl-}k}=\left(\mu-\frac{1}{\Delta\beta_k}\right)^{+},$$

$$\text{and}\quad \mathbf{p}_{\text{cl-}k}=[p_1,p_2,...,p_{|\mathcal{K}_H|}],$$

在第二步中，定义分配给第k个簇中两个设备的功率为:

$$p_{K+1-k}^{\text{cl-}k}=p_k^{\text{cl-}k}=\frac{p_{\text{cl-}k}}{2} \tag{28}$$

算法 2 总结了所提出的 Δ-注水功率分配流程，旨在提高NOMA系统的频谱效率。

算法 2： Δ-注水（修订版），适用于 NOMA系统

输入： $\mathcal{K}_H$，$\mathcal{K}_L$，$\mathrm{P_{RF}}$

NP $\neq\oslash$;
当 *(NP* $\neq\oslash$*)* **时**
  求解式 (27) $\rightarrow$ $\mathbf{p}_{\text{cl-}k}$;
  NP $\leftarrow$ 在$\mathbf{p}_{\text{cl-}k}$中识别空位置;
  $\mathcal{K}_H/\{k\}_{\mathrm{NP}}\leftarrow$ 从$\mathbf{p}_{\text{cl-}k}$ 中排除标记为以下内容的设备 NP
结束
输出： $\mathbf{p}_{\text{cl-}k}=[p_1,p_2,\ldots,p_{|\mathcal{K}_H|}]$

复杂度分析： 在复杂度比较分析中，NOMA系统中用于功率分配的 Δ-WF 算法（算法2）相比经典WF程序（算法1）只需执行两个简单的额外操作：a) 在公式 (27) 中进行 $(\beta_k-\beta_{K+1-k})$ 所示的减法；以及 b) 在公式 (28) 中除以二。此外，NOMA系统与 mMIMO系统需要不同的先验信息，才能相应地执行 PA 流程。

E. 功率分配策略的先验信息 *p*

为实施功率分配策略，基站需要具备先验信息，如表 I 所定义。其中部分必要信息可通过专用导频传输步骤获得，但会产生一定开销，如第二节中所述。此外，还有一个已知的初步步骤，即必须完成所有设备的空间定位和路径损耗估计。具备此类先验信息后，即可执行用户排序和用户配对步骤。

表 I：PA 流程所需的先验信息

PA		β_k	h'_k	K	$\mathcal{K}_H$	$\mathcal{K}_L$	p_k，公式
EPA	大规模多输入多输出	–	–	▲*	–	–	(20)
	NOMA	–	–	▲	–	–	(21), (22)
PICPA	大规模多输入多输出	▲*	–	–	▲*	▲*	(23)
	NOMA	▲	–	–	▲*	▲	(24)、(25)
WF	大规模多输入多输出	▲*	–	–	▲*	▲*	(26)、算法 1
Δ-WF	NOMA	▲	–	–	▲*	▲	(27、28)、算法 2

▲ 所需信息，需要先验* 通过导频开销获得–不需要信息

IV. 数值结果

本节展示了针对非正交多址接入和大规模多输入多输出系统所提出分析的数值评估结果。本节采用的仿真系统及信道参数值见表 II。基站位于小区中心，在典型的非视距信道传播场景中配备了大量M 基站发射天线。同时，设备随机分布在小区区域内，并划分为两个子集， $\mathcal{K}_L$ 和 $\mathcal{K}_H$，如图 2 所示。在所有仿真中，我们采用块衰落模型，将时频资源划分为相干时间间隔 (T)；在每个间隔内，信道保持恒定且频率平坦。相干时间间隔以符号传输周期 (T_{s}) 的整数倍计量。系统和信道场景使用 Matlab 2019 软件进行仿真，运行平台配备一块 Intel HD Graphics 6000 GPU、Intel(R) Dual-Core(TM) I5 CPU @ 1.6 GHz 和 8 GB 随机存取存储器。

表 II：仿真参数

参数	**数值**
基站天线	$M=64$、128 和 256
小区内最大设备数	$K=\zeta\cdot M$ (非正交多址接入) $K=M$ (大规模多输入多输出)
小区负载	$\rho=K/M$
可用总射频功率	$P_{\mathrm{RF}}=1$瓦
NOMA 配对 / 簇	$N=K/\zeta=K/2$
每个簇中的NOMA设备数	$\zeta=2$
每个设备的天线数	1
小区边缘长度	$d_{\max}=350$ 米
强设备位置	$[d_{\min};\ d_1]\in[50;\ 100]$ 米
弱设备位置	$[d_2;\ d_{\max}]\in[150;\ 350]$ 米
阵列增益 MIMO设备	$M-K$
阵列增益 NOMA k_H	$M+1-K/2$
数据信号符号周期	T_{s}
相干时间间隔	$\mathrm{T}=512\cdot T_{\mathrm{s}},\quad \iota=512$
信道	
路径损耗指数	$\xi=3.78$
参考点处的衰减 a d_0	$\beta_0=130$ [dB]
# 蒙特卡洛实现次数	1000

A. 频谱效率比较

本小节通过每次将设备数量增加两个，直至达到负载上限 $\rho = 2$，对大规模多输入多输出和非正交多址接入的频谱效率性能进行分析。结果涵盖 $M = 64$、128 和 256 根基站天线。图 3.(a) 展示了射频功率按照等功率分配策略进行分配时的频谱效率结果，此时每个设备获得相同的功率分配值。当负载为 $\rho < 0.6$ 时，在所有情况下大规模多输入多输出系统的表现均优于非正交多址接入系统。然而，随着设备负载增加，在每种 M 场景下，非正交多址接入系统的频谱效率均高于大规模多输入多输出系统，$\rho > 0.6$。采用零迫非正交多址接入、配备 $M = 256$ 根天线并且 $\rho \approx 0.76$ 时，达到最高平均频谱效率 373 [比特/秒/赫兹]。此外，可以推断，当负载超过 1 时，大规模多输入多输出系统无法工作，这是因为设备满负载时阵列增益降至 0；而非正交多址接入系统则能正常工作，直至负载达到 $M \cdot \zeta$，其中 ζ 为每个簇中的设备数量。

