多天线通信信号处理进展25年

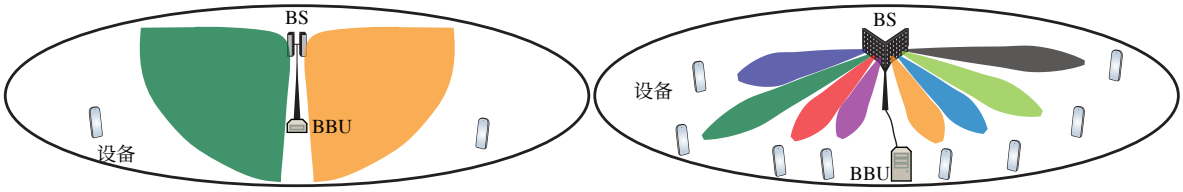
Emil Bj¨ornson、Yonina C. Eldar、Erik G. Larsson、Angel Lozano、H. Vincent Poor

arXiv:2304.02677v1 [eess.SP] 5 Apr 2023

无线通信技术在过去25年中取得了巨大进步，无论是在社会普及程度还是技术复杂性方面。1998年，移动电话仍在逐渐变得小巧且价格亲民，能够在发达国家和发展中国家广泛使用。当时，全球“仅有”3亿移动用户 [1]。蜂窝网络是最早实现私有化的电信市场之一，竞争使设备变成了设计时尚、外观吸引人且可以个性化定制的时尚配饰。服务范围仅限于电话和短信，但这一点具有开创性，因为电信首次实现了人与人之间的通信，而不再局限于地点之间。

如今，全球订户已超过60亿。手机仍是主要的无线设备，但许多方面都已发生变化。带实体键盘的传统功能手机已被大屏触控智能手机取代。如今，电话通话在流量中所占比例微乎其微，绝大部分流量都是承载终端用户应用数据的分组。视频和音频流媒体、社交媒体、游戏以及许多其他应用，占据了流量的绝大部分。新服务不断涌现，巩固了智能手机在我们生活几乎方方面面的核心地位。与此同时，随着社会持续数字化，越来越多由非人类操作的设备接入网络，逐步构成物联网（IoT）。

无线网络在过去几十年里发生了巨大变化，促成了服务提供方面的这场革命，并使得能够适应随之而来的流量急剧增长成为可能。有许多促成因素，包括用于更快发射的新空中接口、用于增强可靠性的信道编码、用于去除冗余的改进源压缩，以及用于减少开销的更精简协议。信号处理是这些改进的核心，但它在多天线通信的发展中发挥了最大的作用。本文讲述了过去25年来主要信号处理进展如何将早期的多天线概念转变为主流技术的故事。因此，这个故事有些任意地从1998年开始。对最先进的信号处理技术的广泛介绍



(a) 典型的 2G 部署。

(b) 典型的第五代移动通信部署。

图1：(a) 1998年部署的第二代移动通信，由固定定向天线组成，每个信号广播至一个扇区。(b) 2023年部署的第五代移动通信，采用天线阵列，可利用多天线的三大主要优势：波束成形增益、空间分集和空间复用。

截至1998年的无线系统相关内容可见于 [2]，以及它与近期教材（例如[3]–[5] 的对比，揭示了这段时间所取得的巨大飞跃。

多天线通信基础知识

传统上，蜂窝网络站点中的基站（基站）配备了与基带单元（基带单元）相连的天线面板，由基带单元负责数字信号处理。这些面板通常又高又窄，包含多个垂直堆叠的辐射单元。通过从这些单元发射相同的信号，利用相长叠加形成一种垂直方向窄、水平方向宽的辐射方向图，以预定义的方式覆盖一片地面区域。如图所示。1（a），其中每个面板的覆盖区域称为一个小区扇区。

当前的基站中，面板已被具有更对称纵横比的天线阵列所取代，从而产生在水平方向和垂直方向上都可能较窄的辐射波束。每个天线单元发射的信号都由基带单元单独控制，而基带单元如今具备强得多的计算能力，可以在时间和频率两个维度上改变所产生波束的物理形状。图 1(b) 展示了这样一种配置，以及每个波束如何足够窄，从而能够瞄准特定用户。当这些阵列用于具有多条间隔较大的传播路径的传播环境时，每个辐射信号都会失去其定向波束形状，转而经过精细调整，使这些路径在预期接收机周围的一小片区域内相干叠加。

天线阵列带来三大类主要优势：

1)波束成形增益：发射波束聚焦于接收机，从而使更大比例的辐射能量到达接收机。同样，多个接收天线可以从选定方向收集更多能量，在接收端增强波束，并聚焦于发射机。总体波束成形增益与发射和接收阵列的规模成正比。

2)**空间分集：**信号通常会通过多条路径在发射机和接收机之间传播，而由此产生的信号副本可能会发生相消叠加。这会导致信号衰落，天线阵列可以通过同时观测多个衰落实现来减轻这一影响。

3) **空间复用：**多个信号可以在不同波束上同时传输，既可以传输给配备多根天线的单个用户，也可以像图 1(b) 所示传输给多个用户。这可以提高业务容量，或带来 复用增益，前提是能够抑制信号之间的干扰。

在上文以及下文中，波束成形增益被定义为接收机处信号功率的增加，但更细致的描述还应进一步包括降低对非预期用户的干扰以及来自非预期用户的干扰 [5, 第 5节。7]。经过精心设计，波束成形可以在增加信号能量与减少干扰之间实现最佳平衡。

1998年的技术前沿

早在1998年之前，人们就已充分认识到天线阵列的一些优势，但当时的技术成熟度等级与今天大不相同。马可尼本人曾于1901年利用波束成形实现跨大西洋无线通信，这一点广为人知。该实验依赖于一个 阵列天线，它通过将多个单元连接到同一信号发生器来实现波束指向性。阵列天线的几何结构决定了辐射信号相长叠加的方向。因此，波束方向是固定的，并在阵列建造和架设时确定。这就是图 1（a）中的2G天线被设计为以固定波束覆盖一个扇区的方式。

与阵列几何结构所决定的波束方向不同的波束方向，可以通过从所有阵元发射相同信号，但施加适当的相移来实现。这一概念最早由费迪南德·布劳恩于1902年通过实验观察到，并由此促成了相控阵技术的发展，该技术自第二次世界大战以来一直用于雷达。相移可以随时间变化，以扫描不同角度方向上的目标。相控阵应用于第二代移动通信的早期现场试验在 1996 [6]进行。将波束指向用户所在位置的可能性，为实现更强的方向性和更高的增益打开了大门，因为波束不再需要覆盖整个扇区。阵列天线与相控阵之间的差异如图2所示，其中还展示了数字天线阵列，这是第五代移动通信中的典型技术，其中每个阵元都连接到一个独立的信号发生器。

在数十年间不断改进相控阵以实现波束成形的同时，利用多个接收天线实现分集也逐渐普及。早在 20 世纪 30 年代，人们就提出了用于信号接收的空间分集 [7] ，其原理直观易懂：

