SCIENTIA SINICA Informationis

纪念《中国科学》创刊 70 周年专刊・评述



Shannon 信息论与未来 6G 技术潜能

尤肖虎1,2

东南大学移动通信国家重点实验室,南京 210096
 网络通信与安全紫金山实验室,南京 211100
 E-mail: xhyu@seu.edu.cn

收稿日期: 2020-04-09; 接受日期: 2020-04-21; 网络出版日期: 2020-09-23

科技部 6G 专项计划、东南大学十大科学技术问题研究计划和紫金山实验室普适通信研究项目资助

摘要 本文从 Shannon 信息论及其扩展形式的角度, 探讨未来 6G 移动通信系统性能提升的潜能. 首 先, 对经典 Shannon 信道容量及传输性能折中理论框架进行了概括, 并探讨了其在当代移动通信系统 中应用的局限性. 其次, 对 Shannon 信息论的多输入多输出 (MIMO) 扩展形式进行了论述, 指出了其 在当代移动通信系统发展中所扮演的基础性作用; 鉴于 Shannon 信息论及其 MIMO 扩展形式均为非 构造性的, 文中还分别讨论了信道容量构造性逼近的两种重要途径 — 信道极化和特征模式无线传输. 然后, 围绕更高的频谱效率与功率效率、更高的可靠性与更低的时延、更高的频段等 6G 技术发展需 求, 从基础理论的角度探讨其性能进一步提升的潜能, 主要结论包括: 通过引入更多的天线、无蜂窝网 络构架创新, 以及分块长度 – 差错概率 – 传输速率 – 最小天线数等关键参数的有效平衡等途径, 未来 6G 技术仍然具有较大的提升潜能, 但需要在系统性能与部署成本之间进行折中, 并有针对性地有效获 取高频段 MIMO 信道容量. 最后本文对涉及未来发展的若干基础性问题进行了总结.

关键词 Shannon 信息论, MIMO, 6G 移动通信, 无蜂窝, 高可靠与低时延

1 引言

20 世纪 50 年代前后, Shannon 及其学生 Gallager 等开创了信息论的数学基础, 他们系统性地 给出了信道最大可能的可靠传输速率 (即信道容量), 以及有限编码长度与可获取速率和传输差错概 率之间折中的理论框架^[1~4], 从而为 70 余年来通信理论与技术的发展奠定了理论基础. 但需指出, Shannon 经典信息论是针对单输入单输出 (SISO) 信道进行的, 而现代移动通信系统基本是以多输入 多输出 (MIMO) 为基础的, 因而上述理论存在着明显的应用局限性.

MIMO 信道是 SISO 信道的拓展形式, 其输入输出由传统的标量信号变为向量信号, 其典型应 用则对应于无线链路收发两端均配置多个天线. MIMO 信息论的起源要归功于 20 世纪 90 年代末期

引用格式: 尤肖虎. Shannon 信息论与未来 6G 技术潜能. 中国科学: 信息科学, 2020, 50: 1377–1394, doi: 10.1360/SSI-2020-0086 You X H. Shannon theory and future 6G's technique potentials (in Chinese). Sci Sin Inform, 2020, 50: 1377–1394, doi: 10.1360/SSI-2020-0086

ⓒ 2020《中国科学》杂志社

Foschini 等和 Telatar 的贡献^[5~7]. 他们的研究表明, MIMO 信道不但具有经典 Shannon 信息论所涉 及的时间和频率自由度, 而且具有空间自由度, 其信道容量可在经典 Shannon 理论的基础上随其空间 自由度的增加而线性增加, 从而拓展了经典 Shannon 信息论. 正因如此, MIMO 无线传输已成为当代 4G 和 5G 提升系统容量和频谱效率的最为重要的核心手段.

5G 提出了"增强宽带、万物互联"的发展理念,通过引入增强型移动宽带 (eMBB)、高可靠与低时延 (uRLLC) 和增强型机器类连接 (eMTC) 等场景与技术,首次将垂直行业应用纳入公众移动通信系统的范畴.除了采用大规模 MIMO 和非正交多址接入 (NOMA) 以提升系统容量和大连接能力之外, 5G 还采用了更短的时隙调度单位、重复发送、多连接等 uRLLC 相关技术^[8],在提高可靠性的同时降低无线传输时延.可以预计,未来 6G 将以此为基础,发展覆盖范围更广、传输速率更高、可靠性与时延等性能指标更为优越的无线传输技术,构建更新一代的普适性、智慧化、全业务移动信息基础设施^[9~11].

纵观移动通信发展历程,5G 已将其技术水平推向一个前所未有的高度.未来 6G 移动通信在系统 容量与频谱效率、可靠性与传输时延等核心技术指标方面是否具有进一步改善的潜能,是学术界和工 业界普遍关心的问题之一.本文试图从 Shannon 信息论框架及其扩展形式这一角度系统性地回答这 一问题.综合结果表明,通过增加收发天线数、引入新型构架,以及寻求分块长度(时延)、差错概率、 传输速率及最小天线数的有效平衡等途径,上述核心技术指标在理论上均具有较大的提升空间,尚不 存在所谓系统性能的"天花板",至少从理论上如此.

Shannon 信息论的相关研究分支众多,成果内容浩瀚. 限于篇幅,本文将重点对经典 Shannon 信息论的基本框架及其 MIMO 扩展形式进行概括. 在此基础上,围绕更高的频谱效率与功率效率、更高的可靠性与更低的时延,以及更高的频段等,从理论上探索未来 6G 技术能力的潜在提升途径.

2 经典 Shannon 通信信息论基本框架及其应用局限性

本节将对经典 Shannon 信息论及传输性能折中的整体框架及重要里程碑进行概述,主要包括 Shannon 信道容量理论、有限编码长度条件下的 Gallager 指数与 Polar 码信道容量逼近、Shannon 与 Polyanskiy 等提出的有限分块长度渐进容量理论等;在此基础上,对经典信息论应用于现代移动通 信系统的局限性进行分析总结.

2.1 Shannon 信道容量理论

Shannon 于 1948 年在其著名的论文 "A mathematical theory of communication" 中给出了一系列 有关通信的基本理论^[1],并首次引入了信道容量的定义. 当传输速率不大于该容量时,则存在某种编 码方式,在编码长度趋于无穷长时,其传输差错概率趋于零. Shannon 在该文中证明了信道容量 *C* 由 下式给出

$$C = \max_{x} \{ H(x) - H(x|y) \},$$
(1)

即信道容量等于发送信号与接收信号之间互信息量的上确界, x 和 y 分别代表发送信号与接收信号, $H(\cdot)$ 和 $H(\cdot|\cdot)$ 分别为随机变量的信息熵和条件信息熵. Gallager 随后证明了当发送信号为 Gauss 分 布时, 上述互信息量为最大.

Shannon 于 1949 年在其另外一篇论文 "Communication in the presence of noise"中, 通过复杂的

推导,给出了加性 Gauss 白噪声信道 (AWGN) 容量的表达式^[2]:

$$C = B \cdot \log_2\left(1 + \frac{P}{\sigma^2}\right) \triangleq B \cdot \log_2(1 + \text{SNR}), \tag{2}$$

其中 B, P 和 σ^2 分别为信号带宽、发送信号功率及噪声方差. 事实上, 当发送信号和噪声均为 Gauss 分布时, 由式 (1) 经过简单的推导便可得到式 (2), 但这丝毫不降低 Shannon 上述奠基性的贡献.

2.2 Gallager 差错指数与 Polar 码容量逼近

Shannon 的信息论是非构造性的,无法由此设计出容量可达 (或传输速率最大化)的通信系统. Shannon 的学生 Gallager 在文献 [1] 的基础上引入了有限长度随机编码理论,将 Shannon 理论向前推 进了重要一步. 运用 Chernoff 界可以证明,对于二进制离散无记忆信道 (B-DMC),若编码分块长度为 n, 且采用随机挑选的编码码字, 其符号平均差错概率 Pe 存在如下两种上限:

其中 R 为编码速率, R₀ 为截止速率 (cutoff rate), 由下式给出:

$$R_0 \triangleq -\log_2 \int_{\mathcal{C}^M} \left(\frac{1}{|\mathcal{M}|} \sum_{x \in \mathcal{M}} \sqrt{p(y|x)} \right)^2 \mathrm{d}y, \tag{4}$$

其中 *M* 为全部码字形成的子空间, $|\mathcal{M}|$ 为码字个数, p(y|x) 为输入输出条件概率. $E_r(R)$ 为更为广泛 意义上的随机编码差错指数, 简称为 Gallager 指数^[4], 并存在 $n \to \infty$ 时, $(\log_2 P_e)/n \to E_r(R)$, 且当 R < C 时, $E_r(R) > 0$.