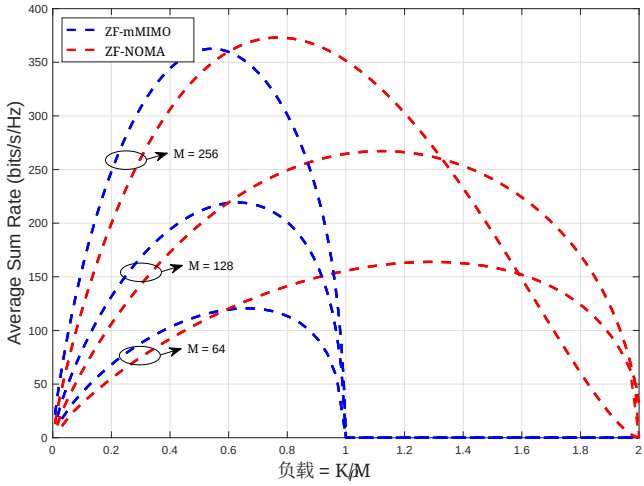
图3.(b)展示了在采用PICPA方法沿基站天线为每台设备分配可用射频功率时，大规模多输入多输出和NOMA系统所达到的频谱效率结果。对于三种基站天线配置 $M = 64$、128和256，在负载 ρ 超过 ≈ 0.62 之前，大规模多输入多输出系统的平均频谱效率最大值均高于NOMA系统。该功率分配技术会向信道条件最差的设备分配更多功率，使频谱效率结果在EPA下达到最大值。

图3.(c)展示了应用于大规模多输入多输出系统的传统注水算法。在这种功率分配方法下，我们指出，在NOMA系统中形成配对不可行。事实上，注水算法能够最大化系统频谱效率，因为它会向信道条件更好的设备分配更多功率。相反，信道条件较差（低于水位）的设备将被排除在服务之外。

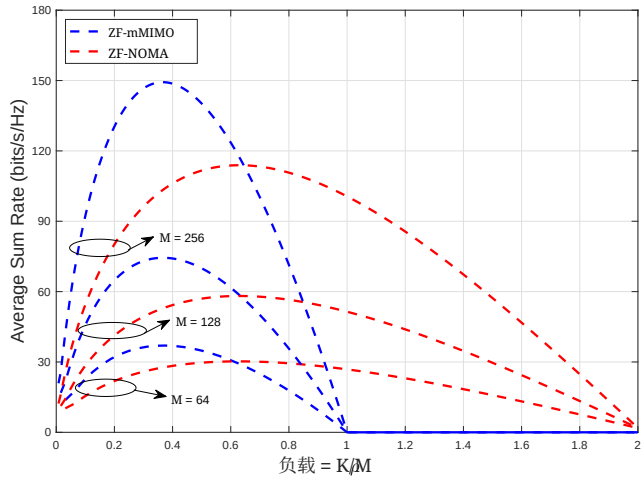
经典注水算法已适配于每次始终丢弃一对设备的 NOMA 系统。当 M 较低时，与大规模多输入多输出中的经典 WF PA 相比，这种适配显著提升了平均和速率。 Δ-WF 功率分配过程保留了 NOMA 系统中的成对簇结构，并向大尺度衰落系数差异较大的簇分配更多功率。图 3.(c) 显示，在 NOMA 系统中采用修改后的 WF 时，在 $\rho = 1$（$K \approx 256$ 台设备）条件下可达到最大平均频谱效率 ≈ 361 [比特/秒/赫兹], 。此外，当基站天线数量较少（$M = 64$或 128）时，NOMA 的峰值高于大规模多输入多输出；例如，对于 $M = 64$ 根天线，大规模多输入多输出的频谱效率峰值出现在负载 $\rho \approx 0.7$时，而 NOMA 的频谱效率峰值出现在 $\rho \approx 1.2$。然而，随着基站天线数量增加，NOMA 的频谱效率优势会减小。

PA 流程后的活跃设备数量。图 4 展示了应用 PA 方法后活跃设备的数量：在 EPA 和 PICPA 算法中，所有设备都保持活跃。然而，在经典WF大规模多输入多输出系统中，当设备数量增加到超过 $\rho \approx 0.25$ 时，一半设备会被淘汰；而在 Δ-WF NOMA PA 中，活跃设备的比例始终更高，例如在 $\rho \approx 1.1$ $M = 256$ 个天线时，最差情况下仍高于 70%。

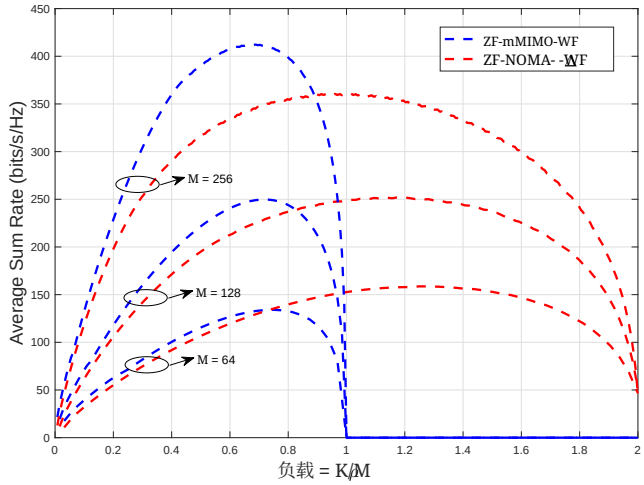
图 5 总结了 NOMA 采用 EPA、mMIMO 采用 WF，以及 NOMA 采用改进的 Δ-WF 时所达到的平均和速率曲面对应的频谱效率×$\rho \times M$ 结果。在初始负载阶段，$\rho < 0.65$，ZF-mMIMO 中的经典 WF PA 在达到该负载（粉色曲面）之前表现更优。当天线数量较少，即 $M = 64$且 ρ 介于 0.7 和 1.8 之间时，应用于 NOMA 的 EPA PA（绿色