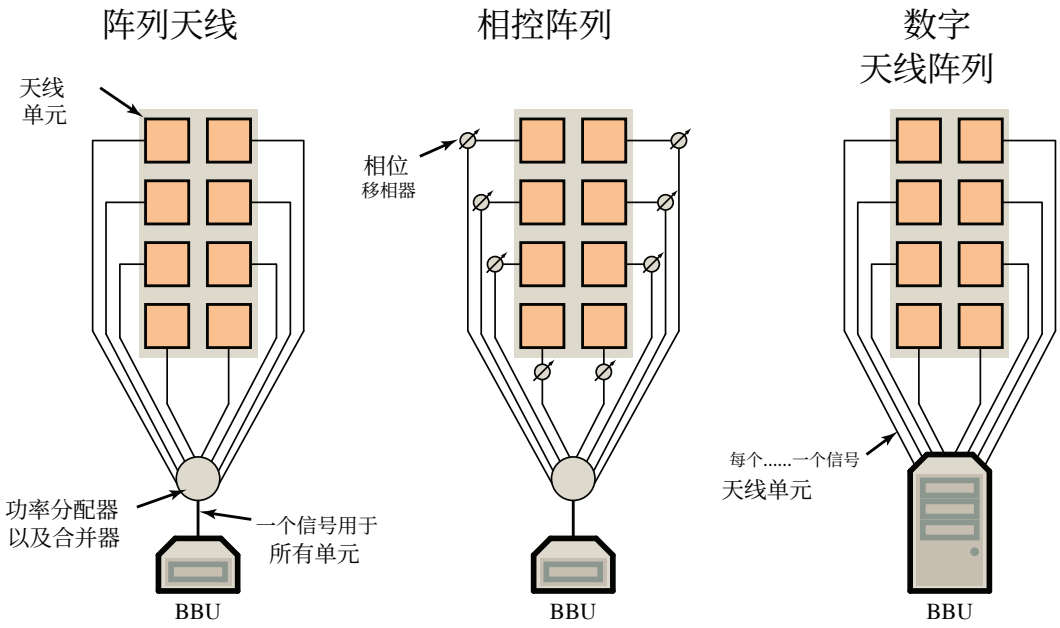


图 2：阵列主要分为三类经典类型：1) 生成固定波束的阵列天线；2) 依靠移相器控制波束方向的相控阵；3)能够完全控制从每个天线单元发射信号的数字天线阵列。

如果同一信号到达多个在物理位置上彼此分离的天线，那么多径传播环境不太可能同时在所有这些天线上造成相消叠加，进而导致信号衰落。通过对多个接收天线接收到的信号进行组合解码，通信的可靠性便能大幅提升。

无线网络当然必须提供从用户到基站的上行链路连接，以及从基站到用户的下行链路连接。因此，在两个链路方向上都需要分集，但发射分集直到20世纪90年代才出现 [8]。1998年，提出了用于两根天线的Alamouti空时分组码 [9] 并在此后不久发表了用于多根发射天线的空时编码更通用框架 [10]。其原理是重复传输一块数据信号符号，同时以预定方式改变空间指向性（例如，使用不同的天线）；接收机在一段时间间隔内收集观测值并对其进行解码。空时码经过精心设计，不仅能够实现解码，还能在高频谱效率（即每秒每赫兹频谱的比特数）、高分集和低复杂度之间取得令人满意的权衡。

总的来说，波束成形增益和空间分集在1998年就已为人所知，人们普遍认为这是天线阵列的两大主要好处：波束成形增益适用于相干阵列，与紧密的天线间距和清晰定义的到达方向相关

以及在天线间经历大体不相关衰落的阵列中，由更大的间距和丰富的多径环境所带来的分集增益。天线阵列的第三个、也是最终最强大的优势——空间复用——当时仍基本未受到关注。然而，其种子早已播撒在有关干扰感知波束成形 [11]以及将多个输入耦合到多个输出的线性信道通信概念的研究工作中 [12]。不同于涉及单个信号副本的波束成形和分集，这些空间复用的先驱技术需要同时在相同带宽上发射和接收不同的信号。尤其具有前瞻性的是采用两根正交极化天线进行发射和接收，随后在一篇研究中得到扩展，该研究提出了配备多天线发射机和接收机的系统，其中包含实现真正空间复用所需的许多要素 [13]。然而，直到1998年之后，所有这些要素才最终就位。

外部技术发展

过去几十年来，三大外部趋势极大地引导并影响了多天线技术的演进。

- 无线流量激增，并按照库珀定律每18个月翻一番；与此同时，新的用户行为和应用也从根本上改变了此类流量的性质。面向电话业务的高效网络必须支持大量同时进行的固定速率连接，而如今的数据网络则力求最大化每台用户设备的比特率（以支持特定应用），以及单位面积的比特率（以容纳大量设备）。

- 集成电路在性能上的指数级提升和尺寸的缩小，催生了片上系统，使无线电设备、存储器和处理器能够集成在一小块硅片上，并具备先进的信号处理能力。1998年，一个具有 $M = 2$ 或 $M = 4$个天线单元的数字天线阵列由 M 个外置天线单元、 M 个射频（RF）单元和一个基带单元组成；如今，5G基站可将 $M = 64$ 个天线单元和射频单元集成在一个机箱内。这一发展也使智能手机能够配备数字阵列，目前配备的是$M = 4$ 个天线单元。

- 信号波形的逐渐变化。当时有多种基于时分多址（TDMA）或码分多址（CDMA）的2G标准。3G的第一版恰好在1998年左右最终确定，完全基于CDMA，它在与竞争的正交频分多址（OFDMA）的较量中胜出。对于4G，最终转向了OFDMA，而第五代移动通信在评估并摒弃了少数几种替代方案后，保留了这种相同的波形。虽然所有波形原则上都兼容

天线阵列，但选择确实会对所需的信号处理算法产生根本性影响。

信号处理进展的五个关键领域

我们已确定了多天线技术从2G到第五代移动通信及以后演进过程中信号处理进展的五个阶段。背景、新解决方案和具体见解将在以下各节中阐述。

从空间分集到空间复用

有人可能会认为，一直追溯到马可尼，波束成形的动机是扩展覆盖范围。反过来，分集的动机是提高可靠性。到1998年，3G之前无线电频谱成本的爆炸性增长带来了一个新的强大需求：提高频谱效率。向高比特率用户应用的转变进一步放大了这一趋势。最大化频谱效率的天线阵列操作模式是空间复用，因此在1998年之后，氛围已经为它最终脱颖而出做好了准备。

空间复用的前提是多输入多输出（MIMO）通信信道，其中每个输入/输出指的是数字天线阵列中的天线单元。有两种MIMO类别：单用户MIMO涉及多天线基站和多天线用户设备，而多用户MIMO包括多天线基站和多个用户设备。