式 (3) 给出了分块长度与平均差错概率之间的内在关系, 在通信理论发展中具有里程碑式的重要 意义. 它表明, 当编码速率满足 *R* < *R*₀ < *C* 时, 即便采用随机编码, 也可在 *n* 充分大时, 使通信系统 的平均差错概率趋于零.

对比式 (3) 中所给出的两种上限, 前者更为简单实用, 后者更为紧致. 图 1 示出了二进制对称信道 (BSC) 条件下的两种上限以及 $E_r(R)$ 与 R_0 和 C 之间的相互关系. 可以看出: (1) $E_r(R)$ 是单调递减 函数, 这意味着随着编码速率 R 的增加, 其平均差错概率 P_e 随 n 增大时的收敛速度将变慢. (2) 对于 固定的 R 值, $E_r(R)$ 随信道容量 C 的增加而变大, 这意味着对于信道容量大的通信系统, 其平均差错 概率 P_e 随 N 增大时的收敛速度更快. (3) 对于固定的 n, $R_0 = E_r(0)$ 增大意味着其平均差错概率 P_e 更小.

截止速率 R₀ 最大化曾经是信道容量逼近研究的核心问题之一,该方面的研究伴随着通信系统构造方式的一系列创新. 学术界曾一度认为信道容量 C 确定后, R₀ 是恒定不变的. Pinsker^[12] 有关级联码的研究改变了人们对 R₀ 的认识; 当信道编码分为内外两级级联码时, 其信道容量 C 保持不变, 而 R₀ 会显著增加.

上述研究诱发了 Polar 码的诞生, 它是目前唯一被严格证明的具有信道容量可达能力的编码方式, 同样具有里程碑式的意义. Arikan^[13,14]提出了一种扩展截止速率 R_0 的信道重构方法. 以下以分块长度 n = 2为例给出其简洁的说明. 首先进行信道 "组合", 如图 2(a)所示; 目的是建立原发送信息 x_1 和 x_2 与重构发送信息 u_1 和 u_2 之间的一一映射, 并通过两个统计独立的 W 信道分别发送. 其次进行信道 "拆分", 如图 2(b)所示; 因引入了上述映射, 原来相互独立的两个信道发生了统计性关联, u_1 和 u_2 的检测问题被转化成为两个条件概率 $p(y_1y_2|u_1)$ 和 $p(y_1y_2u_1|u_2)$ 的计算问题¹⁾, 由此形成了图 2

¹⁾ 在进行 u₂ 检测时, 需事先得到 u₁, 故 Polar 译码需要引入 successive cancelation (SC).







图 2 (网络版彩图) (a) 信道组合与 (b) 拆分示意图 Figure 2 (Color online) Illustrations of (a) channel combing and (b) splitting

中所示的虚拟信道 W_2^- 和 W_2^+ . 基于以上两个条件概率, 不难证明 $C(W_2^-) < C(W) < C(W_2^+)$, 且等 效截止速率满足²)

$$\bar{R}_{0}(W_{2}) \triangleq \frac{1}{2} \left[R\left(W_{2}^{-}\right) + R(W_{2}^{+}) \right] > R_{0}(W) \,. \tag{5}$$

从而可见, 经过上述的信道"组合"与"拆分", 两个虚拟信道的容量发生了两极分化, 且等效截止速率 得以扩展. 以 n = 2 为基础, 采用类似于快速 Fourier 变换的构建过程, 上述结果可以方便地推广至任 意 $n = 2^i$, i = 2, 3... 因 $\bar{R}_0(W_i)$ 为有界单调增序列, 故当 $i \to \infty$ 时, 存在 $\bar{R}_0(W_i) \to C(w)$.

Polar 码是构造性的, 且可以逼近 BSC 信道容量限. 正因如此, Polar 码得到了极大关注, 并被 5G 标准所采纳. 值得一提的是, 我国研究人员为 5G Polar 码发展做出了诸多贡献. 作为一个例子, 作者 所在的研究小组首次提出了灵活的分段 CRC Polar 码构建方法与早停策略^[15,16], 已成为 5G Polar 码 的基本配置方式.

2.3 渐进容量理论

有限分块长度渐进 (asymptotic) 容量理论依然要归功于 Shannon 的早期贡献. 他在 1959 年发表 2) 对于 BSC 信道,存在 0 ≤ *C*(*W*) ≤ 1; 此处假设 0 < *C*(*W*) < 1, 否则无需进行极化处理. 的题为 "Probability of error for optimal codes in a Gaussian channel" 的论文中写道 ^[3], 信道容量的 逆问题在应用中是较为自然的: 对于给定的差错概率, 所需的信道编码分块长度应该是多少呢? 为此, Shannon 针对 AWGN 信道, 从式 (3) 所示的差错指数出发, 经过冗长的推导, 得到了分块长度为 n、编 码速率为 R 时的最优差错概率 $P_e^{\text{opt}}(n, R)$ 的近似表达式, 参见文献 [3] 中式 (9). 为与后续的表述相 一致, 本文将其表达式改写为如下形式:

$$R \approx C - \frac{1}{\sqrt{n}} \sqrt{\frac{2 + \text{SNR}}{2 \cdot \text{SNR}(1 + \text{SNR})}} \Phi^{-1} \left(P_e^{\text{opt}} \right), \tag{6}$$

其中 SNR = P/σ^2 , $\Phi^{-1}(\cdot)$ 为 Gauss 积分逆函数. 该表达式给出了编码速率、信道容量、分块长度, 以 及最优差错概率 $P_e^{\text{opt}}(n, R)$ 之间的渐进关系.

Polyanskiy 等^[17]从更为广泛的角度研究了有限分块长度条件下渐进信道容量的回退问题. 对于 一般性的离散无记忆信道 DMC, 所导出的信道渐进容量被概括如下:

$$R = C - \sqrt{\frac{V}{n}} \Phi^{-1}(P_e) + o(n), \qquad (7)$$

其中 V 为信道色散度 (channel dispersion), 它反映了分块长度为 n 且差错概率为 P_e 时, 最大编码速率与信道容量之间的归一化方差, 其一般性的定义可参见文献 [18]. Polyanskiy 等 ^[17] 利用大数定理及 正态逼近得到了有关参数 V 的计算方法. 具体地, 当信道为 BSC 时, 存在 V = p(1-p), p 为 BSC 信 道的状态转移概率; 当信道为 AWGN 时, 重现了 Shannon 当年的推导结果, 见式 (6).

综合上述讨论, 对于给定的差错概率 *P_e* 以及编码速率 *R*, 所需的信道编码分块长度应大于如下估计值:

$$n^* \approx V \cdot \left[\frac{\Phi^{-1}(P_e)}{C-R}\right]^2,\tag{8}$$

由此明确地回答了 Shannon 当初提出的基本问题. 但需注意, 上述结果仅在 $n \gg 100$ 时相对准确^[17], 这时式 (7) 中的近似项 o(n) 可以忽略.

2.4 经典 Shannon 信息论的应用局限性

Shannon 以朴素的数学工具、直觉化与工程化的思维逻辑, 对上述经典通信信息论基础框架的形成做出了奠基性的贡献. 但应客观地看到, 由于他所处的时代, 上述理论在应用方面存在明显的局限性. 这具体表现在以下方面.

(1) 经典 Shannon 信息论所研究的对象与现实应用渐行渐远. 首先, 经典 Shannon 信息论所涉及 的信道模型过于简单, 基本局限于 AWGN 信道、BSC 信道或 DMC 信道, 而实际移动通信系统的信道 模型大多为 Rayleigh 或 Rice 多径衰落信道. 其次, 经典 Shannon 信息论基本局限于 SISO 形式, 无法 为现代移动通信系统所必备的多天线 MIMO 技术提供理论支撑. 第三, 经典 Shannon 信息论仅考虑 点到点单用户所涉及的基本理论问题, 而多点协作, 特别是多用户 MIMO 已成为现代移动通信技术的 主流发展方向.