(a) EPA



(b) PICPA



(c) ZF-mMIMO 中的经典WF和 ZF-NOMA 中的 Δ-WF

图 3： 在设备负载为 $0 < \rho \leq 2$ 时，采用四种功率分配方法（EPA、PICPA、WF 和 Δ-WF）得到的平均和速率。该平均值基于 1000 个随机设备位置计算得出。

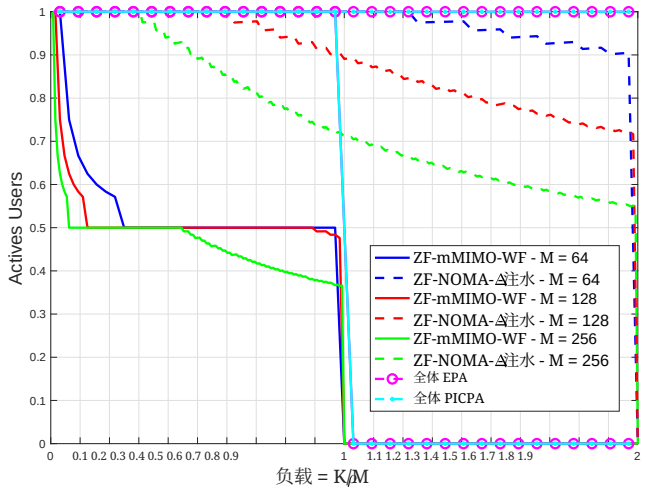


图 4：PA 流程后活跃设备的平均数量与 $0 < \rho \leq 2$ 范围内的设备负载之间的关系。该平均值基于 1000 个随机设备位置计算得出。

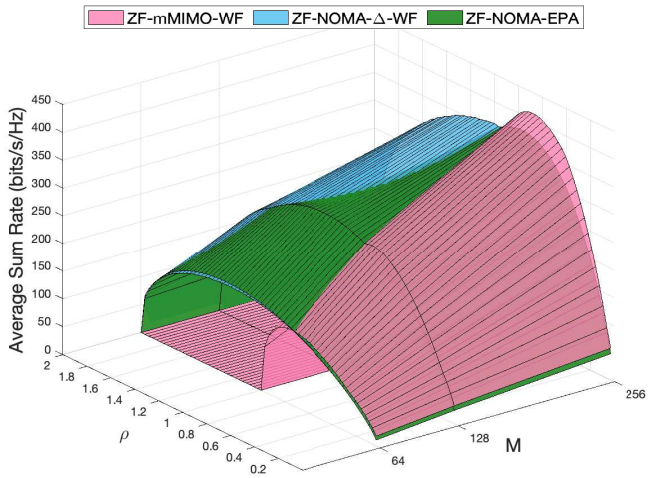


图 5：负载为 ρ and 时的平均和速率 M.

曲面）表现更优。此外，当 $M = 128$且 $0.8 < \rho < 1.6$时，ZF-NOMA-EPA 表现更优（绿色曲面）。当基站中的天线数量较多，即 $M = 256$ 仅在较短的设备负载范围内时，$0.86 < \rho < 0.97$，ZF-NOMA-EPA 的频谱效率表现更优。最后，当 $\rho > 0.97$时，改进的 Δ-WF 表现出具有竞争力的结果（蓝色曲面）。

B. Jain公平性指数

另一项重要分析是考察设备之间的公平性，即了解小区内活跃设备所达到的传输速率差异。对于这项指标，我们使用 [28] and 中所述的 Jain公平性指数，其定义如下：

$$\mathcal{F}_{\mathrm{M}}^{\mathrm{SYST}} = \frac{\left(\sum_{k=1}^{M} R_k\right)^2}{M \sum_{k_1}^{M} R_k^2}. \tag{29}$$

图 6 展示了随着设备负载增加至 $\rho = 2$，采用 EPA、PICPA、WF 和 Δ-WF 功率分配方案的 NOMA 与 mMIMO 系统所能达到的公平性曲线。图 6(a) 展示了采用等功率分配策略时的 Jain 公平性指数；对于 $\rho < 1$，mMIMO 系统的表现优于 NOMA $\mathcal{F}$，另一方面，无论 M 为何，NOMA 在几乎所有设备负载下都能达到 $\mathcal{F}_{\mathrm{M}}^{\mathrm{NOMA}} \approx 0.5$。

图 6(b) 展示了采用 PICPA 方法时的 Jain 公平性指数。尽管此功率分配方案下的频谱效率结果较低，mMIMO 仍取得了最佳公平性结果，其 $\mathcal{F}_{\mathrm{M}}^{m\mathrm{MIMO}}$ 始终保持在 0.85 以上，而 NOMA 仍低于 0.5。

图 6(c) 展示了采用 WF 和 δ-WF 时的 Jain公平性指数。这些算法本身倾向于向信道条件较好的设备分配更多功率，因而会降低设备之间的公平性。在此方法中，可以观察到基站中的天线数量 M 对公平性结果有显著影响。

C. 能效比较

能量效率（EE）是用于比较系统性能的另一项重要性能指标。本节采用了一种功耗模型，其中包括固定电路功耗部分，以及随天线数量 M and 和设备数量 K 变化的部分，并遵循公式 (19)。

表 III [2] 列出了本小节能效分析与比较所采用的参数数值。图 7 展示了在考虑天线的确切数量时，采用 EPA、PICPA、WF 和Δ-WF 功率分配方案的能效性能。图 7.(a) 展示了采用 EPA 时的能效性能。在此方法中，所有设备获得相同的功率。大规模多输入多输出的平均能效比非正交多址接入高约 13% 至 20%。可以观察到，增加基站中的天线数量会增加功耗，从而降低能效。再次，在设备负载为 $0.6 < \rho \leq 2.0$ 时，非正交多址接入的性能优于大规模多输入多输出系统。

表 III：能效分析所采用的参数数值 [2]