可以说，单用户MIMO的主要催化因素是 [14], 从信息论视角出发，旨在设计完美收发机的那项工作。以发射端和接收端数字天线阵列为起点，并且不预先规定如何使用它们，研究发现，如果每个阵列中的天线单元表现出不相关衰落，最优策略就是让每个天线单元辐射一个独立数据信号。这一做法极具创新性，因为它寻求的是利用而非抵消多径传播；正是多径的存在，使接收机能够在每个接收天线上观测到发射信号的不同线性组合，从而可以从中解析出这些发射信号。能够进行空间复用的信号数量随后受限于发射天线单元数量和接收天线单元数量中的较小者。在后续工作中，有人提出了一种用于实现这种空间复用的特定架构，即所谓的分层架构；其突出之处在于，它可以使用现成的编码器和解码器构建，并且不要求发射机了解任何有关信道的信息[15]。更多研究成果逐步巩固了其理论基础 [16]。特别是，从每个天线单元发射一个并发信号的想法，被推广到了

同时从所有阵元发射并发波束。相控阵无法实现这种空间复用，因为它们一次只能生成一个波束，因此数字天线阵列是绝对必要的。

多用户MIMO可追溯到同时从多个用户进行上行接收的信号处理概念 [17] 以及向不同角度方向的用户进行同时下行波束成形 [18]。在此，可空间复用的信号数量不受每个用户天线单元数量的限制，而是受所有用户的限制；即使每个用户只有一个天线单元，也可以向/从每个用户空间复用一个信号。这一主要优势的代价是基站必须仔细安排发射和接收波束，使每个波束与其目标用户的多径特性相匹配，并且它们之间的干扰最小，如图2(b)所示。由此，每个用户都可以连续地在整个系统带宽上传输，而不仅仅是在一个时隙和/或频率子带中，这体现了空间复用的好处。多用户MIMO是经典空分多址 (SDMA) 在多径环境中的推广，用户在空间而非时间或频率上共享信道。有趣的是，SDMA概念比单用户MIMO早了20多年以上 [19],这表明在无线网络中，建立许多同时用户连接长期以来被认为比实现每个连接的高数据速率更重要。

MIMO无论采用单用户还是多用户方式，其潜力都引发了一连串连锁反应，并迅速席卷学术界和业界，到21 世纪初带来了极大的振奋。特别是蜂窝通信标准化机构，即第三代合作伙伴计划 (3GPP)，在3G 后期版本中以有限方式采用了 MIMO，随后从4G开始将其作为设计的重要组成部分。Wi-Fi 中的采用速度更快，首个包含MIMO的版本于2007 年获得认证，并得到包括笔记本电脑、平板电脑和智能手机在内的大量设备的支持。

MIMO 利用图中所示的三种优势维度：3：波束成形增益、空间分集和空间复用。随着时间推移，人们逐渐清楚地认识到这些优势之间的关系。针对不同的阵列配置和信道条件，已经确定了基本的权衡关系。

- 波束成形是空间复用的一种特殊情况，即向单个用户传输单束波束。当信噪比 (SNR) 较低时，这实际上是最优策略；此时最大化信号能量至关重要，而最佳方案是将所有辐射能量集中在最强的波束上。相比之下，在高信噪比下，能量充足，可以分散到多束波束上，直至激活信道和天线数量所允许的尽可能多的波束为最优。这由图3中的蓝色平面表示。

- 空间分集，通常以独立衰落信号副本的数量来粗略量化，

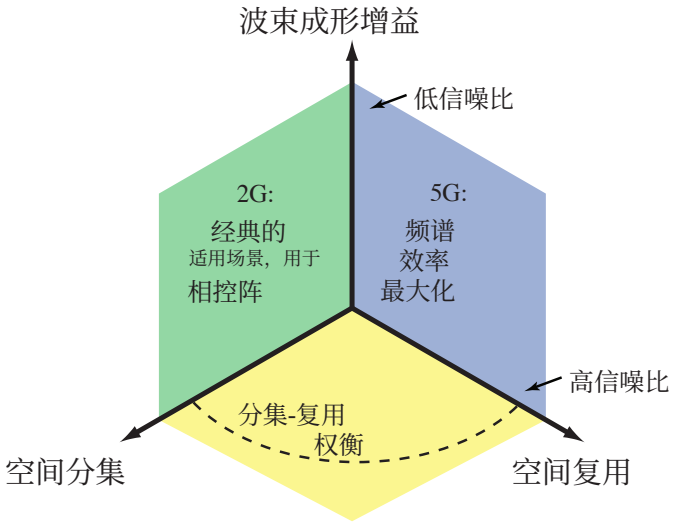


图3：过去几十年中，多天线通信的三种效益维度不断发展。从第二代移动通信到第五代移动通信，系统已从绿色平面转向蓝色平面；也就是说，空间分集逐渐被空间复用所取代。由于空间复用需要较高的信噪比才能在实际中发挥作用，因此在低信噪比下，波束成形增益仍然至关重要。

和空间复用，也就是说并发波束的数量，无法同时达到最大值。这两个量实际上受到一种权衡关系的制约 [20]。在这种分集-复用权衡的极端点上，其中一个量达到最大值，而另一个则处于最小值。在权衡曲线的中间点，各种组合都是可行的，该曲线示意绘制于图3的黄色平面上。

如前所述，波束成形与空间复用之间的选择由信噪比决定，因此由其底层参数（例如发射功率、信道衰减、噪声功率）决定。反过来，分集与复用的混合取决于优先考虑的是提高可靠性还是频谱效率。然而，工作点应整体选择，多年来这导致混合向更少分集和更多复用转变。事实上，随着连续系统代际跨越越来越宽的带宽，越来越多的分集已在频域中获得。额外分集的回报迅速饱和，因此对空间分集的需求已经减弱 [21]。这当然不适用于窄带控制信道或低功率短包物联网通信，在这些情况下空间分集仍然重要，但对于承载蜂窝网络中大部分流量的用户数据信道而言确实如此。

从空间复用到大规模 *MIMO*

MIMO 的基础是在完美信道状态信息（CSI）和丰富散射的前提下开发的，表明通过部署足够多的天线并同时服务许多用户，可以实现任意高的频谱效率。然而，当该技术首次商业化时，实际挑战变得明显。单用户 MIMO 中的空间复用能力常常受到有限散射的限制，而多用户 MIMO 则受不完美 CSI 的制约。大规模 MIMO 是一种源自 [22], 的新型多用户 MIMO，于 2010 年代开发以解决这些问题，现已成为第五代移动通信的核心。

大规模 MIMO 的新特点如下。首先，它要求基站天线的数量远多于空间复用用户的数量。这种设计选择使波束相对较窄（例如，聚焦于目标接收机周围的一小片区域），因此，聚焦于不同用户的波束之间可能几乎没有重叠。此外，凭借这些有利的 条件，即使存在少量干扰，也可以通过低复杂度线性信号处理加以抑制；例如，采用正则化迫零，微调每个波束的聚焦区域，在获得较强波束成形增益的同时降低干扰 [23]。同时，同样由于基站天线数量过多，这些波束提供的有效信道趋于稳定，也就是说，它们变得非常稳定，几乎不受衰落波动的影响。