(2) 经典 Shannon 信息论已无法有效支撑现代移动通信系统的容量分析. 原始 Shannon 信道容量的推导过程仅涉及频率与时间两个维度自由度的利用问题. 与之不同的是, 现代移动通信系统引入了 多天线技术, 其自由度由频率与时间两个维度扩展至频率、时间与空间 3 个维度, 信道容量可随天线 数的增加而线性增加. 可以注意到, 式 (1) 给出的互信息容量模型, 因不涉及具体的信道模型, 仍然具 有普适性; 但式 (2) 给出的 AWGN 信道容量公式需要扩展至多天线情形. 具体的讨论将由 3.1 小节 给出.

(3) 经典 Shannon 信息论所涉及的差错指数与渐进理论等已无法有效支撑现代通信系统有关时 延与可靠性的分析. 经典 Shannon 信息论对信道容量 (速率)、差错概率 (可靠性) 和分块长度 (时延) 以及三者之间的相互关系进行了完备的分析. 但由于其历史局限性, 未考虑现代移动通信系统多天线 配置所引入的实质性影响. 事实上, 多天线与 MIMO 具备天然的空间复用与空间分集增益, 对于给定 差错概率, MIMO 所需的分块长度将会大大小于 SISO; 这意味着经典 Shannon 信息论所涵盖的差错 指数与渐进理论等均需要根据收发两端的天线数进行相应的修正. 具体的讨论将由 4.2 小节给出.

3 Shannon 信道容量理论的 MIMO 扩展形式与性能折中

本节重点讨论多天线 MIMO 信道容量模型. 首先, 多天线 MIMO 是经典 Shannon 信息论更为一般的扩展形式; 其次, 若信道信息是全局性已知的, 则单用户、多用户或多基站联合处理等复杂应用场 景下的信道模型及容量分析均可以用统一的形式加以描述. 本节还给出了 MIMO 信道容量构造性逼 近的两种重要途径: 特征模式 MIMO 无线传输以及 Gallager 指数分析法. 最后讨论了 MIMO 信道的 分集增益、空间复用增益, 以及性能折中的基本原理.

3.1 MIMO 信道容量模型

多天线信道容量理论的发展源于 Foschini 等^[5,6] 最初的研究工作. Telatar 于 1999 年给出了较为 完整的 MIMO 信道容量分析方法^[7],从而为发展更高频谱效率、更大容量的当代移动通信系统奠定 了理论基础.

假设 MIMO 无线链路具有 N_t 个发送天线和 N_r 个接收天线,则该无线链路模型可以一般性地描述为³)

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n},\tag{9}$$

其中 $\mathbf{y} \in \mathcal{C}^{N_r}$, $\mathbf{x} \in \mathcal{C}^{N_t}$, $\mathbf{H} \in \mathcal{C}^{N_r \times N_t}$ 和 $\mathbf{n} \in \mathcal{C}^{N_r}$ 分别代表接收信号、发送信号、信道状态信息矩阵和 Gauss 干扰信号, 并假设 \mathbf{H} 是准静态的, 即在某个特定时间区间内保持不变. 如前所述, Shannon 给出 的信道容量公式 (1) 具有普适性, 若假设发送信号 \mathbf{x} 和干扰噪声 \mathbf{n} 为 Gauss 分布随机向量, 则由式 (1) 容易计算出关于发送信号 \mathbf{x} 和接收信号 \mathbf{y} 的互信息 ^[19], 即信道容量如下:

$$C_{\text{MIMO}}(\boldsymbol{H}) = B \cdot \log_2 \left[\det \left(\boldsymbol{I} + \frac{1}{\sigma^2} \boldsymbol{H} \boldsymbol{Q}_x \boldsymbol{H}^{\dagger} \right) \right], \tag{10}$$

其中, † 代表共轭转置, Q_x 为 x 的自相关矩阵, 不失一般性, 可假设为 $Q_x = \text{diag}[p_1 \cdots p_N]$, $N = \min_{N_t, N_r}$. 运用 Lagrange 乘子法, 可得发送信号 x 的最优功率分配值如下^[7]:

$$p_i = \left[\mu - \frac{\sigma^2}{|\lambda_i|^2}\right]^+, \ |\lambda_i| > 0, \ i = 1, \dots, k, \ k \leqslant N,$$
(11)

且最大可获取信道容量为

$$\max_{\mu} \sum_{i} \left(\log_2\left(\mu \left| \lambda_i \right| \right) \right)^+ \text{ s.t.} \sum_{i} \left(\mu - \frac{1}{\left| \lambda_i \right|} \right) = P, \tag{12}$$

3) 若信道存在多径时延扩展, 经过重新整理后其信道模型仍可等效为式 (9), 参见文献 [19] 中第7章或文献 [20].

其中, λ_i 为 **H** 的特征值, P 为 N_t 个发送信号的功率之和, μ 为 Lagrange 乘子, 通常需要迭代才能得 到其最优解.

为了与 Shannon 最初给出的 AWGN 信道容量公式 (2) 进行对比,此处假设发射端未知信道状态信息只能采用等功率发送,信道 H 的每个元素均为方差为 1 的独立同分布随机变量,且 $N = \min \{N_t, N_r\}$ 充分大.这时运用大数定理,从式 (10) 可近似得到

$$C_{\text{MIMO}}(\boldsymbol{H}) \approx B \cdot N \log_2 \left(1 + \text{SNR}\right).$$
 (13)

由此看到,由于引入了多个收发天线,无线链路的信道容量较 Shannon 最初得到的 AWGN 信道容量 增加了 N 倍,从而扩展了经典 Shannon 信道容量理论.

此处讨论 MIMO 信道的自由度. Shannon 运用 Nyquist 采样定理指出^[2], AWGN 信道的自由度 为 2BT, 其中 T 为通信时长; 这意味着对于时长为 T、带宽为 B 的 AWGN 信道, 最大可以传送 2BT 个独立的符号. 对于 MIMO 信道, 在信道 H 为满秩的条件下, 可以发送 N = min.{N_t, N_r} 个相互独 立的数据流, 并在接收端利用最小均方差 (MMSE) 或最大似然序列检测恢复出所发送的信号; 这意味着 MIMO 信道的自由度为 2NBT, 即最大可传送 2NBT 个独立的符号. 由此可以看出, 经典 Shannon 信息论的自由度仅包含时间与频率两个维度, 而多天线 MIMO 信道的自由度被扩展至时间、频率与 空间 3 个维度.

在以上讨论中, 假设信道 **H** 为满秩, 即 MIMO 信道的空间自由度为 $N = \min . \{N_t, N_r\}$. 实际应 用中, 多个天线之间可能存在某种相关性, 这时 MIMO 信道的空间自由度将取决于 **H** 特征值 λ_i 的 具体分布. 相关讨论将在 3.2 小节给出.

这里进一步讨论多用户及多基站联合 MIMO (MU-MIMO) 的信道容量模型问题. 首先应注意到, 式 (9) 给出的 MIMO 无线链路模型并未事先设定多个天线的相互位置关系. 当收发两端的天线来 自多个用户或多个基站时,式 (9) 可以涵盖多用户及多基站联合 MIMO 信道模型. 但需注意,此时 式 (10) 给出的信道容量是多用户或多基站最大可获取的和速率 (sum-rate). 这时隐含地假设信道状 态信息 *H* 是全局性已知的,且多用户或多基站联合信号处理是全局性优化的 (globally optimized). 此 外还应注意,对于不同的应用场景,信道状态信息 *H* 相应元素的大尺度衰落将有所不同^[19]. 第4节 给出的分布式 MU-MIMO 以及无蜂窝系统模型将是其一种特殊形式.

