本文已提交给 IEEE，供其考虑发表。版权可能在不另行通知的情况下转让，之后此版本可能无法再访问。

arXiv:2301.03575v1 [cs.IT] 9 Jan 2023

多小区大规模多输入多输出系统中增强型移动宽带与超可靠低时延通信的共存

Giovanni Interdonato，会员，*IEEE*，Stefano Buzzi，高级会员，*IEEE*，Carmen D'Andrea，会员，*IEEE*，Luca Venturino，高级会员，*IEEE*，Ciro D'Elia 和 Paolo Vendittelli

摘要—**本文对多小区大规模多输入多输出系统下行链路中增强型移动宽带（eMBB）与超可靠低时延通信（URLLC）的非正交共存进行了严格分析。我们提出了一个统一的信息论框架，将遍历状态下 eMBB 频谱效率（SE）的无限码长分析，与基于失配译码及所谓鞍点近似的 URLLC 错误概率有限块长分析相结合。在假设用户仅掌握统计信道状态信息（CSI）的条件下，研究将打孔（PUNC）和叠加编码（SPC）作为应对业务间干扰的两种备选下行链路共存策略。随后，我们评估了不同预编码技术和功率控制方案下 eMBB 与 URLLC 的性能，并考虑了基站处不完美的信道状态信息、基于导频的估计开销、导频污染、空间相关信道、无线帧结构以及 URLLC 激活模式的特征。仿真结果表明，在许多运行场景中，SPC 优于 PUNC：它在为 eMBB 提供更高频谱效率的同时，也能以较高概率满足 URLLC 的目标可靠性。此外，在 URLLC 业务负载较高时，PUNC 可能导致 eMBB 服务中断。不过，在无法充分缓解多用户干扰的场景中，为了维持 URLLC 性能，PUNC 则不可或缺。**

索引词—**增强型移动宽带、误差概率、大规模多输入多输出、失配译码、网络可用性、非正交多址接入、打孔、鞍点近似、频谱效率、叠加编码、超可靠低时延通信。**

一、引言

WITH 随着移动应用生态系统的兴起，以及智能设备的数据处理和存储能力不断提升，各种异构业务应运而生，对数据速率、时延、可靠性和海量连接提出了严苛且各不相同的通信要求。第三代合作伙伴计划（3GPP）将这些要求及相关用例归纳为三类宏观业务，即增强型移动宽带 （eMBB）、超可靠低时延通信（URLLC）和海量机器类通信 （mMTC） [1]。增强型移动宽带业务需要高峰值数据速率和稳定的连接，涵盖大多数日常应用：娱乐、多媒体、通信、协作、地图服务、网页浏览等。超可靠低时延通信业务要求单向无线电时延为 1 毫秒，成功概率达到 99.999%，适用于实时和时间关键型应用，例如自动驾驶、自动化控制、增强现实、视频和图像处理等。海量机器类通信业务支持海量各类设备之间的连接，应用包括智能电网、交通管理系统、环境监测等。

第五代移动通信开始以增强型移动宽带业务的形式在各地推出，其本质上类似于更快版本的 LTE；而海量机器类通信和超可靠低时延通信的需求仍在不断完善，并将在未来十年内逐步实现，尽管世界许多地区已经开展了一些实验活动[1]。学术研究和工业标准化目前正关注此类异构业务的不同共存机制，显然已逐渐偏离最初的切片网络愿景 [2]。网络切片基本上意味着为异构业务分配正交资源（存储、计算、无线通信等），以保证它们相互隔离。从广义上说，这种方法通常称为正交多址接入（OMA）。作为正交资源分配的一种有趣替代方案，非正交多址接入（NOMA）正日益受到重视，尤其是在无线接入网（RAN）通信资源分配方面。传统的 RAN 切片方法是在时域和/或频域中分离增强型移动宽带、海量机器类通信和超可靠低时延通信业务，而 NOMA 则依赖高效的共存策略，使异构业务共享相同的时频资源，并在功率域和空间域中彼此区分。在这方面，常采用异构正交多址接入（H-OMA）这一术语 [2] ，以区分异构业务的正交资源分配与同类型业务的正交资源分配，后者称为 OMA。（对于 H-NOMA 与 NOMA，也适用同样的区分。）

大规模多输入多输出 [3]–[5] 是一种利用基站（BS）上大量共址天线，在相同无线资源上同时为多个用户提供相干服务的技术

本研究得到意大利企业与意大利制造部（前身为 MISE）“智慧城市出行管理”项目（5G-SUMMA）、第二轴“新兴技术支持”的资助。G. Interdonato、S. Buzzi、C. D'Andrea、L. Venturino 和 C. D'Elia 任职于意大利卡西诺与南拉齐奥大学电气与信息工程系，邮编 03043，卡西诺，意大利。他们同时隶属于意大利国家电信大学间联合会（CNIT），邮编 43124，帕尔马，意大利。P. Vendittelli 任职于 TIM S.p.A.，邮编 20133，米兰，意大利。S. Buzzi 还隶属于米兰理工大学，邮编 20133，米兰，意大利。通讯作者：Giovanni Interdonato。

[1]例如，意大利前经济发展部的资助计划，以及欧洲国家、欧盟、美国、中国和日本的相关计划。

。用户通过波束成形技术在空间域进行复用，从而实现高方向性传输与接收。使用大量天线还会带来有利传播，进一步降低多用户干扰；同时也会产生信道硬化，从而减小有效信道增益的随机波动。因此，无需采用复杂的信号处理技术来处理多用户干扰。这种激进的空间复用，加上大规模多输入多输出技术固有的实用性和可扩展性，可实现高水平的能效和频谱效率、空间分集、链路可靠性与连接能力。

大规模多输入多输出研究的重点一直是提高用户数据速率，从而满足增强型移动宽带的需求。近年来，一些研究凸显了大规模多输入多输出能够为超可靠低时延通信带来的显著优势 [6]–[8]，即通过降低中断概率和错误概率来提高链路可靠性。更高的可靠性意味着更少的重传，进而降低时延。海量机器类通信也能从大规模多输入多输出技术中受益 [7], [9]，通过利用高能效来延长设备的电池寿命。此外，有利传播支持对海量机器类通信设备进行激进的空间复用，有助于设备检测和随机接入过程。

A. 相关工作

异构业务之间的共存最初是在单天线基站为多个异构用户提供服务的系统中进行研究的。在 [2], Popovski 等人提出了首个可处理的通信理论模型，用以捕捉增强型移动宽带、超可靠低时延通信和海量机器类通信业务流量的关键特征。（这些特征总结于表 I。）具体而言， [2] 分析了单小区模型下的两种场景：(*i*)面向超可靠低时延通信和增强型移动宽带的切片，以及(*ii*)面向海量机器类通信和增强型移动宽带的切片。 [10]研究了超可靠低时延通信与增强型移动宽带的下行复用，旨在满足超可靠低时延通信业务流量的服务质量要求，同时最大化增强型移动宽带业务流量的效用，并对物理层的运行进行抽象。 [11] 针对多小区网络的上行链路提出了基于打孔技术的超可靠低时延通信与增强型移动宽带业务流量共存机制，其中假设采用无衰落的简化 Wyner信道模型。对于多用户多输入多输出系统， [12] 提出了一种基于零空间的空间抢占式调度器，用于联合调度超可靠低时延通信和增强型移动宽带业务流量，以实现跨目标优化；该方法在保证关键超可靠低时延通信服务质量（QoS）的同时，最大化增强型移动宽带的遍历容量。该方法利用基站的空间自由度来满足超可靠低时延通信的解码要求，而不损害增强型移动宽带用户的性能。 [13] 针对分布式架构开展了类似研究，为雾网络下行链路构建了联合用户关联和资源分配问题，并考虑了面向物联网（IoT）应用的超可靠低时延通信与增强型移动宽带业务共存。研究提出了层次分析法，用于设定各项业务的优先级，并建立双边匹配博弈，其中在雾网络基础设施与物联网设备之间建立稳定关联。

增强型移动宽带与超可靠低时延通信之间的共存最受关注 [14]–[18],，主要通过三种替代技术进行处理，以下按复杂度降序列出：

- 连续干扰消除 (SIC)，接收机通过该方法迭代地解码并移除累积接收信号中特定业务的贡献。该方法要求接收机能够获取信道状态信息 (CSI)，以便执行多阶段译码，并逐步降低干扰水平，直至达到所需的成功译码概率。

- 打孔 (PUNC)，其核心是避免业务间干扰。在下行链路中，每当发射机需要发送超可靠低时延通信信号时，增强型移动宽带信号便会在该超可靠低时延通信传输所占用的信道使用次数内被丢弃。在上行链路中，只要接收机能够检测到超可靠低时延通信传输的存在（例如通过能量检测），便会使用擦除译码器丢弃增强型移动宽带信号。

- 叠加编码 (SPC)，发射机通过该方法直接发送增强型移动宽带信号和超可靠低时延通信信号的线性组合。在接收机处，无论是上行链路还是下行链路，业务间干扰都被视为不相关噪声（TIN）。同样，该方法要求接收机能够检测到非目标传输的存在。

在 [14]中，分析了共享模拟前传链路的云无线接入网架构上行链路中超可靠低时延通信与增强型移动宽带业务的共存，并考虑了连续干扰消除、打孔和将干扰视为噪声。该研究基于标准蜂窝模型，对加性高斯噪声链路和有限小区间干扰条件下，异构正交多址接入与 H-NOMA 两种方案中的超可靠低时延通信和增强型移动宽带业务流量性能进行了信息论研究。主要结论是，与异构正交多址接入相比，非正交多址能够实现更高的增强型移动宽带可达速率，同时保证超可靠低时延通信以低速率可靠传输，并将接入时延降至最低。此外，采用连续干扰消除的 H-NOMA 性能最佳；与采用容量受限的数字前传链路的情况不同，将干扰视为噪声始终优于打孔。 [11] 开展了类似分析，研究了云无线接入网的上行链路和下行链路，不采用模拟前传，但考虑了衰落、超可靠低时延通信发射机缺乏信道状态信息、增强型移动宽带发射机的速率自适应以及有限前传容量等实际因素。Abreu 等人在 [16]分析了增强型移动宽带业务流量的异构正交多址接入和 H-NOMA 方案，以及上行链路中的免授权超可靠低时延通信，并考虑了采用最小均方误差接收机（有无连续干扰消除）以及瑞利衰落信道的情况。由此得到的中断概率和可达速率表明，只有在采用连续干扰消除且信噪比足够高时，或在某些超可靠低时延通信负载较低的情况下，将干扰视为噪声才大多有益。否则，异构正交多址接入能够同时支持两种业务更高的负载。近期， [17] 提出了一种方法，以提升上行链路中超可靠低时延通信所能支持的负载，涵盖存在增强型移动宽带业务流量时的异构正交多址接入和 H-NOMA

表 I 第五代移动通信用例的特征

	eMBB	超可靠低时延通信	mMTC
特征	高速率、中等可靠性	低时延、超高可靠性、低速率	低速率、大规模连接
业务流量	大载荷、多个设备	小载荷、少量设备	小载荷、海量设备
激活模式	稳定	间歇性	间歇性
时间跨度	长时间、多种资源	短时隙	长时间、多种资源
频率跨度	单种/多种资源	多种资源、多样性	单种资源
调度	以防止接入冲突	以确保高可靠性	不可行
随机接入	如有需要	以支持间歇性	基本原理
目标	最大化数据速率	满足时延和可靠性要求	最大化可支持的到达率
可靠性要求	$\sim 10^{-3}$	$\sim 10^{-5}$	$\sim 10^{-1}$
应用	视频流媒体、增强现实，娱乐	互联工厂、交通安全，au-自动驾驶车辆 ，远程医疗	物联网、低功耗传感器，智慧城市

业务流量，并表明 H-NOMA 在确保两种业务均满足可靠性要求方面更具优势。 [18], [19] 对下行链路进行了类似分析，将最优资源分配策略与 H-NOMA 相结合，以满足增强型移动宽带和超可靠低时延通信的服务质量约束，并假设增强型移动宽带的信道状态信息完美已知，超可靠低时延通信的信道状态信息为统计已知。

上述研究用于表征增强型移动宽带用户和超可靠低时延通信用户性能的信息论框架，出于不同原因，无法应用于大规模多输入多输出场景。基于块衰落信道模型，在遍历（无限码长）区间确定增强型移动宽带用户的速率（或频谱效率）是合理的，因为增强型移动宽带用户的码字跨越了无限多个相互独立的衰落实现。然而，对于准静态衰落场景下超可靠低时延通信用户的性能，使用中断容量会导致错误概率评估不准确；其分析包含无限码长假设，这一点已由 [8]证明。此外，在使用导频估计上行信道时，中断容量分析无法体现信道状态信息获取开销的影响。作为替代方案，研究者已针对传统蜂窝网络 [18], [19],、集中式大规模多输入多输出网络 [20],[21] 以及无小区大规模多输入多输出网络 [22], 中的超可靠低时延通信提出有限码长分析，并采用 [23],提出的信息论界和工具，例如广为人知的正态近似。然而， [8] 的研究证明，正态近似在超可靠低时延通信所关注的低错误概率区域并不准确（$<10^{-4}$），尤其是在基站天线数量增加以及存在不完美信道状态信息时。值得注意的是，Ostman等人在 [8]中提出了更严谨的有限码长信息论框架，该框架采用失配译码 [24], 和鞍点近似 [25] 来评估集中式大规模多输入多输出系统中超可靠低时延通信用户的错误概率。该框架最初针对 [26]–[28], 中的无线衰落信道提出，能够考虑线性信号处理、不完美信道状态信息、瞬时信道估计误差，以及包括多用户干扰在内的加性不相关噪声。然而， [8] 中的分析仅限于超可靠低时延通信场景，尚未在统一的信息论框架下研究其与增强型移动宽带的共存。

B. 贡献

我们的贡献可概括如下。

- 我们研究多小区大规模多输入多输出系统下行链路中增强型移动宽带与超可靠低时延通信的非正交复用，并提出统一的信息论框架：该框架结合无限码长分析评估增强型移动宽带的频谱效率，并结合有限块长分析评估超可靠低时延通信的错误概率。

- 与以往不恰当地采用中断容量分析或正态近似得到的错误概率来评估超可靠低时延通信性能的研究不同，本文的有限码长信息论分析依赖于 [8],中建立的结果和工具，其中假设采用失配接收机和鞍点近似，但该文未研究超可靠低时延通信与增强型移动宽带之间的共存。

- 所提出的统一框架支持两种可选的共存策略：PUNC 和 SPC。前者通过避免业务间干扰来保障超可靠低时延通信的可靠性，后者则接受干扰以维持增强型移动宽带业务。此外，该分析框架还考虑了基站通过上行导频传输获取信道状态信息时的不完美信道状态信息获取、导频污染和导频开销、空间相关信道，以及用户缺乏信道状态信息的情况。

- 我们通过数值评估，考察 PUNC 和 SPC 在不同预编码方案（即最大比、正则化迫零和多小区最小均方误差）以及不同功率分配策略（即等功率分配、加权分数功率分配，以及使整个网络中的信干噪比乘积最大化的最优功率分配）下的性能。我们还探讨了增强型移动宽带与超可靠低时延通信在各种场景下的共存情况，包括时分双工无线帧的不同配置，以及超可靠低时延通信随机激活模式的不同设置。

- 我们的全面仿真结果表明，在所考虑的大多数运行场景中，SPC 明显优于 PUNC。SPC 的主要局限在于会造成多用户干扰，而采用正则化零迫和多小区最小均方误差通常可以克服这一问题；不过，这些方法都依赖于

高质量的信道状态信息获取。当由于复杂度或硬件约束而无法采用这些预编码技术时，可以通过微调所提出的加权分数功率分配参数来满足超可靠低时延通信的可靠性要求。反之，如果通过预编码消除干扰的效果不佳，例如导频污染严重或多用户干扰过强，则必须采用 PUNC，才能保障超可靠低时延通信的性能。

- 超可靠低时延通信用户之间的导频污染尤其具有破坏性。因此，我们设计了一种优先考虑超可靠低时延通信用户的导频分配策略。在我们的方法中，我们主要为超可靠低时延通信用户分配唯一的正交导频，仅允许增强型移动宽带用户之间复用导频。在可行的情况下，我们在小区内分配正交导频，以避免小区内导频污染；如果上行训练长度足够长，则可确保所有用户获得彼此正交的导频。