其次，大规模 MIMO 专为高效利用资源的 CSI 获取而设计。主要的估计原理是从每个天线单元分别发射预定义的导频信号，然后根据接收单元对这些导频信号的观测来估计信道系数。大规模 MIMO 采用时分双工（TDD），在上行链路和下行链路之间交替使用相同的带宽。由于互易性，两种方向上的信道相同，因此只需估计一个方向上的信道系数即可。具体而言，上行链路和下行链路所需的 CSI 均通过上行导频获取。因此，所需的导频资源取决于复用用户的数量，而与基站天线的数量无关。相比之下，许多早期的多用户MIMO商用实现采用频分双工（FDD），上下行链路的信道完全不同；也有一些采用不利用互易性的 TDD。此时，下行链路操作要求基站发射与其天线数量相同的导频，但在某些特定传播场景中，信道可以用少量角度参数化，因此不在此限。此外，每个用户都需要量化其信道估计值并反馈给基站。在典型的第五代移动通信设置中，若有 $M = 64$ 根基站天线，为 $K = 8$ 个用户提供空间复用服务，则 FDD 方案所需的导频数量将增加至 $M/K = 8$ 倍，CSI反馈量也会相应增加。

在具有众多复杂传播环境的情况下，TDD 操作尤其有用。

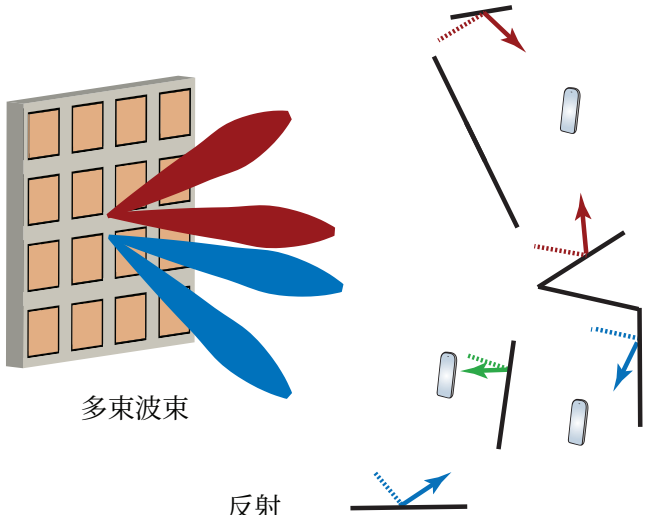


图 4：传播信道由大量镜面反射或漫反射物体构成。要以足够的精度估计此类信道，从而实现多用户MIMO通信，极具挑战性；但大规模 MIMO 中的 TDD CSI 估计方法能够应对这一问题。

每位用户的路径，例如图 4 所示的路径，其中最优的下行链路发射会将用户的信号能量分散到多个方向，以匹配反射物体。通过上行导频进行CSI获取可以自动捕捉这些细微特征，无需任何先验信道知识或阵列校准。在 FDD运行中，除了需要多得多的导频资源外，反馈量化还会丢失关键的信道细节。在蜂窝网络中，必须谨慎地在各小区间复用导频信号，以避免 导频污染现象，即基站会无意中朝向邻近小区中共享导频的用户进行波束成形。这在TDD 运行中尤其令人担忧，因为上行链路估计误差也会影响下行链路。在过去十年中，人们开发了大量信号处理和资源分配方案来缓解导频污染 [3], [4]。

大规模 MIMO 为实际信号处理设计提供了坚实的基础。虽然到 2000 年代已经存在针对多用户MIMO 的复杂信息论 [24], 它主要局限于具有完美 CSI 的场景。由于 CSI 质量是多用户MIMO 性能的主要限制因素，这限制了现有理论的实际实用性。得益于对线性信号处理的依赖，大规模 MIMO 分析成功处理了不完美 CSI 和硬件不完善，从而得出了严谨且数学上简洁的频谱效率表达式，这些表达式不仅能准确预测实际性能，而且可作为系统优化的有效工具（例如，导频分配、功率控制和波束成形）。大规模 MIMO 理论不仅将多用户 MIMO 变成了实际可行的技术；其分析优雅性也

拓展了信息论在无线通信领域的教学方式 [3], [4]。

如前所述，TDD模式下的上行链路-下行链路互易性对于大规模 MIMO至关重要。只要信道冲激响应保持不变，空中传播就具有互易性；也就是说，双工过程须在信道相干时间内完成。然而，收发器硬件在发射和接收之间通常并不具备互易性，例如本地振荡器不匹配就会导致这种情况。此类硬件非互易性需要通过校准程序来解决：该程序通过不时进行相互测量，使阵列内的天线实现相位同步。

如今，5G基站几乎全部采用时分双工频段下的大规模 MIMO配置，其中最常见的是配备 $M = 32$ 或 $M = 64$ 根天线的阵列。起初，人们担心信号处理会导致能耗过高，但后来这一担忧消除了。如今已有专用片上系统，能够以合理的能耗为整个基带单元（包括大规模 MIMO）实现高效的信号处理算法。

追求更高频段下的更大带宽

长期以来，比特率一直是无线技术用户最熟悉的性能指标，因此无线技术不断演进，以支持更高的比特率。单个设备所能享有的比特率等于频谱效率与频谱带宽的乘积。因此，除了通过单用户MIMO提升比特率外，随着新频段的分配，比特率也逐步提高。第二代移动通信网络通常可使用载波频率约为20 MHz的频谱，载波频率约为1 GHz；而当前的第五代移动通信网络主要覆盖100 MHz的3.5-GHz频段，标准支持的带宽则超过500 MHz。

无线电频谱是一种有限的自然资源，由众多技术共享，其中也包括本文所讨论的民用无线通信领域以外的技术。虽然一些低于6GHz的频段已从过时技术重新分配给蜂窝网络，但对新增带宽的追求不可避免地促使系统转向越来越高的频率。尤其是，名义上从30 GHz开始的毫米波频段，如今已纳入第五代移动通信。少数电信运营商已推出第一代毫米波技术，而大多数运营商则在等待硬件成熟，并等待3.5-GHz频段变得拥堵。

自由空间中点源辐射信号的场强随距离增加而衰减，这种衰减与频率无关。随后，从此类电场中接收的功率与接收机的接收孔径成正比；由于天线单元的尺寸随波长缩短而减小，因此载波频率升高时，需要增加天线单元以维持所需的接收孔径。此外，蜂窝网络中的信道条件会持续变得更具挑战性。

随着频率升高，挑战也随之加剧，因为散射和衍射减弱，穿透损耗增大，而这些都需要波束成形增益。因此，多天线技术在毫米波频率下至关重要。

与此同时，实现变得困难，这不仅是因为增加了硬件组件以及数字信号处理中的巨大维度。当转向更高频率和更宽的带宽，即更快的采样率时，功率放大器效率以及模数转换器（ADC）的功耗也是需要进一步关注的问题。信号处理领域探索了两种主要方式来应对硬件和算法的复杂性 [25]。