3.2 MIMO 信道的构造性容量逼近

首先引入具有 MIMO 信道容量逼近能力的特征模式 (eigen mode) 无线传输方式. 为此将信道 H进行 SVD 分解, 得到 $H = V^{\dagger}\Lambda U$, Λ 为 H 的特征值 λ_i 组成的对角阵, 即 $\Lambda = \text{diag}[\lambda_1 \cdots \lambda_k, 0 \cdots 0]$, 并由此把式 (9) 改写成为

$$Vy = \Lambda Ux + Vn \ \vec{x} \ \tilde{y} = \Lambda \tilde{x} + \tilde{n}, \tag{14a}$$

其中 $\tilde{y} = Vy$, $\tilde{x} = Ux$ 及 $\tilde{n} = Vn$, 且上式可写为标量形式:

$$\tilde{y}_i = \lambda_i \tilde{x}_i + \tilde{n}_i, \quad i = 1, \dots, k,$$
(14b)

其中 k 代表了 MIMO 信道的空间自由度, 且容易验证 ñ 仍为 Gauss 白噪声. 这提示我们, 若把发送 信号和接收信号分别映射到信道 H 的特征空间 U 和 V, 则 MIMO 信道可以被分解为 k 个独立平行 的 AWGN 信道. 若进一步根据式 (11) 分配发射功率, 即可实现 MIMO 信道容量的最大化.

图 3 示出了特征模式 MIMO 无线传输系统构造示意图. 因其系统模型与式 (10)~(12) 完全一致, 故在理论上是信道容量可达的. 但需注意,该系统模型较为理想化,需要在系统发送端确知信道的瞬





时信息 H,这在实际应用中通常难以做到,或需要较多的无线资源,并以较短的时间从系统接收端反馈该信息.

解决以上难题的一个有效途径是由东南大学研究团队提出的统计特征模式 MIMO 无线传输^[20~23], 其优点在于系统发送端仅需已知信道的统计信息 $E_H{H^{\dagger}H}$,并具有较低的系统实现复杂度.为此, 将上述瞬态容量 $C_{\text{MIMO}}(H)$ 的逼近目标改为遍历性容量 $E_H{C_{\text{MIMO}}(H)}$.因 $E_H{C_{\text{MIMO}}(H)}$ 难以 闭式求解,需在引入紧致上界^[20,22] 的基础上,将问题转化为可求解的凸优化问题.其中一个较为简捷 的做法是利用 Jensen 不等式^[20],得到其遍历性容量的紧致上界:

$$E_{\boldsymbol{H}} \{ C_{\text{MIMO}} (\boldsymbol{H}) \} \triangleq B \cdot E_{\boldsymbol{H}} \left\{ \log_2 \left[\det \left(\boldsymbol{I} + \frac{1}{\sigma^2} \boldsymbol{H} \boldsymbol{Q}_x \boldsymbol{H}^{\dagger} \right) \right] \right\}$$
$$\leqslant B \cdot \log_2 \left[\det \left(\boldsymbol{I} + \frac{1}{\sigma^2} \boldsymbol{Q}_x E_{\boldsymbol{H}} \left\{ \boldsymbol{H}^{\dagger} \boldsymbol{H} \right\} \right) \right]. \tag{15}$$

该上界的推导基于以下几个事实:存在恒等式 det(I + AB) = det(I + BA),基于 H 的 SVD 分解可 将 MIMO 等效为 N 个平行的独立信道, log(1 + x) 为凸函数以及统计学中常见的 Jensen 不等式.该 上界的最优求解与式 (11) 和 (12) 相类似,仅需用 $E_H{H^{\dagger}H}$ 替换原有的 $H^{\dagger}H$.分析结果表明,统 计特征模式 MIMO 传输可以较为完美地逼近 MIMO 的遍历性容量 ^[22],且因统计信道信息的时间变 化尺度远大于瞬态信息,其本身具有低计算复杂度和低信道反馈开销等特点 ^[20].关于系统的具体构造 方法、信道非理想信道条件下的容量逼近迭代检测及相关实验测试结果等,可参见文献 [20,23].此外 还应注意到,因未对 MIMO 信道特征设置任何前提条件,上述统计特征模式 MIMO 无线传输具有广 泛的普适性.

通过对截止速率 R₀ 或 Gallager 指数的计算, 分析 MIMO 无线传输的最大可获取速率, 是评判 MIMO 信道容量逼近的另外一种可行途径 ^[24,25]. 这类分析推导较为复杂, 一般只能给出较为松弛的 传输性能上下界, 但仍不失其重要性. 观察式 (4) 给出的定义可知, R₀ 与具体信道形式无关, 因而适用 于 MIMO 无线传输可达速率的评估. 但因涉及 MIMO 分块发送信号子空间上的概率积分与求和, R₀ 的求解较为困难. 文献 [24] 借助于 MIMO 信道成对差错概率的近似表达式, 给出了一种关于 MIMO 信道截止速率 R₀ 的上下界; 其对 2 × 2 MIMO 信道的分析结果表明, 若信道状态信息确知, 则其信道 容量在非渐进 (低信噪比) 区域是可获取的. 文献 [25] 给出了等功率发射条件下 MIMO 信道 Gallager 指数的表达式, 及平均差错概率与分块长度、信道相干时间和传输速率之间的关系; 运用这些表达式, 可以综合评估 MIMO 无线传输速率逼近信道容量时所能得到的差错概率, 并可在分块长度 (时延) 和差错概率 (可靠性) 之间寻求传输性能的折中.

3.3 MIMO 信道的分集增益与空间复用增益及性能折中

MIMO 信道可同时具备空间复用增益和多天线分集增益,并可在两者之间寻求折中,这意味着 MIMO 系统可在传输速率与可靠性方面寻求某种平衡.

•讨论 MIMO 信道的多天线分集增益. 在经典文献中, 多天线技术可以有效对抗移动通信系统所固有的信道衰落特性, 提升无线传输的可靠性. 以单发多收 (SIMO) 系统为例, 若接收端采用多天线最大比合并, 则随着接收端天线数 N_r 不断增大, 其信道衰落特性将产生硬化 (channel hardening) 效应, 最终趋于等效的 AWGN 信道 ^[26], 其误码率特性可近似表达为 $P_e \propto \text{SNR}^{-N_r}$. 该结果表明, 天线数的增加对于改善无线传输的可靠性极为有效. 进一步, 若采用 N_t 个天线重复发送同一信息, 并用 N_r 个天线进行合并接收, 则由此所形成的 MIMO 系统的误码率可近似表达为 $P_e \propto \text{SNR}^{-N_tN_r}$. 为衡量多天线系统的分集特性, 给出其分集度的一般性定义如下 ^[27]:

$$d = -\lim_{\text{SNR}\to\infty} \frac{\log_2 P_e(\text{SNR})}{\log_2 \text{SNR}}.$$
(16)

它等同于高信噪比条件下系统误码率相对于 SNR 的衰减指数.显然,对于上述 SIMO 和 MIMO 系统, 其分集度分别为 N_r 和 N_tN_r.

•讨论 MIMO 信道的空间复用增益. 如同 3.1 小节所述, 对于独立同分布的 MIMO 信道 *H*, 其空间自由度为 *N* = min.{*N_t*, *N_r*}, 系统可以相互独立地发送 *N* 个数据流, 这意味着 MIMO 信道的空间 复用增益为 *N*. 为在一般意义上衡量 MIMO 信道的空间复用增益, 此处参照文献 [27] 给出其空间复 用增益的定义如下:

$$r = \lim_{\text{SNR}\to\infty} \frac{R(\text{SNR})}{\log_2 \text{SNR}},\tag{17}$$

其中 *R*(SNR) 为某一特定 MIMO 系统构架的最大可获取速率. 该式等同于高信噪比条件下系统最大可获取速率与对数 SNR 之间的比例因子. 显然, 对于 3.1 小节所述的独立同分布 MIMO 信道 *H*, 运用式 (13) 和 (17), 可以得知其空间复用增益为 *N*.

• 在较为理想的假设条件下,讨论 MIMO 信道分集增益与空间复用增益的性能折中. 在上述讨论 中, MIMO 信道先后被分别用于分集增益的最大化和空间复用增益的最大化;前者分集增益最大,但 由于只能发送单一数据流,故传输速率最低;而后者正好与之相反. 可以直观地猜测,如果将空间复用 增益 (即拟发送的独立数据流个数) 限定为 $r < N = \min . \{N_t, N_r\}$,则 MIMO 系统的分集增益将会得 以增加,从而提高传输可靠性. 还可进一步直观地猜测,若 MIMO 信道的发送数据块足够长,在空间 复用增益限定为 $r < N = \min . \{N_t, N_r\}$ 时, MIMO 系统的分集增益应为 $(N_r - r + 1)(N_t - r + 1)$;这是 因为,在 N_t 个发送天线和 N_r 个接收天线中,分别有 r 个发送天线和 r 个接收天线被用于空间复用, 其余则被用于获取分集增益. 文献 [27] 给出了上述猜想的证明,其假设的充分条件是发送数据块长度 需满足 $n > N_t + N_r - 1$,且在该数据块内信道 **H** 保持不变.