C. 论文结构

本文其余部分安排如下。第二节介绍多小区大规模MIMO系统的系统模型，包括上行训练的说明，以及一个适用于数据传输阶段的统一框架，该框架兼顾打孔和叠加编码技术。第三节分别针对增强型移动宽带和超可靠低时延通信的性能评估，介绍无限块长度范围和有限块长度范围内的信息论分析。第四节详述用于应对增强型移动宽带用户与超可靠低时延通信用户共存的预编码技术和功率分配策略。第五节给出仿真结果并进行讨论，第六节则讨论本研究的主要发现。

D. 符号说明

向量和矩阵分别用粗体小写字母和粗体大写字母表示。花体大写字母表示集合，而 $\mathbb{C}$ 和 $\mathbb{R}$ 分别表示复数集和实数集。$\mathsf{E}\{\cdot\}$ 表示期望算子，$\Pr\{\cdot\}$ 表示某个集合的概率。x^+表示正部函数，即 $x^+ = \max\{x, 0\}$，而 $\lfloor\cdot\rfloor$ 表示向下取整函数。自然对数记为 $\log(\cdot)$，$Q(\cdot)$表示高斯 Q 函数。$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$表示均值为 $\boldsymbol{\mu}$、协方差矩阵为 $\boldsymbol{\Sigma}$的圆对称复高斯分布。上标 $(\cdot)^{\mathsf{T}}$、$(\cdot)^*$ 和 $(\cdot)^{\mathsf{H}}$ 分别表示转置、共轭和共轭转置（厄米）算子。$\mathrm{tr}(\mathbf{A})$表示矩阵 A 的迹，而$\|\mathbf{a}\|$ 表示向量 a 的 ℓ_2范数。符号 $[\mathbf{A}]_{:,i}$表示矩阵 A 的第 i列。$\mathbf{I}_N$ 表示大小为 $N \times N$的单位矩阵。表 II 列出了本文系统模型中使用的符号定义。

表 II 系统模型符号说明

符号	描述	符号	描述
L	小区数量	K	每小区用户数量
M	基站天线数	K_{u}	每小区超可靠低时延通信用户数
α	$K_{\mathrm{u}}/K \in (0,\ 1)$	K_{e}	每小区增强型移动宽带用户数
τ_{c}	TDD帧长度	$\mathcal{K}_j^{\mathrm{u}}$	小区 j中的超可靠低时延通信用户集合
τ_{p}	上行训练长度	$\mathcal{K}_j^{\mathrm{e}}$	小区 j中的增强型移动宽带用户集合
τ_{d}	下行数据传输长度	T	TDD帧中的时隙数
$\mathbf{h}_{lk}^j$	小区 l中基站 j 与用户 k 之间的信道，向量 $\mathbb{C}^M$		
$\widehat{\mathbf{h}}_{lk}^j$	$\mathbf{h}_{lk}^j$的估计值	$\widetilde{\mathbf{h}}_{lk}^j$	估计误差 $\mathbf{h}_{lk}^j - \widehat{\mathbf{h}}_{lk}^j$
$\mathbf{R}_{lk}^j$	$\mathbf{h}_{lk}^j$的相关矩阵	β_{lk}^j	的平均信道增益 $\widehat{\mathbf{h}}_{lk}^j$
$\mathbf{C}_{lk}^j$	$\widehat{\mathbf{h}}_{lk}^j$的相关矩阵	f	导频复用因子
p_{jk}^{p}	上行导频功率	$\rho_j^{\max}$	基站 j的最大发射功率
$\mathcal{P}_{jk}$	与小区 j中的用户 k 使用相同导频的所有用户集合		
A_{jk}^t	如果小区 j 中的超可靠低时延通信用户 k 在时隙 t处于活动状态，则为 1，否则为 0		
a_{u}	抽取 A_{jk}^t的伯努利分布参数		
基站 j 在信道使用 n中向增强型移动宽带用户 k 传输的$\varsigma_{jk}^{\mathrm{e}}[n]$ 数据			
$\varsigma_{ji}^{\mathrm{u}}[n]$ 基站 j 在信道使用 n中向超可靠低时延通信用户 i 发送的数据			
$\mathbf{w}_{jk}$	基站 j 向其用户 k使用的预编码向量 $\mathbb{C}^M$		
σ_{u}^2	上行噪声方差	ρ_{ji}^{u}	向超可靠低时延通信用户 i发送的下行功率
σ_{d}^2	下行噪声方差	ρ_{jk}^{e}	向增强型移动宽带用户 k发送的下行功率
g_{jk}^{li}	基站 l 使用$\mathbf{w}_{li}$ 向小区 j中的用户 k 发送的预编码下行信道		
$\widehat{g}_{jk}^{li}$	超可靠低时延通信码字长度的 g_{jk}^{li} n_{d} 估计值		
$\epsilon_{jk}^{\mathrm{dl}}$	下行错误概率 η^{dl} 下行网络可用性		
ν	表征分数功率分配（FPA）的指数		
ω	用于调节分配给超可靠低时延通信用户的功率的 FPA 权重		

II. 系统模型

考虑一个包含L 个小区的多小区大规模多输入多输出系统，每个小区均由一座位于小区中心、配备 M 根共址天线的基站提供服务。每个小区覆盖面积为 $D \times D$ 平方千米的正方形区域，并为 K 个用户提供服务。满足 $M \gg K$，因此可以利用空间自由度有效地抑制干扰。K个用户中有 $0 \leq \alpha \leq 1$的比例请求超可靠低时延通信业务，例如智能交通系统中蜂窝车联网（C-V2X）应用场景中的车辆，或“工业4.0”工厂自动化应用场景中的机器。设 $K_{\mathrm{u}} = \alpha K$ 为每小区的超可靠低时延通信用户数，则 $K_{\mathrm{e}} = K - K_{\mathrm{u}}$ 为每小区的增强型移动宽带用户数。小区 j 中增强型移动宽带用户和超可靠低时延通信用户的索引集合分别记为 $\mathcal{K}_j^{\mathrm{e}}$ 和 $\mathcal{K}_j^{\mathrm{u}}$。

A. 时分双工协议与帧结构

所考虑的系统采用时分双工（TDD）模式，以便进行信道状态信息获取并限制估计开销。此外，我们假设射频链经过完美校准，因此信道具有互易性。利用信道互易性，基站在上行链路中获取的信道估计随后用于下行链路，以设计发射预编码向量。由于在大多数传播环境中，具有足够大天线阵列的共址大规模多输入多输出系统均满足信道硬化特性，因此我们假设用户不估计下行链路信道，而仅依靠统计信道状态信息的知识即可可靠地解码下行数据。因此，时分双工协议包括三个阶段：(*i*) 基于导频的上行训练、(*ii*) 上行数据传输和(*iii*) 下行数据传输。

时频资源按TDD帧组织，每个TDD帧包含一组子载波和时间采样点，假设其上的信道响应在频率上平坦且在时间上不变。TDD帧必须容纳

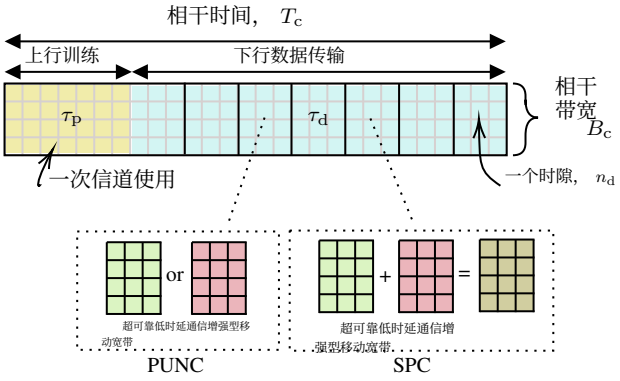

图1. 假设不存在上行数据传输阶段的TDD帧示意图，并展示打孔（PUNC）和叠加编码（SPC）操作下的资源分配。

上述协议阶段并支持所有用户，因此其大小设计为与网络中最小用户的相干块大小相匹配。如图1所示，TDD帧由 $\tau_c = T_c B_c$ 个采样点（或信道使用次数）组成，其中 T_c 为相干时间， B_c 为相干带宽。 τ_c 个信道使用次数中的 τ_p 个用于上行CSI获取，其余信道使用次数则用于上行链路和下行链路数据传输。由于本文仅关注下行链路运行，且不失一般性，我们假设 $\tau_d = \tau_c - \tau_p$ 为下行数据传输阶段的长度。下行数据传输阶段又被划分为 T 个长度相等的时隙。按照遍历状态下的常见假设，eMBB传输跨越多个TDD帧（理论上可达无限多个），期间信道实现根据块衰落模型相互独立地演变。为评估增强型移动宽带用户所达到的频谱效率，我们考察单个TDD帧，并采用无限块长度范围 [4], [5]下的信息论界和工具。相比之下，超可靠低时延通信传输受时间限制，以满足极其严格的时延要求，且仅允许跨越一个时隙。因此，一个时隙中的信道使用次数等于超可靠低时延通信码字长度。我们假设超可靠低时延通信用户采用随机激活模式。在一个TDD帧内，超可靠低时延通信用户可能在多个时隙中处于激活状态。为表征超可靠低时延通信传输的错误概率，我们分别考察TDD帧中的每个时隙，并采用 [8]中介绍的有限码长信息论界和工具。

B. 信道模型与上行训练

小区 l 中第 k个用户与小区 j 中基站之间的信道响应表示为 M维复向量 $\mathbf{h}_{lk}^{j}$。我们假设信道服从相关瑞利衰落，即 $\mathbf{h}_{lk}^{j} \sim \mathcal{CN}\left(\mathbf{0}_M, \mathbf{R}_{lk}^{j}\right)$，其中 $\mathbf{R}_{lk}^{j} \in \mathbb{C}^{M\times M}$是半正定空间相关矩阵。相应的平均信道增益（或大尺度衰落系数）为 $\beta_{lk}^{j} = \mathrm{tr}(\mathbf{R}_{lk}^{j})/M$。假设基站已知大尺度衰落量。

在上行训练阶段，每个用户都会发送一个占用 τ_p 次信道使用的导频序列。导频序列 小区 j 中用户 k 的导频序列记为 $\boldsymbol{\phi}_{jk} \in \mathbb{C}^{\tau_p}$。所有导频序列均选自一个由 τ_p 个两两正交导频组成的集合，因此，两个导频的内积要么为 τ_p（当序列相同时），要么为 0（当序列彼此正交时）。请注意，由于分配给训练的 TDD 帧比例有限，而且信道状态信息获取开销会显著降低频谱效率，因此在整个网络中重复使用导频可能不可避免。导频复用会产生额外干扰，称为 导频污染 [3]，它会降低获取的信道状态信息质量，并使信道估计之间产生相关性。基站 j 接收到的累积上行链路信号记为$\mathbf{Y}_j^{\mathsf{p}} \in \mathbb{C}^{M\times\tau_p}$，其表达式为

$$\mathbf{Y}_j^{\mathsf{p}} = \sum_{k=1}^{K} \sqrt{p_{jk}^{\mathsf{p}}}\mathbf{h}_{jk}^{j}\boldsymbol{\phi}_{jk}^{\mathsf{T}} + \sum_{\substack{l=1\\ l\neq j}}^{L}\sum_{i=1}^{K} \sqrt{p_{li}^{\mathsf{p}}}\mathbf{h}_{li}^{j}\boldsymbol{\phi}_{li}^{\mathsf{T}} + \mathbf{N}_j^{\mathsf{p}}, \quad (1)$$

其中， p_{jk}^{p} 为发射导频功率， $\mathbf{N}_j^{\mathsf{p}}$ 为加性接收机噪声，其元素独立同分布，服从 $\mathcal{CN}\left(0,\sigma_{\mathrm{u}}^2\right)$，其中 σ_{u}^2 为上行链路中的接收机噪声方差。为了估计用户 k 在其所属小区中的信道 $\mathbf{h}_{jk}^{j}$，基站 j 将 $\mathbf{Y}_j^{\mathsf{p}}$ 与已知导频序列 $\boldsymbol{\phi}_{jk}$ 相关，得到

$$\begin{aligned}\mathbf{y}_{jjk}^{\mathsf{p}} &= \mathbf{Y}_j^{\mathsf{p}}\boldsymbol{\phi}_{jk}^{*}\\ &= \sqrt{p_{jk}^{\mathsf{p}}}\tau_p\mathbf{h}_{jk}^{j} + \sum_{\substack{i=1\\ i\neq k}}^{K}\sqrt{p_{ji}^{\mathsf{p}}}\mathbf{h}_{ji}^{j}\boldsymbol{\phi}_{ji}^{\mathsf{T}}\boldsymbol{\phi}_{jk}^{*}\\ &\quad + \sum_{\substack{l=1\\ l\neq j}}^{L}\sum_{i=1}^{K}\sqrt{p_{li}^{\mathsf{p}}}\mathbf{h}_{li}^{j}\boldsymbol{\phi}_{li}^{\mathsf{T}}\boldsymbol{\phi}_{jk}^{*} + \mathbf{N}_j^{\mathsf{p}}\boldsymbol{\phi}_{jk}^{*}. \quad (2)\end{aligned}$$

在（2）中，最右侧等式右边的第二项表示小区内导频污染项，而第三项则量化了小区间导频污染。传统的导频分配策略是在同一小区内为用户分配彼此正交的导频，并在不同小区间复用导频序列 [5]。这是一种合理的选择，因为小区内导频污染通常比小区间导频污染更严重。令 $\tau_p = fK$ ，其中 f 称为导频复用因子。重要的是，为避免危及超可靠低时延通信传输的超高可靠性，我们假设在可行的情况下（即当 $\tau_p > LK_e$ 时），为网络中的所有超可靠低时延通信用户分配唯一且彼此正交的导频序列。总而言之，我们提出的导频分配策略主要旨在避免超可靠低时延通信用户受到导频污染，其次是避免小区内导频污染。最后，如果 τ_p 足够大，即 $\tau_p \geq LK$，那么就能保证为所有用户分配彼此正交的导频。定义集合

$$\mathcal{P}_{jk} = \left\{(l,i) : \boldsymbol{\phi}_{li} = \boldsymbol{\phi}_{jk},\, l = 1,\ldots,L,\, i = 1,\ldots,K\right\}, \quad (3)$$

包括所有与小区 j 中用户 k 使用相同导频的用户（及其对应小区）的索引。因此，我们可以将（2）改写为

$$\mathbf{y}_{jjk}^{\mathsf{p}} = \sqrt{p_{jk}^{\mathsf{p}}}\tau_p\mathbf{h}_{jk}^{j} + \tau_p\sum_{(l,i)\in\mathcal{P}_{jk}\setminus(j,k)}\sqrt{p_{li}^{\mathsf{p}}}\mathbf{h}_{li}^{j} + \mathbf{N}_j^{\mathsf{p}}\boldsymbol{\phi}_{jk}^{*}. \quad (4)$$

处理后的上行链路信号 $\mathbf{y}^{\mathsf{p}}_{jjk}$ 是估计 $\mathbf{h}^{j}_{jk}$ 的充分统计量 。已知空间相关矩阵后，基站 j 可以根据观测值 $\mathbf{y}^{\mathsf{p}}_{jjk}$，计算 $\mathbf{h}^{j}_{jk}$ 的最小均方误差（MMSE）估计值，记为 $\widehat{\mathbf{h}}^{j}_{jk}$，即 [5]

$$\widehat{\mathbf{h}}^{j}_{jk}=\sqrt{p^{\mathsf{p}}_{jk}}\mathbf{R}^{j}_{jk}\mathbf{\Psi}^{j}_{jk}\mathbf{y}^{\mathsf{p}}_{jjk} \tag{5}$$

其中

$$\mathbf{\Psi}^{j}_{jk}=\left(\sum_{(l,i)\in\mathcal{P}_{jk}}p^{\mathsf{p}}_{li}\tau_{\mathrm{p}}\mathbf{R}^{j}_{li}+\sigma^{2}_{\mathrm{ul}}\mathbf{I}_{M_j}\right)^{-1}. \tag{6}$$