第一种选择是通过将收发器设计为相控阵和数字天线阵列的混合来减少射频单元的数量，特别是转换器。由此产生的混合模拟-数字天线阵列如图所示。5，其中每一列是一个相控子阵列，分别连接到基带单元，以便可以发送和接收不同的信号。每个子阵列可以形成一个单一波束，然后可以通过这些波束的不同线性组合应用空间复用。如果信道具有少量传播路径，每个子阵列可以将波束聚焦在其中一条路径上，并且可以用更少的硬件组件实现数字天线阵列的通信性能。在图中所示的多径场景中4，五个不同的波束方向足以有效通信。混合天线阵列可以采取其他形式，但通常涉及半模拟波束成形实现，其天线单元多于数字端口 [26]。放弃数字天线阵列范式需要付出若干代价。只能传输与数字端口数量相同的波束，并且在宽带系统中波束成形保真度受损，因为所有子载波必须使用相同波束的组合。信道估计变得更加复杂，因为每个相控阵必须扫描与其单元数量相同的波束方向，以激励所有信道维度。

降低射频单元数量的另一种方法是采用降低 ADC 分辨率的数字天线阵列，如图 5 所示。ADC 的能耗随分辨率呈指数增长，因此，将传统的每样本 15 位降至例如每样本 5 位，就有可能大幅降低能耗。然而，由于包含 M 个阵元、配备 b位 ADC 的阵列在每个采样周期内总共采集 bM 位数据，即使 b 较小，ADC 的总位数仍可能相当可观，这也解释了为何仍能保持较高的频谱效率 [27]。上行链路大规模 MIMO 的极端情形（$b = 1$）恰好在分析上易于处理 [28], ，这推动了能够补偿由此产生的量化失真的信号处理算法的发展。其下行链路对应方案则采用低分辨率数模转换器（DAC） [29]。

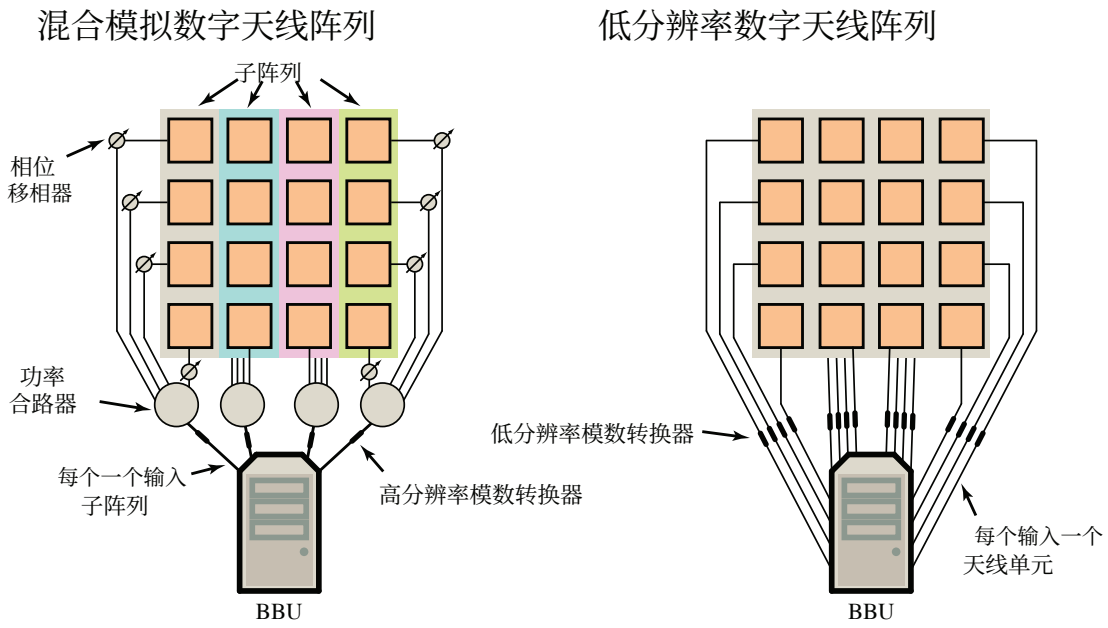


图5：传统数字天线阵列在毫米波频率实现时会产生高能耗。有两种主要方法可以规避这一点：1）使用由多个相控子阵列组成的混合模拟-数字天线阵列来减少组件数量；2）设计具有简化组件的数字天线阵列，例如低分辨率模数转换器。

降低位分辨率同样适用于混合天线阵列；如果针对当前通信任务，对量化前的模拟信号合成和数字后处理进行适当优化，信号内容就更适合采用低比特表示 [30]。

初始的5G毫米波产品基于混合阵列，但有迹象表明低分辨率方法最终可能成为首选解决方案 [31]。

降维的进一步机遇

朝着具有更多天线单元和更宽带宽的阵列联合演进，这需要更高的采样率，因此必须避免过度设计收发器。作为扩大传统小型数字阵列的替代方案，图 5 展示了两种在增加天线单元数量的同时降低每个单元硬件复杂度的方法。这两种方法都利用了大规模 MIMO 的理念，即天线单元数量多于多路复用波束，这使得可以降低波束成形精确度，因为波束非常窄，用户间干扰本来就很低。通过利用信道结构（例如角度、频率和时间域中的稀疏性），或利用系统设计要解决的特定任务，可以进一步降维。

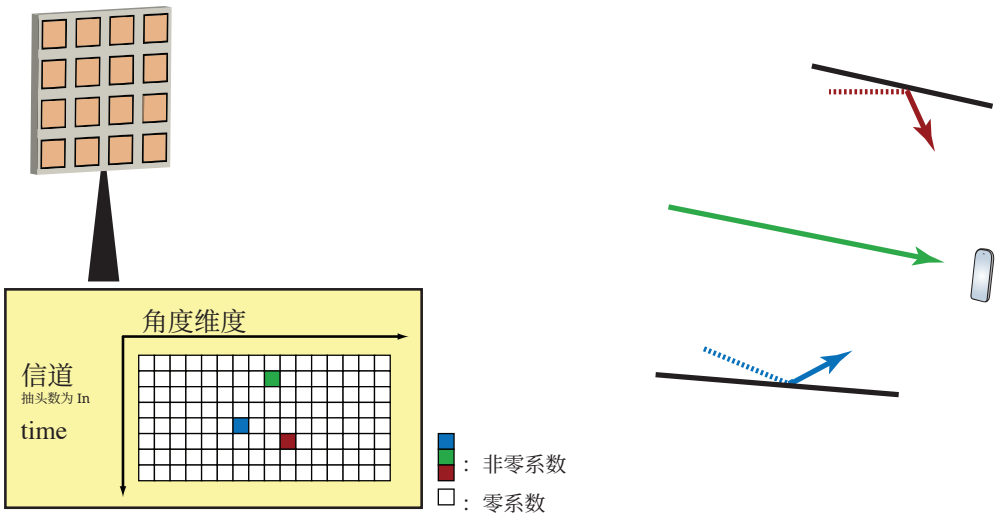


图 6：在具有较大频谱带宽和大量天线的 mmWave 系统中，信道可能在联合时间-角度域中呈现稀疏性。在此示例中，存在 $N = 3$ 条路径，它们在时间和角度上均各不相同。可以利用稀疏冲激响应结合压缩感知技术来降低采样率，从而降低功耗。