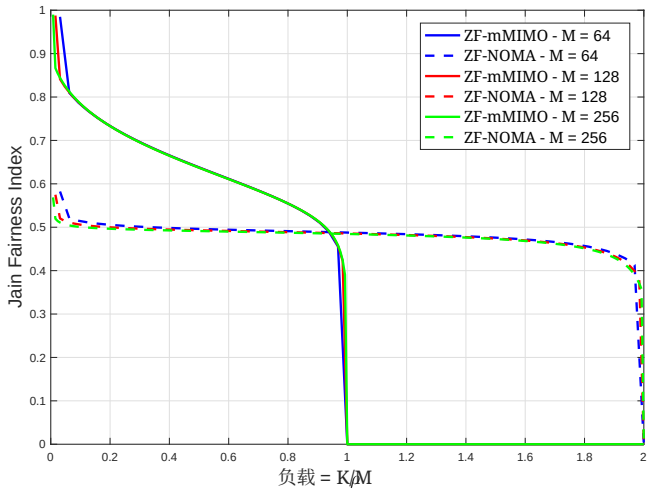
参数	数值
回传基础设施	$P_0 = 2$ 瓦
单振荡器	$P_{syn} = 2$W
编码与调制	每台设备 $P_{COD} = 4$W
解码与解调	每台设备 $P_{DEC} = 0.5$ W
接收功率	$P_{RX} = 0.3$ W/设备
发射功率	$P_{TX} = 1$ W/天线
功率放大器效率	$\varpi = 0.3$
每焦耳运算次数	每焦耳$\mathrm{L} = 10^9$ 次运算

图 7.(b) 展示了采用 PICPA 时的能效性能。在该方法中，分配给信道系数较差的设备的功率更多，因此 NOMA 和 M-MIMO 系统的能效性能均有所下降，其最大值分别为 0.25 和 0.16 [比特/瓦]，对应大规模多输入多输出和 NOMA。大规模多输入多输出达到的最大能效通常比 NOMA 高约 50%。然而，当设备负载为 $\rho \geq 0.65$ 时，NOMA 的能效性能优于大规模多输入多输出。

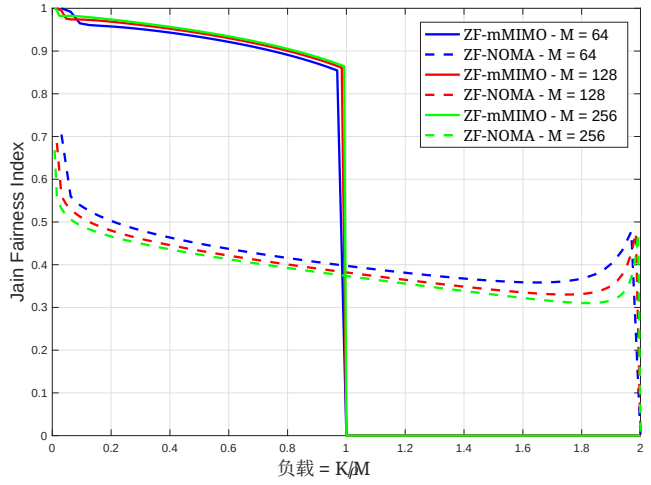
图 7.(c) 展示了大规模多输入多输出采用经典WF以及 NOMA 采用 Δ-注水算法时的能效性能。可以确认，在其稳定运行的范围内，即 $0 < \rho < 1$，大规模多输入多输出的能量效率更高。请注意，对于不同数量的基站天线，大规模多输入多输出达到的最大能效约比 NOMA 高 70%。最后，只有在设备负载较高时，$\rho > 0.95$，NOMA 的能效才会高于大规模多输入多输出。

在所有分析的系统场景中，采用经典 WF PA 流程的大规模多输入多输出可实现更高的最大能效。对于 $\rho < 1$，大规模多输入多输出的能效表现优于非正交多址接入。另一方面，非正交多址接入可服务的设备数量更多（是大规模多输入多输出的两倍）。