估计误差为 $\widetilde{\mathbf{h}}^{j}_{jk}=\mathbf{h}^{j}_{jk}-\widehat{\mathbf{h}}^{j}_{jk}$，其相关矩阵为

$$\mathbf{C}^{j}_{jk}=\mathsf{E}\left\{\widetilde{\mathbf{h}}^{j}_{jk}(\widetilde{\mathbf{h}}^{j}_{jk})^{\mathsf{H}}\right\}=\mathbf{R}^{j}_{jk}-p^{\mathsf{p}}_{jk}\tau_{\mathrm{p}}\mathbf{R}^{j}_{jk}\mathbf{\Psi}^{j}_{jk}\mathbf{R}^{j}_{jk}.$$

由此可知，$\widetilde{\mathbf{h}}^{j}_{jk}$ 和 $\widehat{\mathbf{h}}^{j}_{jk}$ 是相互独立且服从如下分布的随机变量

$$\widetilde{\mathbf{h}}^{j}_{jk}\sim\mathcal{CN}\left(\mathbf{0}_M,\mathbf{C}^{j}_{jk}\right),$$
$$\widehat{\mathbf{h}}^{j}_{jk}\sim\mathcal{CN}\left(\mathbf{0}_M,\mathbf{R}^{j}_{jk}-\mathbf{C}^{j}_{jk}\right).$$

C. 下行链路传输

在下行链路传输阶段，每个基站向其小区内的所有活跃用户传输有效载荷数据。设 A^{t}_{jk} 为一个系数：若小区 j中的 URLLC用户 k 在第 t 个时隙发生超可靠低时延通信传输，则该系数为 1，否则为 0。该系数用于建模超可靠低时延通信用户的随机激活模式，其服从参数为 a_{u} 的伯努利分布 $A^{t}_{jk}\sim\mathcal{B}$伯努利($a_{\mathrm{u}}$)。为处理下行链路中增强型移动宽带用户与超可靠低时延通信用户的共存，我们考虑两种传输技术：（i）打孔和（ii）叠加编码。采用打孔时，只要基站在某个时隙触发了超可靠低时延通信传输，该时隙内的所有增强型移动宽带传输都会被丢弃。不过，在帧中没有超可靠低时延通信用户活跃的其余时隙，仍可保证增强型移动宽带业务。采用叠加编码时，增强型移动宽带传输会在所有时隙进行；每当触发超可靠低时延通信传输时，每个基站都会对增强型移动宽带信号和超可靠低时延通信信号进行线性组合。

接下来详细介绍的分析框架具有普适性，也就是说，只需针对任意基站 j 和时隙 t设置系数，即可适用于上述两种传输技术：

$$\tilde{A}^{t}_{j}=\begin{cases}\left(1-\sum\limits_{i\in\mathcal{K}^{\mathsf{u}}_{j}}A^{t}_{ji}\right)^{+}, & \text{for puncturing,}\\ 1, & \text{for superposition coding.}\end{cases}$$

若 k 分别为增强型移动宽带用户或超可靠低时延通信用户，则令 $\varsigma^{\mathsf{e}}_{jk}[n]$ 或 $\varsigma^{\mathsf{u}}_{jk}[n]$ 表示基站j 在任意信道使用次数 n上发送给用户 k 的数据信号符号。我们假设$\varsigma^{\mathsf{s}}_{jk}[n]\sim\mathcal{CN}(0,1)$，其中 $\mathsf{s}=\{\mathsf{e},\mathsf{u}\}$。一个时隙包含 n_{d}次信道使用，其中 $n_{\mathrm{d}}=\lfloor\tau_{\mathrm{d}}/T\rfloor$，且其长度等于超可靠低时延通信码字的长度。数据信号符号使用M维预编码向量 $\mathbf{w}_{jk}$进行预编码，该向量是以下内容的函数

基站在上行训练期间获取的信道状态信息。另有 $\mathsf{E}\left\{\|\mathbf{w}_{jk}\|^2\right\}=1$。基站 j 在时隙 t 的任意信道使用次数 n 上发送的数据信号为

$$\mathbf{x}^{t}_{j}[n]=\tilde{A}^{t}_{j}\sum_{k\in\mathcal{K}^{\mathsf{e}}_{j}}\sqrt{\rho^{\mathsf{e}}_{jk}}\mathbf{w}_{jk}\varsigma^{\mathsf{e}}_{jk}[n]+\sum_{i\in\mathcal{K}^{\mathsf{u}}_{j}}A^{t}_{ji}\sqrt{\rho^{\mathsf{u}}_{ji}}\mathbf{w}_{ji}\varsigma^{\mathsf{u}}_{ji}[n], \tag{7}$$

其中 $n=1$、…、n_{d}；ρ^{e}_{jk}和 ρ^{u}_{ji}分别是基站 j 用于其eMBB用户 k 和URLLC用户 i的下行传输功率，且满足以下每基站功率约束

$$\mathsf{E}\left\{\left\|\mathbf{x}^{t}_{j}[n]\right\|^2\right\}=\tilde{A}^{t}_{j}\sum_{k\in\mathcal{K}^{\mathsf{e}}_{j}}\rho^{\mathsf{e}}_{jk}+\sum_{i\in\mathcal{K}^{\mathsf{u}}_{j}}A^{t}_{ji}\rho^{\mathsf{u}}_{ji}\le\rho^{\max}_{j}, \tag{8}$$

其中 $j=1$、…、L，$\rho^{\max}_{j}$是基站 j的最大发射功率。用户 k 在小区j 中，于时隙 t 的任意信道使用 n 期间接收到的数据信号记为$y^{t,\mathsf{s}}_{jk}[n]$，其中 $\mathsf{s}=\{\mathsf{e},\mathsf{u}\}$。与传统的大规模多输入多输出系统的运行方式一致，我们假设用户无法获得瞬时下行信道状态信息，而是依赖对其下行预编码信道的均值近似。当发生信道硬化时，这种近似是准确的。若用户 k 是小区 j 中的eMBB用户，即 $k\in\mathcal{K}^{\mathsf{e}}_{j}$，则其在时隙 t 的任意信道使用 n 期间接收到的数据信号可写为下一页顶部的式 (9)，其中 $w_{jk}[n]\sim\mathcal{CN}\left(0,\sigma^2_{\mathrm{d}}\right)$ 是方差为 σ^2_{d}的独立同分布接收机噪声；此外，我们定义了 $g^{li}_{jk}=(\mathbf{h}^{l}_{jk})^{\mathsf{H}}\mathbf{w}_{li}$，即小区 l中的基站使用面向其用户 i的预编码向量时，该基站与小区 j中的第 k个用户之间的预编码下行（标量）信道。若用户 k 是小区 j 中的URLLC用户，则其在时隙 t 的任意信道使用 n 期间接收到的数据信号可写为下一页顶部的式 (10)。式 (9)强调，用户 k仅知道下行信道的统计信道状态信息，即 $\mathsf{E}\left\{g^{jk}_{jk}\right\}$。式 (9) 中的第二项表示由于缺乏瞬时信道状态信息而产生的自干扰，称为波束成形增益不确定性。下文中，小区内跨业务干扰项和小区内同业务干扰项分别表示由小区 j中的URLLC用户和eMBB用户造成的干扰。这可能强于其他小区中eMBB用户（即同业务）和URLLC用户（即跨业务）造成的小区间干扰。式 (10)也对URLLC用户 k 在小区 j 中的各种信号贡献作了类似区分。此处，用户缺乏瞬时信道状态信息这一点将在下一节中重点说明。

III. 性能分析

在本节中，我们评估增强型移动宽带用户和超可靠低时延通信用户的下行性能。对于增强型移动宽带用户，我们采用遍历状态 [4], [5], [29]下建立的无限码长信息论结果来考察频谱效率（SE）。通过对式（9）中的信号模型应用常用的硬化界技术 [4], [5] ，并将所有干扰源视为不相关噪声，可以得到可达的下行频谱效率，即遍历下行容量的下界。具体而言，

$$y_{jk}^{t,\mathsf{e}}[n] = \underbrace{\mathsf{E}\left\{g_{jk}^{jk}\right\}\tilde{A}_j^t\sqrt{\rho_{jk}^{\mathsf{e}}}\varsigma_{jk}^{\mathsf{e}}[n]}_{\text{desired signal}} + \underbrace{\left(g_{jk}^{jk}-\mathsf{E}\left\{g_{jk}^{jk}\right\}\right)\tilde{A}_j^t\sqrt{\rho_{jk}^{\mathsf{e}}}\varsigma_{jk}^{\mathsf{e}}[n]}_{\text{self-interference}} + \underbrace{\sum_{i\in\mathcal{K}_j^{\mathsf{u}}} g_{jk}^{ji}A_{ji}^t\sqrt{\rho_{ji}^{\mathsf{u}}}\varsigma_{ji}^{\mathsf{u}}[n]}_{\text{intra-cell inter-service interference}}$$
$$+ \underbrace{\sum_{i\in\mathcal{K}_j^{\mathsf{e}}\setminus\{k\}} g_{jk}^{ji}\tilde{A}_j^t\sqrt{\rho_{ji}^{\mathsf{e}}}\varsigma_{ji}^{\mathsf{e}}[n]}_{\text{intra-cell intra-service interference}} + \underbrace{\sum_{\substack{l=1\\l\neq j}}^{L}\sum_{i\in\mathcal{K}_l^{\mathsf{e}}} g_{jk}^{li}\tilde{A}_l^t\sqrt{\rho_{li}^{\mathsf{e}}}\varsigma_{li}^{\mathsf{e}}[n]}_{\text{inter-cell intra-service interference}} + \underbrace{\sum_{\substack{l=1\\l\neq j}}^{L}\sum_{i\in\mathcal{K}_l^{\mathsf{u}}} g_{jk}^{li}A_{li}^t\sqrt{\rho_{li}^{\mathsf{u}}}\varsigma_{li}^{\mathsf{u}}[n]}_{\text{inter-cell inter-service interference}} + \underbrace{w_{jk}[n]}_{\text{noise}} \tag{9}$$

$$y_{jk}^{t,\mathsf{u}}[n] = \underbrace{g_{jk}^{jk}A_{jk}^t\sqrt{\rho_{jk}^{\mathsf{u}}}\varsigma_{jk}^{\mathsf{u}}[n]}_{\text{desired signal}} + \underbrace{\sum_{i\in\mathcal{K}_j^{\mathsf{u}}\setminus\{k\}} g_{jk}^{ji}A_{ji}^t\sqrt{\rho_{ji}^{\mathsf{u}}}\varsigma_{ji}^{\mathsf{u}}[n]}_{\text{intra-cell intra-service interference}} + \underbrace{\sum_{\substack{l=1\\l\neq j}}^{L}\sum_{i\in\mathcal{K}_l^{\mathsf{u}}} g_{jk}^{li}A_{li}^t\sqrt{\rho_{li}^{\mathsf{u}}}\varsigma_{li}^{\mathsf{u}}[n]}_{\text{inter-cell intra-service interference}} + \underbrace{\sum_{l=1}^{L}\sum_{i\in\mathcal{K}_l^{\mathsf{e}}} g_{jk}^{li}\tilde{A}_l^t\sqrt{\rho_{li}^{\mathsf{e}}}\varsigma_{li}^{\mathsf{e}}[n]}_{\text{inter-service interference}} + \underbrace{w_{jk}[n]}_{\text{noise}} \tag{10}$$

任意小区 j中的增强型移动宽带用户 k 的可达下行频谱效率为

$$\mathsf{SE}_{jk}^{\mathsf{e}} = \frac{\tau_{\mathrm{d}}}{\tau_{\mathrm{c}}}\frac{1}{T}\sum_{t=1}^{T}\log_2(1+\mathsf{SINR}_{jk}^{t,\mathsf{e}}),\ \text{[bits/s/Hz]}\ , \tag{11}$$

其中 $\tau_{\mathrm{d}}/\tau_{\mathrm{c}}$ 表示估计开销，

$$\mathsf{SINR}_{jk}^{t,\mathsf{e}} = \frac{\tilde{A}_j^t\rho_{jk}^{\mathsf{e}}\left|\mathsf{E}\left\{g_{jk}^{jk}\right\}\right|^2}{\sum\limits_{l=1}^{L}\sum\limits_{i=1}^{K}\varrho_{li}^t\,\mathsf{E}\left\{|g_{jk}^{li}|^2\right\}-\tilde{A}_j^t\rho_{jk}^{\mathsf{e}}\left|\mathsf{E}\left\{g_{jk}^{jk}\right\}\right|^2+\sigma_{\mathrm{d}}^2}\ , \tag{12}$$

是用户 $k\in\mathcal{K}_j^{\mathsf{e}}$ 的有效信干噪比，其中期望是针对随机信道实现取的，且

$$\varrho_{li}^t = \begin{cases} A_{li}^t\rho_{li}^{\mathsf{u}}, & \text{if } i\in\mathcal{K}_l^{\mathsf{u}}\ , \\ \tilde{A}_l^t\rho_{li}^{\mathsf{e}}, & \text{if } i\in\mathcal{K}_l^{\mathsf{e}}\ . \end{cases} \tag{13}$$

式（11）所示的可达频谱效率表达式适用于任意预编码方案、任意信道估计器以及任意信道分布。重要的是，该表达式考虑了异构业务之间任意共存技术的选择，即打孔或叠加编码。上述无限码长分析基于块衰落信道模型的假设，因此每个增强型移动宽带码字长度无限，并跨越大量相互独立的衰落实现。这一假设不适用于超可靠低时延通信场景。对于超可靠低时延通信用户，我们采用 [8] 中建立的有限码长信息论结果，以时隙为粒度对下行误差概率进行非渐近分析。首先，将式（10）改写为

$$y_{jk}^{t,\mathsf{u}}[n] = g_{jk}^{jk}q_{jk}[n] + z_{jk}[n],\quad n=1,\ldots,n_{\mathrm{d}}, \tag{14}$$

其中 $q_{jk}[n]=A_{jk}^t\sqrt{\rho_{jk}^{\mathsf{u}}}\varsigma_{jk}^{\mathsf{u}}[n]$，且

$$z_{jk}[n] = \sum_{i\in\mathcal{K}_j^{\mathsf{u}}\setminus\{k\}} g_{jk}^{ji}q_{ji}[n] + \sum_{i\in\mathcal{K}_j^{\mathsf{e}}} g_{jk}^{ji}\tilde{A}_j^t\sqrt{\rho_{ji}^{\mathsf{e}}}\varsigma_{ji}^{\mathsf{e}}[n] + \sum_{\substack{l=1\\l\neq j}}^{L}\left(\sum_{i\in\mathcal{K}_l^{\mathsf{u}}} g_{jk}^{li}q_{li}[n] + \sum_{i\in\mathcal{K}_l^{\mathsf{e}}} g_{jk}^{li}\tilde{A}_l^t\sqrt{\rho_{li}^{\mathsf{e}}}\varsigma_{li}^{\mathsf{e}}[n]\right) + w_{jk}[n]\ . \tag{15}$$

然而，小区 j 中的超可靠低时延通信用户 k 无法获取 g_{jk}^{jk}，而是仅利用其均值 $\widehat{g}_{jk}^{jk}=\mathsf{E}\left\{(\mathbf{h}_{jk}^{j})^{\mathsf{H}}\mathbf{w}_{jk}\right\}$ 进行数据解码，并将该均值视为准确值。当信道硬化成立时，该估计值是准确的。注意，在时隙 t 中传输长度为 n_{d} 的超可靠低时延通信码字期间，预编码信道 g_{jk}^{jk} 是频率平坦且时不变的。此外，对于同一 TDD 帧中各时隙内基站 j 向用户 k 进行的任何其他传输，g_{jk}^{jk} 均保持不变。给定所有信道和预编码向量，有效噪声项 $\{z_{jk}[n]\in\mathbb{C};\ n=1,\ldots,n_{\mathrm{d}}\}$ 是条件独立同分布的随机变量，方差为 σ_{jk}^2，即 $\mathcal{CN}\left(0,\sigma_{jk}^2\right)$，其表达式为

$$\sigma_{jk}^2 = \sum_{i\in\mathcal{K}_j^{\mathsf{u}}\setminus\{k\}} A_{ji}^t\rho_{ji}^{\mathsf{u}}|g_{jk}^{ji}|^2 + \sum_{i\in\mathcal{K}_j^{\mathsf{e}}}\tilde{A}_j^t\rho_{ji}^{\mathsf{e}}|g_{jk}^{ji}|^2 + \sum_{\substack{l=1\\l\neq j}}^{L}\left(\sum_{i\in\mathcal{K}_l^{\mathsf{u}}} A_{li}^t\rho_{li}^{\mathsf{u}}|g_{jk}^{li}|^2 + \sum_{i\in\mathcal{K}_l^{\mathsf{e}}}\tilde{A}_l^t\rho_{li}^{\mathsf{e}}|g_{jk}^{li}|^2\right) + \sigma_{\mathrm{d}}^2\ . \tag{16}$$