除了与分辨率呈指数关系外，ADC 的能耗还与采样率成正比。已提出大量基于亚奈奎斯特采样和压缩感知的想法，通过利用许多场景中存在的各种信道结构来降低采样率 [32]。在具有少量 N 传播路径的毫米波信道中，无论带宽如何，信道冲激响应中可能只有大约 N 个非零抽头。这种信道响应中的时域稀疏性可用于降低采样率 [33]。这些 N 路径在角度域中也可能是不同的，因为基站通常部署在环境杂波之上的高处；反射仅发生在每个用户周围的局部区域，在远处的基站处形成较小的角度扩展。图 6 说明了具有 $N = 3$ 条不同路径的此类场景，支持三个多路复用信号。视距路径具有明显较短的延迟，而其余两条路径具有相似的延迟，但在角度域中可清晰区分。联合信道稀疏性由时间-角度矩阵中的三个彩色条目表示，绝大多数条目不包含传播路径。将这些稀疏性形式与现代压缩感知工具相结合，可大幅降低采样率。

许多数据服务呈现间歇性活动模式；在与某个基站关联的数千台设备中，每个时隙内只有一小部分设备需要传输数据。信号处理可使这些设备无需先行接入流程即可高效传输。关键在于为每台设备分配一个唯一但非正交的导频序列，并

利用用户域中的稀疏性，以及通过天线阵列获得的大量空间采样，实现用户识别和信道估计 [34]。联合用户与数据检测问题也已采用压缩感知方法进行研究 [35]。

商业大规模 MIMO 产品已经利用了一些基本的信道稀疏性；例如，许多天线上的接收信号可能会被转换到相同数量的角度维度。功率很小的维度会在数字上行链路处理的早期阶段被丢弃，以缩小剩余计算的维度。然而，更激进的压缩感知解决方案尚未付诸实施。

基于机器学习的算法改进

当代信号处理中最活跃的领域之一是机器学习 (ML)。虽然这是一门已有数十年历史的学科，但近年来大量数据和处理能力的可用性大大增强了其潜力，将许多信号处理任务的实现从更传统的模型驱动算法转变为数据驱动算法。这种转变也正在无线通信信号处理的背景下发生，传统上该领域非常依赖（且成功地）基于模型。有几个趋势正在推动这种转变。一个趋势是，随着大量物联网和机器类型连接与人类宽带连接共存，无线网络流量变得越来越难以准确建模，从而使网络运行难以优化。另一个趋势是，天线阵列和其他传感器在智能手机和其他联网设备上变得无处不在，因此可用于学习的数据量正在急剧膨胀。第三个趋势是，分布在整个无线网络中的处理能力正在迅速增长，从而催生了雾计算和边缘计算等范式。

机器学习与通信在无线网络优化领域正日益融合。这是机器学习非常自然的应用场景，因为此类网络的运行涉及大量机器学习擅长处理的任务，包括信道估计、信号检测和数据解码等推理任务，以及路由、接入控制和资源分配等决策任务。基于机器学习的解决方案能够捕捉现有算法所依赖模型中被忽略的实际特征。然而，为确保机器学习算法能够在现有方法的基础上实现改进，务必审慎地启动训练过程。

该模型辅助机器学习范式提供了一种结构化方式，将信号处理领域的经典知识迁移到新的机器学习算法中 [36]。图7 展示了如何将现有的迭代算法转化为增强型机器学习算法。现有算法获取初始输入信号，并对其进行迭代处理/更新，直到预定义的终止准则为止

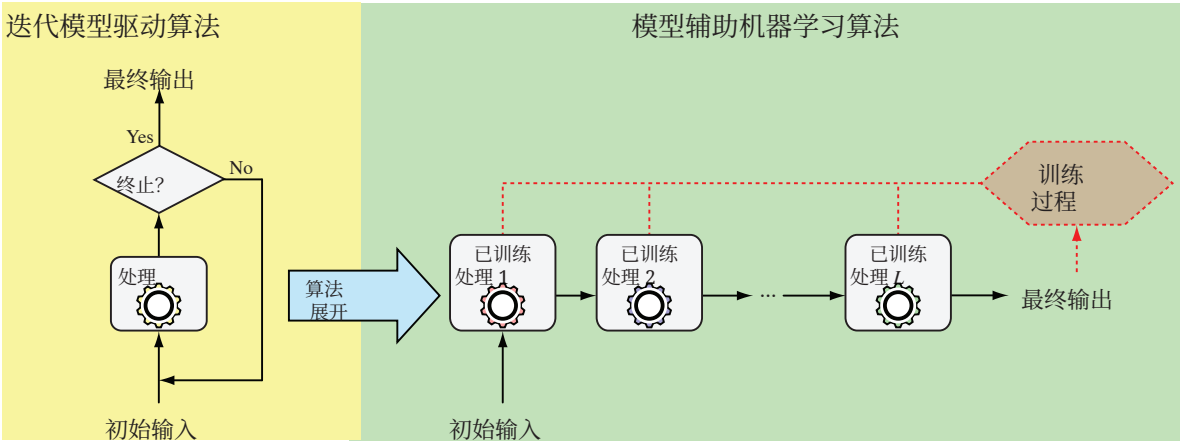


图 7：许多用于多天线通信优化的传统模型算法都是迭代算法。假设其中一种算法可以表示为一个迭代处理循环，并持续运行直至满足终止准则。通过算法展开可以开发出增强型机器学习算法；也就是说，将迭代过程写成 L 个独立的处理层，并通过数据辅助训练过程对这些层进行微调。

满足要求，此时即可获得最终输出。具体处理通常通过基于模型的算法设计来实现。算法并非表示为循环， L 算法的迭代可以表示为一系列 L 相同的处理层。如果采用数据驱动的训练过程对这些处理层进行微调，使其不再必须完全相同，那么得到的便是一种机器学习算法，并且可以保证其性能优于原始的基于模型的算法。这一过程称为算法展开或深度展开，近年来已被用于增强各种多天线任务，包括信号检测 [37]以及下行链路多用户MIMO的波束成形优化 [38]。

展望未来

过去25年来，多天线技术已从用于波束成形和分集合并的基础设计，发展成为主流技术，通过大规模空间复用来提升第五代移动通信网络的容量。如今，MIMO技术已应用于所有上市的智能手机和基站。快速且高效的信号处理算法推动了这一飞跃，目前也正助力低功耗第五代移动通信毫米波收发器的发展：这类收发器以数字处理取代高分辨率硬件组件。近期的理论突破，包括基于机器学习的算法，必将继续推动这项技术进步。