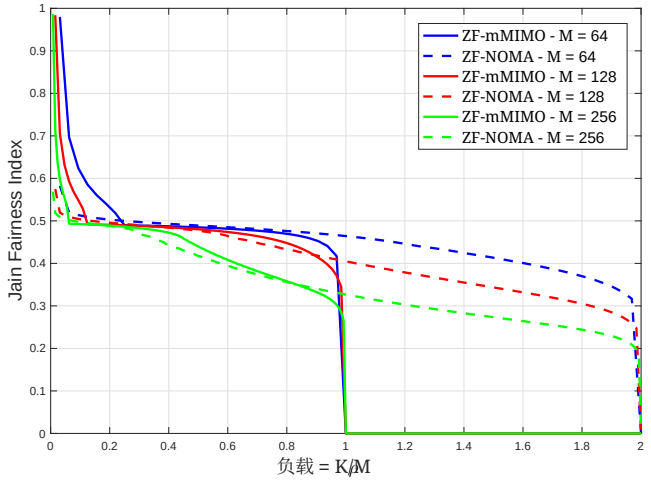
图 8 以曲面图总结了大规模多输入多输出采用 WF 时的最佳能效结果，且在整个



(a) EPA

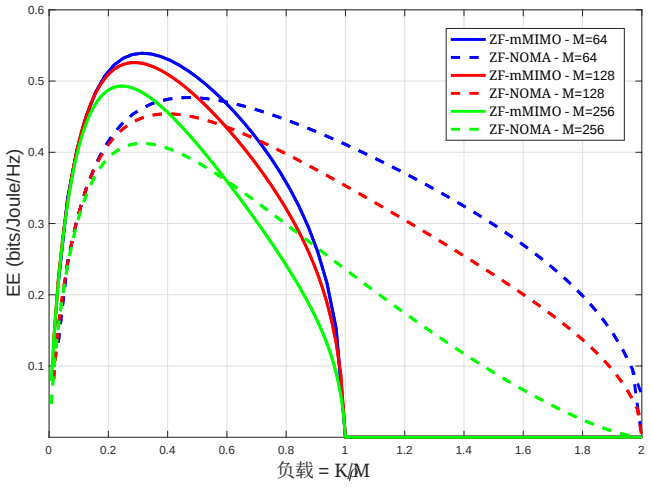


(b) PICPA

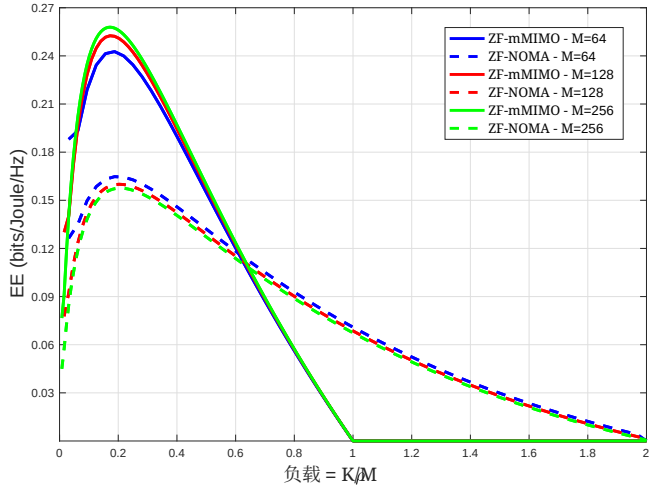


(c) 注水和 Δ-注水

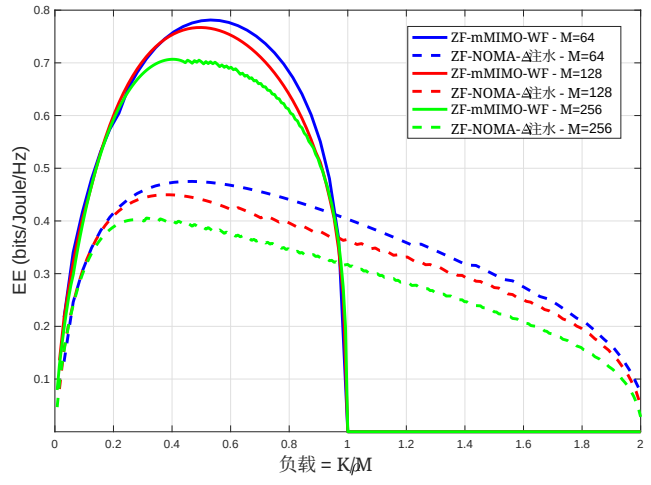
图 6： 三种功率分配方法下非正交多址接入和大规模多输入多输出系统的公平性：(a) EPA；(b) PICPA；(c) 注水和 Δ-注水。平均值基于 1000 个随机设备位置计算。



(a) EPA



(b) PICPA



(c) 注水和 Δ-注水

图 7： 采用三种功率分配方法时非正交多址接入与大规模多输入多输出的能效对比：(a) EPA；b) PICPA；(c) 注水和 Δ-注水。随机设备位置下 1000 次模拟得到的平均能效。

其运行稳定的负载范围。对于设备负载$\rho > 1$，当负载低于$\rho = 2$时，非正交多址接入的能效较低。此外，考虑到基站天线数量最少的情况，采用 EPA 的非正交多址接入优于采用改进型 Δ-注水的非正交多址接入；然而，随着基站天线数量增加，采用 Δ-注水的非正交多址接入可获得略高的能效。

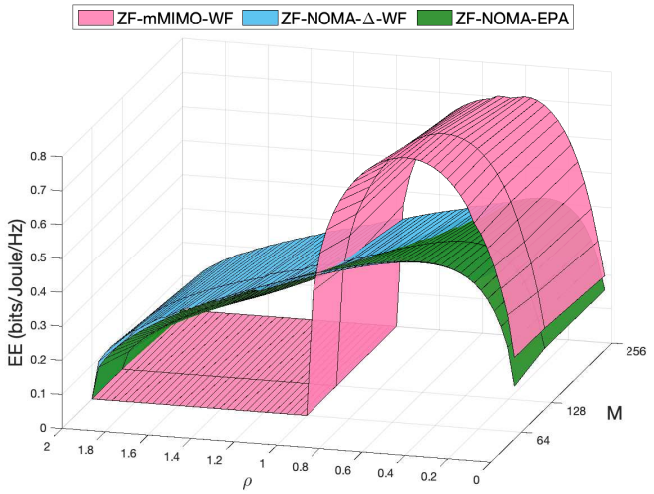


图8：负载为 ρ 且 M 为给定值时的能效。

D. 频谱效率和能效曲线下面积

为进行公平比较，可以考察设备负载变化范围内的平均频谱效率和能效，并将其按天线归一化；非正交多址接入和大规模多输入多输出系统均可达到这些指标。因此，我们考察图3和图7中相应的频谱效率曲线和能量效率曲线下面积，即：

$$\mathcal{S}_{\mathrm{M}}^{\mathrm{SYST}} = \frac{1}{M} \cdot \int_0^{\rho=2} \overline{\mathrm{SE}}(\rho)\, d\rho \qquad \left[\frac{\text{bits/antenna}}{\text{s} \cdot \text{Hz}}\right]$$

and

$$\mathcal{E}_{\mathrm{M}}^{\mathrm{SYST}} = \frac{1}{M} \cdot \int_0^{\rho=2} \overline{\mathrm{EE}}(\rho)\, d\rho \qquad \left[\frac{\text{bits/antenna}}{\text{Joule} \cdot \text{Hz}}\right],$$

其中，$\overline{\mathrm{SE}}(\rho)$表示平均总体系统总速率，$\overline{\mathrm{EE}}(\rho)$表示在特定设备负载 ρ下实现的平均总体系统能量效率。因此，通过比较图3和图7中频谱效率曲线和能量效率曲线下相应的面积，我们得到了图9。