为了确定所传输的码字

$$\mathbf{q}_{jk} = [q_{jk}[1],\ldots,q_{jk}[n_{\mathrm{d}}]]^{\mathsf{T}}\ ,$$

小区 j 中的用户 k 采用失配缩放最近邻（SNN）解码器 [30], ，该解码器通过应用以下规则从码本 $\mathcal{C}$ 中选择码字 $\widetilde{\mathbf{q}}_{jk}$:

$$\widehat{\mathbf{q}}_{jk} = \underset{\widetilde{\mathbf{q}}_{jk}\in\mathcal{C}}{\arg\min}\left\|\mathbf{y}_{jk}^{t,\mathsf{u}} - \widehat{g}_{jk}^{jk}\widetilde{\mathbf{q}}_{jk}\right\|^2\ , \tag{17}$$

其中$\mathbf{y}_{jk}^{t,\mathsf{u}}=[y_{jk}^{t,\mathsf{u}}[1],\ldots,y_{jk}^{t,\mathsf{u}}[n_{\mathrm{d}}]]^{\mathsf{T}}\in\mathbb{C}^{n_{\mathrm{d}}}$,为接收数据向量。

令 $\epsilon_{jk}^{\mathsf{dl}}=\mathsf{Pr}\{\widehat{\mathbf{q}}_{jk}\neq\mathbf{q}_{jk}\}$ 表示采用 SNN 解码时，小区 j 中的超可靠低时延通信用户 k 所经历的下行误差概率。通过采用标准随机编码方法，可得到 $\epsilon_{jk}^{\mathsf{dl}}$ 的一个上界 [31],

$$\epsilon_{jk}^{\mathsf{dl}}\leq\mathsf{E}_{g_{jk}^{jk}}\left\{\mathsf{Pr}\left\{\sum_{n=1}^{n_{\mathrm{d}}}\imath_s(q_{jk}[n],\mathbf{y}_{jk}^{t,\mathsf{u}}[n])\leq\log\frac{m-1}{r}\middle|g_{jk}^{jk}\right\}\right\},\tag{18}$$

其中，$m=2^b$是长度为 n_{d}、传递 b 个信息比特的码字数量，r 是一个在区间 $[0,1]$ 上均匀分布的随机变量，而 $\imath_s(q_{jk}[n],\mathbf{y}^{t,\mathsf{u}}_{jk}[n])$ 是广义信息密度，其定义如下

$$\begin{aligned}&\imath_s(q_{jk}[n],\mathbf{y}_{jk}^{t,\mathsf{u}}[n])\\&\quad=-s\left|\mathbf{y}_{jk}^{t,\mathsf{u}}[n]-\widehat{g}_{jk}^{jk}q_{jk}[n]\right|^2+\frac{s|\mathbf{y}_{jk}^{t,\mathsf{u}}[n]|^2}{1+s\rho_{jk}^{\mathsf{u}}|\widehat{g}_{jk}^{jk}|^2}\\&\qquad+\log(1+s\rho_{jk}^{\mathsf{u}}|\widehat{g}_{jk}^{jk}|^2)\,,\end{aligned}\tag{19}$$

对于所有 $s>0$均成立。在式（18）中，期望是对 g_{jk}^{jk} 的分布取的，而概率是相对于下行链路数据信号 $\{q_{jk}[n]\}_{n=1}^{n_{\mathrm{d}}}$、有效加性噪声 $\{z_{jk}[n]\}_{n=1}^{n_{\mathrm{d}}}$以及随机变量 r计算的。式（18）中上界的评估需要进行极其繁重的数值计算：首先求出概率，然后通过对 s进行优化，在数值上收紧上界，使其达到 URLLC应用场景所要求的低错误概率目标。

幸运的是，利用 [8, Th. 2] 中给出的鞍点近似，我们可以对 (18) 的右侧进行可靠的闭式近似，从而显著减轻计算负担。

鞍点近似之所以存在，是因为 $-\imath_s(q_{jk}[n],\mathbf{y}^{t,\mathsf{u}}_{jk}[n])$ 的矩生成函数在由 [8,附录 B] 给出的值 $\varepsilon<0<\bar{\varepsilon}$ 所界定的零点邻域内存在三阶导数。

$$\underline{\varepsilon}=-\frac{\sqrt{(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})^2+4\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)}+\zeta_{\mathsf{a}}-\zeta_{\mathsf{b}}}{2\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)}\,,\tag{20}$$

$$\bar{\varepsilon}=\frac{\sqrt{(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})^2+4\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)}-\zeta_{\mathsf{a}}+\zeta_{\mathsf{b}}}{2\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)}\,,\tag{21}$$

其中

$$\zeta_{\mathsf{a}}=s(\rho_{jk}^{\mathsf{u}}|g_{jk}^{jk}-\widehat{g}_{jk}^{jk}|^2+\sigma^2)\,,\tag{22}$$

$$\zeta_{\mathsf{b}}=\frac{s}{1+s\rho_{jk}^{\mathsf{u}}|\widehat{g}_{jk}^{jk}|^2}(\rho_{jk}^{\mathsf{u}}|g_{jk}^{jk}|^2+\sigma^2)\,,\tag{23}$$

$$\mu=\frac{s^2\left|\rho_{jk}^{\mathsf{u}}|g_{jk}^{jk}|^2+\sigma^2-(g_{jk}^{jk})^*\widehat{g}_{jk}^{jk}\rho_{jk}^{\mathsf{u}}\right|^2}{\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1+s\rho_{jk}^{\mathsf{u}}|\widehat{g}_{jk}^{jk}|^2)}\,.\tag{24}$$

鞍点近似取决于 $-\imath_s(q_{jk}[n],\ \mathbf{y}^{t,\ \mathsf{u}}_{jk}[n])$ 的累积生成函数，该函数定义为

$$\upsilon(\varepsilon)=\log\mathsf{E}\left\{e^{-\varepsilon\imath_s(q_{jk}[n],\mathbf{y}_{jk}^{t,\mathsf{u}}[n])}\right\}\,,\tag{25}$$

其一阶导数 $\upsilon'(\zeta)$ 和二阶导数 $\upsilon''(\zeta)$，对于所有 $\varepsilon\in(\underline{\varepsilon},\ \bar{\varepsilon})$

$$\begin{aligned}\upsilon(\varepsilon)=&-\varepsilon\log(1+s\rho_{jk}^{\mathsf{u}}|\widehat{g}_{jk}^{jk}|^2)\\&-\log(1+(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})\varepsilon-\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)\varepsilon^2)\end{aligned}\tag{26}$$

$$\begin{aligned}\upsilon'(\varepsilon)=&-\log(1+s\rho_{jk}^{\mathsf{u}}|\widehat{g}_{jk}^{jk}|^2)\\&-\frac{(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})-2\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)\varepsilon}{1+(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})\varepsilon-\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)\varepsilon^2}\end{aligned}\tag{27}$$

$$\begin{aligned}\upsilon''(\varepsilon)=&\left[\frac{(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})-2\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)\varepsilon}{1+(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})\varepsilon-\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)\varepsilon^2}\right]^2\\&+\frac{2\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)}{1+(\zeta_{\mathsf{b}}-\zeta_{\mathsf{a}})\varepsilon-\zeta_{\mathsf{a}}\zeta_{\mathsf{b}}(1-\mu)\varepsilon^2}.\end{aligned}\tag{28}$$

令 $m=e^{n_{\mathrm{d}}R}$ 为某个严格为正的传输速率 $R=(\log\,m)/n_{\mathrm{d}}$，并令 $\varepsilon\in(\underline{\varepsilon},\bar{\varepsilon})$ 为方程 $R=-\upsilon'(\varepsilon)$ 的解。令 I_{s} 为广义互信息 [30]，其定义为 $I_{\mathsf{s}}=\mathsf{E}\{\imath_s(q_{jk}[1],v_{jk}[1])\}=-\upsilon'(0)$。最后，考虑由 $R_{\mathsf{s}}^{\mathsf{cr}}=-\upsilon'(1)$ 给出的临界速率 [31, Eq. (5.6.30)] 。于是，错误概率上界 [8] 有三种可能的鞍点近似。

若 $\varepsilon\in[0,1]$，则 $R_{\mathsf{s}}^{\mathsf{cr}}\leq R\leq I_{\mathsf{s}}$ 且

$$\begin{aligned}&\mathsf{Pr}\left\{\sum_{n=1}^{n_{\mathrm{d}}}\imath_s(q_{jk}[n],\mathbf{y}_{jk}^{t,\mathsf{u}}[n])\leq\log\frac{e^{n_{\mathrm{d}}R}-1}{r}\right\}\\&\quad\approx e^{n_{\mathrm{d}}[\upsilon(\varepsilon)+\varepsilon R]}\left[\Psi_{n_{\mathrm{d}},\varepsilon}(\varepsilon)+\Psi_{n_{\mathrm{d}},\varepsilon}(1-\varepsilon)\right]\,,\end{aligned}\tag{29}$$

其中

$$\Psi_{n_{\mathrm{d}},\varepsilon}(\ell)\triangleq e^{\frac{1}{2}n_{\mathrm{d}}\ell^2\upsilon''(\varepsilon)}Q\left(\ell\sqrt{n_{\mathrm{d}}\upsilon''(\varepsilon)}\right)\,.\tag{30}$$

若 $\varepsilon>1$，则 $R<R_{\mathsf{s}}^{\mathsf{cr}}$且

$$\begin{aligned}&\mathsf{Pr}\left\{\sum_{n=1}^{n_{\mathrm{d}}}\imath_s(q_{jk}[n],\mathbf{y}_{jk}^{t,\mathsf{u}}[n])\leq\log\frac{e^{n_{\mathrm{d}}R}-1}{r}\right\}\\&\quad\approx e^{n_{\mathrm{d}}[\upsilon(1)+R]}\left[\widetilde{\Psi}_{n_{\mathrm{d}}}(1,1)+\widetilde{\Psi}_{n_{\mathrm{d}}}(0,-1)\right]\,,\end{aligned}\tag{31}$$

其中

$$\begin{aligned}\widetilde{\Psi}_{n_{\mathrm{d}}}(\ell_1,\ell_2)\triangleq&e^{n_{\mathrm{d}}\ell_1\left[R_{\mathsf{s}}^{\mathsf{cr}}-R+\frac{1}{2}\upsilon''(1)\right]}\\&\times Q\left(\ell_1\sqrt{n_{\mathrm{d}}\upsilon''(1)}+\ell_2\frac{n_{\mathrm{d}}(R_{\mathsf{s}}^{\mathsf{cr}}-R)}{\sqrt{n_{\mathrm{d}}\upsilon''(1)}}\right).\end{aligned}\tag{32}$$

若 $\varepsilon<0$，则 $R>I_{\mathsf{s}}$ 且

$$\begin{aligned}&\mathsf{Pr}\left\{\sum_{n=1}^{n_{\mathrm{d}}}\imath_s(q_{jk}[n],\mathbf{y}_{jk}^{t,\mathsf{u}}[n])\leq\log\frac{e^{n_{\mathrm{d}}R}-1}{r}\right\}\\&\quad\approx 1-e^{n_{\mathrm{d}}[\upsilon(\varepsilon)+\varepsilon R]}\left[\Psi_{n_{\mathrm{d}},\varepsilon}(-\varepsilon)-\Psi_{n_{\mathrm{d}},\varepsilon}(1-\varepsilon)\right].\end{aligned}\tag{33}$$

与传统使用的正态近似 [23]相比，在超可靠低时延通信大规模多输入多输出场景下，鞍点近似更加准确，因为前者刻画了错误概率随超可靠低时延通信码字长度和传输速率要求 R变化的指数衰减，即误差指数，而后者仅利用 *Berry-Esseen*中心极限定理 （正态近似中所用）刻画误差指数项之后的乘法因子。正态近似的公式直接涉及广义互信息，I_{s}，但不

R，只有在 I_{s} 接近 R时才准确。超可靠低时延通信并不处于这种运行场景，因为为了实现极低的错误概率目标，R 通常低于 I_{s} 。通过鞍点近似得到下行误差概率的近似上界后，我们计算下行网络可用性 [8], η^{dl}，其表达式为

$$\eta^{\mathsf{dl}} = \mathsf{Pr}\left\{\epsilon_{jk}^{\mathsf{dl}} \leq \epsilon_{\mathrm{target}}^{\mathsf{dl}}\right\} \tag{34}$$

它衡量在存在干扰用户的情况下，任意用户 k 在小区 j中满足目标错误概率$\epsilon_{\mathrm{target}}^{\mathsf{dl}}$ 的概率。错误概率定义中的期望是针对小尺度衰落和有效加性噪声取的，并给定一次大尺度衰落实现；而网络可用性定义中的概率则是针对大尺度衰落（即路径损耗、阴影衰落等）计算的。式（34）所示的网络可用性表达式适用于任意预编码方案、任意信道估计器以及任意信道分布。重要的是，该表达式涵盖异构业务之间的任意共存技术选择，即打孔或叠加编码。

IV. 预编码与功率控制

预编码方案和下行功率分配的选择会深刻影响增强型移动宽带用户的频谱效率以及超可靠低时延通信用户的网络可用性。为便于比较，本文考虑三种预编码方案和三种功率分配策略。面向小区 j 中用户 k 的预编码向量的一般表达式为

$$\mathbf{w}_{jk} = \frac{\mathbf{v}_{jk}}{\|\mathbf{v}_{jk}\|}\ , \tag{35}$$

其中，分母用于使预编码向量的平均功率归一化为单位值，随后对$\mathbf{v}^{jk}$ 进行表征。

多小区最小均方误差（M-MMSE）：

$$\mathbf{v}_{jk}^{\mathsf{M-MMSE}} = \left[\left(\sum_{l=1}^{L}\widehat{\mathbf{H}}_l^j\mathbf{P}_l(\widehat{\mathbf{H}}_l^j)^{\mathsf{H}} + \boldsymbol{\Upsilon}^j + \sigma_{\mathrm{u}}^2\mathbf{I}_M\right)^{-1}\widehat{\mathbf{H}}_j^j\mathbf{P}_j\right]_{:,k}$$

其中，$\mathbf{P}_l$ =diag($pli,\ldots,p_{lK}$) $\in\mathbb{R}^{K\times K}$是一个矩阵，其对角元素为小区 l中所有用户的上行发射功率，$\boldsymbol{\Upsilon}^j = \sum_{l=1}^{L}\sum_{i=1}^{K}p_{li}\mathbf{C}_{li}^j$，且 $\widehat{\mathbf{H}}_l^j = [\widehat{\mathbf{h}}_{l1}^j \ldots \widehat{\mathbf{h}}_{lK}^j]$。多小区最小均方误差预编码能够实现近乎最优的下行频谱效率，但要求每个基站获取多小区系统中所有用户的信道状态信息和统计信道状态信息。此外，预编码向量的计算需要进行矩阵求逆$M\times M$，对于大型基站阵列而言可能计算量很大。尽管不具实用性，多小区最小均方误差预编码仍将作为基准。

正则化迫零（RZF）：

$$\mathbf{v}_{jk}^{\mathsf{RZF}} = \left[\widehat{\mathbf{H}}_j^j\left((\widehat{\mathbf{H}}_j^j)^{\mathsf{H}}\widehat{\mathbf{H}}_j^j + \sigma_{\mathrm{u}}^2\mathbf{P}_j^{-1}\right)^{-1}\right]_{:,k}.$$