的确，我们预计多天线通信的发展历程将继续延伸。数据流量的持续增长，唯有通过部署越来越多的天线和越来越大的带宽才能满足。大规模 MIMO理念要求基站天线的数量 M 与活跃用户的数量 K成比例增长。由于正则化迫零等算法的复杂度与 MK^2 [4] 成正比，因此 K 和 M 都线性增长意味着复杂度将呈立方增长。此外，当带宽超过1 GHz时，采样率将接近现有处理器的时钟速度，使得实现更具挑战性。前文所述的压缩感知算法或许适合应对这些挑战，但信号处理领域仍有望取得许多新的进展。

当向基站添加越来越多的天线时，实际的尺寸和重量限制可能使得新的部署原则成为必要，超越塔上盒子的范式。一种有前景的方法是将天线分布在多个物理位置，同时保留相干发射和接收处理，这一概念源于小区协作和网络 MIMO 思想，以及远程射频单元的概念，其当前体现称为无小区大规模 MIMO [39]。除了更强的波束成形增益外，分布式天线部署可以提供改进的空间复用能力以及对环境中大型物体阴影的宏观分集。将基带计算从基站站点转移到边缘云计算机的当前趋势将促进这种部署方法的采用。

在智能手机统治无线生态系统二十年之后，其他设备如扩展现实眼镜预计将占据中心舞台。新服务将会出现，用户期望无线网络提供的比特率、延迟和可靠性将有更新的标准。其他性能指标可能会出现，以指导未来的技术发展，特别是与可持续性、环境影响和部署成本相关，以及数字化与远离数字化的世界地区之间的数字鸿沟。

除了本文所涵盖的信号处理进展之外，还有两个新兴研究课题建立在多天线技术之上。第一个是集成通信与传感 [40],，它探索如何同时使用大规模天线阵列进行精确的雷达传感、定位和通信。部署大量天线用于通信目的，似乎自然可以成为其他受益于无线测量的应用的催化剂。另一个相关研究方向是智能表面 [41], ，其信号反射特性可以通过具有可编程阻抗模式的超材料来控制。这些可重构智能表面提供了一种被动波束成形，特别有助于改善无线信道上的传播条件。

作者

Emil Bj¨ornson （emilbjo@kth.se）是瑞典斯德哥尔摩皇家理工学院（KTH）无线通信教授。他撰写了三本关于MIMO技术的教科书，获得了23000次引用，并发布了大量仿真代码。他获得了2018年和2022年IEEE马可尼无线通信论文奖、2019年EURASIP早期职业奖、2019年IEEE通信学会Fred W. Ellersick奖、2019年IEEE信号处理杂志最佳专栏奖、2020年皮埃尔-西蒙·拉普拉斯早期职业技术成就奖、2020年CTTC早期成就奖以及2021年IEEE ComSoc RCC早期成就奖。他是IEEE会士。

Yonina C. Eldar （yonina.eldar@weizmann.ac.il）是以色列雷霍沃特魏茨曼科学研究所数学与计算机科学系教授，她在该所领导生物医学工程与信号处理中心。她还是麻省理工学院和博德研究所的访问教授，以及杜克大学的兼职教授，并曾担任斯坦福大学访问教授。她是以色列科学与人文学院院士和 EURASIP 会士。她因研究和教学卓越获得多项奖项，包括 IEEE 信号处理学会技术成就奖、IEEE/AESS Fred Nathanson 纪念雷达奖以及 IEEE Kiyo Tomiyasu 奖。她领导以色列高等教育机构促进性别公平委员会。她是 IEEE 会士。

Erik G. Larsson （erik.g.larsson@liu.se）是瑞典林雪平大学教授。他合著了教材大规模 *MIMO* 基础（剑桥大学出版社，2016 年）。他获得的奖项包括 2015 年 IEEE 通信学会 Stephen O. Rice 通信理论奖、2017 年 IEEE 通信学会 Leonard G. Abraham 奖、2018 年 IEEE 通信学会最佳教程论文奖以及 2019 年 IEEE 通信学会 Fred W. Ellersick 奖。他的研究兴趣包括无线通信、统计信号处理和网络。他是 IEEE 会士。

Angel Lozano (angel.lozano@upf.edu)是庞培法布拉大学的教授。他于1999年获得斯坦福大学博士学位，1999年至2008年间任职于贝尔实验室（朗讯科技，现为诺基亚），并于2005年至2008年间担任哥伦比亚大学副教授。他的论文曾获得多项奖项，包括2009年StephenO. Rice奖、2016年Fred W. Ellersick奖，以及2016年通信学会与信息论学会联合论文奖。他还曾获得ERC高级资助，资助期为2016年至2021年，并于2017年入选高被引作者。他是剑桥大学出版社于2019年出版的教材《MIMO通信基础》的合著者。他是IEEE会士。

H. Vincent Poor (poor@princeton.edu) 是普林斯顿大学的迈克尔·亨利·斯特雷特大学教授，其研究兴趣包括信息论、机器学习和网络

科学，以及它们在无线网络、能源系统和相关领域的应用。他在这些领域的出版物中包括最近的一本书机器学习与无线通信(剑桥大学出版社,2022)。他是美国国家工程院和科学院的成员，并于2017年获得IEEE亚历山大·格雷厄姆·贝尔奖章。他是IEEE终身会士。

参考文献

[1] 国际电信联盟，《世界电信发展报告：移动蜂窝》，技术报告第 5 号，1999年。[2] H. V. Poor 和 G. W. Wornell，《无线通信：信号处理视角》，Prentice-Hall 信号处理丛书。Prentice Hall PTR，1998年。[3] 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年。[4] E. Bj¨ornson、J. Hoydis 和 L. Sanguinetti，“大规模 MIMO 网络：频谱效率、能效与硬件效率”，《信号处理基础与趋势》，第 11 卷，第 3–4 期，第 154–655 页，2017年。[5] R. W. Heath Jr. 和 A. Lozano，《MIMO 通信基础》。英国剑桥：剑桥大学出版社，2018年。[6] S. Anderson、U. Forssen、J. Karlsson、T. Witzschel、P. Fischer 和 A. Krug，“采用自适应天线的爱立信/曼内斯曼 GSM 现场试验”，载于《先进 TDMA 技术与应用 IEE 专题讨论会》，1996年。[7] H. O. Peterson、H. H. Beverage 和 J. B. Moore，“R.C.A. Communications, Inc. 的分集电话接收系统”，《IRE》，第 19 卷，第 4 期，第 562–584 页，1931年。[8] A. Wittneben，“数字同步广播的基站调制分集”，载于《IEEE 车辆技术会议（VTC）》，1991年，第 848–853 页。[9] S. Alamouti，“一种简单的无线通信发射分集技术”，《IEEE通信领域精选期刊》，第 16 卷，第 8 期，第 1451–1458 页，1998年。[10] V. Tarokh、N. Seshadri 和 A. Calderbank，“高速率无线通信的空时码：性能准则与码构造”，《IEEE信息论汇刊》，第 44 卷，第 2 期，第 744–765 页，1998年。[11] J. H. Winters，“存在同信道干扰时数字移动无线电中的最优合并”，《IEEE通信领域精选期刊》，第 2 卷，第 4 期，第 528–539 页，1984年。[12] J. Salz，“通过交叉耦合线性信道的数字传输”，《贝尔实验室技术期刊》，第 64 卷，第 6 期，第 1147–1159 页，1985年。[13] J. Winters，“瑞利衰落环境中具有分集的无线通信系统容量”，《IEEE通信领域精选期刊》，第 5 卷，第 5 期，第 871–878 页，1987年。[14] G. J. Foschini 和 M. J. Gans，“多天线用于衰落环境时无线通信的极限”，《无线个人通信》，第 6 卷，第 3 期，第 311–335 页，1998年。[15] P. W. Wolniansky、G. J. Foschini、G. D. Golden 和 R. A. Valenzuela，“V-BLAST：在丰富散射无线信道上实现极高速率的数据架构”，《URSI国际信号、系统与电子专题讨论会（ISSSE）》，第 295–300 页，1998年9月。[16] E. Telatar，“多天线高斯信道的容量”，《欧洲电信汇刊》，第 10 卷，第 6 期，第 585–595 页，1999年。[17] J. H. Winters，“多用户室内无线系统的最优合并”，《IEEE通信汇刊》，第 35 卷，第 11 期，第 1222–1230 页，1987年。