从频谱效率的角度来看，在采用等功率分配策略时，当基站天线数量为 $M = 64$ 时，可获得更高的频谱效率曲线下面积比增益：

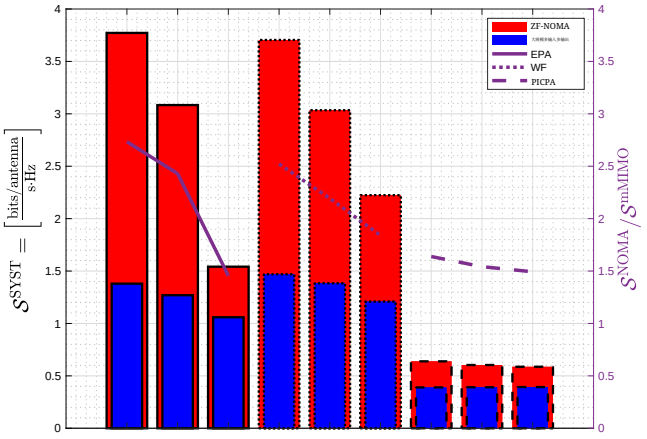
$$\mathcal{S}_{\mathrm{M}=64}^{\mathrm{NOMA}} \approx 2.7 \cdot \mathcal{S}_{\mathrm{M}=64}^{m\mathsf{MIMO}}.$$

请注意，随着天线数量 M 增加，上述比值会降低。同样地，在采用注水策略时，增益趋势保持不变。相比之下，在采用功率迭代协方差矩阵预编码与分配策略时，该比值基本保持不变。

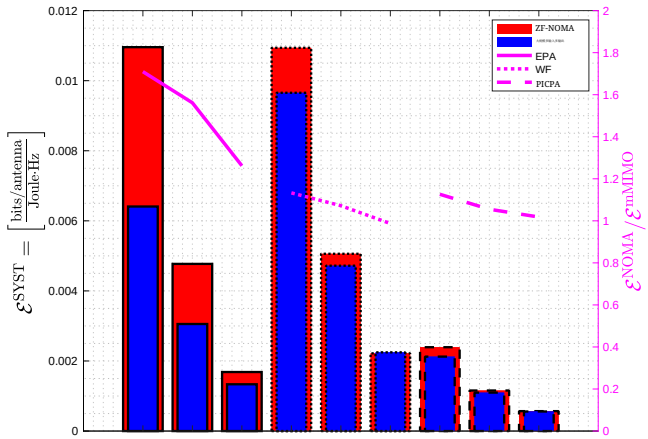
此外，从能效的角度来看，在等功率分配策略下，当基站配备 $M = 256$ 根天线时，可获得更高的比值：

$$\mathcal{E}_{\mathrm{M}=256}^{\mathrm{NOMA}} \approx 1.8 \cdot \mathcal{E}_{\mathrm{M}=256}^{m\mathsf{MIMO}}.$$

由此可以得出，在几乎所有场景中，在广泛的设备负载范围 $0 < \rho \leq 2$ 内，非正交多址接入的频谱效率和能量效率均高于大规模多输入多输出；平均而言，其能量效率大约高出 80%，频谱效率高出 170%。



(a) $\mathcal{S}$ - 频谱效率曲线下面积的柱状图。



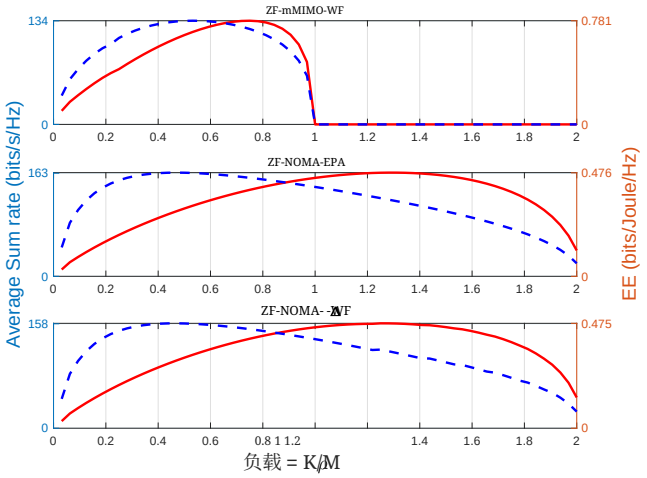
(b) $\mathcal{E}$ - 能量效率曲线下面积的柱状图。

图 9：采用三种功率分配方法时，非正交多址接入和大规模多输入多输出系统的曲线下面积：(a) 频谱效率曲线；(b) 能量效率曲线。平均值是基于 1000 个随机设备位置计算得出的。

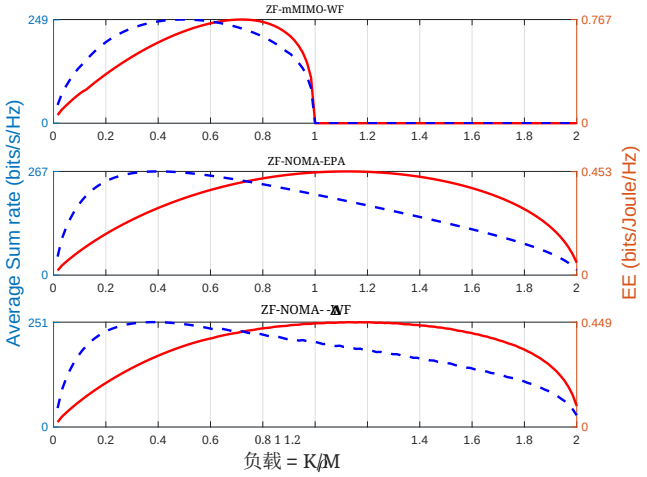
E. 资源效率（频谱效率与能效权衡）

本文从频谱效率与能效权衡（即资源效率（RE））的角度分析了非正交多址接入和大规模多输入多输出，并考虑设备负载最高增至 2 的情况。从图 10 中可以直观地看出，对于每种基站天线配置 M，能够最大化频谱效率-能效权衡的最佳设备负载范围。左侧纵轴表示平均频谱效率，右侧纵轴表示平均能效。表IV汇总了能够最大化频谱效率-能效权衡的最佳设备负载，并显示功率分配后的活跃用户百分比及 Jain公平性指数。图 10.(a)展示了 $M = 64$ 时的结果：采用 EPA 的非正交多址接入在设备负载最高时实现了频谱效率-能效权衡中的最高频谱效率；另一方面，采用经典WF的大规模多输入多输出实现了最高的频谱效率-能效权衡，但活跃设备的比例约为 0.5。图 10.(b)展示了 $M = 128$ 时的结果：采用 EPA 的非正交多址接入在设备负载最高时实现了频谱效率-能效权衡；相比之下，采用经典WF的大规模多输入多输出在设备负载较低时，以一半的活跃设备实现了更高的频谱效率和能效权衡值。图 10.(c)展示了M=256 时的结果：采用 Δ-注水的非正交多址接入在设备负载最高时实现了频谱效率-能效权衡；而大规模多输入多输出采用注水再次实现了更高的频谱效率-能效权衡，活跃设备比例为47%。由此可以证明，