与 M-MMSE 相比，正则化迫零预编码只要求每个基站估计其用户的信道。此外，计算正则化迫零预编码向量的计算复杂度更低，因为需要求逆的矩阵大小为 $K\times K$。然而，正则化迫零仅能抑制小区内干扰，而不像 M-MMSE 那样为用户提供抵御小区间干扰和信道估计误差的保护机制。

最大比率（MR）： $\mathbf{v}_{jk}^{\mathsf{MR}} = \widehat{\mathbf{h}}_{jk}^j$. 这是计算成本最低、但性能最差的预编码方案。MR仅旨在最大化期望信号的功率，不提供任何干扰抑制机制。MR将作为性能的下界。

合理分配下行功率，对于满足超可靠低时延通信的严格可靠性要求以及提升增强型移动宽带用户的频谱效率至关重要。接下来，我们将介绍三种功率分配方案，这些方案考虑了基站的功率预算、采用的eMBB-URLLC共存策略，以及下行链路运行中基站已知的URLLC激活模式。

等功率分配（EPA）： 其做法是设定

$$\rho_{ji}^{\mathsf{u}} = \rho_j^{\mathsf{max}}\frac{A_{ji}^t}{\tilde{A}_j^t K_{\mathrm{e}} + \sum\limits_{k\in\mathcal{K}_j^{\mathrm{u}}}A_{jk}^t}\ ,\ i\in\mathcal{K}_j^{\mathrm{u}} \tag{36}$$

$$\rho_{jk}^{\mathsf{e}} = \rho_j^{\mathsf{max}}\frac{\tilde{A}_j^t}{\tilde{A}_j^t K_{\mathrm{e}} + \sum\limits_{i\in\mathcal{K}_j^{\mathrm{u}}}A_{ji}^t}\ ,\ k\in\mathcal{K}_j^{\mathrm{e}} \tag{37}$$

使式（8）中的每基站功率约束以等号成立，并为每个用户分配相同的功率份额，而不考虑其信道条件和服务要求。

加权分数功率分配（FPA）： 其功率设置如下

$$\rho_{ji}^{\mathsf{u}} = \frac{\omega\rho_j^{\mathsf{max}}A_{ji}^t(\beta_{ji}^j)^{\nu}}{(1-\omega)\tilde{A}_j^t\sum\limits_{k\in\mathcal{K}_j^{\mathsf{e}}}(\beta_{jk}^j)^{\nu} + \omega\sum\limits_{u\in\mathcal{K}_j^{\mathrm{u}}}A_{ju}^t(\beta_{ju}^j)^{\nu}}\ ,\ i\in\mathcal{K}_j^{\mathrm{u}} \tag{38}$$

$$\rho_{jk}^{\mathsf{e}} = \frac{(1-\omega)\rho_j^{\mathsf{max}}\tilde{A}_j^t(\beta_{jk}^j)^{\nu}}{(1-\omega)\tilde{A}_j^t\sum\limits_{e\in\mathcal{K}_j^{\mathsf{e}}}(\beta_{je}^j)^{\nu} + \omega\sum\limits_{i\in\mathcal{K}_j^{\mathrm{u}}}A_{ji}^t(\beta_{ji}^j)^{\nu}}\ ,\ k\in\mathcal{K}_j^{\mathsf{e}} \tag{39}$$

其中，权重 $\omega\in(0,1)$ 用于调节分配给超可靠低时延通信用户的下行功率，而 ν 则根据平均信道增益确定功率控制策略。令$\nu>0$即可实现机会式功率分配，此时信道条件较好的用户会分配到更多功率。相反，令 $\nu<0$则可兼顾公平性，此时信道条件较差的用户会分配到更多功率。若 $\omega\in(0.5,1)$，则分配给超可靠低时延通信用户的功率份额大于分配给增强型移动宽带用户的份额；若$\omega\in(0,0.5)$，则情况相反。请注意，若 $\nu=0$ 且 $\omega=0.5$，则 FPA 会退化为等功率分配（EPA）。

最大化信干噪比乘积的最优功率分配（OPA）：

这些功率是该优化问题的解

$$\underset{\{\rho_{jk}^{\mathsf{s}}\}}{\text{maximize}} \quad \prod_{j=1}^{L}\prod_{k=1}^{K} \mathsf{SINR}_{jk}^{t,\mathsf{s}} \tag{40a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \varrho_{jk}^{t} \le \rho_{j}^{\max}\ , \forall j, \tag{40b}$$

其中，上标 $\mathsf{s}=\mathsf{e}$ 表示用户 $k \in \mathcal{K}_j^{\mathrm{e}}$，$\mathsf{s}=\mathsf{u}$ 则表示其他情况，ϱ_{jk}^{t} 由（13）给出。为避免（40）中的符号更加繁杂，我们指出，非活跃用户的信干噪比被虚设为 1，以保持优化问题的表述形式。该功率分配策略将所有用户都视为增强型移动宽带用户，因此，如果给定时隙中没有超可靠低时延通信用户处于活跃状态，该策略通过最大化多小区系统总频谱效率的一个下界，便是最优的。尽管将（12）中的信干噪比表达式用于超可靠低时延通信用户时并无意义，我们仍可将由（40）得出的超可靠低时延通信功率启发式地代入误差概率分析，并通过考察性能来为这种做法提供依据。原则上，所有考虑的功率分配方案都按时隙级运行，以便根据超可靠低时延通信激活模式调整功率系数。幸运的是，这些方案只依赖统计信道状态信息，因此可以预先计算部分功率系数；若传播环境没有宏观变化，也可以在多个时隙／帧内保持功率分配不变。与等功率分配和 FPA 方案不同，最优功率分配方案要求基站之间进行一定程度的协作：基站必须将统计信道状态信息发送给中央处理单元（例如主基站），由其计算所有用户的信干噪比并求解优化问题，再将待用功率系数反馈给基站。这会给超可靠低时延通信用户带来无法容忍的时延。此外，求解问题（40）虽然可以高效地作为几何规划 [5, Th. 7.2], 来完成，但在一个时隙内完成求解并不现实，尤其是在用户密集的网络中。因此，最优功率分配方案的实际用途有限，但可用于基准比较。

V. 仿真结果

本节展示并讨论仿真结果，深入分析不同设置下增强型移动宽带与超可靠低时延通信的共存。具体而言，我们重点探讨不同因素对性能的影响，例如传输技术与预编码方案、功率控制策略、不完美信道状态信息与估计开销、导频污染、TDD帧中时隙的长度和数量，以及超可靠低时延通信激活模式的特征。

我们的仿真场景由一个包含 $L=4$ 个小区的多小区大规模多输入多输出系统构成。每个小区覆盖的名义面积为 500×500 平方米，并由一座位于小区中心的基站提供服务。该基站配备一个均匀线性阵列（ULA），其中包含 $M=100$ 个等间距的半波长天线单元。我们按照 [5, 第 4.1 节3]采用环绕拓扑。用户在覆盖区域内均匀随机分布，但与基站的距离至少为 25 米。此外，我们假设超可靠低时延通信用户随机均匀分布在基站周围面积为 125×125 平方米的区域内。用户位置的一次随机实现决定了一组大尺度衰落系数，并构成一个网络快照。对于给定的网络快照，活跃 eMBB用户可实现的下行频谱效率根据式（11）计算，而超可靠低时延通信用户的下行误差概率则根据近似式（29）至（33）计算。随后，通过大量网络快照绘制频谱效率的累积分布函数（CDF）和网络可用性。信道相关矩阵根据广为采用的局部散射空间相关模型 [5, 第 2 节6],生成。我们假设散射仅局限于用户周围，且散射位置在时延扩展为 25° 度 [8]的范围内均匀随机分布。平均信道增益根据适用于 2 GHz载波的 3GPP 宏小区非视距模型计算， [32], 并以 dB 表示为

$$\beta_k = -35.3 - 37.6\log_{10}\left(\frac{d_k}{1\ \mathrm{m}}\right) + F_k$$

对于距离其基站为 d_k 的任意用户 k，其中 $F_k \sim \mathcal{N}\left(0,\ \sigma_{\mathsf{sh}}^2\right)$ 将对数正态阴影衰落建模为标准差为 $\sigma_{\mathsf{sh}}=4$ dB 的独立同分布随机变量。传输带宽为 20 MHz，上行链路和下行链路的接收机噪声功率均为 -94 dBm。此外，我们令 $\rho_j^{\max}=46$ dBm，$j=1,\ \ldots,\ L$，所有用户的导频和有效载荷数据的上行发射功率均为 23 dBm。我们假设超可靠低时延通信数据包包含 $b=160$ 位，对应的传输速率为 $R=b/n_{\mathrm{d}}$，适用于工厂自动化应用场景，例如运动控制，并符合低时延要求 [33,附录 A]。最后，在不失一般性的前提下，我们设定 $\tau_{\mathrm{u}}=0$，因为我们只关注下行性能。除非另有说明，我们考虑长度为 $\tau_{\mathrm{c}}=580$ 次信道使用的 TDD帧，其由 $T_{\mathrm{c}}=2$ ms 和 $B_{\mathrm{c}}=290$ kHz 给出，支持最高 67.50 km/h 的用户移动速度。

在第一组仿真中，我们考虑如下设置：$K=20$，$\alpha=0.2$，$a_{\mathrm{u}}=10^{-0.5}$，$\tau_{\mathrm{p}}=80$ （无导频污染），$T=5$ 个时隙，每个时隙长度为 $n_{\mathrm{d}}=100$ 次信道使用。图2绘出了在叠加编码（上图）和打孔技术（下图）下，采用不同预编码和功率分配策略时，每个“活跃”eMBB用户可实现的下行频谱效率累积分布函数。在上述假设下，其他条件相同时，SPC的性能远优于PUNC。正如预期，采用最优功率分配的M-MMSE可获得最高频谱效率，但无论采用何种预编码方案，等功率分配的表现都几乎与之相当。RZF在M-MMSE和MR之间实现了出色且实用的权衡。这些结果表明，由于M-MMSE和RZF分别提供了完全和部分的干扰抑制机制，系统近似处于无干扰场景。对于FPA策略，在这些仿真中，我们选择 $\nu=0.5$ 以促进机会式功率分配，并选择 $\omega=0.6$ 以优先保障超可靠低时延通信用户。这种选择不利于eMBB用户，也解释了在采用SPC时，FPA在所考虑策略中的性能为何最差。

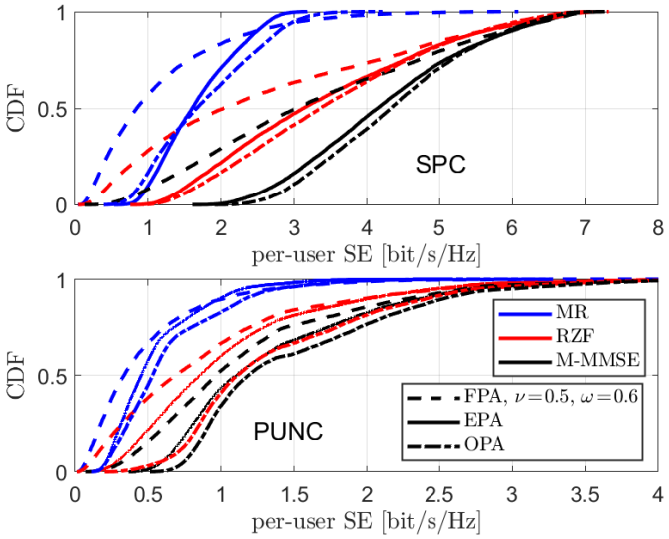

图2。不同传输、预编码和功率分配策略下，每个活跃增强型移动宽带用户可实现的下行频谱效率的累积分布函数。设置：$K=20$、 $\alpha=0.2$、 $a_{\mathrm{u}}=10^{-0.5}$、 $\tau_{\mathrm{p}}=80$、 $T=5$、 $n_{\mathrm{d}}=100$。

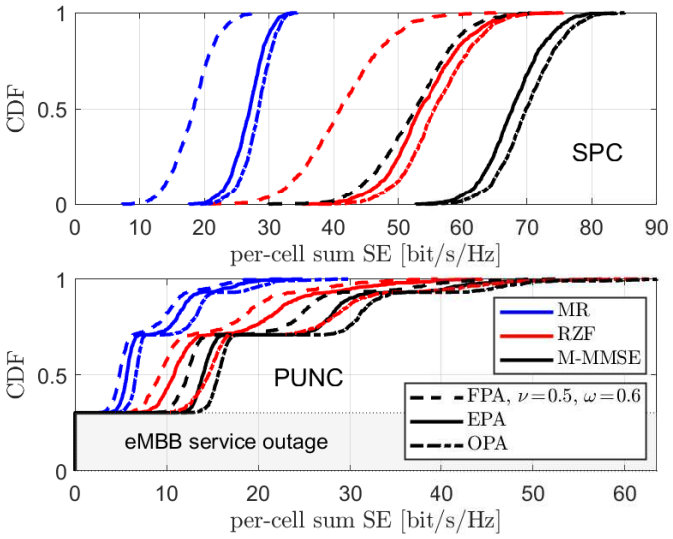

图3。不同传输、预编码和功率分配策略下，每小区可实现的下行总频谱效率的累积分布函数。设置： $K=20$、 $\alpha=0.2$ $a_{\mathrm{u}}=10^{-0.5}$ $\tau_{\mathrm{p}}=80$ $T=5$ $n_{\mathrm{d}}=100$ 、、、、。

图3所示结果也得出了相同结论，其中展示了相应的每小区总频谱效率累积分布函数。在这些图中，我们重点强调采用PUNC时可能出现的eMBB 业务中断。我们将PUNC操作下的eMBB服务中断定义为

$$\varsigma_{\mathrm{out}} = \mathsf{Pr}\left\{ \sum_{k \in \mathcal{K}_j^{\mathrm{e}}} \mathsf{SE}_{jk}^{\mathrm{e}} = 0 \right\}, \quad j = 1, \ldots, L \ ,$$

其中，概率是相对于大尺度衰落计算的。基站在一个TDD帧内无法为其增强型移动宽带用户提供任何业务的概率，取决于超可靠低时延通信用户的激活模式以及每帧时隙数。我们将在后文详细讨论这一问题。在

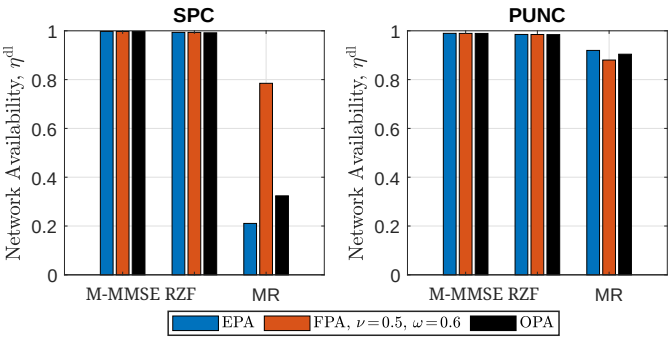

图4。不同传输、预编码和功率分配策略下的网络可用性。设置： $K=20$、 $\alpha=0.2$、 $a_{\mathrm{u}}=10^{-0.5}$、 $\tau_{\mathrm{p}}=80$、 $T=5$ $n_{\mathrm{d}}=100$ 、。

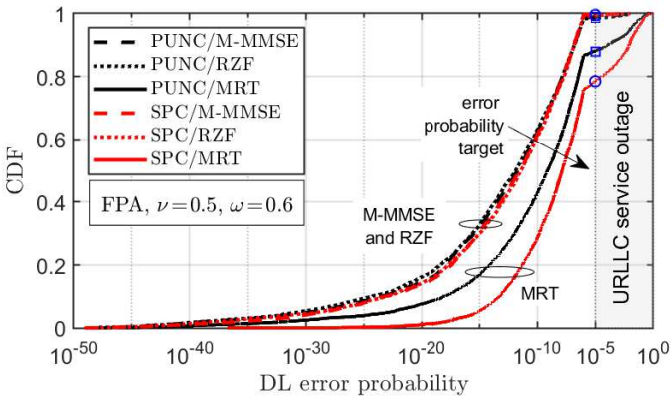

图5. 不同传输和预编码策略下的下行用户误差概率。设置：等功率分配， $K=20$, $\alpha=0.2$, $a_{\mathrm{u}}=10^{-0.5}$, $\tau_{\mathrm{p}}=80$, $T=5$ $n_{\mathrm{d}}=100$,.