[18] S. C. Swales、M. A. Beach、D.J. Edwards和J.P.McGeehan，“蜂窝陆地移动无线电系统中多波束自适应基站天线的性能增强”，IEEE《车辆技术汇刊》，第39 卷，第 1 期，第56–67 页，1990 年。[19] Y. Tsuji和Y. Tada，“SDMA/TDMA 卫星通信系统同步突发的发射相位控制系统”，美国专利，1974 年，3995111。[20] L.Zheng 和D.Tse，“分集与复用：多天线信道中的基本权衡”，IEEE《信息论汇刊》，第49卷，第 5 期，第 1073–1096 页，2003 年。[21] A. Lozano 和N. Jindal，“现代 MIMO 系统中的发射分集与空间复用”，IEEE《无线通信汇刊》，第9卷，第1 期，第 186–197 页，2010年。[22] T. L. Marzetta，“具有无限数量基站天线的非协作蜂窝无线通信”，IEEE《无线通信汇刊》，第9卷，第11 期，第3590–3600 页，2010 年。[23] M.Joham、W. Utschick 和 J.Nossek，“MIMO 通信系统中的线性发射处理”，IEEE《信号处理汇刊》，第53卷，第 8 期，第 2700–2712 页，2005 年。[24] H. Weingarten、Y.Steinberg 和 S. S.Shamai，“高斯多输入多输出广播信道的容量区域”，IEEE《信息论汇刊》，第52卷，第 9期，第 3936–3964 页，2006 年。[25] R. W. Heath、N. Gonzalez-Prelcic、S. Rangan、W.Roh 和A. M. Sayeed，“毫米波MIMO 系统信号处理技术概述”，IEEE《信号处理精选主题期刊》，第10卷，第3 期，第 436–453 页，2016 年。[26] R. M´endez-Rial、C.Rusu、N. Gonz´alez-Prelcic、A.Alkhateeb和R. W.Heath，“毫米波通信的混合MIMO 架构：移相器还是开关？”，IEEEAccess，第 4 卷，第 247–267 页，2016年。[27] C. Moll´en、J. Choi、E. G. Larsson和R. W.Heath，“粗量化大规模MIMO 的可达上行链路速率”，载于IEEE国际声学、语音与信号处理会议（ICASSP），2017 年。[28] Y. Li、C. Tao、G. Seco-Granados、A. Mezghani、A.L. Swindlehurst 和 L. Liu，“单比特大规模 MIMO 系统的信道估计与性能分析”，IEEE《信号处理汇刊》，第 65卷，第15 期，第 4075–4089 页，2017年。[29] S. Jacobsson、G.Durisi、M. Coldrey和C.Studer，“大规模 MU-MIMO-OFDM下行链路的低分辨率 DAC线性预编码”，IEEE《无线通信汇刊》，第18 卷，第3 期，第 1595–1609页，2019 年。[30] N.Shlezinger 和Y. C. Eldar，“面向任务的量化及其在 MIMO接收机中的应用”，《通信与信息系统》，第 20 卷，第131–162 页，2020年。[31] K. Roth、H.Pirzadeh、A. L. Swindlehurst 和J. A. Nossek，“面向多个用户和不完美 CSI的低分辨率 ADC混合波束成形与数字波束成形比较”，IEEE《信号处理精选主题期刊》，第 12卷，第3期，第 484–498页，2018年 6 月。[32] Y. C.Eldar，《采样理论：超越带限系统》。剑桥大学出版社，2015年。[33] Z. Gao、L. Dai、S. Han、C.-L.I、Z. Wang 和 L. Hanzo，“面向下一代无线通信的压缩感知技术”，IEEE《无线通信》，第25卷，第 3 期，第 144–153 页，2018 年。[34] L. Liu、E. G. Larsson、W. Yu、P. Popovski、C. Stefanovic 和E. deCarvalho，“免授权大规模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 5期，第 88–99 页，2018 年。[35] J.Yuan、Q. He、M. Matthaiou、T.Q. S. Quek和 S. Jin，“混合 ADC 分布式大规模 MIMO 中面向物联网大规模连接的研究”，IEEE《物联网期刊》，第7 卷，第3 期，第 1841–1856 页，2020 年。[36] V.Monga、Y. Li 和 Y.C. Eldar，“算法展开：用于信号与图像处理的可解释、高效深度学习”，IEEE《信号处理杂志》，第 38卷，第 2 期，第18–44 页，2021 年。[37] N. Samuel、T.Diskin 和A. Wiesel，“学习检测”，IEEE《信号处理汇刊》，第 67 卷，第 10 期，第 2554–2564 页，2019 年。

[38] L. Pellaco, M. Bengtsson, and J. Jald´en, “加权 MMSE 波束成形算法的无矩阵求逆深度展开，”*IEEE* 通信学会开放期刊, 卷. 3, 页. 65–81, 2022.[39] ¨O. T. Demir, E. Bj¨ornson, and L. Sanguinetti, “以用户为中心的无小区大规模 MIMO 基础,” 信号处理基础与趋势, 卷. 14, 期. 3-4, 页. 162–472, 2021.[40] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, “通信与雷达联合感知的信号处理技术综述,” *IEEE* 信号处理精选主题期刊, 卷. 15, 期. 6, 页. 1295–1315, 2021.[41] E. Bj¨ornson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, “可重构智能表面：具有无线应用的信号处理视角,” *IEEE* 信号处理杂志,卷.39, 期. 2, 页. 135–158, 2022.