以上讨论所涉及的假设条件较为理想,但式 (16) 和 (17) 不失一般性.更为广泛意义上的无线传 输性能折中框架将在 4.2 小节中进一步讨论.

4 未来 6G 核心技术指标与性能提升

本节将以第 2,3 节分别给出的经典 Shannon 信息论、MIMO 扩展形式及性能折中理论框架为基础,讨论进一步提升 6G 移动通信系统核心技术指标的理论途径与潜能,包括更高的频谱效率与功率效率、更高的可靠性与更低的时延、更高的频段与容量等.

4.1 更高的频谱效率与功率效率

伴随着技术的不断进步,移动通信系统的无线资源利用率不断提升,其频谱效率已从第二代移动 通信系统 (如 GSM)的 0.2 bps/Hz 左右提高至 5G 移动通信系统的 50 bps/Hz 或更高.可以预见,未 来 6G 将继续延续这一重要发展趋势^[9].

如第3节所述,如若增加无线链路收发两端的天线数,所构成的 MIMO 信道容量在理论上可以随 之线性增加.但不幸的是,由于用户终端的物理尺寸受限,能够安装的有效天线数有限.从 3.1 小节的 讨论可知,点到点无线链路的信道容量受限于收发两端的最小天线数 $N = \min \{N_t, N_r\}$,这意味着单 个点到点无线链路的频率效率 $\eta = C_{\text{MMO}}/B$ 将难以进一步得到大幅提升.

5G 移动通信系统引入了多用户 MIMO (MU-MIMO), 将移动通信系统的频率利用率提升到了一 个新的高度. 其原理基于以下基本事实: 尽管单个用户终端的天线数受限, 但若将多个用户联合起来, 可构成一个基站与多个用户之间的点对多点 (P-2-MP) 联合处理 MU-MIMO 系统. 进一步, 若基站一 侧所有的信道状态信息均已知⁴), 则式 (9)~(12) 的 MIMO 信道及容量模型依然成立. 但需注意, 这时 用户一侧的天线数为各用户天线数的总和, 且信道容量为各个用户的和速率 (sum-rate). 当基站一侧 的天线数足够多时, 便可在相同的频率上同时支持足够多的用户数. 这意味着进一步增加基站的天线 数⁵), 形成超大规模 MIMO, 便可使未来 6G 系统的频谱利用率进一步得以提高. 但需注意, 当基站天 线阵列部署的物理尺寸受限时, 过于密集的天线部署将存在严重的互偶效应 ^[28], 相关 MIMO 信道容 量存在某种理论极限 ^[29]. 采用连续孔径多天线技术, 构成连续孔径相控 (continuous aperture phased, CAP) MIMO ^[30] 或全息 (holographic) MIMO ^[31], 是在天线阵列孔径受限条件下进一步挖掘信道容量 及频谱效率的可能技术途径.

需要特别指出, MU-MIMO 的引入还改变了有关移动通信系统频谱利用率的常用评估规则. 以下 分别以 *M* 个用户的 FDMA (frequency division multiple access) 和 MU-MIMO 为例进行对照说明. 假设两种情形下用户在小区内的位置均符合独立同分布特征, 且每个用户占用相同的带宽. 对于 *M* 个 FDMA 用户, 因每个用户占用不同的频率资源, 故 *M* 个用户总的频谱利用率等同单个用户. 对 于 MU-MIMO, 所有 *M* 个用户叠加在相同的频段上, 同时与一个基站进行通信; 且在理想条件下, 每 个用户的通信速率在统计意义上与 FDMA 用户相同; 故 *M* 个用户的 MU-MIMO 的总频谱利用率是 FDMA 的 *M* 倍. 以上例子说明, 对于 MU-MIMO, 其频谱利用率需要在多用户小区层面 (而非单个用 户层面) 综合进行评估.

5G 移动通信系统所采用的多用户 MIMO (MU-MIMO) 还可以进一步扩展至多基站、多用户联 合处理情形,从而构成多点对多点 (MP-2-MP) 形式的无蜂窝 (cell-free) 移动通信系统 ^[19,32,33],并由 此进一步提升整个系统的频谱利用率.图 4 示出了传统蜂窝系统与无蜂窝系统的对比.以上行链路为 例给出如下说明.对于无蜂窝系统,由于引入了多小区联合处理,多个用户在多个小区的覆盖范围内 构成了 MP-2-MP 分布式 MU-MIMO,所有的用户和基站可同时同频工作;这时,式 (9)~(12) 给出的 MIMO 信道容量模型依然适用,区别在于基站一侧的多个天线分别来自不同的小区.与之不同的是, 由于传统蜂窝系统不具备多小区联合处理能力,因而无法获得 MU-MIMO 空间复用增益.

无蜂窝系统另外一个潜在的优越特性是消除了传统蜂窝构架在小区频率复用方面的限制,其小区频率复用因子等于 1. 这意味着无蜂窝系统不再受频率资源静态分配的限制,可实现真正意义上的跨

⁴⁾ 基站可以通过信道估计获取 MU-MIMO 上行链路的全局信道状态信息. 对于 TDD 系统,可以基于上下行链路的对偶性,利用上行链路的信道状态信息对 MU-MIMO 下行链路进行预编码处理. 对于 FDD 系统,可以利用上行链路 信道估计的统计信息,对 MIMO 下行链路进行预编码处理.

⁵⁾ 例如,从现有 5G 系统所配置的 64 至 192 个等效天线阵元,增加至 512 或更多.



图 4 (网络版彩图) (a) 传统蜂窝系统构架与 (b) 无蜂窝系统构架的对比

 ${\bf Figure} \ {\bf 4} \quad {\rm (Color \ online)} \ {\rm Comparison} \ {\rm of} \ {\rm (a)} \ {\rm cellular \ and} \ {\rm (b)} \ {\rm cell-free \ system \ architectures}$

小区、全动态的频率资源调配,从而为构建资源调配灵活、频谱利用率更高的 6G 移动通信系统带来 全新的可能.

构建低能耗、环境友好的移动通信网络将是未来 6G 发展的一个重要研究方向. 限于篇幅, 此处仅 讨论无蜂窝系统基站一侧总发射功率的有效利用问题. 不失一般性, 假设基站一侧发射天线总数 N_t 充 分多, 下行链路 M 个用户均配置单天线, 且所处的位置符合随机独立同分布, 这时式 (9) 中的信道 H 为独立同分布.类似于式 (13) 的推导,利用大数定理,可以得到基站一侧总发射功率的标度率 (scaling law) 如下^[19]:

$$P_{\rm BS} \approx \frac{\left(2^{\eta/M} - 1\right)}{\overline{\rm SNR}} \approx \frac{\eta}{\overline{\rm SNR} \cdot M}, \ N > M, \ \square \ M \ \widehat{\rm Thetat},$$
(18)

其中为 $\eta = C/B$ 为频谱效率, $\overline{SNR} \triangleq \sigma_h^2/\sigma^2$ 为归一化信噪比, σ_h^2 为信道 **H** 各个元素的方差. 该结果 与文献 [34] 中有关 MIMO 发射功率效率的直观论述基本一致. 由此可看到, 基站一侧所需的总发射功率与系统的频谱效率成正比, 与归一化信噪比成反比; 且当基站一侧的天线数充分多时, 增加用户一侧总的天线数将有利于基站总发射功率的减少.

综上可知,未来移动通信系统的频谱效率和功率效率尚不存在理论上的限制.通过构建更多天线的 MU-MIMO 蜂窝系统或无蜂窝系统,可以获得比 5G 更为优越的频谱和功率效率性能,所需要付出的代价是系统部署成本的相应增加.