图3所示的设置下，增强型移动宽带业务中断相当严重，约为30%。

图4展示了当 $\epsilon_{\mathrm{target}}^{\mathrm{dl}} = 10^{-5}$ 时所达到的下行网络可用性，从而转而考察超可靠低时延通信（URLLC）的性能。尽管执行SPC时会受到增强型移动宽带用户造成的干扰，M-MMSE和正则化零迫（RZF）都能够提供接近1的网络可用性，与PUNC相当，这表明它们具有出色的抑制干扰和支持高可靠性的能力。相反，使用等功率分配（EPA）或最优功率分配（OPA）（对增强型移动宽带用户而言是最优的）方案时，最大比合并（MR）在SPC下的性能较差。值得注意的是，我们为FPA方案选择的参数使SPC/MR组合表现良好。图4所示的网络可用性数值是根据错误概率得出的，其累积分布函数如图5所示。为了更好地理解其含义，网络可用性由每用户误差概率的累积分布函数与表示误差概率目标值的垂直线的交点给出，如图5所示（蓝色圆圈标记）。通过这组仿真，我们得出结论：采用M-MMSE或RZF时，SPC在频谱效率方面明显优于PUNC，同时仍能提供极高的网络可用性。如果最大比合并（MR）是唯一可行的选择（例如，由于严格的复杂度或硬件约束），那么在适当设置设计参数 ν 和 ω后，采用FPA的SPC是一种有效选择，既能保持较高的网络可用性，又能避免

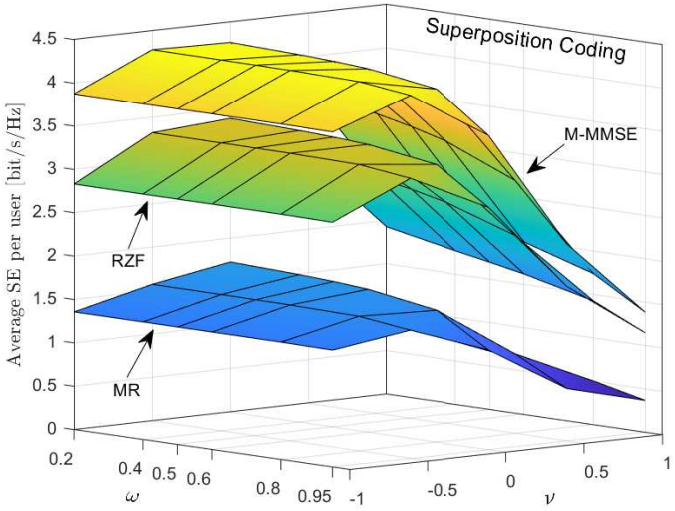

图 6. 在不同预编码方案和 ν、ω取值下，采用FPA的SPC所实现的平均用户频谱效率。该平均值取自200个网络快照。设置： $K=20$、 $\alpha=0.2$、 $a_{\mathrm{u}}=10^{-0.5}$、 $\tau_{\mathrm{p}}=80$、 $T=5$、$n_{\mathrm{d}}=100$ 。

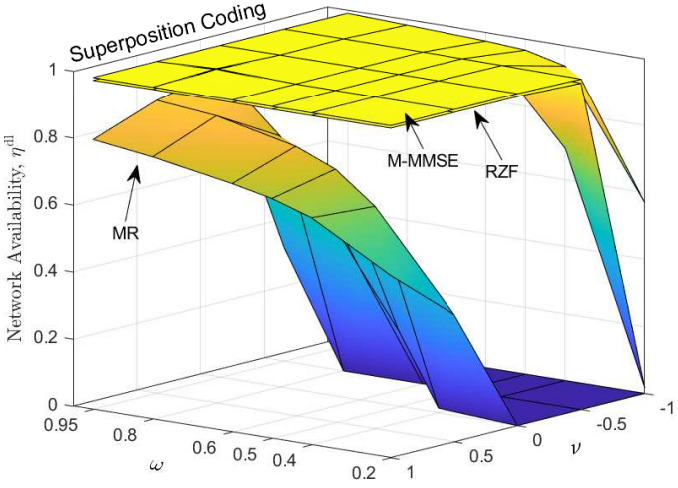

图 7. 采用不同预编码方案及不同取值时，FPA 与 SPC 所实现的网络可用性 ν, ω。参数设置： $K=20$, $\alpha=0.2$, $a_{\mathrm{u}}=10^{-0.5}$ $\tau_{\mathrm{p}}=80$ $T=5$ $n_{\mathrm{d}}=100$,,,。

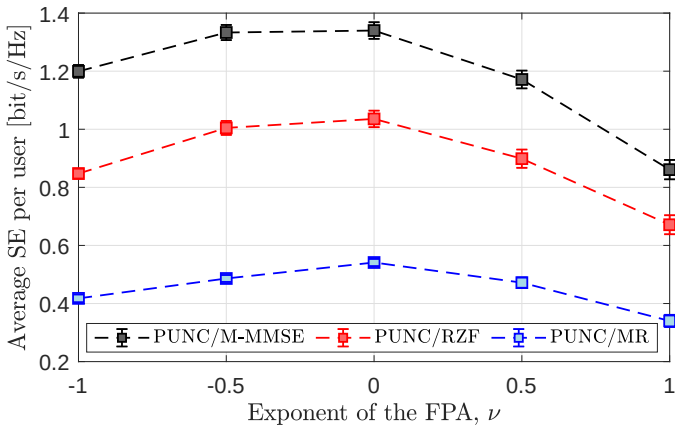

图 8. 对于不同的预编码方案和 ν 取值，采用 FPA 的 PUNC 所实现的用户平均频谱效率（含 95%置信区间）。平均值基于 200 个网络快照计算。设置： $K=20$, $\alpha=0.2$, $a_{\mathrm{u}}=10^{-0.5}$,$\tau_{\mathrm{p}}=80$ $T=5$ $n_{\mathrm{d}}=100$,,。

任何增强型移动宽带业务中断。

为此，我们现在着重研究如何恰当地选择 ν 和 ω。采用与第一组仿真相同的设置，在图 6 中，我们绘制了采用SPC和不同预编码方案时的平均用户频谱效率，并考察FPA取 ν 和 ω 时的变化。从eMBB用户的角度来看，最好将 ω设为较小值，并使 ν 处于区间$[-0.5,0]$内。前者显而易见，后者则需要进一步讨论。事实上，请回想， ν 取正值时，可以为信道条件更好的用户分配更多功率。由于我们假设超可靠低时延通信用户均匀分布在基站周围较小的区域内，因此他们很可能比大多数eMBB用户离基站更近。因此， ν 取负值会提高公平性并改善eMBB用户的性能。 ω 和 ν 的取值都较大时，会使功率分配过度偏向超可靠低时延通信用户，从而降低eMBB用户的频谱效率。

相反，如图 7 所示，在 SPC 运行中， ω 和 ν 的值都较小时，超可靠低时延通信用户的网络可用性会崩溃。尽管如此，除非 $\nu=-1$，M-MMSE 和正则化零迫都能将网络可用性维持在接近 1 的水平，而最大比合并则对这些参数的调整相当敏感。采用 SPC 时，抑制多用户干扰至关重要。正则化零迫虽然无法处理小区间干扰，却在性能与实用性之间实现了出色的权衡。对 FPA 方案的参数进行微调后，采用最大比合并时可获得令人满意的性能。只要无法通过预编码技术消除多用户干扰，FPA 就是应对该干扰的一种有效的启发式替代方案。

使用 PUNC 与 FPA 时，设置 ω 已毫无意义，因为所考虑的时隙中仅发生超可靠低时延通信传输。因此，在图 8 和图 9 中，我们关注每用户的平均频谱效率和网络可用性，因为只有 ν 会变化。对于这两种情况，我们都发现等功率分配，即 $\nu=0$，是理想的。就增强型移动宽带用户的频谱效率而言， ν 的负值有助于获得较低的频谱效率（例如，每用户 95% 可能达到的频谱效率），从而提升用户间的公平性；而 ν 的较大正值则以贪婪方式追求峰值频谱效率，忽视较低的频谱效率。因此，若目标是平均频谱效率，$\nu=0$ 是合理的选择，尤其是在多用户干扰被部分或完全消除时。就超可靠低时延通信用户的网络可用性而言，只要采用 M-MMSE 或正则化零迫，$\nu\in[-1,1]$ 的任何取值都很可靠；而每当 ν 取非中性值时，最大比合并的性能都会相对受损。推测同时在同一时隙处于活动状态的超可靠低时延通信用户数量（由所选的 α 和 a_{u} 决定）不会造成显著的多用户干扰。

接下来，我们评估性能随TDD帧中的时隙数 T 和时隙大小 n_{d} 的变化情况，而时隙大小又决定了超可靠低时延通信码字长度。在这组仿真及后续内容中，根据前述结果，我们省略最大比合并（MR）的结果，仅考虑采用 $\nu=0$ 和 $\omega=\alpha$ 的 FPA。图10展示了三种不同配置下每小区总频谱效率的累积分布函数：

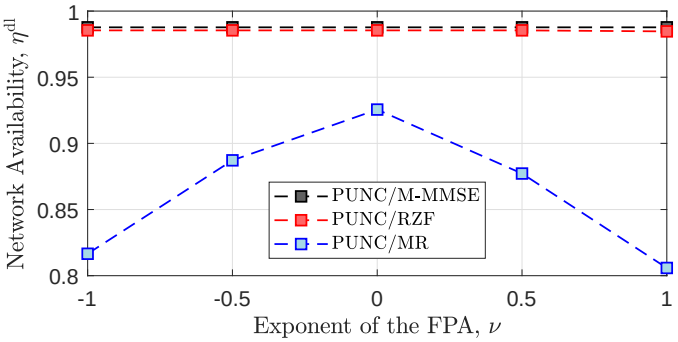

图 9. 不同预编码方案和 ν 取值下，PUNC 采用 FPA 实现的网络可用性。设置：$K=20$, $\alpha=0.2$, $a_{\mathrm{u}}=10^{-0.5}$, $\tau_{\mathrm{p}}=80$ $T=5$ $n_{\mathrm{d}}=100$,,。

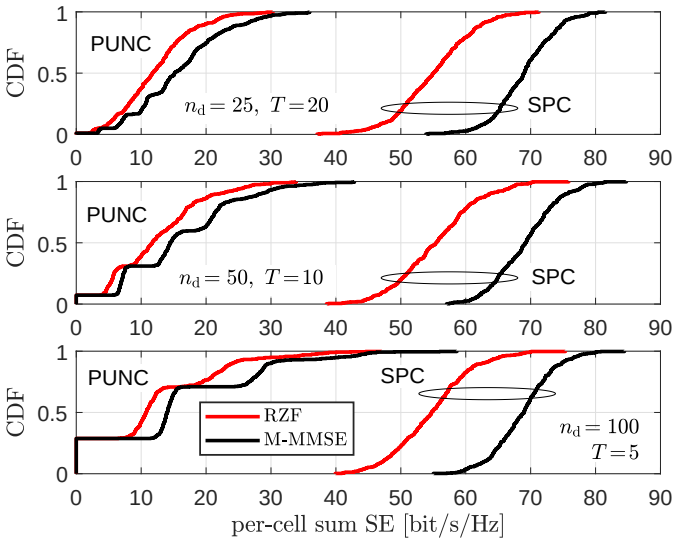

图 10. 随着每帧时隙数变化，不同传输和预编码策略下可实现的每小区下行总频谱效率的累积分布函数。设置：FPA， $\nu=0$ 和 $\omega=0.2$, $K=20$, $\alpha=0.2$, $a_{\mathrm{u}}=10^{-0.5}$, $\tau_{\mathrm{p}}=80$ 。

(*i*) $n_{\mathrm{d}}=25$, $T=20$; (*ii*) $n_{\mathrm{d}}=50$, $T=10$; 以及 (*iii*) $n_{\mathrm{d}}=100$, $T=5$。使用 SPC 时，TDD帧结构对增强型移动宽带用户的频谱效率影响不大。相反，在采用 PUNC 时，帧结构会显著影响每小区总频谱效率。事实上，增加每帧时隙数会降低增强型移动宽带服务中断的概率，因为这增加了增强型移动宽带用户找到没有超可靠低时延通信用户活动的时隙的机会。图10中的结果支持这一观点：当 $T=20$、$T=10$ 和 $T=5$ 时，增强型移动宽带服务中断率分别为 0.01、0.0725 和 0.2875。另一方面，时隙数较少时，增强型移动宽带用户可能会持续活动更长时间，从而获得更高的频谱效率。这解释了为什么随着 T 减小，每小区总频谱效率的变化幅度会增大。

时隙长度会直接影响超可靠低时延通信用户的性能。如图 11 所示，网络可用性会随着时隙长度（即超可靠低时延通信码字长度）的增加而大幅提高。事实上，超可靠低时延通信码字长度决定了超可靠低时延通信用户的传输速率 $R=b/n_{\mathrm{d}}$，因此码字越短，

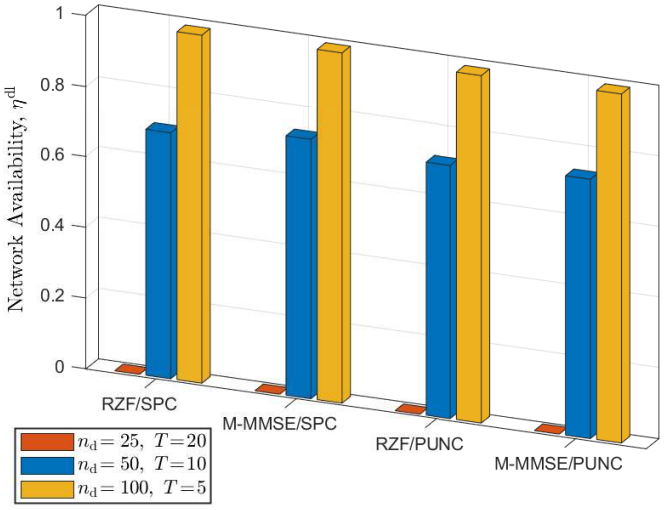

图 11. 随着时隙长度变化，不同传输和预编码策略下的网络可用性。设置：FPA， $\nu=0$ 和 $\omega=0.2$, $K=20$, $\alpha=0.2$, $a_{\mathrm{u}}=10^{-0.5}$, $\tau_{\mathrm{p}}=80$。

要可靠实现的速率要求就越高，进而错误概率也越大。[2] 同样，SPC 是总体上能为增强型移动宽带用户和超可靠低时延通信用户提供最佳性能的技术，因为其主要局限——所造成的多用户干扰——可通过采用基于干扰抑制的预编码方案来克服。最后，尽管让超可靠低时延通信传输跨越多个信道使用次数有利于提高网络可用性，但时延要求要求将传输限定在特定时间范围内。

现在，我们将重点转向导频污染和估计开销对性能的影响。固定TDD帧长度和每帧时隙数后，我们改变上行训练长度，从而改变可用正交导频的数量，并相应调整每个时隙的长度。图12展示了不同运行场景下每小区平均总频谱效率随上行训练长度的变化。在这些仿真中，我们假设 $K=20$、$\alpha=0.2$、$\tau_{\mathrm{c}}=580$ 和 $T=5$。τ_{p} 较小时，信道估计开销较低，但导频污染程度较高，从而降低预编码的有效性。我们的导频分配方案通过在可用时为超可靠低时延通信用户分配专属导频，来保障其性能；否则，导频将随机分配，污染会无差别地影响所有用户。根据所选参数，此场景下可能同时活跃的超可靠低时延通信用户最多为16个。因此，当 $\tau_{\mathrm{p}}=10$ 时，导频会随机分配，导致小区内和小区间导频污染，使每小区总频谱效率较低：采用SPC时约为30 bit/s/Hz，采用PUNC时则低于10 bit/s/Hz。当 $\tau_{\mathrm{p}}=20$时，性能会进一步恶化，因为超可靠低时延通信用户的保护机制此时会被触发，增强型移动宽带用户只能共享4个正交导频。随着 τ_{p}增大，小区内导频