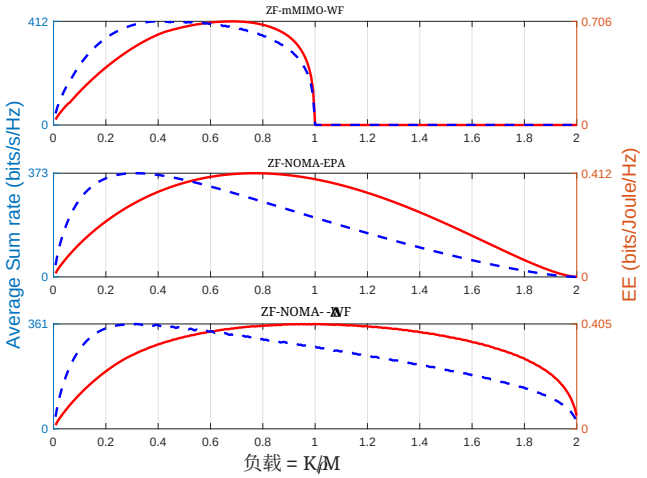
增加基站中的天线数量会提高频谱效率，但会降低能效。



(a) $M = 64$ 基站天线



(b) $M = 128$ 基站天线



(c) $M = 256$ 基站天线

图10：当 $M = 64$、128 和 256 时的频谱效率-能效权衡点。

五、结论与未来工作

本研究对下行单小区中采用三种天线配置的基站所部署的大规模多输入多输出与非正交多址接入进行了频谱效率与能效对比分析。从频谱效率的角度来看，采用经典注水算法的大规模多输入多输出在低负载和中等负载下表现更佳。另一方面，当

表IV：频谱效率与能效权衡

	ρ	SE	EE	活跃用户	$\mathcal{F}$
M = 64					
mMIMO-WF	0.652	131.15	**0.762**	.50	.485
NOMA-EPA	**0.875**	**147.45**	0.428	1.0	.485
NOMA-Δ-注水	0.844	143.48	0.441	1.0	.475
M = 128					
	ρ	SE	EE	活跃用户	$\mathcal{F}$
mMIMO-WF	0.625	**243.21**	**0.745**	.50	.475
NOMA-EPA	**0.734**	241.54	0.411	1.0	.49
NOMA-Δ-注水	0.703	226.32	0.405	.98	.45
M =256					
	ρ	SE	EE	活跃用户	$\mathcal{F}$
mMIMO-WF	0.578	**404.84**	**0.691**	.47	.43
NOMA-EPA	0.523	344.70	0.380	1.0	.495
NOMA-Δ-WF	**0.594**	333.31	0.375	.85	.398

系统负载较高时，$\rho > 0.6$非正交多址接入在 $0.6 < \rho \leq 2$范围内表现更佳。

应用于 NOMA 系统的 PA 方法（EPA、PICPA 和 Δ-WF）会产生不同的频谱效率性能。事实上，当信道硬化条件完全满足且基站天线数量增加（$M = 128$ 和 256）时，采用所提出的 Δ-WF 算法可获得最佳频谱效率结果，但正如预期的那样，公平性指数会受到影响。

从能效角度来看，在 $K < M$ 下采用三种 EPA、PICPA 和注水 PA 方法时，大规模多输入多输出取得了更好的结果。不过，NOMA 可在更高的系统负载下运行，即 $K < 2M - 1$。

就频谱效率曲线下面积和能效曲线这两项指标而言，NOMA系统分别在 $\mathcal{S}$ 和 $\mathcal{E}$ 方面取得了更好的结果，这是因为它能够服务比大规模多输入多输出更多的用户。这些数值结果证实了 NOMA 能够在设备高负载下运行。另一方面，NOMA 不可能实现较高的公平性。

从频谱效率与能效权衡的角度来看，大规模多输入多输出取得了最佳结果，这是因为其能效更优；在所有 M 设置中，设备负载为 $\rho = 0.6$ 时始终如此。

NOMA系统具有令人瞩目的特性，作为设计未来无线通信世代的一项有前景的技术，已受到广泛研究。未来的研究可以探索混合非正交多址系统以及速率分割多址（RSMA）等替代技术，以提升大规模多输入多输出系统的整体能效。

致谢

本研究部分获得巴西国家科学技术发展委员会（CNPq）项目资助，项目编号为310681/2019-7；部分获得CAPES-巴西-财务代码001资助，以及隆德里纳州立大学-巴拉那州政府（UEL）的资助。

参考文献

[1] E. G. Larsson、O. Edfors、F. Tufvesson 和 T. L. Marzetta，《面向下一代无线系统的大规模多输入多输出》，IEEE《通信杂志》，第 52 卷，第 2 期，第 186–195 页，2014 年。[2] E. Bjornson、L. Sanguinetti、J. Hoydis 和 M. Debbah，《面向能量效率的多用户多输入多输出设计：何时大规模多输入多输出才是答案？》，2014 年 IEEE 无线通信与网络会议（WCNC），第 242–247 页，2014 年 4 月。[3] Z. Ding 和 H. V. Poor，《采用有限反馈的大规模多输入多输出非正交多址接入设计》，IEEE《信号处理快报》，第 23 卷，第 629–633 页，2016 年 5 月。[4] K. Senel、H. V. Cheng、E. Bjornson 和 E. G. Larsson，《非正交多址接入能在大规模多输入多输出中发挥什么作用？》，IEEE《信号处理选题期刊》，第 4553 卷，c 期，第 1–16 页，2019年。[5] A. Anwar、B.-C. Seet、M. A. Hasan 和 X. J. Li，《非正交多址在不同无线网络中的应用综述》，《电子学》，第 8 卷，第 1355 页，2019 年 11 月。