4.2 更高的可靠性与更低的时延

5G NR 首次将高可靠与低时延 (uRLLC) 技术指标引入公众移动通信系统. 一方面, 为将无线传输时延降低至 1 ms 以下, 引入了更短的时隙结构、免许可接入认证, 以及移动边缘计算等技术; 另一方面, 为将无线数据包差错概率降低至 10⁻⁶ 以下, 又引入了时域或频域重复发送、多点发送 (multi-TRP) 或多连接等^[8]. 总体看来, 由于未能从信息论的角度寻求传输时延、可靠性、传输速率及天线数的最优平衡, 在性能和资源的有效利用等方面仍有较大的提升空间^[11].

以下以 Shannon 信息论及其 MIMO 扩展形式为基础,从理论的角度探讨未来 5G 演进及 6G 发展过程中进一步提升 uRLLC 性能指标的途径与技术潜能.

(1) 厘清 5G NR 高可靠无线传输相关理论基础, 以等同性的单输入多输出 (SIMO) 数学模型支撑 5G NR 高可靠无线传输的进一步完善与发展. 现有 5G NR 所采用的时域或频域重复发送、多点发送 或多连接等技术, 均是利用信道的多重分集特性来提升无线传输的可靠性, 所付出的代价是无线资源 消耗或硬件设备的成倍增加; 若采用时间域重复发送, 其传输时延还将成倍增加. 在现代移动通信系 统设计中, 通常假设不同数据块的信道状态信息之间是相互统计独立的. 在此条件下, 重复发送、多点 或多连接发送可以统一用 SIMO 模型加以描述^[35], 其分集度为同一数据块的冗余发送次数 *d*, 其数据 块的无线传输差错概率与 *R*·SNR^{-d} 成正比, *R* 为编码速率, SNR 为单次 (或单个) 数据发送的信噪 比⁶⁾. 因此, SIMO 信道的理论结果对于 5G NR 的高可靠无线传输具有特殊的指导意义. 此处列举一个简单的例子, 对于差错概率为 10⁻² 量级的无冗余数据发送, 若 SNR = 10 dB, 采用 3 次 (或 3 个) 冗余发送便可将其差错概率降低为 10⁻⁶ 量级. 采用更多次数的冗余发送, 将进一步提升其传输可靠性.

(2) 寻求 MIMO 信道的分集度与空间复用增益折中, 探究 5G 向 6G 演进发展中更具应用潜能的 uRLLC 技术途径. 大规模天线阵列已成为 5G 乃至未来 6G 系统所必备的核心技术, 这为寻求系统传 输速率和传输可靠性之间的性能折中带来极大的便利. 以大规模 MIMO 下行链路为例说明. 假设发送 端和接收端分别配置 N_t 和 N_r 个天线, $N_t \gg N_r$, 且所涉及的信道 **H** 为独立同分布. 从 3.1 和 3.3 小 节可知, 该 MIMO 信道的空间自由度为 N_r , 最大传输速率为 $B \cdot N_r \log_2(1 + \text{SNR})$, 系统的分集度为 $N_t - N_r + 1$, 传输差错率正比于 $\text{SNR}^{-(N_t - N_r + 1)}$. 与为寻求更高的传输可靠性, 考虑选择更低的传输速 率, 即采用 $r < N_r$ 个独立的数据流进行发送; 这时, MIMO 信道的无线传输速率为 $B \cdot r \log_2(1 + \text{SNR})$, 分集度为 $(N_t - r + 1)(N_r - r + 1)$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 传输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 行输差错率正比于 $\text{SNR}^{-(N_t - r + 1)(N_r - r + 1)}$, 代表 M (名) 无式传输系统构架, 仅需调

⁶⁾ 这里简单地假设, 每次冗余发送的 SNR 是等同的. 更为一般情形下的性能分析, 见文献 [35].

节发送端的数据流及预编码,便可方便地在传输速率与可靠性之间寻求灵活的折中,且可以直接推广 至 MU-MIMO.从理论上来说,当用户侧固定其配置,且基站侧大规模天线阵元数足够多时,可以在不 增加各个数据流传输时延的条件下,获得充分高的无线传输可靠性.

(3) 探索无蜂窝等新型网络架构在 uRLLC 无线传输方面的潜在优势,并将点到点的 5G NR uRLLC 无线传输拓展至未来 6G 无线网络覆盖的 QoS 保障研究. 如同 4.1 小节所述,无蜂窝构架具有 P2MP 或 MP2MP 联合处理能力,多连接 (SIMO) 和点到点 MIMO 均是其特例,且可实现多小区范围内的时间、频率和空间自由度全动态调配,因而在跨小区 uRLLC 无线传输设计和性能折中方面具有独特的技术优势.此外,还可针对无蜂窝网络构架,综合考量多小区无线覆盖的非均匀性及相应的 QoS 保障能力.例如,文献 [19] 第 4 章给出了无蜂窝网络构架典型场景下的区域中断容量 (area outage capacity),用于衡量整个覆盖范围内无线传输速率小于指定门限值的概率.在此基础上,通过进一步引入传输时延 (分块长度)、差错概率等约束指标,可以计算覆盖范围内 QoS 指标得到保障的概率,从而设计出在无线网络层面 uRLLC 性能可预见的 6G 移动通信系统.

(4) 探索更为一般意义上的无线传输性能折中,包括传输时延、可靠性、传输速率和所需最小天线 数等,以寻求 6G 系统 uRRLC 无线传输性能的进一步提升和部署成本的有效控制. 云端控制的智能 制造等工业应用对未来移动通信系统提出了更短时延 (如 0.1~0.3 ms) 和更高可靠性 (如 10⁻⁹~10⁻¹²) 的要求. 考虑到传输时延与可靠性在系统设计上是一对相互冲突的技术指标,依靠经典 Shannon 信息 论寻求两者的最优折中已难以满足上述需求,需要引入多天线、并在 MIMO 无线传输性能折中层面 寻求新的突破. 如同 3.3 小节指出, MIMO 无线传输具备天然的多天线分集能力,对于给定的码流速 率和传输分块长度,可以提供更高的传输可靠性;换言之,更短的时延⁷⁾所带来的可靠性的下降,将可 以由增加天线数、特别是基站侧的天线数来弥补. 因此,未来 6G 系统的 uRLLC 最优设计应建立传输 时延、可靠性、传输速率和所需最小天线数 (即射频通道最低成本) 之间最优折中理论框架的基础上. 该方面的研究尽管较为有限,但已经取得了可喜的进展;例如,文献 [18] 给出了 MIMO 信道条件下的 最大可获取传输速率、分块长度与差错概率之间的一般性关系;文献 [36] 从 Gallager 指数的角度出 发,研究了不同天线条件下分块长度与传输可靠性之间的关系;文献 [37] 给出了传输速率和差错概率 约束条件下,最小天线数的计算求解方法等. 对于上述 4 个技术指标之间的更为一般性的显性关系描述,则有待于今后更为深入的研究.

4.3 更高的频段

2019 年世界无线电大会已为 5G 乃至更为长远的移动通信发展明确了可用的毫米波与太赫兹频 段^[11],标志着移动通信系统开始步入更高的频段,可供利用的频率带宽将大幅度增加.这一发展趋势 看似与 Shannon 信息论及其 MIMO 扩展形式并无直接关联,实则不然.这是由于随着载波频率的不 断升高,毫米波与太赫兹电波传播特性将愈来愈接近"准光学"特性,基于富散射 (rich scattering) 的 经典统计信道模型将难以适用,因而需要有针对性地重新分析评估与之相关的 MIMO 信道容量及影响因素.以下将会看到,由于信道模型的显著变化,影响毫米波与太赫兹 MIMO 信道容量的因素也将 发生根本性的改变.