[2](18) 中的随机编码联合界将错误概率定义为平均广义信息密度小于传输速率要求的概率。

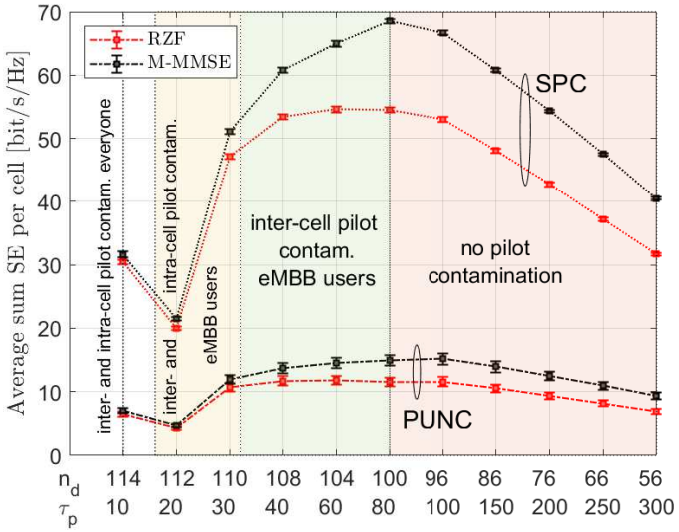

图12. 不同传输和预编码策略下的每小区平均频谱效率（含95%置信区间），随 τ_{p}（及 n_{d}）变化的情况。平均值取自200个网络快照。设置：采用 FPA，$\nu=0$ 和 $\omega=0.2$，$K=20$，$\tau_{\mathrm{c}}=580$，$T=5$，$\alpha=0.2$，$a_{\mathrm{u}}=10^{-0.5}$，，。

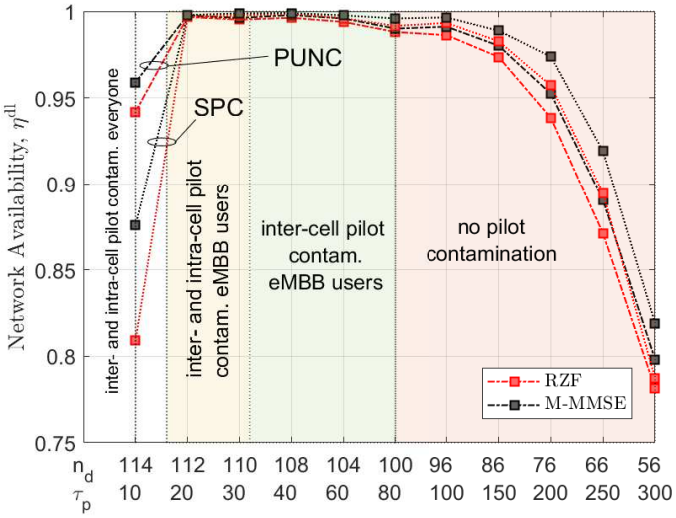

图13. 在不同传输和预编码策略下，随着 τ_{p}（以及 n_{d}）变化时的网络可用性。设置：采用 $\nu=0$ 和 $\omega=0.2$的FPA，$K=20$，$\tau_{\mathrm{c}}=580$，$T=5$，$\alpha=0.2$，$a_{\mathrm{u}}=10^{-0.5}$，，。

污染主要通过为同一小区内的增强型移动宽带用户分配正交导频来减少。如果 $\tau_{\mathrm{p}}\geq 32$，则可避免小区内导频污染，而增强型移动宽带用户之间仅受到小区间干扰的影响。每小区总频谱效率持续增长，直至 $\tau_{\mathrm{p}}=80$；此时，网络中的所有用户均被分配相互正交的导频，完全避免导频污染所带来的收益超过了估计开销增加造成的损失。显然，进一步增大 τ_{p}并不会带来任何信道估计方面的收益，反而会使估计开销变得高昂，并大幅降低每小区总频谱效率。最后请注意，当发生小区内和小区间导频污染时，正则化零迫和M-MMSE的性能基本相同，因为这两种方案抑制多用户干扰的能力都较弱。

对于超可靠低时延通信用户，当 $\tau_{\mathrm{p}}<16$ 时，导频污染会严重影响网络可用性，尤其是在采用SPC的情况下；尽管较长的时隙会降低速率要求，这一点仍然成立，如图13所示。超可靠低时延通信用户之间的导频污染具有破坏性，主要是因为这些用户可能彼此距离很近且靠近基站，因而会受到强烈的相互干扰；当它们的信道估计相关时，这种干扰无法消除。因此，我们提出在导频分配中优先考虑超可靠低时延通信用户的方法，在技术上是合理的。此外，增加估计开销会大幅降低网络可用性，因为分配给数据传输的资源会减少，也就是说时隙长度会缩短；而且，如前所述，超可靠低时延通信用户的速率要求会提高。

接下来，我们研究随机激活模式和每帧潜在活跃的超可靠低时延通信用户数量对性能的影响。图14展示了在采用不同传输和预编码方案以及 $\nu=0$ 和 $\omega=\alpha$的FPA时，a_{u} 和 α 变化情况下的每小区平均总频谱效率。请注意，将 ω 与 α 成比例增加是合理的

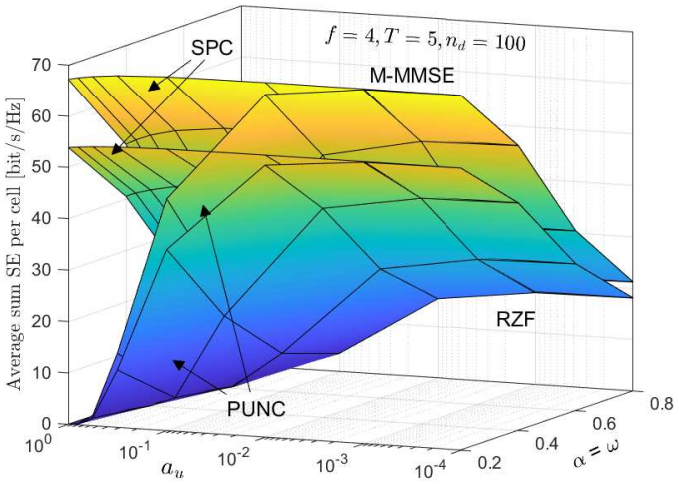

图14. 在不同传输和预编码策略下，随着 a_{u} 和 α 变化时的每小区平均频谱效率。平均值基于200个网络快照计算。设置：采用 $\nu=0$ 和 $\omega=\alpha$的FPA，$K=20$，$\tau_{\mathrm{c}}=580$，$f=4$，$T=5$，$n_{\mathrm{d}}=100$。

随着分配给越来越多潜在活跃的超可靠低时延通信用户的功率增加，SPC 的性能表现为……，尤其是在 a_{u} 取值较大时。在这些仿真中，我们假设两种 TDD帧配置：(i) $f=4$，$T=5$，$n_{\mathrm{d}}=100$，以及 $(ii)f=3$，$T=8$，$n_{\mathrm{d}}=65$（其结果则见图 15）。首先，我们观察到，采用所考虑的 TDD帧配置可以获得相近的每小区平均总频谱效率：使用第二种帧配置时，导频污染会使增强型移动宽带用户的性能略有下降。当 $a_{\mathrm{u}}\geq 10^{-2}$ 时，PUNC 的性能会收敛到 SPC 的性能，因此在激活模式稀疏时确实如此，这符合预期。同样，在第二种场景（图 15）中，正则化零迫与 M-MMSE 之间的性能差距缩小，因为小区间导频污染降低了 M-MMSE 抑制多用户干扰的能力。当 a_{u} 取值较大时，PUNC 会导致增强型移动宽带服务中断，而

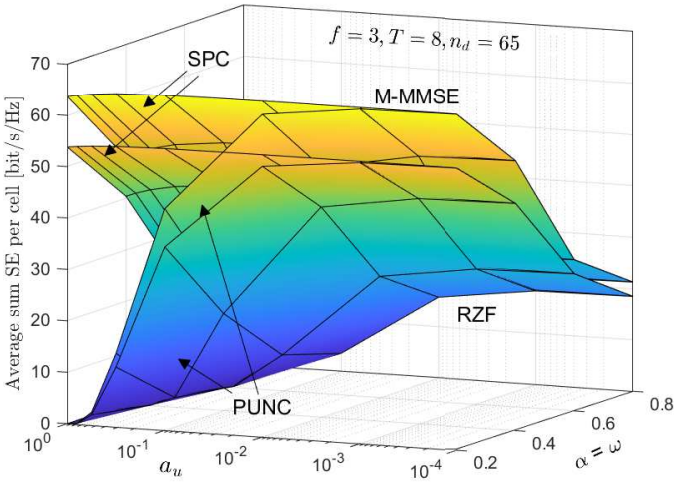

图15. 不同传输和预编码策略下，随着 a_{u} 和 α 变化的每小区平均频谱效率。该平均值基于200个网络快照计算。设置：采用FPA， $\nu=0$ 和 $\omega=\alpha$、 $K=20$、 $\tau_{\mathrm{c}}=580$、 $f=3$、 $T=8$、$n_{\mathrm{d}}=65$ 。

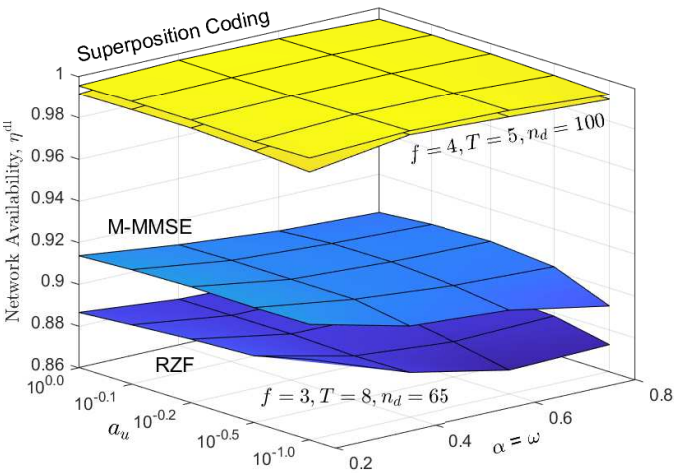

图16. 不同预编码策略下，随着 a_{u} 和α 变化的网络可用性。该平均值基于200个网络快照计算。设置：采用SPC和FPA， $\nu=0$ 和 $\omega=\alpha$、 $K=20$、 $\tau_{\mathrm{c}}=580$。考虑了两种TDD帧配置。

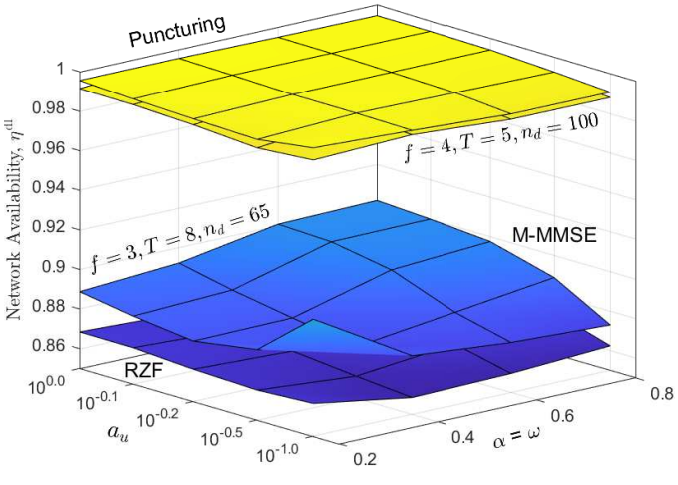

图17. 不同预编码策略下，随着 a_{u} 和α 变化的网络可用性。设置：采用PUNC和FPA， $\nu=0$ 和 $\omega=\alpha$、 $K=20$、 $\tau_{\mathrm{c}}=580$。考虑了两种TDD帧配置。

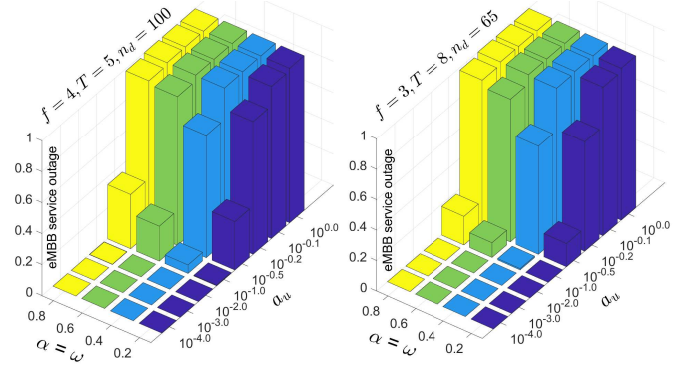

图18. 不同预编码策略下，增强型移动宽带服务中断随 a_{u} 和α 变化的情况。设置：PUNC和FPA，采用 $\nu=0$ 和 $\omega=\alpha$、 $K=20$、 $\tau_{\mathrm{c}}=580$。考虑两种TDD帧配置。

SPC 仍能消除超可靠低时延通信用户干扰，并提供出色的频谱效率。最后，我们观察到，如果 80% 的用户请求超可靠低时延通信业务，那么与情形$\alpha=0.2$相比，增强型移动宽带用户的性能会下降近三分之一。这一结果主要是由 FPA 方案中所选的 ω值造成的；该方案旨在随着超可靠低时延通信用户数量增加而优先保障其性能。

如图16所示，对于SPC而言，图17所示，对于PUNC而言，两种所考虑的TDD帧配置在网络可用性方面所取得的性能存在明显差异。在这两种情况下，缩短时隙长度都会导致约10%的性能损失，而导频污染只影响增强型移动宽带用户。使用PUNC时，这种性能差距略为明显，因为基站的全部功率都分配给了超可靠低时延通信用户，从而导致更强的相互干扰。总体而言，第一种TDD帧配置对于所考虑的各种传输和预编码策略、随机的超可靠低时延通信激活模式以及超可靠低时延通信用户负载，都表现得相当稳健。

这组仿真还需分析的最后一个方面是：采用 PUNC 时，增强型移动宽带服务中断概率如何随a_{u} 和 α 变化。这将完整展现 PUNC 对增强型移动宽带用户而言在哪些工作点是有效选择，也将进一步说明合理构造 TDD 帧的重要性。如图 18 所示，采用 PUNC 时，使用 $T=8$ 个时隙的 TDD 帧配置相较于采用 $T=5$的配置，能更有效地避免增强型移动宽带服务中断。例如，当 $a_{\mathrm{u}}=10^{-1}$ 和 $\alpha=0.8$或 $\alpha=0.6$时，将帧中用于数据传输的部分划分为 8 个时隙，可使增强型移动宽带服务中断率比采用 5 个时隙的情况降低一半。总体而言，只有在超可靠低时延通信业务负载较低的场景中，并且合理构造 TDD 帧、能够容忍增强型移动宽带性能在每小区总频谱效率或增强型移动宽带服务中断方面适度下降时，PUNC 才能与 SPC 竞争。另一方面，SPC 依赖于能够抑制多用户干扰的预编码方案，而这又利用了可用的空间自由度

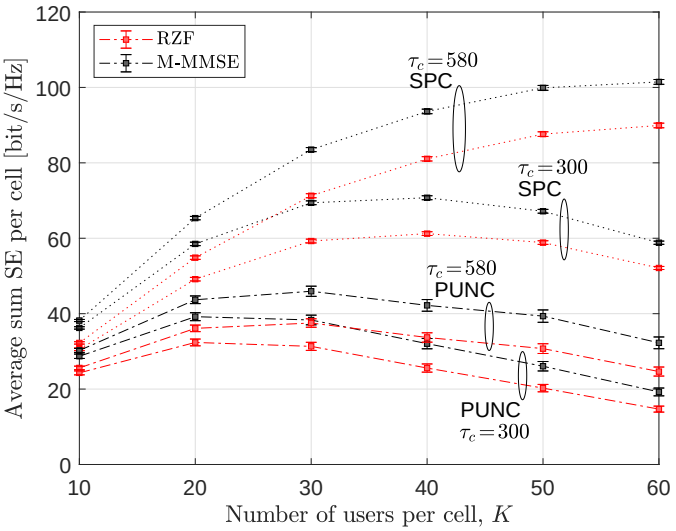

图 19. 不同传输和预编码策略下的每小区平均频谱效率（含 95%置信区间），随 K 和 τ_c 变化。平均值基于 200 个网络快照计算。设置：FPA，$\nu=0$ 和 $\omega=0.2$，$\alpha=0.2$,$a_u=10^{-1}$ $f=3$ $T=5$,,。