[6] M. S. Ali、E. Hossain 和 D. I. Kim，“用于下行多用户 MIMO 系统的非正交多址接入（NOMA）：用户聚类—5，第 565–577 页，2017年。[7] M. Zeng、A. Yadav、O. A. Dobre、G. I. Tsiropoulos 和 H. V. Poor，“多用户簇中 MIMO-NOMA 与 MIMO-OMA 的容量比较 ”，IEEE 通信领域精选期刊，第 35 卷，第 2413–2424 页，2017 年 10 月。[8] S. M. R. Islam、M. Zeng、O. A. Dobre 和 K.-S. Kwak，“下行链路NOMA系统的资源分配：关键技术与待解决问题”，IEEE无线通信，第 25 卷，第 40–47 页，2018 年 4 月。[9] X. Chen、F.-K. Gong、G. Li、H. Zhang 和 P. Song，“大规模MIMO-NOMA系统中的用户配对与配对调度”，IEEE通信快报，第 22 卷，第 788–791 页，2018 年 4 月。[10] H. V. Cheng、E. Bjornson 和 E. G. Larsson，“基于训练的多用户MIMO系统中 NOMA 的性能分析”，IEEE 无线通信汇刊，第 17 卷，第 372–385 页，2018 年 1 月。[11] L. Dai、B. Wang、M. Peng 和 S. Chen，“结合无线信息与能量同时传输的毫米波大规模MIMO-NOMA混合预编码”，IEEE通信领域精选期刊，第 37 卷，第 131–141 页，2019 年 1 月。[12] B. Makki、K. Chitti、A. Behravan 和 M.-S. Alouini，“NOMA综述：现状与开放研究挑战”，IEEE通信学会开放期刊，第 1 卷，2019 年 12 月，第 179–189 页，2020。[13] M. R. Usman、A. Khan、M. A. Usman、Y. S. Jang 和 S. Y. Shin，“下行链路非正交多址接入（NOMA）中理想与非理想串行干扰消除的性能”，2016 年电气与信息系统智能绿色技术国际会议（ICSGTEIS），第 102–106 页，2016 年 10 月。[14] E. Bjornson 和 E. G. Larsson，“无线通信系统的能效究竟能有多高？”，会议记录——阿西洛马会议：信号、系统与计算机，第 2018-Octob 卷，第 1252–1256 页，2019。[15] Y. Zhang、H. M. Wang、T. X. Zheng 和 Q. Yang，“非正交多址接入中的能效传输设计”，IEEE车辆技术汇刊，第 66 卷，第 3 期，第 2852–2857 页，2017年。[16] S. Mounchili 和 S. Hamouda，“通过配对距离分辨率与功率控制提升 NOMA 系统的海量连接能力”，IEEE车辆技术汇刊，第 69 卷，第 4 期，第 4093–4103 页，2020。[17] F. Rezaei、C. Tellambura、A. Tadaion 和 A. R. Heidarpour，“采用三种线性预编码器的无蜂窝大规模MIMO-NOMA速率分析”，IEEE通信汇刊，第 68 卷，第 6 期，第 3480–3494 页，2020。[18] H. Zhang、H. Zhang、W. Liu、K. Long、J. Dong 和 V. C. Leung，“太赫兹MIMO-NOMA系统中的能效用户聚类、混合预编码与功率优化”，arXiv，第 38 卷，第 9 期，第 2074–2085 页，2020。[19] O. Maraqa、A. S. Rajasekaran、S. Al-Ahmadi、H. Yanikomeroglu 和 S. M. Sait，“面向未来无线网络使能技术的速率最优功率域NOMA综述”，IEEE通信综述与教程，第 22 卷，第 4 期，第 2192–2235 页，2020年。[20] S. Rajoria、A. Trivedi 和 W. W. Godfrey，“大规模MIMO无线回传赋能的双层异构网络中的 NOMA 和速率优化”，AEU——国际电子与通信期刊，第 132 卷，2020 年 10 月，第 153626 页，2021。[21] D. Kim 和 M. Choi，“分布式天线系统中的非正交多址接入：最大最小公平性与最大和速率”，IEEE Access，第 9 卷，第 69467–69480 页，2021。[22] I. Budhiraja、N. Kumar、S. Tyagi、S. Tanwar、Z. Han、D. Y. Suh 和 M. J. Piran，《面向 5G 及后续技术的NOMA变体系统综述》，第 9 卷。IEEE，2021。[23] F. Zafari、A. Gkelias 和 K. K. Leung，“室内定位系统与技术综述”，IEEE通信综述与教程，第 21 卷，第 3 期，第 2568–2599 页，2019。[24] E. M. Mohamed，“基于用户位置的下行毫米波 NOMA 用户联合选择与波束成形”，IET 通信，第 14 卷，第 8 期，第 1234–1240 页，2020。[25] T. L. Marzetta、E. G. Larsson、H. Yang 和 H. Q. Ngo，《大规模MIMO基础》，剑桥大学出版社，2016。[26] H. Yang 和 T. L. Marzetta，“大规模天线系统中共轭波束成形与迫零波束成形的性能”，IEEE通信领域精选期刊，第 31 卷，179，2013 2，第 172 页——。[27]S. Boyd 和 L. Vandenberghe，“数值线性代数背景”，*[在线]* ——网址：https://web.stanford.edu/class/ee364a/lectures/num-lin-alg.pdf。[28] J. L. Jacob 和 T. Abrão，“确保用户最大公平性的非正交多址接入系统优化”，Transactions on 《新兴电信技术》，第 31 卷，第 4 期，第 1–14 页，2020 年。