为简单起见,下面以等间隔阵元分布的线性均匀阵 (ULA) 为例进行论述. 假设点对点无线链路收 发两端分别配置天线阵元数为 N_r 和 N_t 的 ULA, 且收发两端分别用导向向量 (steering vector) u_r 和

⁷⁾ 对于分块长度 1 ≪ n ≤ N_t + N_r - 1 情形, 分集增益和复用增益将发生退化^[27], 但基本结论仍然成立.

 u_t 对 ULA 进行波束成形处理,

$$\boldsymbol{u}_{r}\left(\boldsymbol{\theta}_{r}\right) = \frac{1}{\sqrt{N_{r}}} \left[1, \frac{\exp\left(-j2\pi\sin\left(\boldsymbol{\theta}_{r}\right)\right)}{\upsilon}, \dots, \frac{\exp\left(-j2\pi N_{r}\sin\left(\boldsymbol{\theta}_{r}\right)\right)}{\upsilon}\right]^{\mathrm{T}}, \\ \boldsymbol{u}_{t}\left(\boldsymbol{\theta}_{t}\right) = \frac{1}{\sqrt{N_{t}}} \left[1, \frac{\exp\left(-j2\pi\sin\left(\boldsymbol{\theta}_{t}\right)\right)}{\upsilon}, \dots, \frac{\exp\left(-j2\pi N_{t}\sin\left(\boldsymbol{\theta}_{t}\right)\right)}{\upsilon}\right]^{\mathrm{T}},$$

其中 θ_r 和 θ_t 分别为天线阵的到达角 (AoA) 和离开角 (AoD), v 为载波波长. 当信道仅存在直达传播 路径 (LoS) 时, 则所形成的 $N_t \times N_r$ 维等效信道为 ^[38]

$$\boldsymbol{H} = \sqrt{N_r N_t} \,\beta \,\boldsymbol{u}_r \left(\theta_r\right) \boldsymbol{u}_t^{\dagger} \left(\theta_t\right), \tag{19}$$

其中 β 为信道增益因子.显然式 (19) 给出的信道 **H** 秩为 1. 根据 3.1 小节的讨论,该 MIMO 信道的 空间自由度为 1,故仅能发送和接收单个数据流.

当信道存在 *i* = 1, 2, ..., *L* 个反射路径时, 式 (19) 演变为以下波束域信道模型或称为有限维 MIMO 信道模型^[38]

$$\boldsymbol{H} = \sqrt{\frac{N_r N_t}{L}} \sum_{i=1}^{L} \beta_i \boldsymbol{u}_r \left(\theta_{r,i}\right) \boldsymbol{u}_t^{\dagger} \left(\theta_{t,i}\right) \triangleq \boldsymbol{U}_r \boldsymbol{\Lambda} \boldsymbol{U}_t^{\dagger}, \qquad (20)$$

其中 $U_r \triangleq [u_r(\theta_{r,1}) \cdots u_r(\theta_{r,L})], U_t \triangleq [u_t(\theta_{t,1}) \cdots u_t(\theta_{t,L})], \Lambda 为 \sqrt{N_r N_t / L \beta_i}$ 组成的对角阵. 不失一般性, 假设 β_i 为复 Gauss 变量. 文献 [38] 指出, 当 L 充分大时, 式 (20) 趋于第 3 节所讨论的独立同分布模型. 因此, 式 (20) 可以统一地描述 "准光学"无线信道和经典的富散射信道.

应注意到,有限维信道容量分析与导向向量 u_r 和 u_t 的选择密切相关,在一般情形下较为复 杂^[39].借鉴文献 [38] 一般性信道投影方法,将 u_r 和 u_t 限定为 Fourier 变换正交基的子集,再将 式 (20) 代入 3.1 小节给出的式 (15),并假设发射端天线等功率分配,可容易得到关于式 (20) 的遍历性 容量的紧致上界如下⁸⁾:

$$E\left\{C_{\text{MIMO}}\left(\boldsymbol{H}\right)\right\} \leqslant B \cdot L \cdot \log_2\left(1 + \text{SNR}\frac{N_r E\left\{|\boldsymbol{\beta}|^2\right\}}{L}\right),\tag{21}$$

其中 $E\{|\beta|^2\}$ 可由 χ^2 分布计算出. 上式表明: (1) 增加接收端的天线数将有助于提高有限维 MIMO 信 道的容量; (2) 由于模型的原因, 发送天线数 N_t 的影响被体现在为信道增益因子 β 中, 它与 N_t 成正 比; (3) 对于 L 个反射路径所带来的影响, 一方面由此增加了信道的空间自由度, 而另一方面又降低了 系统的有效信噪比. 为综合考虑上述影响, 将式 (21) 带入式 (16), 则容易证明, 上述有限维 MIMO 信 道的空间复用增益等于反射路径数 L, 而非第 3 节给出的天线配置数 min.{ N_t, N_r }. 更为一般条件下 的论述可参见文献 [39].

最后考虑将上述点到点无线链路推广至点对多点 (对应于 MU-MIMO) 情形. 对于仅存在直达经的 LoS 信道,若发送天线数为 N_t ,用户数为 $M, M \leq N_t$,则可以通过 M 个相互正交的向导向量 $u_t(\theta_{t,i}), i = 1, \ldots, M$,形成指向 M 个用户的赋形波束^[32].因用户之间不存在相互干扰,该问题是一个较为典型的和速率最优功率指配问题,可以借鉴 3.1 小节给出的方法求解.对于存在反射路径的有限维信道模型,依然可以采用上述方法,但需要考虑用户之间的相互干扰. 若 M 个用户的联合信道信息已知,则可将式 (20) 直接推广至多用户情形.这时的信道容量对应于多用户和速率,其信号处理与一般的 MU-MIMO 并无差异,但信道容量的闭合解析将变得十分复杂,有待于学术界进一步深入研究.

⁸⁾ 这里隐含地假设 L ≤ min.{N_t, N_r}, 对于毫米波和太赫兹信道, 反射路径较少, 该假设一般均成立.

5 结束语

毋容置疑, Shannon 及其同事们在 20 世纪 50 年代前后所发展的经典信息论框架奠定了 70 余年 来的通信理论与技术基础. 但应客观地看到, Shannon 无法预见移动通信已经从最初的 SISO 系统发 展至目前主流的 MIMO 系统. 如果未考虑研究对象的变化, 而将 Shannon 最初的信息论简单地套用 到当代移动通信系统中, 便会产生似是而非的结论.

经典 Shannon 信息论研究时间 – 频率二维自由度的最优利用问题, 而其 MIMO 扩展形式是研究 "时间 – 频率 – 空间"三维自由度的最优利用问题.显然, 由于时间和频率属于不可再生的资源, 经典 Shannon 信息论所涉及的 SISO 信道存在某种可能的极限.与之不同的是, MIMO 信道的空间自由度 可以通过增加收发两端的天线配置而人为地增加, 在某种程度上属于"可再生、可延展"的无线资源. 这意味着未来移动通信系统在特定的理论假设条件下尚不存在所谓的性能极限.

经典 Shannon 信息论除了给出 SISO 信道容量外,还给出了最大可获取速率与分块长度及差错概率之间的折中理论框架,后者对发展 uRLLC 无线传输技术极为重要,但需拓展至 MIMO 情形方能适用于未来移动通信系统.对此,学术界应给予充分的重视.

从本文的论述中,还可以总结出有益于未来 6G 发展的若干规律性的认识: (1) 天线数的增加总体 上有利于无线传输性能的提高,包括更高的频谱效率、更低的发射功率、更高的可靠性等; (2) 天线数 的增加虽然并不能直接降低传输时延,但可增加可获取速率与可靠性,并由此缓解"传输速率(容量) - 分块长度(时延) - 差错概率(可靠性)"三者之间的相互制约; (3) 通过引入类似于无蜂窝形式的网 络构架创新,将更加有利于 MIMO 信道空间自由度的"可再生、可延展"式构建与有效利用; (4) 6G 核心技术指标的进一步提升,无不伴随着系统部署成本和复杂度的上升,因而寻求系统性能与实施成 本的有效平衡将是未来 6G 移动通信发展的主要挑战之一.