基站与所获取信道状态信息的高准确度。

最后，我们通过改变用户总数和 TDD 帧长度来评估性能。图 19 展示了不同传输和预编码策略下的每小区平均总频谱效率；每小区用户数 K从 10 增至 60，并考虑两种不同的 TDD 帧长度，即 580 和 300 次信道使用。与 580 次信道使用的情况相比，较短的帧可能对应更短的相干时间、更窄的相干带宽以及更高的用户移动性。不过，较短的帧意味着可分配给数据传输和上行训练的资源更少。在这些仿真中，我们假设采用 FPA，且$\nu=0$和 $\omega=0.2$, $\alpha=0.2$, $a_u=10^{-1}$, $T=5$ 以及导频复用因子 $f=3$。此外，由于 $\tau_p=fK$和 τ_c 固定，对于 K 的每个取值，上行训练配置和时隙长度各不相同，即分别为 τ_p 和 n_d。从图 19 可以看出，每小区平均总频谱效率随 K增大而提高，这说明采用 M-MMSE 和正则化零迫的 SPC 具有出色的用户空间复用能力。当 $\tau_c=580$时，K大于 60 后，每小区平均总频谱效率趋于饱和；而在 $\tau_c=300$ 时，这一现象约在 40 处出现，此时信道估计开销对频谱效率造成了沉重负担。PUNC 的性能远逊于 SPC，因为它分配给增强型移动宽带用户的资源较少，而且随着 K 增加，性能差距也会扩大，因为每小区超可靠低时延通信用户的数量会按比例增长。因此，增大K 会使打孔时隙更容易出现，这不仅会占用增强型移动宽带用户的资源、降低其频谱效率，还会增加增强型移动宽带服务中断率，如表 III 所示。请注意，只要 T 保持不变，改变 τ_c 不会影响增强型移动宽带服务中断率。

表 III 和表 IV 展示了不同传输与预编码策略以及不同 K 取值下的网络可用性，同时也强调了 τ_p 和 n_d 如何相应变化以满足 TDD 帧长度。具体而言，表 III 展示了采用 $\tau_c=580$时的性能，而

表 III 网络可用性与增强型移动宽带服务中断，$\tau_c=580$

K	τ_p	n_d	η^{dl}				ζ中断
			SPC		PUNC		PUNC
			M-MMSE	RZF	M-MMSE	RZF	
10	30	110	0.9989	0.9966	1	0.9989	0.0012
20	60	104	0.9988	0.9957	0.9944	0.9906	0.0038
30	90	98	0.9988	0.9950	0.9934	0.9893	0.0225
40	120	92	0.9969	0.9885	0.9881	0.9819	0.0625
50	150	86	0.9864	0.9787	0.9790	0.9672	0.1050
60	180	80	0.9807	0.9697	0.9728	0.9601	0.1737

表 IV 网络可用性与增强型移动宽带服务中断，$\tau_c=300$

K	τ_p	n_d	η^{dl}				ζ中断
			SPC		PUNC		PUNC
			M-MMSE	RZF	M-MMSE	RZF	
10	30	54	0.7936	0.7683	0.7844	0.7534	0.0012
20	60	48	0.6786	0.6353	0.6905	0.6685	0.0038
30	90	42	0.4796	0.4296	0.5646	0.5435	0.0225
40	120	36	0.1813	0.1457	0.3192	0.3192	0.0625
50	150	30	0.0021	0	0.0250	0.0250	0.1050
60	180	24	0	0	0	0	0.1737

表 IV 展示了采用 $\tau_c=300$时的性能。采用 $\tau_c=580$ 的 TDD 帧可在每小区用户数达到 60（其中包括 12 个 URLLC 用户）时，使网络可用性保持在 96% 以上，且适用于所有考察的传输与预编码技术。这意味着，这样的资源量足以出色地支持所考察的 URLLC 用户负载及其激活模式。相反，表 IV 所示采用$\tau_c=300$的 TDD 帧所能支持的网络可用性要低得多；在 $K\geq 50$时甚至接近于零（或等于零）。这突显了网络可用性对 TDD 帧长度、进而对可用资源量的敏感程度。值得注意的是，随着K 增加，网络可用性呈下降趋势；对于 PUNC，这种下降较为缓和，主要归因于较短的 URLLC 码字长度；而对于 SPC，下降则很显著，主要归因于多用户干扰的增加。事实上，表 IV 的结果清楚地证实，在 $K\geq 20$时，PUNC 比 SPC 更稳健。

VI. 结论

本文研究了多小区大规模MIMO系统下行链路中异构业务的非正交复用，即增强型移动宽带（eMBB）与超可靠低时延通信（URLLC）的复用。eMBB和URLLC的特性截然不同，需求也各异。eMBB传输涉及跨越多个无线帧的大载荷，并要求较高的频谱效率。相比之下，URLLC用户会间歇性地在极短时间内传输小载荷，同时要求低时延和达到 10^{-5} 量级的成功概率。这种异构性要求采用有效的资源分配策略，使eMBB与URLLC能够和谐共存。首先，我们提出了一个统一的信息论框架，用于评估无限块长度遍历状态下eMBB的频谱效率（SE），以及非渐近有限块长度范围内URLLC的错误概率。两种

分析均考虑了基站（BS）通过上行导频传输获取信道状态信息（CSI）时的不完美性、导频污染和导频开销、空间相关信道，以及用户端缺少CSI的情况。其次，我们将所提出的框架加以推广，以适用于两种可选的共存策略：打孔（PUNC）和叠加编码（SPC）。前者通过避免业务间干扰来保护URLLC的可靠性，后者则接受此类干扰，以维持eMBB业务。第三，我们通过数值分析评估了在不同预编码和功率分配方案下，采用PUNC和SPC所能达到的性能，并考察了不同的时分双工无线帧配置和URLLC随机激活模式。仿真结果表明，当充分利用基站可用的空间自由度，采用基于干扰抑制的预编码方案，并获取高质量信道状态信息时，可以显著消除SPC操作引起的多用户干扰，使eMBB频谱效率远高于PUNC，同时确保URLLC的错误概率仍处于同样优异的水平。然而，如果这些条件不成立，例如严重的导频污染导致信道估计值质量下降，或自由度不足以处理多个用户间的干扰，那么为维持URLLC性能，打孔就成为必要操作，尽管这可能导致eMBB服务中断。不同于以往研究中不恰当地通过中断容量分析或正态近似得到的错误概率来评估URLLC性能，本文的有限块长度信息论分析依赖于失配接收机和鞍点近似，后者适用于大规模MIMO运行中的URLLC场景。未来可将海量机器类通信（mMTC）纳入共存策略，并在所提出的广义框架中研究上行链路。最后，研究分布式用户中心系统（如无小区大规模MIMO）中异构业务的非正交复用也是一个很有吸引力的未来研究方向；这类系统能够提供用户邻近性、宏分集和泛在连接 [34]–[36]。

参考文献

[1] *IMT愿景——2020*年及以后国际移动通信发展的框架与总体目标，国际电信联盟无线电通信部门标准 M.2083-0，2015年。[2] P. Popovski、K. F. Trillingsgaard、O. Simeone 和 G. Durisi，“面向增强型移动宽带、超可靠低时延通信和海量机器类通信的第五代移动通信无线网络切片：通信理论视角”，*IEEE Access*，第 6 卷，第 55 765–55 779 页，2018年。[3] T. L. Marzetta，“具有无限数量基站天线的非协作蜂窝无线通信”，*IEEE Trans. Wireless Commun.*，第 9 卷，第 11 期，第 3590–3600 页，2010年。[4] T. L. Marzetta、E. G. Larsson、H. Yang 和 H. Q. Ngo，大规模多输入多输出基础。剑桥大学出版社，2016年。[5] E. Bj¨ornson、J. Hoydis 和 L. Sanguinetti，“大规模多输入多输出网络：频谱效率、能效与硬件效率”，*Foundations and Trends® in Signal Processing*，第 11 卷，第 3–4 期，第 154–655 页，2017年。[6] P. Popovski、J. J. Nielsen、˘C. Stefanovi´c、E. d. Carvalho、E. Str¨om、K. F. Trillingsgaard、A. Bana、D. M. Kim、R. Kotaba、J. Park 和 R. B. Sørensen，“面向超可靠低时延通信的无线接入：原则与基本组成”，*IEEE Network*，第 32 卷，第 2 期，第 16–23 页，2018年3月。

[7] A.-S. Bana、E. de Carvalho、B. Soret、T. Abr˜ao、J. C. Marinello、E. G. Larsson 和 P. Popovski，“面向物联网（IoT）连接的大规模多输入多输出”，《物理通信》，第 37， 100859, 卷，第 . 页。2019 [ ]。在线。可访问：http://www.sciencedirect.com/science/article/ pii/S1874490719303891

[8] ¨ J. Ostman、A. Lancho、G. Durisi 和 L. Sanguinetti，“大规模多输入多输出中的超可靠低时延通信：有限码长下的分析与设计”，IEEE无线通信汇刊，第20卷，第10期，第6387–6401页，2021年10月。[9] E. Bj¨ornson、E. de Carvalho、J. H. Sørensen、E. G. Larsson 和 P. Popovski，“高密度大规模多输入多输出系统中用于导频分配的随机接入协议”，IEEE无线通信汇刊，第16卷，第4期，第2220–2234页，2017年4月。[10] A. Anand、G. De Veciana 和 S. Shakkottai，“第五代无线网络中超可靠低时延通信与增强型移动宽带业务的联合调度”，载于IEEE计算机通信会议（INFOCOM）论文集，2018年4月，第1970–1978页。[11] R. Kassab、O. Simeone、P. Popovski 和 T. Islam，“雾无线架构中超可靠业务与宽带业务的非正交复用”，IEEE Access，第7卷，第13 035–13 049页，2019年。[12] A. A. Esswie 和 K. I. Pedersen，“多用户第五代网络中超可靠低时延通信与增强型移动宽带共存的机会式空间抢占调度”，IEEE Access，第6卷，第38 451–38 463页，2018年。[13] S. F. Abedin、M. G. R. Alam、S. M. A. Kazmi、N. H. Tran、D. Niyato 和 C. S. Hong，“雾网络中面向超可靠及增强型移动宽带物联网应用的资源分配”，IEEE通信汇刊，第67卷，第1期，第489–502页，2019年1月。[14] A. Matera、R. Kassab、O. Simeone 和 U. Spagnolini，“采用模拟前传的云无线接入网中的非正交增强型移动宽带-超可靠低时延通信无线接入”，Entropy，第20卷，第9期，2018年。[15] M. Alsenwi、N. H. Tran、M. Bennis、A. Kumar Bairagi 和 C. S. Hong，“增强型移动宽带-超可靠低时延通信资源切片：一种风险敏感方法”，IEEE通信快报，第23卷，第4期，第740–743页，2019年4月。[16] R. Abreu、T. Jacobsen、G. Berardinelli、K. Pedersen、N. H. Mahmood、I. Z. Kovacs 和 P. Mogensen，“上行链路中宽带业务流量与免授权超可靠通信的复用”，载于IEEE车辆技术会议（VTC-Spring）论文集，2019年4月，第1–6页。[17] E. N. Tominaga、H. Alves、R. D. Souza、J. Luiz Rebelatto 和 M. Latva-aho，“非正交多址接入与网络切片：增强型移动宽带与超可靠低时延通信的可扩展共存”，载于IEEE车辆技术会议（VTC-Spring）论文集，2021年4月，第1–6页。[18] F. Saggese、M. Moretti 和 P. Popovski，“面向增强型移动宽带用户和超可靠低时延通信用户的下行频谱切片功率最小化”，IEEE无线通信汇刊，第21卷，第12期，第11 051–11 065页，2022年12月。[19] M. Almekhlafi、M. A. Arfaoui、C. Assi 和 A. Ghrayeb，“第六代无线网络中超可靠低时延通信与增强型移动宽带业务复用的联合资源与功率分配”，载于IEEE国际通信会议（ICC）论文集，2021年6月，第1–6页。[20] J. Zeng、T. Lv、R. P. Liu、X. Su、Y. J. Guo 和 N. C. Beaulieu，“通过大规模多用户多输入多输出，在阴影衰落下实现超可靠低时延通信”，IEEE物联网期刊，第7卷，第1期，第234–246页，2020年1月。[21] H. Ren、C. Pan、Y. Deng、M. Elkashlan 和 A. Nallanathan，“面向大规模多输入多输出赋能的超可靠低时延通信工业物联网网络的联合导频与有效载荷功率分配”，IEEE精选通信领域期刊，第38卷，第5期，第816–830页，2020年5月。[22] A. A. Nasir、H. D. Tuan、H. Q. Ngo、T. Q. Duong 和 H. V. Poor，“超可靠低时延通信有限码长区间内的无小区大规模多输入多输出”，IEEE无线通信汇刊，第20卷，第9期，第5861–5871页，2021年9月。[23] Y. Polyanskiy、H. V. Poor 和 S. Verdu，“有限码长区间内的信道编码速率”，IEEE信息论汇刊，第56卷，第5期，第2307–2359页，2010年5月。[24] J. Scarlett、A. Martinez 和 A. Guill´en i F`abregas，“失配译码：错误指数、二阶速率和鞍点近似”，IEEE信息论汇刊，第60卷，第5期，第2647–2666页，2014年5月。[25] A. Martinez 和 A. Guill´en i F`abregas，“随机编码界的鞍点近似”，载于信息论与应用研讨会（ITA）论文集，2011年2月，第1–6页。[26] W. Yang、G. Durisi、T. Koch 和 Y. Polyanskiy，“有限码长下的准静态多天线衰落信道”，IEEE信息论汇刊，第60卷，第7期，第4232–4265页，2014年7月。[27] G. Durisi、T. Koch、J. ¨Ostman、Y. Polyanskiy 和 W. Yang，“瑞利衰落多天线信道上的短分组通信”，IEEE通信汇刊，第64卷，第2期，第618–629页，2016年2月。[28] J. ¨Ostman、G. Durisi、E. G. Str¨om、M. C. Coc。kun 和 G. Liva，“分块无记忆衰落信道上的短分组：导频辅助

或非相干传输？”*IEEE*通信汇刊，第67卷，第2期，第1521–1536页，2019年2月。[29] D. Tse 和 P. Viswanath，无线通信基础。剑桥大学出版社，2005年。[30] A. Lapidoth 和 S. Shamai，“衰落信道：完美的‘边信息’究竟需要达到何种程度？”*IEEE*信息论汇刊，第48卷，第5期，第1118–1134页，2002年5月。[31] R. G. Gallager，信息论与可靠通信。美国纽约州纽约：约翰·威立父子出版公司，1968年。[32] 3GPP，演进型通用陆地无线接入物理层方面的进一步演进（第*9*版）。3GPP TS 36.814，2017年3月。[33] 3第三代合作伙伴计划3GPP，垂直行业中信息物理控制应用的服务要求。3GPP TS 22.104 v. 17.2.0，2019年12月。[34] S. Buzzi 和 C. D’Andrea，“以用户为中心的通信与面向第五代蜂窝网络的无小区大规模多输入多输出之比较”，载于*WSA 2017*；第*21*届国际*ITG*智能天线研讨会。VDE，2017年，第1–6页。[35] G. I nterdonato、E. Bj¨ornson、H. Q. Ngo、P. Frenger 和 E. G. Larsso n，“无处不在的无小区大规模MIMO通信”，*EURASIP*无线通信与网络期刊，第2019卷，第1期，第197页，2019年。[36] S. Buzzi、C. D’ Andrea、A. Zappone 和 C. D’Elia，“以用户为中心的第五代蜂窝网络：资源分配及与无小区大规模多输入多输出方案的比较”，*IEEE*无线通信汇刊，第19卷，第2期，第1250–1264页，2020年2月。