参考文献 –

- 1 Shannon C E. A mathematical theory of communication. Bell Syst Tech J, 1948, 27: 379–423
- 2 Shannon C E. Communication in the presence of noise. Proc IRE, 1949, 37: 10–21
- 3 Shannon C E. Probability of error for optimal codes in a Gaussian channel. Bell Syst Tech J, 1959, 38: 611-656
- 4 Gallager R G. Information Theory and Reliable Communication. Hoboken: Wiley, 1968
- 5 Foschini G J, Gans M J. On limits of wireless communications in a fading environment when using multiple antennas. Wirel Personal Commun, 1998, 6: 311–335
- 6 Shiu D S, Foschini G J, Gans M J, et al. Fading correlation and its effect on the capacity of multielement antenna systems. IEEE Trans Commun, 2000, 48: 502–513
- 7 Telatar E. Capacity of multi-antenna Gaussian channels. Eur Trans Telecomm, 1999, 10: 585–595
- 8 3GPP. Study on physical layer enhancements for NR ultra-reliable and low latency case. TR 38.824, 2019. https:// www.3gpp.org/
- 9 Bi Q. Ten trends in the cellular industry and an outlook on 6G. IEEE Commun Mag, 2019, 57: 31-36
- 10 Zhang L, Liang Y C, Niyato D. 6G visions: mobile ultra-broadband, super internet-of-things, and artificial intelligence. China Commun, 2019, 16: 1–14
- 11 You X H, Yin H, Wu H Q. 6G and wide-area IoT. Chinese J Int Things, 2020, 4: 3-11 [尤肖虎, 尹浩, 邬贺铨. 6G 与 广域物联网. 物联网学报, 2020, 4: 3-11]
- 12 Pinsker M S. On the complexity of decoding. Problemy Peredachi Inform, 1965, 1: 84–86

- 13 Arikan E. Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Trans Inform Theory, 2009, 55: 3051–3073
- 14 Arikan E. On the origin of polar coding. IEEE J Sel Areas Commun, 2016, 34: 209–223
- 15 Zhou H Y, Zhang C, Song W Q, et al. Segmented CRC-aided SC list polar decoding. In: Proceedings of the 83rd Vehicular Technology Conference, 2016
- 16 Zhang C, Zhou H Y, You X H. Means for segmented CRC-aided SC list polar decoding. CN Patent, CN 105337696 A, 2016-02-17 [张川, 周华羿, 尤肖虎. 基于分段 CRC 校验的极化解码方法. 中国发明专利, CN 105337696 A, 2016-02-17]
- 17 Polyanskiy Y, Poor H V, Verdú S. Channel coding rate in the finite blocklength regime. IEEE Trans Inform Theory, 2010, 56: 2307–2359
- 18 Potter C, Kosbar K, Panagos A. On achievable rates for MIMO systems with imperfect channel state information in the finite length regime. IEEE Trans Commun, 2013, 61: 2772–2781
- 19 You X H, Wang D M, Wang J Z. Distributed MIMO and Cell-Free Mobile Communication. Beijing: Science Press, 2019 [尤肖虎, 王东明, 王江舟. 分布式 MIMO 与无蜂窝移动通信. 北京: 科学出版社, 2019]
- 20 Gao X Q, You X H, Jiang B, et al. Unifying eigen-mode MIMO transmission. Sci China Ser F-Inf Sci, 2009, 52: 2269–2278
- You X H, Chen G A, Chen M, et al. Toward beyond 3G: the FuTURE project in China. IEEE Commun Mag, 2005,
 43: 70–75
- 22 Gao X Q, Jiang B, Li X, et al. Statistical eigenmode transmission over jointly correlated MIMO channels. IEEE Trans Inform Theory, 2009, 55: 3735–3750
- 23 Wang D M, Gao X Q, You X H. Low complexity turbo receiver for multi-user STBC block transmission systems. IEEE Trans Wirel Commun, 2006, 5: 2625–2632
- 24 Shin H, Win M Z. Gallager's exponent for MIMO channels: a reliability-rate tradeoff. IEEE Trans Commun, 2009, 57: 972–985
- 25 Ho P K M, Yar P K, Kam P Y. Cutoff rate of MIMO systems in Rayleigh fading channels with imperfect CSIR and finite frame error probability. IEEE Trans Veh Technol, 2009, 58: 3292–3300
- 26 Proakis J. Digital Communications. 4th ed. New York: McGraw-Hill, 2001
- 27 Zheng L Z, Tse D N C. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. IEEE Trans Inform Theory, 2003, 49: 1073–1096
- 28 Wallace J W, Jensen M A. Mutual coupling in MIMO wireless systems: a rigorous network theory analysis. IEEE Trans Wirel Commun, 2004, 3: 1317–1325
- 29 Murata K, Honma N. On potential channel capacity of massive MIMO array within small finite space. In: Proceedings of IEEE International Symposium on Antennas & Propagation, Florida, 2013. 868–869
- 30 Sayeed A M, Behdad N. Continuous aperture phased MIMO: a new architecture for optimum line-of-sight links.
 In: Proceedings of IEEE International Symposium on Antennas & Propagation, Monterey, 2011. 293–296
- 31 Pizzo A, Marzetta T L, Sanguinetti L. Spatial characterization of holographic MIMO channels. 2020. ArXiv: 1911.04853v1
- Ngo H Q, Ashikhmin A, Yang H, et al. Cell-free massive MIMO versus small cells. IEEE Trans Wirel Commun, 2017, 16: 1834–1850
- 33 Buzzi S, D'Andrea C. Cell-free massive MIMO: user-centric approach. IEEE Wirel Commun Lett, 2017, 6: 706–709
- 34 Rusek F, Persson D, Lau B K, et al. Scaling up MIMO: opportunities and challenges with very large arrays. IEEE

1392

Signal Process Mag, 2013, 30: 40–60 $\,$

- 35 Wolf A, Schulz P, Dorpinghaus M, et al. How reliable and capable is multi-connectivity? IEEE Trans Commun, 2019, 67: 1506–1520
- 36 Xue J, Ratnarajah T, Zhong C J, et al. Reliability analysis for large MIMO systems. IEEE Wirel Commun Lett, 2014,
 31: 553-556
- 37 Makki B, Svensson T, Coldrey M, et al. Finite block-length analysis of large-but-finite MIMO systems. IEEE Wirel Commun Lett, 2019, 8: 113–116
- 38 Brady J, Behdad N, Sayeed A M. Beamspace MIMO for millimeter-wave communications: system architecture, modeling, analysis, and measurements. IEEE Trans Antenn Propag, 2013, 61: 3814–3827
- 39 Yang X, Li X, Zhang S L, et al. On the ergodic capacity of mmWave systems under finite-dimensional channels. IEEE Trans Wirel Commun, 2019, 18: 5440–5453

Shannon theory and future 6G's technique potentials

Xiaohu YOU 1,2

National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China;
 Purple Mountain Laboratories for Network and Communication Security, Nanjing 211100, China

E-mail: xhyu@seu.edu.cn

Abstract From the perspective of Shannon theory and its extensions, this paper is devoted to evaluate the technique potentials of the future 6G mobile communication system. First, the classic Shannon theory framework, including the performance tradeoff between block length, data rate, and reliability, is summarized, and the limitations of its application in the contemporary mobile communication system are addressed. Second, the multiple-input-multiple-output (MIMO) extension of classic Shannon theory is described, which plays fundamental roles in the development of contemporary mobile communication systems. Moreover, because Shannon theory and its MIMO extension are nonconstructive in nature, two kinds of constructive capacity-approaching mechanisms, namely, channel polarization and eigen-mode wireless transmission, are also introduced. Furthermore, aiming at higher spectrum efficiency and power efficiency, higher reliability and lower latency, and higher frequency band, which are essential indicators of future 6G, the technique potentials are theoretically discussed from the perspective of Shannon theory framework. It reveals that by introducing more antennas together with the innovation of cellfree network architecture and by making an effective balance between block length, error probability, data rate, and the minimum number of antennas, future 6G technology still has great potential to be improved. However, a compromise between system performance and deployment cost must be made, and the special features of MIMO channels in higher frequency bands must be carefully utilized. Finally, several fundamental issues related to future 6G development are summarized.

Keywords Shannon theory, MIMO, 6G mobile communications, cell free, uRLLC



Xiaohu YOU received his M.S. and Ph.D. degrees in electrical engineering from Southeast University, Nanjing, China, in 1985 and 1988, respectively. Since 1990, he has been working with National Mobile Communications Research Laboratory at Southeast University, where he is currently the director of the Lab. From 1999 to 2002, he was the principal expert of the C3G Project. From 2001 to 2006, he was the principal

expert of the China National 863 Beyond 3G FuTURE Project. From 2013 to 2019, he was the principal investigator of China National 863 5G Project. Now he is the secretary general of the FuTURE Forum, vice chair of the China IMT-2020 Promotion Group, and vice chair of the China National Mega Project on New Generation Mobile Network.